Imaging the Universe in Three Dimensions: Astrophysics with Advanced Multi- Wavelength Imaging Devices

W. van Breugel, J. Bland-Hawthorn

November 18, 1999

U.S. Department of Energy

Lawrence
Livermore
National
Laboratory

DISCLAIMER

This document was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor the University of California nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or the University of California, and shall not be used for advertising or product endorsement purposes.

This work was performed under the auspices of the U. S. Department of Energy by the University of California, Lawrence Livermore National Laboratory under Contract No. W-7405-Eng-48.

This report has been reproduced directly from the best available copy.

Available electronically at <http://www.doc.gov/bridge>

Available for a processing fee to U.S. Department of Energy
And its contractors in paper from
U.S. Department of Energy
Office of Scientific and Technical Information
P.O. Box 62
Oak Ridge, TN 37831-0062
Telephone: (865) 576-8401
Facsimile: (865) 576-5728
E-mail: reports@adonis.osti.gov

Available for the sale to the public from
U.S. Department of Commerce
National Technical Information Service
5285 Port Royal Road
Springfield, VA 22161
Telephone: (800) 553-6847
Facsimile: (703) 605-6900
E-mail: orders@ntis.fedworld.gov
Online ordering: <http://www.ntis.gov/ordering.htm>

OR

Lawrence Livermore National Laboratory
Technical Information Department's Digital Library
<http://www.llnl.gov/tid/Library.html>

Table of Contents

Preface	xxx
Conference Organization	xxx
Conference Participants	xxx

Part 1. Keynote Speaker

Instrumentation and Astrophysics: How did we get to be so Lucky? . .	XX
<i>M. Harwit</i>	

Part 2. Galaxies and Cosmology

Section A. Imaging

Galaxy Evolution and Large Scale Structure (Review)	XX
<i>R. Ellis</i>	
3-D Imaging (Review)	XX
<i>J. Bland-Hawthorn</i>	
Critical Comparison of 3-D Imaging Approaches (Review)	XX
<i>C. L. Bennett</i>	
100 Million Galaxies in Technicolor	XX
<i>A. Szalay</i>	
Concept for a 8-m Telescope with a 3 Degree Field at f/1.25: The Dark Matter Telescope	XX
<i>R. Angel, M. Lesser, R. Sarlot, E. Dunham</i>	
High-Redshift Gravitationally Lensed Galaxies and Tunable Filter Imag- ing	XX
<i>P. C. Hewett, S. J. Warren, J. P. Willis, J. Bland-Hawthorn & G. F. Lewis</i>	
The Power of Differential Observing	XX
<i>C. G. Tinney</i>	
A Four-Color CCD Imaging System for the REACT Network (Poster) .	XX
<i>A. H. Diercks, F. A. Harrison, S. R. Kulkarni & C. W. Stubbs</i>	

Section B. Spectroscopy

Anglo-Australian Tunable Littrow Articulated Spectrograph (Review)	XX
<i>K. Taylor, G. Robertson & S. Barden</i>	
Faint Object 2-D Spectroscopy: Promise and Limitations	XX
<i>M. M. Roth, T. Becker & J. Schmoll</i>	
Probing the Metallicity Evolution of Disk Galaxies via Imaging Spectroscopy	XX
<i>D. .R. Garnett</i>	
Science & Surveys with Integral Field Spectrographs (Review)	XX
<i>R. L. Davies</i>	
Fabry-Perot Observations of Spiral Galaxies: The Impact on Mass Distribution (Poster)	XX
<i>S. Blais-Ouellette, C. Carignan & P. Amram</i>	
New Improved Photometric Redshifts of Galaxies in the HDF (Poster)	XX
<i>H. Furusawa, K. Shimasaku, M. Doi & S. Okamura</i>	
Tunable Filter and Broadband-Selected Redshift Surveys (Poster)	XX
<i>D. H. Jones</i>	
SAURON: Integral-field Spectroscopy of Galaxies (Poster)	XX
<i>B. W. Miller, M. Bureau, E. Verolme, P. T. de Zeeuw, R. Bacon, Y. Copin, E. Emsellem, R. L. Davies, R. F. Peletier, J. R. Allington-Smith, C. M. Carollo & G. Monnet</i>	

Part 3. Obscured Systems**Section A. Infrared-Wavelength 3-D**

Multiple Deployable Integral Field Units in the Near-IR (Review)	XX
<i>G. S. Wright</i>	
Supermassive Black Hole Searches with 3-D Spectroscopy	XX
<i>R. Bacon, E. Emsellem, Y. Copin & G. Monnet</i>	
The Nature of Obscured Nuclear Starbursts: M 82	XX
<i>N. M. Förster-Schreiber</i>	
BEAR Imaging FTS: High Resolution Spectroscopy in Infrared Emission Lines	XX
<i>J. P. Maillard</i>	
CIRPASS: A Near-IR Integral Field Spectrograph	XX
<i>I. Parry, A. J. Dean, R. S. Ellis, D. King, C. D. Mackay, R. G. McMahon, S. R. Medlen, J. M. Pritchard & A. N. Ramaprakash</i>	
Kinematics of Ionized Gas in the Central Parsec of the Galaxy from High-Resolution Spectroscopy of the Brackett- γ Line (Poster)	XX
<i>M. Morris & J. P. Maillard</i>	

The New Palomar Integral Field Spectrograph	XX
<i>T. Murphy, Jr., K. Matthews & B. T. Soifer</i>	
Near-IR Integral-Field Spectroscopy with Adaptive Optics	XX
<i>N. Thatte, S. Anders, F. Eisenhauer, M. Tecza, S. Mengel, A. Eckart, R. Genzel, G. Monnet & D. Bonaccini</i>	
NGC 4388: Imaging and Kinematics of the Nuclear Region through Inte- gral Field Spectroscopy (Poster)	XX
<i>S. Ciroi, V. Afanasiev, S. Dodonov, M. Radovich, P. Rafanelli, G. Richter & S. Tempurin</i>	
Integral Field Spectroscopy of Ultraluminous Infrared Galaxies (Poster)	XX
<i>L. Colina, S. Arribas & K. Borne</i>	
Imaging Spectropolarimetry: The Nature of Hyperluminous Infrared Galax- ies (Poster)	XX
<i>D. C. Hines</i>	
3-D Spectroscopic Observations of the IRLG Mrk266 (Poster)	XX
<i>T. Ishigaki, M. Yoshida, K. Aoki H. Ohtani, H. Sugai, T. Hayashi, S. Ozaki, T. Hattori & M. Ishii</i>	
Merging Signatures in the Core of Mkn938 (Poster)	XX
<i>P. Rafanelli, A. Rifatto, V. Afanasiev, S. Dodonov, K. Birkle, P. Böhm, C. Cannavacciuolo, G. Richter & A. and J. Vennik</i>	
 Section B. Millimeter-Wavelength 3-D	
Imaging Spectroscopy at mm-Wavelengths (Review)	XX
<i>N. Z. Scoville & A. I. Sargent</i>	
Submm Continuum Surveys for Obscured Galaxies (Review)	XX
<i>I. Smail, R. Ivison, A. Blain, J. P. Kneib & F. Owen</i>	

Part 4. Outflows from Active Galaxies and Young Stars

Section A. Active Galaxies

Ionized Outflows in 3-D: Insights from Herbig-Haro Objects & Applica- tions to Nearby AGN (Review)	XX
<i>G. Cecil</i>	
Fabry-Perot Imaging Spectroscopy of Starburst and AGN Winds	XX
<i>S. Veilleux</i>	
Deep TTF Imaging of Powerful Radio Galaxies	XX
<i>C. Tadhunter, M. Villar-Martin & R. Morganti</i>	
3-D Observations of Nearby Active Galaxies with the IFU TIGER	XX
<i>P. Ferruit, E. Pécontal & A. S. Wilson</i>	

Integral Field Spectroscopy with Optical Fibers: Studies of the Circum-nuclear Regions of Nearby Active Galactic Nuclei (Review)	XX
<i>S. Arribas & E. Mediavilla</i>	
Infrared 3-D Observations of Nearby Active Galaxies	XX
<i>R. Maiolino, N. Thatte, A. Alonso-Herrero, D. Lutz & A. Marconi</i>	
The Kyoto Multimode 3-D Spectrographs	XX
<i>H. Ohtani, T. Ishigaki, T. Hayashi, S. Ozaki, T. Hattori, H. Sugai & M. Ishii</i>	
Integral Field Spectroscopy with Gemini and Space Telescopes	XX
<i>J. Allington-Smith, R. Content, R. Haynes & D. Robertson</i>	
Integral Field Spectroscopy of NGC 2992 (Poster)	XX
<i>B. García-Lorenzo, S. Arribas & E. Mediavilla</i>	
Tridimensional Spectroscopy of the Interacting System NGC7592 (Poster)	XX
<i>T. G. Hattori, M. Yoshida, H. Ohtani, T. Ishigaki, H. Sugai, T. Hayashi, S. Ozaki, & M. Ishii</i>	
First Results with a Wide-Field Near-Infrared Integral Field Unit (Poster)	XX
<i>T. M. Herbst</i>	
The Disturbed Structure of the Nuclear Region of Mkn298 (Poster)	XX
<i>A. Rifatto, P. Rafanelli, V. Afanasiev, S. Dodonov, K. Birkle, P. Böhm, T. Boller, M. Radovich, G. Richter & J. Vennik</i>	
3-D Spectroscopy of High-z Radio Galaxies with the IFUs TIGER and OASIS (Poster)	XX
<i>B. Rocca-Volmerange, G. Adam, P. Ferruit & E. Moy</i>	
Kinematics of the Ionized Gas in Irregular Galaxies: IC 1613 and NGC 4449 (Poster)	XX
<i>M. Valdez-Gutiérrez, M. Rosado, L. Georgiev, C. Lacey, J. Borissova, R. Kurtev</i>	
Section B. Young Stars	
Molecular Hydrogen Imaging of Star-Forming Regions	XX
<i>S. K. Ramsay-Howat, A. Chrysostomou, M. G. Burton & P. W. J. L. Brand</i>	
Probing the Origin of Mass-Loss in Young T-Tauri Stars	XX
<i>C. Dougados, C. Lavalley & S. Cabrit</i>	
 Part 5. Quasars	
QSO's and the 3-D Universe (Review)	XX
<i>B. Boyle</i>	
Recent Spectroscopy of the 3C 48 Host Galaxy and a Simple Image Slicer Design	XX
<i>A. Stockton & G. Canalizo</i>	

Rugate Filters: Quasars Beyond $z \sim 7$	XX
<i>S. Cianci, J. Bland-Hawthorn & J. O'Byrne</i>	
Have Most High-Redshift Quasars Been Overlooked?	XX
<i>C. Wolf, K. Meisenheimer, H. -J. Röser, S. V. W. Beckwith, R. Focken-brock, H. Hippelein, B. von Kuhlmann, S. Phleps & E. Thommes</i>	
A Very Extended Ionized Nebula around the Quasar MR2251-178 (Poster)	XX
<i>P. L. Shopbell, S. Veilleux & J. Bland-Hawthorn</i>	
Tunable Filter Imaging of a Quasar Field at $z = 0.9$ (Poster)	XX
<i>J. C. Baker</i>	

Part 6. 3-D in Space

Section A. The Next Generation Space Telescope

The NGST Science Program (Review)	XX
<i>H. S. Stockman & J. C. Mather</i>	
An Imaging Fourier Transform Spectrometer for the Next Generation Space Telescope	XX
<i>J. R. Graham</i>	
IFMOS: Integral Field Multi Object Spectrograph for the NGST	XX
<i>O. Le Fèvre, E. Prieto, J. Allington-Smith, R. Bacon, R. Content, S. Cristiani, R. Davies, B. Delabre, R. Ellis, G. Monnet, E. Pécontal, W. Posselt, N. Thatte, P. T. de Zeeuw, P. van der Werf</i>	
The Tunable Filter Program for NGST	XX
<i>S. Satyapal, M. A. Greenhouse, R. Barclay, D. Amato, B. Arritt, R. Barry, C. Holt, J. Kuhn, T. Higelman, N. Fonneland & L. Lesnya</i>	
A Multi-Object Spectrometer using Micro Mirror Arrays	XX
<i>J. W. MacKenty & M. Stiavelli</i>	
Programmable 2-Dimensional Microshutter Arrays (Poster)	XX
<i>S. H. Moseley, R. K. Fettig, A. S. Kutyrev, C. W. Bowers, R. A. Kimble, J. Orloff & B. E. Woodgate</i>	
Concept of an Imaging FTS and Broad-band Imager for NGST (Poster)	XX
<i>J. P. Maillard & J. C. Renault</i>	
Beyond the NGST (Next Generation Space Telescope)	XX
<i>M. C. Abrams, D. D. Gregory & R. Hookman</i>	

Section B. SOFIA and SIRTf

LISA: A Low Resolution MIR/FIR Imaging Spectrograph for SOFIA	XX
<i>A. Krabbe & J. Wolf</i>	

SPICA: IR-Camera for SOFIA (Poster)	XX
<i>J. Wolf, A. Krabbe & D. Rabanus</i>	
3-D Spectroscopy with SIRTIF (Poster)	XX
<i>L. Armus</i>	

Part 7. Special Techniques and Technologies

Section A. Advanced Detectors

CCD Inovations	XX
<i>J. L. Tonry & G. A. Luppino</i>	
The Astronomical Potential of Optical STJs	XX
<i>M. A. C. Perryman & A. Peacock</i>	
5,000 by 5,000 Spatial by 15,000 Spectral Resolution Elements: First Astronomical Observations with a Novel 3-D Detector	XX
<i>C. U. Keller</i>	
A Novel Interferometer Spectrometer for Sensitive Stellar Radial Velocimetry	XX
<i>D. J. Erskine & J. Ge</i>	

Section B. Image Slicers

Image Slicing with Infrared Fibers	XX
<i>J. E. Larkin, A. Quirrenbach & J. R. Graham</i>	
Advanced Image Slicers from the Laboratory to NGST	XX
<i>R. Content</i>	
Focussing Image Slicers: Refractive and Reflective (Poster)	XX
<i>E. H. Richardson, A. Moore, T. Tillemann & D. Crampton</i>	
3-D Spectroscopy with Fibers: Extraction Methods and their Limitations (Poster)	XX
<i>T. Becker, M. M. Roth, J. Schmoll</i>	
SIMON: The Infrared Multi-Object Spectrometer of the Observatoire du Mont Mégantic (Poster)	XX
<i>R. Doyon, D. Nadeau, and Ph. Vallée</i>	

Section C. Grisms and Gratings

Tunable Gratings: Imaging the Universe in 3-D with Volume-Phase Holographic Gratings (Review)	XX
<i>S. C. Barden, J. B. Williams, J. A. Arns & W. S. Colburn</i>	

Development of Grisms and Immersion Gratings for the Spectrographs of the Subaru Telescope (Poster)	XX
<i>N. Ebizuka, H. Kobayashi and Y. Hirahara, M. Wakaki, K. Kawaguchi, T. Sasaki & M. Iye</i>	
Adaptive Optics High Resolution Spectroscopy: Present Status and Future Direction (Poster)	XX
<i>J. Ge, D. Ciarlo, P. Kuzmenko, C. Alcock, B. Macintosh, K. Cook, C. Max, R. Angel, N. Woolf, M. Lloyd-Hart & J. Najita</i>	
Slit Mask Exchange in a Cryogenic IR Multi-Object Spectrograph (Poster)	XX
<i>P. R. Hastings & G. S. Wright</i>	
An Image Slicer for the UKIRT Imager Spectrometer (Poster)	XX
<i>M. Wells, P. Hastings & S. Ramsay-Howat</i>	
PMAS: An Integral Field Spectrophotometer for the Calar Alto Observatory 3.5m Telescope (Poster)	XX
<i>M. M. Roth, W. Altmann, S. -M. Bauer, T. Becker, F. Dionies, T. Fechner, T. Hahn, U. Laux, E. Popow, J. Schmoll & D. Wolter</i>	
Section D. Filters	
An OH Airglow Supression Spectrograph with Multi-Object Feeder	XX
<i>T. Maihara & F. Iwamuro</i>	
Recent Developments in Fabry-Perot Interferometers	XX
<i>K. A. R. B. Pietraszewski</i>	
OSIRIS: A Tunable Filter Spectrograph for the GTC 10m Telescope (Poster)	XX
<i>J. Cepa, J. Bland-Hawthorn, J. J. González & the OSIRIS Consortium</i>	
Multi-notch and Multi-band Interference Coatings (Poster)	XX
<i>S. Cianci, J. Bland-Hawthorn & J. O'Byrne</i>	
The GSFC Lyot Filter for Night Time Astronomy (Poster)	XX
<i>P. Palunas, B. E. Woodgate, A. C. Danks, D. A. Klimesmith & R. B. Dunn</i>	
Part 8. Conference Summary	
Conference Summary	XX
<i>B. E. Woodgate</i>	
Author index	xxx

THE UNIVERSITY OF
THE STATE OF NEW YORK
IN SENATE

PREFACE

These proceedings arose from an international conference – “*Imaging the Universe in Three Dimensions: Astrophysics with Advanced Multi-Wavelength Imaging Devices*” – which was held under the auspices of the Lawrence Livermore National Laboratory in Walnut Creek, California, during March 29 – April 1, 1999.

Modern physics and astrophysics share a long and complex history founded on the discoveries of the early astronomers. It is difficult to pinpoint a time when observations of the night sky became a science. Some historians cite the ancient timekeepers or the not so Ancient Greeks (ca. 600–300 B.C). Remarkably, the origins of optical technology date back to remote antiquity: a mirror in excellent condition was unearthed near the pyramid of Sesostri II (ca. 1900 B.C.). It is plausible that the first optical observations date to Egypt during the 1st millenium. The great physicist Alhazen (ca. 965–1039) wrote the treatise *Treasury of Optics* which was to lay the foundations for later developments in Europe.

Six hundred years on, a refracting telescope patented in The Hague by Hans Lippershey inspired Galileo Galilei to build his own. As all physicists and astronomers know today, the night time observations of Galileo, Kepler and others were to profoundly influence the development of physical laws.

But Freeman Dyson (1999; *Nature*, 400, 27) reminds us that ‘modern’ science really began with James Bradley who performed the first high-precision measurements in 1729. A deep understanding of experimental technique allowed him to achieve astrometric measurements with 6-figure accuracy, orders of magnitude ahead of what had gone before. Bradley even recorded temperature and barometric pressure whenever he made an observation. This lead to a slew of discoveries, most notably, the first direct demonstration that the Copernican view was correct. He went on to discover aberration, i.e., the displacement of an image of a star on the sky due to the speed of the Earth in its orbit around the Sun. Bradley’s value for the speed of light *is within one percent of the modern value*. A hundred and fifty years later, Michelson & Morley attempted to measure the second-order aberration of light due to the Earth’s motion, and their null result paved the way to General Relativity.

A clear lesson – a recurrent theme in Harwit’s book ‘*Cosmic Discovery*’ – is that science benefits profoundly from the few individuals who devote their lives to improving and perfecting technical apparatus. This book pays homage to the great experimentalists, both past and present.

At the close of the 2nd millenium, we are witnessing yet another technological leap with vastly improved sensitivities provided by modern telescopes, large-area detectors, and routine adaptive optics correction. The primary science drivers of our day always emphasize the importance of surveying large numbers of sources. During the last decade, this has lead to the development of ‘3-dimensional’ instruments.

There are numerous examples of 3-D instruments which provide simultaneous spatial and spectral coverage over large fields and wavelength ranges (e.g., 2 spatial, 1 spectral). These include multi-beam imaging cameras and spectrographs, integral field spectrographs, tunable (e.g., Lyot, acousto-optic, liq-

uid crystal) filters, Fabry-Perots, and Fourier transform spectrographs. New 'enabling' technologies that are already extending the capabilities of these instruments include adaptive optics, micro-mirrors, staircases (image slicers), 3-D photon detectors, superconducting tunnel junction devices, sky suppression techniques, tunable gratings, and so on. Future spectrographs will incorporate arrays of fore optics (e.g., multiple integral field units), which can be deployed over a wide field, or adaptable feeds, which allow arbitrary formatting of the spectral-spatial dimensions at the detector.

Depending on their spatial- and spectral-multiplexing capabilities and resolution, 3-D instruments will revolutionize the way astronomy is being conducted. Broad-band, multi-color, wide-field cameras are already used for deep surveys of many types of objects such as distant galaxies, quasars, galaxy clusters, and for dark matter searches; fiber, multi-slit, and integral field spectrographs provide the spectroscopic capability to obtain large numbers of spectra of these objects simultaneously; Fabry-Perots and Fourier transform spectrographs provide high spectral resolution needed for detailed kinematic and emission-line mapping studies. Tunable filters are a particularly exciting recent development in that an arbitrary bandpass can be chosen at an arbitrary wavelength, allowing for cosmological and imaging studies of emission lines over a range of redshifts. Remote sensing 3-D instruments may be useful for some astronomical applications, including studies of the Sun and the Solar System, and the search for planets near other stars.

With the many new instruments currently under design for major ground- and space-based telescopes, we concluded that it was timely to bring people who are designing and building these instruments together with astronomers to help the cross fertilization of ideas, techniques and 3-D data analysis tools. To forge a cohesive science case that covers as much of the new developments in 3-D instrumentation as possible, over a large wavelength range (UV - optical - infrared - sub-mm wavelengths), we limited the scientific focus primarily to the formation and evolution of galaxies and galaxy activity and related topics (with a few notable exceptions, as required to illustrate some novel techniques). Included are young stellar objects (YSOs) because of their usefulness in helping understand starburst systems, and the many similarities between the ionization and outflows in YSOs and the jets and outflows in active galaxies.

To ensure a broad representation at the conference across the 3-D instrumentation spectrum as well as a balanced scientific program, we enlisted the help of an International Advisory Committee (IAC, co-chaired by R. Genzel and G. Monet) and a Scientific Organizing Committee (SOC, co-chaired by J. Bland-Hawthorn and W. van Breugel). Their advice has been invaluable, and has made the conference program a great success. The proceedings of the conference presented here include 18 invited review talks, covering scientific opportunities and instrumentation challenges, 23 talks highlighting some of the first results obtained with 3-D instruments, 19 contributed talks, and 43 poster presentations. We expect that the volume will serve as a reference work on advanced multi-wavelength 3-D imaging.

Of the Local Organizing Committee special mention must be made of J. Tweed and D. McCown for their tireless help with most organizational aspects, including accounting, logistical support, communications, and website design

and updates. The cover art was drawn by M. McInnis (LLNL, University Relations Program), from a sketchy outline by W. van Breugel. The meeting was generously sponsored by the Lawrence Livermore National Laboratory, the Institute of Geophysics and Planetary Physics, the California Space Institute, Queensgate Instruments Ltd., and ITT Aerospace. The work by W. van Breugel at IGPP/LLNL is performed under the auspices of the US Department of Energy under contract W-7405-ENG-48.

W. van Breugel & J. Bland-Hawthorn
Livermore, California, & Sydney, Australia, October 1999

Picture Credit The picture on the front cover attempts to catch the spirit of the meeting: scientific discovery, enabled by various social commodities, lifts the spirits of scientists and society into 3-D space. It was inspired by Martin Harwit's wonderful keynote speech on 'how did we get to be so lucky?'. A touch of realism was provided by the conference hotel, which announced that our meeting was about 'Imagining the Universe in Three Dimensions'.

100-100000
100-100000
100-100000
100-100000

100-100000
100-100000
100-100000
100-100000

Conference Organization

International Advisory Committee

Alcock, Charles (*LLNL, USA*)
Barden, Samuel (*NOAO, USA*)
Beckwith, Steven (*STScI, USA*)
Davies, Roger (*U. Durham, UK*)
Ekers, Ron (*ATNF, Australia*)
Genzel, Reinhard (*MPE, Germany*) - Co-chairman
Giacconi, Ricardo (*ESO, Germany*)
Karoji, Hiroshi (*Subaru, Japan*)
Le Fevre, Olivier (*LAS/Marseille, France*)
Lena, Pierre (*U. Paris 7 & Obs. de Paris, France*)
Maillard, Jean-Pierre (*CFHT, France*)
McLean, Ian (*UC Los Angeles, USA*)
Miller, Joseph (*UC Santa Cruz, USA*)
Monnet, Guy (*ESO, Germany*) - Co-chairman
Morris, Simon (*DAO, Canada*)
Mould, Jeremy (*MSSSO, Australia*)
Ohtani, Hiroshi (*Kyoto U., Japan*)
Perryman, Michael (*ESA/ESTEC, The Netherlands*)
Taylor, Keith (*AAO, Australia*)
Tonry, John (*U. Hawaii, USA*)
Welch, Jack (*UC Berkeley, USA*)
Woodgate, Bruce (*NASA/GSFC, USA*)
Wright, Gillian (*UKATC/Edinburgh, UK*)

Scientific Organizing Committee

Bland-Hawthorn, Joss (*AAO, Australia*) - (Co-chairman)
Boyle, Brian (*AAO, Australia*)
Cecil, Gerald (*U. North Carolina / NOAO, USA*)
Cook, Kem (*LLNL, USA*)
di Serego Alighieri, Sperello (*Arcetri Obs., Italy*)
Dickinson, Mark (*STScI, USA*)
Ellis, Richard (*IoA Cambridge, UK*)
Fosbury, Robert (*ESO, Germany*)

Freeman, Ken (*MSSSO, Australia*)
Graham, James (*UC Berkeley, USA*)
Guhathakurta, Puragra (*UC Santa Cruz, USA*)
Heckman, Timothy (*JHU, USA*)
Hutchings, John (*DAO, Canada*)
Mather, John (*NASA/GSFC, USA*)
McMahon, Richard (*IoA Cambridge, UK*)
Meisenheimer, Klaus (*MPIfA, Germany*)
Miley, George (*Leiden Obs., The Netherlands*)
Quirrenbach, Andreas (*UC San Diego, USA*)
Rowan-Robinson, Michael (*Imperial College London, UK*)
Silk, Joseph (*UC Berkeley, USA*)
Staveley-Smith, Lister (*ATNF, Australia*)
Tadhunter, Clive (*Sheffield U., UK*)
van Breugel, Wil (*LLNL, USA*) - (Co-chairman)
Wilson, Andrew (*U. Maryland, USA*)

Local Organizing Committee

Bland-Hawthorn, Joss (*Visiting Astronomer, LLNL, USA*)
Labov, Simon (*LLNL, USA*)
Macintosh, Bruce (*LLNL, USA*)
McCown, Donna (*LLNL, USA*)
Tweed, Jan (*LLNL, USA*)
van Breugel, Wil (*LLNL, USA*) - Chairman
Wurtz, Ron (*LLNL, USA*)

Sponsors

Lawrence Livermore National Laboratory
Institute of Geophysics and Planetary Physics
The California Space Institute
Queensgate Instruments Ltd.
ITT Aerospace

Participant List

Abrams, Marc, ITT, USA <mcabrams@itt.com>
Allington-Smith, Jeremy, U. Durham, UK <j.r.allington-smith@durham.ac.uk>
Angel, Roger, Steward Obs., USA <rangel@as.arizona.edu>
Armus, Lee, SIRTf, USA <lee@ipac.caltech.edu>
Arribas, Santiago, IoA Canarias, Spain <sam@ll.iac.es>
Bacon, Roland, CRA Lyon, France <bacon@obs.univ-lyon1.fr>
Baker, Joanne, UC Berkeley, USA <jcb@astro.berkeley.edu>
Barden, Samuel, NOAO, USA <barden@noao.edu>
Becker, Thomas, AIP, Germany <??>
Bennett, Charles, LLNL, USA <bennett2@llnl.gov>
Blais-Ouellette, Sebastien, U. Montreal, Canada
<blaisous@astro.umontreal.ca>
Bland-Hawthorn, Joss, AAO, Australia <jbh@aaoepp.aao.gov.au>
Boyle, Brian, AAO, Australia <director@aaoepp.aao.gov.au>
Cabrera, Blas, Stanford U., USA <cabrera@leland.stanford.edu>
Canizares, C., MIT, USA <crc@space.mit.edu>
Cecil, Gerald, U. North Carolina / NOAO, USA <gcecil@noao.edu>
Cianci, Sonia, U. Sydney, Australia <scianci@physics.usyd.edu.au>
Ciroi, Stefano, AIP, Germany <sciroi@aip.de>
Content, Robert, U. Durham, UK <Robert.Content@durham.ac.uk>
Cook, Kem, LLNL, USA <kcook@llnl.gov>
Crampton, David, DAO, Canada <David.Crampton@hia.nrc.ca>
Davies, Roger, U. Durham, UK <Roger.Davies@durham.ac.uk>
Diercks, Alan, U. Washington, USA <diercks@astro.washington.edu>
Dougados, Catherine, Obs. Grenoble, France
<dougados@laog.obs.ujf-grenoble.fr>
Doyon, Rene, U. Montreal, Canada <doyon@astro.umontreal.ca>
Ebizuka, Noboru, CRL, Japan <ebizuka@crl.go.jp>
Edelstein, Jerry, UC Berkeley, USA <jerrye@ssl.berkeley.edu>
Ellis, Richard, IoA Cambridge, UK <rse@ast.cam.ac.uk>
Erskine, David, LLNL, USA <erskinel@llnl.gov>
Ferruit, Pierre, U. Maryland, USA <pierre@astro.umd.edu>
Förster-Schreiber, Natascha, Saclay, France <??>
Furusawa, Hisanori, U. Tokyo, Japan <furusawa@astron.s.u-tokyo.ac.jp>
Gabriel, Carlos, ESA-VILSPA, Spain <cgabriel@iso.vilspa.esa.es>
Garcia-Lorenzo, Begonia, IoA Canarias, Spain <cgabriel@ll.iac.es>

Garnett, Don, Steward Obs., USA <dgarnett@as.arizona.edu>
Ge, Jian, LLNL, USA <ge1@llnl.gov>
Geha, Marla, LLNL, USA <mgeha@llnl.gov>
Genzel, Reinhard, MPE, Germany <genzel@mpe-garching.mpg.de>
Graham, James, UC Berkeley, USA <jgraham@astro.berkeley.edu>
Harwit, Martin, Cornell U., USA <martin.harwit@ibm.net>
Hastings, Peter, UKATC, UK <prh@roe.ac.uk>
Hattori, Takashi, Kyoto U., Japan <hattori@kusastro.kyoto-u.ac.jp>
Herbst, Tom, MPIA, Germany <herbst@mpia-hd.mpg.de>
Hewett, Paul, IoA Cambridge, UK <phewett@ast.cam.ac.uk>
Hines, Dean, Steward Obs., USA <dhines@as.arizona.edu>
Ivison, Rob, U. College London, UK <rji@star.ucl.ac.uk>
Keller, Christoph, National Solar Obs., USA <ckeller@noao.edu>
Kimble, Randy, NASA/GSFC, USA <kimble@ccd.gsfc.nasa.gov>
Korpela, Eric, UC Berkeley, USA <korpela@ssl.berkeley.edu>
Krabbe, Alfred, German Aerospace Cntr., Germany <krabbe@dlr.de>
Kutyrev, Alexander, NASA/GSFC, USA <kutyrev@stars.gsfc.nasa.gov>
Labov, Simon, LLNL, USA <slabov@llnl.gov>
Larkin, James, UC Los Angeles, USA <larkin@virgo.astro.ucla.edu>
LeFèvre, Olivier, CNRS, France <lefevre@astrsp-mrs.fr>
Luppino, Gerard, U. Hawaii, USA <ger@hokupa.ifa.hawaii.edu>
Macintosh, Bruce, LLNL, USA <bmac@igpp.llnl.gov>
Maihara, Toshinori, Kyoto U., Japan <maiara@cr.scphys.kyoto-u.ac.jp>
Maillard, Jean-Pierre, IAP-CNRS, France <maillard@iap.fr>
Maiolino, Roberto, Obs. Arcetri, Italy <maiolino@arcetri.astro.it>
Martin, Chris, Caltech, USA <cmartin@srl.caltech.edu>
Max, Claire, LLNL, USA <max1@llnl.gov>
Miller, Brian, Leiden Obs., The Netherlands <bmillier@strw.leidenuniv.nl>
Miller, Joseph, Lick Obs., USA <miller@ucolick.org>
Monnet, Guy, ESO, Germany <gmonnet@eso.org>
Morris, Mark, DAO, Canada <Simon.Morris@hia.nrc.ca>
Murphy, Tom, Caltech, USA <tmurphy@mop.caltech.edu>
Nelson, Cailin, LLNL, USA <cailin@physics.berkeley.edu>
Ohtani, Hiroshi, Kyoto U., Japan <ohtani@kusastro.kyoto-u.ac.jp>
Palunas, Povilas, NASA/GSFC, USA <palunas@gsfc.nasa.gov>
Parry, Ian, IoA Cambridge, UK <irp@ast.cam.ac.uk>
Perryman, Michael, ESA-ESTEC, The Netherlands
<mperryman@astro.estec.esa.nl>

Pietraszewski, Chris, Queensgate Instruments, UK
<Chris.Pietraszewski@queensgate.com>
Poglitsch, Albrecht, MPE, Germany <alpog@ifi.mpe-garching.mpg.de>
Prieto, Eric, LAS, France <eric.prieto@astrsp-mrs.fr>
Puxley, Phil, Gemini, USA <ppuxley@gemini.edu>
Quirrenbach, Andreas, UC San Diego, USA <aquirrenbach@ucsd.edu>
Rafanelli, Piero, U. Padua, Italy <piraf@pd.astro.it>
Ramsay-Howat, Suzanne, UKATC, UK <skr@roe.ac.uk>
Rifatto, Agatino, Obs. Capodimonte, Italy <rifatto@na.astro.it>
Rigopoulou, Dimitra, MPE, Germany <dar@mpe.mpg.de>
Roth, Martin, AIP, Germany <mmroth@aip.de>
Satyapal, Shobita, NASA/GSFC, USA <satyapal@stars.gsfc.nasa.gov>
Scoville, Nick, Caltech, USA <nzs@radio.caltech.edu>
Shopbell, Patrick, U. Maryland, USA <pls@astro.umd.edu>
Smail, Ian, U. Durham, UK <Ian.Smail@durham.ac.uk>
Spinrad, Hyron, UC Berkeley, USA <spinrad@bigz.berkeley.edu>
Stanford, Adam, LLNL, USA <adam@igpp.llnl.gov>
Stern, Daniel, UC Berkeley, USA <dan@bigz.berkeley.edu>
Stiavelli, Massimo, STScI, USA <mstiaavel>
Stockman, Peter, STScI, USA <stockman@stsci.edu>
Szalay, Alex, JHU, USA <szalay@jhu.edu>
Tacconi, Linda, MPE, Germany <linda@hethp.mpe-garching.mpg.de>
Tadhunter, Clive, U. Sheffield, UK <c.tadhunter@sheffield.ac.uk>
Taylor, Keith, AAO, Australia <kt@aaoepp.aao.gov.au>
Thatte, Niranjana, MPE, Germany <thatte@mpe.mpg.de>
Tinney, Chris, AAO, Australia <cgt@aaoepp.aao.gov.au>
Tonry, Joel, U. Hawaii, USA <jt@avidya.ifa.hawaii.edu>
Ullom, Joel, LLNL, USA <ullom1@llnl.gov>
Valdes, Francisco, NOAO, USA <FValdes@noao.edu>
Valdez-Gutierrez, Margarita, INAOE, Mexico <mago@inaoep.mx>
van Bemmell, Ilse, ESO, Germany <ivbemmell@eso.org>
van Breugel, Wil, LLNL, USA <wil@igpp.llnl.gov>
Veilleux, Sylvain, U. Maryland, USA <veilleux@astro.umd.edu>
Walton, Nicholas, ING, UK <naw@ing.iac.es>
Wilson, Andrew, U. Maryland <wilson@astro.umd.edu>
Wishnow, Edward, LLNL, USA <wishnow@ssl.berkeley.edu>
Wolf, Christian, MPIA, Germany <cwolf@mpia-hd.mpg.de>
Wolf, Juergen, German Aerospace Cntr., Germany <juergen.wolf@dlr.de>

Woodgate, Bruce, NASA/GSFC, USA <woodgate@s2.gsfc.nasa.gov>

Wurtz, Ron, LLNL, USA <wurtz@llnl.gov>

Instrumentation and Astrophysics: How did we get to be so lucky?

Martin Harwit
Cornell University

Abstract. This has been a fabulous conference, showing the enormous influence that instrumentation is having on the directions we are currently pursuing in astronomy. To see where the future might take us, I thought I might review how we got to be where we are now – how we got to be so lucky.

1. A Golden Age

The exciting discoveries of recent decades have come about through a wealth of new techniques. As we built instruments of ever-greater acuity, we stumbled upon a dazzling range of phenomena that revealed a Universe rich beyond anything we could have imagined. Gamma rays displayed the most violent outbursts witnessed in the Universe. X-rays led to the discovery of extremely hot ionized gases pervading clusters of galaxies. Infrared radiation showed the existence of galaxies undergoing huge bursts of star formation, while radio telescopes revealed radiation emitted, long ago, when the Universe was thirty thousand times younger than today (Harwit 1981).

In the spirit of these successes we press on, knowing that we must continue to probe the Universe in as many ways as we can, to whatever depth we are able. This is the spirit of this conference, — it is why we are here. If we knew more about the Universe, we might take a different, more targeted approach, guided to key truths by astrophysical theory. But we know we cannot. Our theorists are supposed to be our body guards protecting us against cosmic surprises. In this we know they are failing. Instead of relying on them, we have opted, rather, for an increasingly heavy barrage of instrumentation to conquer the Universe — a massive straightforward extrapolation of the approach that yielded so many great advances in recent decades. The question is, “Will this approach succeed?” To answer this, we may look at the history of other straightforward extrapolations.

2. Jules Verne

Over the weekend of March 20, 1999, Bertrand Piccard and Brian Jones completed the first circumnavigation of the Earth by balloon. Most remarkable was its prediction a century and a quarter earlier. In 1873, when Jules Verne wrote “Around the World in Eighty Days,” balloon-

ing seemed the most obvious way for flying around the world. Today the flight of the Breitling Orbiter-3 is considered more of a stunt. After 1903 and the flight of the Wright brothers, the airplane took over. Aeronautics underwent a discontinuity. Predictions about the balloon's technological future suddenly were all wrong. The balloon had given way to the airplane.

When we went to the Moon, a flight that Jules Verne had predicted even earlier, in 1865, it involved another discontinuity. Verne wanted to launch his spacecraft with a cannon. The actual flight required rockets — an entirely new technology which had the merits of sufficiently low g-forces that astronauts could survive. The barrel of a cannon to launch humans into space with an accelerating force as low as a tolerable 2g would be a surprising two thousand miles long — perhaps possible, but unwieldy in practice. The cannon had given way to the rocket. The desire to go to the Moon had stayed the same.

Going ahead just as always in the past is not invariably the easiest approach. What is more important is keeping ones ultimate goals clearly in sight, and then taking the technological path of least resistance. Jules Verne had laid out the goals. We just followed a different path — an important lesson to keep in mind.

So much for extrapolation. But there is also a matter of cost.

3. Cost

When our colleagues in high-energy physics began constructing the Superconducting Supercollider, they were extending previously successful accelerator technology. By the time the project was stopped, a huge circular tunnel many miles long had been dug beneath a Texas landscape. Was this their equivalent of a 2000 mile cannon barrel? I have the greatest respect for these physicists. We need to heed their plight to avoid their fate.

The natural question for us becomes "How much further can present trends in astronomy, and their projections now before the current U.S. "Decadal Review", be continued into the future before we, too, are turned back?

Astronomical missions that cost more than two billion dollars are beyond the limits of even international collaborations. Yet some of the more ambitious enterprises we now have in mind will hit this barrier if we expect to launch them in the foreseeable future. The main downfall of the Superconducting Supercollider was its cost.

Many among us still remember LDR, the Large Deployable Reflector for submillimeter astronomy, proposed a couple of decades ago for launch into space. We still have not built it, and if we did, today, it would certainly still break the bank. It too was too costly for its time, with its 10 to 30 meter aperture, depending on who was talking, and its demanding technology. Our aims in submillimeter astronomy, however, have not been stopped by this setback. Our goals are still firmly in place. Submillimeter astronomy is healthy and going forward.

4. The Rise in Data Rates

A further challenge today is the enormous rise in data rates. We keep building increasingly large detector arrays, with higher sensitivities, dynamic ranges, and spectral, spatial, and time resolution. Many space missions now on the drawing boards could ideally use gigabyte per second transmission rates. But our telemetry systems are a thousand times slower. This is a real worry.

We speak with confidence of compressing data to reduce the transmission problem. On currently planned missions the projected data gathering rates already exceed the transmission rate by one or two orders of magnitude. First we build an instrument to collect huge amounts of data, then we throw most of it away. Where observations are marred by varieties of cosmic ray glitches or other unpredictable sources of noise, data compression is likely to be wasteful.

We speak with similar confidence about laser based telemetry. But these systems do not yet exist, and I know of no astronomer who has proposed building a large light collector — a laser receiving station — at an ideal cloudless site optimized solely to receive the torrent of data we expect to soon be reaching us from an armada of spacecraft at the second Lagrangian point, L2. If we are serious about gathering massive data streams in space, we will need far higher capacity transmission systems. A laser transmitter at L2 would require a one-meter dish — which is not much, but it would also need 0.1 arcsecond pointing stability so its beam would not wander too far around the receiving station on the ground. Ideally, three such receiving stations would need to be deployed around the globe for continuous data reception.

Most astronomers assume that all such problems will soon be overcome. I am not so confident. To see why, we need to look at the origins of our past successes — just how we got to be so lucky — so we might once again take a path of least resistance as we move ahead.

5. Brief Recapitulation of the History of Radio Astronomy

Let us begin by recollecting how some of our prime advances and discoveries came about:

Radar was one of the most powerful weapons systems developed during World War II. Invented just before the onset of hostilities and further developed throughout the conflict, radar enabled British and German intelligence to see the approach of enemy aircraft long before they had crossed the British Channel. Fighter planes could be scrambled in time, and the approaching airplanes shot down.

On two successive mornings in the winter of 1942, an alert shook up defense forces all over England. Years later, James Stanley Hey, in charge of trouble shooting the British wartime radar network, reported on this incident in a 1946 letter to the journal *Nature*. He wrote (Hey 1946):

“It is now possible to disclose that, on one occasion during the War, Army equipments observed solar radiation of the order of [a hundred

thousand] times the power expected from the Sun . . . this abnormally high intensity . . . occurred on February 27 and 28, 1942. . . the main evidence that the disturbance was . . . of solar origin was obtained by the bearings and elevations measured independently by the [radar] receiving sets, sited in widely separate parts of Great Britain . . . [Hull, Bristol, Southampton, Yarmouth] . . .”

Unknown to the British, the observatory at Meudon had detected strong solar flares at the time, and it soon became evident that solar radio emission was enhanced during periods of solar activity.

Shortly after the War, Hey and his coworkers also discovered that the galaxy now catalogued as Cygnus A emits an enormous radio flux.

Right after the War, also, Martin Ryle gathered discarded British military radar equipment to set up a radio astronomy research group at Cambridge; Bernard Lovell did the same at Manchester; and Jan Oort used former German radar equipment to start a radio astronomy program in the Netherlands.

Thus was radio astronomy born. The next twenty-five years would revolutionize astrophysics with discoveries of quasars, cosmic masers, pulsars, superluminal sources, and the microwave background.

6. The Age of the Rockets

Let me turn to a different story. Over a period of a dozen years starting in the early 1930s, Wernher von Braun and a huge army of technical experts had painstakingly learned how to build the powerful military V-2 rockets with which Hitler hoped to win the War. In 1945, the U.S. military captured many of these rockets and brought them home for testing. Onboard the test flights, Richard Tousey and Herbert Friedman of the U. S. Naval Research Laboratory placed the first ultraviolet and X-ray sensors to conduct observations of the Sun.

The V-2s became the basis for America's post-war rocket-astronomical discoveries and also the foundation on which an entire U. S. rocket industry would arise.

The Soviet Union had similarly gathered their own German experts and had also begun to develop a powerful rocket industry. With the launch of Sputnik in 1957, they exhibited an impressive ability to accurately place satellites into Earth orbit, and showed that their rockets could now reach any place on Earth with great precision and presumably with significant nuclear warheads. The high ground of space gave them the ultimate means of surveillance anywhere on Earth.

To counter this advance, the United States created a crash program to develop both more powerful rockets and exquisitely incisive surveillance techniques. The Space Race of the 1960s had begun!

7. The Military as Pioneering Astronomers

This race, though in the public's mind a contest to reach the Moon first, was far more a scramble to gain military ascendancy. Perhaps the most phenomenal discovery this brought about came in 1968. So secret was the finding that it remained classified till 1973.

You may recall the Atmospheric Test Ban Treaty, signed in the Kennedy – Khrushchev era, which had finally banned the testing of nuclear bombs above ground to eliminate the radioactive fallout they created. The problem the treaty posed was that an antagonist could learn a great deal about the yield of a weapon tested underground by sensing the seismic waves it generated.

The United States worried that the Soviet Union might strive to seek greater secrecy by exploding their test devices not underground but, surreptitiously, at great distances out in space. To detect such bursts, the U. S. designed the “Vela” project. It involved several gamma-ray sensing satellites in Earth-orbit, stationed so at least one satellite would always be positioned to detect a nuclear explosion anywhere in space.

The first gamma-ray burst the Vela satellites detected must have been a tremendous shock to the military. And then others followed, one every few months. Once the military realized that these bursts were not due to Soviet nuclear bombs, but to unimaginably vaster explosions somewhere out in space, they declassified them and published their findings in the *Astrophysical Journal Letters* (Klebesadel *et al.* 1973). The news stunned the astronomical community. It has taken us twenty-five years to decide the bursts are extragalactic and their power is staggering.

8. Theorists Play the Same Game

Astronomical observers were not alone in working with military support. Throughout World War II and the Cold War, theorists were heavily involved (Thorne 1994).

Physicists and astrophysicists who have straddled the fence between academic and military work read like a Who's Who of theorists: J. Robert Oppenheimer in 1939 had done seminal work on neutron stars and black holes. Hans Bethe that same year explained how stars shine by converting hydrogen into helium. Edward Teller similarly had made important contributions to astrophysics before the War. John Archibald Wheeler pioneered the study of black holes. All were major figures on the American nuclear bomb projects. In Nazi Germany, the young astrophysicist Carl Friedrich von Weizsäcker had worked toward a German atomic bomb under Werner Heisenberg. In the Soviet Union, Yakov Borisovich Zel'dovich, perhaps the most prolific theoretical astrophysicist of the post-war era, was deeply involved in the design of Soviet nuclear bombs. And the father of the Soviet hydrogen bomb, Andrei Dmitrievich Sakharov, who later became celebrated for his efforts at establishing world peace, also made fundamental contributions to cosmology. Accompanying these giants came an army of less well known theorists who worked on infla-

tion, gamma-ray bursts, supernovae, and other outbursts of interest to astronomers – and to the military, which, in the classified sphere, funded a good fraction of the basic theory as well as experiments to determine cross-sections needed to calculate stellar opacities and nuclear reaction rates in stars, needed incidentally also for designing nuclear weapons.

9. The Inconspicuous Military Influence

The military's efforts, however, have not always been this obvious, and most of us tend to remain unaware of the true extent to which they pervade virtually all aspects of our work. Europeans, in particular, tend to be under the impression that military influences are largely an American phenomenon. But the Infrared Space Observatory, the most successful infrared astronomical mission of the decade, launched by the European Space Agency in 1995, is a clear counter example. The infrared camera on board, constructed primarily in France, had a detector array especially constructed for the mission by the French defense establishment. The Short Wavelength Spectrometer built primarily in the Netherlands, incorporated previously classified infrared detectors provided through the efforts of a co-investigator at the U. S. Air Force. Yet most of the many hundreds of astronomers who have carried out observations with these instruments remain unaware of those contributions, and would heatedly deny that this space mission had any military ties. SIRTf, of course, will have vastly greater sensitivity than ISO and incomparably larger detector arrays — all military hand-me-down devices.

10. Ground-Based Optical Astronomy

Our ground-based colleagues world-wide have similarly benefitted from adaptive optics now installed on all the most powerful optical telescopes (courtesy of our hosts here at Livermore).

Long shrouded in military secrecy and researched at significant cost to the U. S. defense establishment, these techniques have in recent years been made available to the astronomical community worldwide. A large international user community is now in place, advancing the art in a field of significant military interest.

11. The Role of Astronomers

If we now ask how we in astronomy have come so far, we must truthfully answer, somewhat like Isaac Newton, and say "By standing on the shoulders of military/industrial giants."

The role that we astronomers have played in all these efforts has been to build upon technical advances wrought by others. This is quite natural. In typical years, Department of Defense and Department of Energy budgets for research and development have exceeded the astronomy and space science research budgets of NASA and the NSF, combined, by at

least an order of magnitude. These larger-scale efforts have been central to our success. Industrial research particularly on computing, which lies at the heart of the 3-D approach — the topic of this meeting — has also played a major role. Where military or industrial support did not exist and we had to go ahead on our own, progress has been much slower. That is why the Gravitational Probe GP-B has already taken more than thirty years to develop. Neither industry nor the military have had an interest in general relativity. Nor have they shown an interest in interferometers placed in space. So, we might have to start out in interferometry in space by solving a whole range of technical problems of formation flying and station keeping. These are expensive propositions. Our limited astronomy budgets may not stretch sufficiently far, and we cannot expect these budgets to greatly increase. This means that we cannot, except in rare cases of uncommon importance, ask for support of missions that require capabilities which significantly outstrip commonly available military or industrial capacities. Gravitational wave detection is one of these few urgent exceptions. Hence we are deeply involved in the LIGO project. But should we, for example, spend our own limited budgets on devising and perfecting laser telemetry or delicate station keeping systems in space? If others will build these for us because they have practical industrial or military applications, our path of least resistance might be to wait for such technologies to eventually arrive. These are choices we should be considering as we plan astronomy's future.

12. A Societal Connection

Astronomy is a luxury that society can afford only when it has provided people with life's primary needs — food, shelter, security. This is why we see astronomy flourishing precisely where society has produced equipment we can adapt and adopt. In recent decades, security has played a dominant role, and so we inherited sensitive detection techniques, sophisticated computers, and powerful launch vehicles. But increasingly preoccupations in the United States have shifted to health and the environment. To the extent that our search for the origins of life fits into astronomy, we may expect that tools developed for medical and biotechnologies would, in the next few decades, provide insights into the most primitive organisms and their macromolecular components. We may have to search for these deep in the oceans, or far beneath the surface of the earth, where the hyperthermophiles reside, or look for trace constituents in carbonaceous chondrites; or fossils in ancient terrestrial rocks, or the atmospheres of extrasolar planets. The next generation of astronomers could be dedicating itself to a vast search for the origins of life using tools totally foreign to most of us today but common, by then, to biomedical industry. This might be one way to follow a path of least resistance. It might not be the path we would be planning today — but then, many of the most-striking astronomical advances of recent decades were not planned by us either.

13. Return to the Decadal Review

As far as data-intensive efforts go, the 3-D approach we are discussing at this meeting may persist as the prime thrust of research for some time to come. The deciding factor will be the direction that industrial and military research take. As long as Moore's law, with its doubling of computing capabilities every eighteen months continues to hold, the 3-D approach will flourish. When we reach the point where this law breaks down or limitations on data transmission rates are reached, we should be ready to take a long hard look at our discipline's future. This may not be too many years from now.

14. Summary

Our field as a whole cannot get out too far ahead of technologies developed by others to meet basic societal needs.

All the generations of scientists who preceded us shared in one certainty, namely that society could provide them only with limited support, and that they would have to make use of whatever tools were ready at hand to forge a new view of the cosmos. Where they asked for more than society could easily afford progress inevitably was slow and often faltered. It is up to us to find the necessary means. This may involve steady progress as currently planned in 3-D instrumentation, or perhaps radically new directions — discontinuities that may ultimately prove more useful in unpredictable ways.

As long as people all over the world continue to invent new ways to improve their lot, new technologies will also open up for our search through the Universe. If we remain sufficiently astute to take advantage of these, to learn more about the Cosmos and answer questions that generations of people have posed since antiquity, we will be both serving society and working with means that society can afford. This is how we've succeeded these past few decades. That's how we came to be so lucky!

I thank the organizers of this conference for their generous invitation to attend. I am pleased to also acknowledge research support from NASA.

References

- Harwit, M. 1981, "Cosmic Discovery — The Search, Scope and Heritage of Astronomy," Basic Books, New York.
- Hey, J.S. 1946, "Solar Radiations in the 4 - 6 meter Radio Wavelength Band," *Nature* 157, 47.
- Klebesadel, R.W., Strong, I.B. & Olson, R.A. 1973, "Observations of Gamma-Ray Bursts of Cosmic Origin," *Astrophysical Journal* 182, L85.
- Thorne, K.S. 1994, "Black Holes & Time Warps — Einstein's Outrageous Legacy," W. W. Norton & Co, New York.

Galaxy Formation and Large Scale Structure

Richard S. Ellis

Institute of Astronomy, University of Cambridge CB3 0HA, England

Abstract.

Galaxies represent the visible fabric of the Universe and there has been considerable progress recently in both observational and theoretical studies. The underlying goal is to understand the present-day diversity of galaxy forms, masses and luminosities in the context of theories for the growth of structure. Popular models predict the bulk of the galaxy population assembled recently, in apparent agreement with optical and near-infrared observations. However, detailed conclusions rely crucially on the choice of the cosmological parameters. Although the star formation history has been sketched to early times, uncertainties remain, particularly in connecting to the underlying mass assembly rate. I discuss the expected progress in determining the cosmological parameters and address the question of which observations would most accurately check contemporary models for the origin of the Hubble sequence. The new generation of ground-based and future space-based large telescopes, equipped with instrumentation appropriate for studying the resolved properties of distant galaxies, offer the potential of substantial progress in this field.

1. Introduction

Despite a dramatic surge of interest in observational cosmology (Figure 1), our physical understanding of how the local diverse population of galaxies formed remains unclear. Although semi-analytical models based on hierarchical growth in a cold dark matter dominated Universe are able to reproduce many basic observables such as the joint distributions of magnitude, colour and redshift to faint limits, and the spatial distribution of sources selected in different ways at various look-back times, we seek more fundamental ways of making progress and, in particular, of finding observational tests of galaxy formation models immune from uncertainties introduced by our poor knowledge of the cosmological parameters and the power spectrum of mass fluctuations. In this brief non-specialist review, I assess the renewed enthusiasm for independently determining the cosmological parameters and emphasise how the new generation of ground-based and future space-based large telescopes, equipped with instrumentation appropriate for studying the resolved properties of distant galaxies, might then approach the question of understanding the origin of the Hubble sequence of morphological types.

Figure 1. The rapid growth in publications in Ap J and AJ containing the keyword *observational cosmology*. The dates of some significant events are included. Keyword statistics cannot be reliably used prior to 1975.

2. Models for the Growth of Structure

The most popular model for the origin and development of structure in the expanding Universe is the cold dark matter (CDM) model (White & Rees 1978, Blumenthal et al 1984) whose predictive power for both the assembly of mass structures and the history of galaxy formation has been exploited via both numerical simulations (Frenk et al 1985) and semi-analytical calculations (Kauffmann et al 1993, Cole et al 1994, 1999).

CDM posits that the bulk of gravitating mass in the Universe is non-baryonic in form and which detaches from the expanding plasma prior to recombination. A present-day galaxy can be regarded as a peak of radiating baryonic matter embedded within a more extensive dark matter halo. Although the growth of mass structures is governed by the merging hierarchy of dark matter halos (Press & Schechter 1985), the onset of galaxy formation depends on the interplay between gas cooling and star formation which relies on more detailed astrophysics. Further development e.g. details of the formation of the Hubble sequence remains unclear. CDM should therefore be envisioned as a theory of structure formation underpinned by a fundamental assumption - the presence of copious amount of non-interacting dark matter - upon which more detailed astrophysical ingredients must be added.

Even in the context of understanding large scale structure, the CDM picture represents a working framework rather than a specific well-defined model. Within hierarchical dark-matter dominated models are variants which differ in the assumed cosmological parameters (Ω_{mass} , Ω_{baryon} , Ω_{Λ} , H_0 etc) and the spectrum of initial fluctuations, characterised by the slope n ($n=1$ in most in-

flationary models) and absolute normalisation (the latter quoted in terms of the variance of mass fluctuations on $8h^{-1}$ Mpc scales, σ_8). None of these key parameters is known with certainty, including perhaps even H_0 , but three combined sets consistent with much of the available data are in popular use (see Table 1) (adapted from Thomas et al 1998). Each represents a variant of the original 'Standard' Λ CDM which fails to jointly fit the absolute value of microwave background fluctuations found by COBE and the present-day number density of rich clusters (Wright et al 1992).

Table 1. The CDM Family

CDM Variant	Ω_{mass}	Ω_{baryon}	Ω_Λ	H_0	n	σ_8
Standard Λ CDM	1.0	0.050	0.0	50	1	0.61
Open Λ CDM	0.3	0.026	0.0	70	1.3	0.85
Lambda Λ CDM	0.3	0.026	0.7	70	1	1.30
Tilted τ CDM	1.0	0.050	0.0	50	0.7	0.68

The key observables which have, over the past few years, been used to reject Λ CDM and support some or all of the others include the absolute normalisation of large scale fluctuations in the microwave background (Bunn & White 1997), the local number density of rich clusters of galaxies (White et al 1993) and the power spectrum of the local galaxy distribution (Peacock 1997). Broadly speaking, the CDM variants are underconstrained by these observables if the cosmological parameters are allowed a free range. However, structure *evolution* over $0 < z < 2$, during which time the local galaxy population might have assembled, depends sensitively on the values of Ω_M and Ω_Λ and thus if it became possible to independently constrain these parameters, systematic observations of distant galaxies would provide less ambiguous information on physical aspects of galaxy formation. The bottom line is that it will be hard to find a shortcut to substantial progress in understanding galaxy formation without comparable effort to nailing down the cosmological parameters.

3. Progress in Structure Formation and Cosmology

Should we believe the growing optimism that we will soon resolve many of the key variables that govern the history of large scale structure, including the elusive cosmological parameters (e.g. Tytler 1997, Turner 1999) - or are we witnessing a rerun of the enthusiasm that swept observational astronomy in the 1970's (e.g. Gunn & Tinsley 1975, Gott et al 1976)? I have selected four basic observational initiatives and good progress can be expected for each of them in the next decade.

3.1. The Local Distribution of Galaxies

The local spectrum of mass fluctuations has, until now, been largely constrained by the abundance of rich clusters quantified via σ_8 (Eke et al 1996) (see Table 1). However, clusters of galaxies cannot be easily selected on an uniform basis in

terms of their total mass. Cluster mass estimates based on galaxy dynamics are affected by substructures whose nature is poorly defined even when hundreds of velocities are available. Likewise, masses based on X-ray luminosities or temperatures rely on structural data which is in short supply. The power spectrum $P(k)$ ¹ of a large uniformly-selected galaxy redshift survey constrains the mass distribution on large scales with the disadvantage that galaxies may be biased tracers of the underlying matter. When coupled with the amplitude of angular fluctuations in the microwave background, joint constraints on various cosmological parameters become available (see Eisenstein et al 1998).

Very large, statistically-complete, redshift surveys are now underway (Gunn & Weinberg 1995, Colless 1999) motivated by the need to constrain the power spectrum and topology of the galaxy distribution as well as the peculiar velocity field, either statistically or through specific surveys targetting individual classes of galaxies (Colless et al 1999). At the time of writing the largest local survey is the Anglo-Australian 2dF Galaxy Redshift Survey (2dFGRS – Taylor 1995, Colless 1999) which has $\simeq 50,000$ redshifts to $b_J=19.5$ from a currently-incomplete magnitude-limited sample within one $75^\circ \times 5-15$ declination strip in each Galactic hemisphere. An important development with these surveys is the generation of 'mock catalogues' (Cole et al 1998) based on numerical simulations incorporating bias prescriptions, observational selection criteria and the instrumental characteristics of the actual surveys (Figure 2).

Deeper surveys can, in principle, constrain the correlation function of galaxies at large look-back times, in order to track the evolution of structure. In practice however, there are formidable obstacles. Surveys such as 2dFGRS are magnitude-limited and, faintward, k-corrections increase the spread in redshift considerably so that the mean depth $\bar{z}$ is not strongly dependent on magnitude limit (Ellis et al 1996). A $B=25$ wide-field 8-metre telescope survey is likely to have a significant volume overlap with one surveyed to $B=22$ by 2dF itself (although the galaxies sampled would of course be intrinsically less luminous).

If galaxies can be selected to lie in specific redshift intervals through judicious photometric selection (e.g. the Lyman dropout technique) then *angular* correlations can be more effective (Steidel et al 1996). The difficulty is relating restricted populations to those observed locally. However, this problem remains in pure magnitude-limited samples since k-corrections distort the observed mix of types each of which may have different spatial distributions. A number of groups are now undertaking multi-band panoramic surveys with a view to utilising photometric techniques to examine angular correlations as a function of inferred redshift (Brunner et al 1999, Marzke 1999, Figure 3). Such analyses will necessitate complex modelling to account for the subtle selection effects involved (Kauffmann et al 1999).

In biased galaxy formation, early star-forming systems, e.g. Lyman break galaxies, should be more strongly clustered (or biased) than their later versions forming at lower redshift. Bias can be tested within a given dataset by correlating the clustering amplitude with luminosity or preferably mass. Luminosity-dependent clustering is tentatively claimed for the Lyman break galaxies (Steidel

¹ $k = 2\pi/\lambda$ is the Fourier wavenumber of a particular comoving length scale, λ Mpc.

Figure 2. (Top) The redshift-space distribution of galaxies so far in the Anglo-Australian 2dF Galaxy Redshift Survey (Colless 1999). (Bottom) That expected in τ CDM from the mock catalogue produced by Cole et al (1998). The mock catalogue incorporates many of the technical limitations of the 2dF instrument and details of the survey strategy as well as the manner in which galaxies are biased.

Figure 3. A new generation of faint panoramic optical and infrared imaging surveys is motivated by the desire to constrain evolution in the angular clustering of galaxies as a function of epoch. The image represents a 13×13 arcmin² field limited taken with four mosaiced exposures using CIRS1 - a 4×1024^2 HgCdTe array (Beckett et al 1998). Analyses of angular correlations for IR-selected samples partitioned according to photometric redshift will necessitate consideration of both k -correction and luminosity-dependent biases.

et al 1999) and interpretation of this signal will depend on deriving the dynamical masses of individual examples (Pettini et al 1999). In the next few years we can thus expect a huge influx of $z < 0.3$ redshift survey data, the completion of several deep panoramic optical-infrared imaging surveys and detailed studies of selective populations of galaxies at $z > 1$. These data will be influential in understanding biased galaxy formation and constraining the power spectrum $P(k)$ and recent growth in clustering.

3.2. Angular Fluctuations in the Microwave Background

Proponents of the two major microwave background satellites of the next decade (the Microwave Anisotropy Probe (MAP, <http://map.gsfc.nasa.gov>) and particularly the Planck Surveyor <http://astro.estec.esa.nl/Planck/>) promise precision measurement of all cosmological parameters and characteristics of the initial power spectrum. Is such optimism simply a reflection of the fact that the theoretical modelling is well ahead of the observational data?

A good assessment of the immediate opportunities can be found by considering evidence for the location of a possible 'Doppler peak' in the spectrum of angular fluctuations delineated by various experiments (Figure 4). The sought-after feature is the sound horizon at recombination and its angular scale measures the angular diameter distance at $z \approx 1100$ and hence the spatial curvature

Figure 4. Angular spectrum of fluctuations in the microwave background as constrained by various experiments (see legend) leading to the suggestion of a ‘Doppler peak’ whose angular scale constrains the spatial curvature $\Omega_M + \Omega_\Lambda$ to be flat (after Efstathiou et al 1999)

representing the sum of the contributing energy densities $\Omega_M + \Omega_\Lambda$. Efstathiou et al (1999)² analysed the extant data and found:

$$\Omega_M + \Omega_\Lambda \equiv 0.88 \pm 0.06(1\sigma)$$

Considerable progress is expected in the next few years in confirming or otherwise this result which, at the moment, involves combining measurements made with different instruments. The Boomerang project³ should be the first to trace the fluctuation spectrum through the putative Doppler peak with a single instrument.

3.3. Supernovae and The Cosmic Deceleration

The traditional apparent magnitude redshift relation for brightest cluster galaxies led to conflicting claims on the cosmic deceleration parameter q_0 ⁴ (Gunn & Oke 1974, Kristian et al 1977). Uncertain corrections for luminosity evolution in giant ellipticals even led to early claims for a cosmic acceleration (Gunn & Tinsley 1975)! A more recent infrared study (Aragón-Salamanca et al 1998) introduced a further correction for the assembly of stellar mass from mergers since $z \simeq 1$. First-ranked cluster members are evidently complex evolving sys-

²See also the independent analyses by Lineweaver (1998) and Tegmark (1998).

³<http://astro.caltech.edu/~lgg/boom/boom.html>

⁴The deceleration parameter $q_0 = -\ddot{R}R/\dot{R}^2$, where $R(t)$ is the scale factor, is no longer commonly used as it maps onto Ω_M and Ω_Λ in a redshift-dependent manner.

tems whose properties will have to be understood in considerable detail before any believable constraints on the cosmic deceleration emerge.

It is thus only reasonable to be skeptical about recent claims (Garnavich et al 1998, Perlmutter et al 1999) for a cosmic acceleration based on the apparent magnitude - redshift relation for distant Type Ia supernovae (SNIa, (Figure 5). May their poorly-understand progenitors not evolve in some as yet undiscovered way with redshift?

There are, however, fundamental differences between the first-ranked cluster galaxy and SNIa campaigns. Foremost, supernovae are *thermodynamical events* not a restricted class of objects drawn in a (possibly biased) manner from a wider statistical population. Secondly, the information obtained for each distant SN event is impressive and includes (i) time-dependent spectra for day-by-day comparison with that of local examples constraining possible compositional and energetic differences, (ii) light curves in various photometric bands useful for clarifying correlations between the rest-frame peak luminosity and light curve shapes, and (iii) the morphological properties of the host galaxies and the location of the SN event within them. With larger samples, analyses can be executed for subsamples representing, e.g. the most luminous examples, those found in dust-free environments or with old stellar populations. The absence of an increased scatter in Figure 5 at large redshift is highly significant. If progenitor evolution, host galaxy dust or other systematic effects were distorting the curve towards a non-zero Λ , then such effects would have to be orchestrated to be similar across a wide range of environments and galaxy evolutionary histories.

The SNIa data do not yet span a sufficiently large redshift range to independently constrain Ω_M and Ω_Λ and thus measures deceleration over a particular redshift range, effectively $\Omega_M - \Omega_\Lambda$. Contrary to popular perception, the strongest result is the rejection of the matter-dominated Einstein-de Sitter model *not* the presence of a non-zero cosmological constant (the latter being only a 3σ effect with the current data). Only by combining the SN constraint with that of spatial curvature derived from the location of the microwave background Doppler peak is strong evidence for a cosmic acceleration inferred (Figure 6, Efstathiou et al 1999).

What are the weak points in the SNe analyses? Notwithstanding the near-constant scatter with redshift referred to above, uniformly distributed dust extinction (arising within all host galaxies at $z > 0.5$ or in the intergalactic medium - Aguirre 1999) and progenitor evolution (c.f. Riess et al 1999) remain a concern. Fortunately, we can expect to address both convincingly in the next few years via much larger samples from which carefully-selected subsets can be compiled, and by extending the analysis beyond $z \simeq 1$ where the effects of a non-zero Λ should *decrease* in comparison with that of a uniform absorbing medium. The challenge, as always, lies in designing the most effective facility to conduct the necessary survey. Considering the impact of the SNe Ia results, it is surely high time to consider more ambitious instrumentation dedicated to locating them and their multi-band follow-up.

Figure 5. The Hubble diagram (top) for Type Ia supernovae published by the Supernova Cosmology Project (Perlmutter *et al.* 1999). The dispersion in rest-frame blue apparent magnitude, corrected for light curve width - peak luminosity correlations, remains remarkably tight at all redshifts (inset). Distant supernovae are $\simeq 0.5$ mag too faint to be consistent with a spatially-flat matter-dominated Universe (bottom) ($[\Omega_M : \Omega_\Lambda] = [1, 0]$) and provide tantalising evidence for a non-zero Ω_Λ . Individual outliers (unfilled circles) with incomplete or unsatisfactory data do not affect this conclusion.

Figure 6. The location of the Doppler peak in the microwave background (which constrains $\Omega_M + \Omega_\Lambda$ - Figure 4) can be combined with complementary constraints on $\Omega_M - \Omega_\Lambda$ from the Type Ia supernova data (Figure 5) (Efsthathiou et al 1999). The resulting solution supports a spatially flat ($\Omega_K=1-(\Omega_M + \Omega_\Lambda)=0$) Universe with $\Omega_M=0.3$, $\Omega_\Lambda=0.7$.

4. Gravitational Lensing and the Mean Mass Density

The synergy achieved between the SNe and microwave background data is, at first sight, reassuring (Ostriker & Steinhardt 1995). However, the conclusions lead cosmology into difficult territory. What is the physical origin of a non-zero Λ and why is its energy density anywhere near that of the gravitating mass? (Steinhardt et al 1999)?

Not only do we seek verification via at least one further constraint on Figure 6 but also, in understanding galaxy formation, the relative distribution of baryonic and non-baryonic matter and precise constraints on Ω_M are important goals. The most fundamental assumption of CDM is the presence of non-baryonic matter and, short of its direct detection, comparing the baryonic and non-baryonic components is an essential test. Mass distributions on cosmological scales can be constrained with ambitious wide field imaging surveys utilising the weak gravitational lensing of background field galaxies (Narayan & Bartelmann 1996, Mellier 1998,).

Gravitational lensing is a fascinating phenomenon. However, even its most ardent supporters would admit it has yet to provide convincing constraints on the distribution of non-baryonic matter on scales upward of 1 Mpc. Work has concentrated where the signals are strongest, i.e. the cores of rich clusters. Here weak shear and strong lensing (multiply-imaged arcs) studies have confirmed high M/L ratios consistent with $\Omega_M \simeq 0.3$ inferred already from dynamical and X-ray studies (Smail et al 1997, Allen 1998).

Powerful techniques have now been developed to derive the projected distribution of lensing mass from the weak shear seen across the field of a CCD detector. The methods assume that, statistically, the intrinsic ellipticity and red-

Figure 7. The extension of weak gravitational lensing as a probe of mass distributions outside rich clusters is an imminent prospect with panoramic imaging cameras. The image shows Kaiser et al's (1999) mass reconstruction in a field containing two X-ray clusters. Dark material connecting such systems may be revealed through weak lensing techniques.

shift distribution of the background sources are well-behaved (Kaiser & Squires 1996, Kaiser et al 1998, Schneider et al 1998). In practice, the techniques are limited by instrumental and seeing corrections, particularly for ground-based data. Comparative studies utilising different techniques for the same rich cluster observed independently with different telescopes (Erben et al 1999) suggest mass distributions can be extracted reliably to sensitivity levels equivalent to those needed to see coherent shear from larger-scale structures *outside* rich clusters. Potential targets for weak lensing studies therefore now extend to individual filaments whose existence is inferred from redshift surveys or otherwise (e.g. Kaiser et al 1998, Figure 7).

There is great potential because lensing offers a bias-free probe of mass fluctuations on various scales (Jain & Seljak 1997). CCD cameras with upwards of 8000^2 pixels make surveying for mass structures independently of their baryonic fraction a practical prospect provided the systematics can be controlled. Correlation studies of the shear drawn from many randomly-chosen independent fields can provide a global measure of the mass power spectrum on a particular scale (the 'cosmic shear' - Mould et al 1994) whereas ultimately one can hope to define projected mass maps and interpret these statistically for Ω_M by correlating with similar data from redshift surveys (Figure 2).

In conclusion, there are four observational programmes each of which already offers complementary constraints on the cosmological parameters and power spectrum of mass fluctuations. More importantly, the future prospects for each are excellent. By eliminating the cosmological uncertainties which un-

derpin all theories for structure formation, we can hope to move forwards to a physical picture of how galaxies form and evolve.

5. The Origin of the Hubble Sequence

Although originally introduced for taxonomic purposes, the Hubble sequence of morphological types defines an axis along which there are strong physical trends. For stellar populations the sequence represents the locus of the ratio of the current star formation rate to a long-term average (Struck-Marcell & Tinsley 1978). There are near-monotonic dependences in gas content, rotational support, bulge/disk ratio and central density (de Vaucouleurs 1977, Efstathiou & Silk 1983). Most importantly, there are suggestive variations in the population mix as a function of environment (Dressler 1980, Whitmore et al 1993).

The challenge to infer the evolutionary history of this sequence has been revolutionised by Hubble Space Telescope's (HST) ability to resolve kpc scale structures at most redshifts of interest. However, surprisingly little progress has been made in equivalent 2-D spectroscopic follow-up of these distant sources with ground-based telescopes. This is one of the key areas of importance in the next decade.

5.1. Monolithic and Hierarchical Theories of Galaxy Formation

The continuous merging of dark matter halos and delayed onset of star formation in the hierarchical picture has led to a profound challenge to classical views for the origin of the Hubble sequence. Prior to CDM, the traditional viewpoint (e.g. Eggen et al 1962, Sandage 1986) associated the high stellar density, low specific angular momentum and homogeneous old stellar populations of spheroidal galaxies with rapid dissipationless collapse at high redshift. Dissipative collapse of gas clouds at later epochs would lead to rotationally-supported disks destined to become local spirals. Giant ellipticals forming at high redshift via 'monolithic' collapse prompted numerous searches for luminous primaeval galaxies (Djorgovski 1996).

In hierarchical pictures however, the first stellar systems will be small and numerous, reflecting the abundance of appropriately-sized halos at early times. Rapid merging of these initial systems delivers a population of young bulges. Continued gas cooling produces stable disks around these bulges some of which will later merge to form large spheroidal galaxies (Baugh et al 1995).

The differences between the monolithic and hierarchical pictures are greatest for the spheroidal galaxies. Merger-driven morphological evolution occurs earlier in the denser environments (Kauffman 1995) and thus strong differential effects are also expected in the sense that field spheroidals will always be younger and more inhomogeneous than their clustered equivalents. Although evidence continues to accumulate suggesting that many local ellipticals formed via the merger of disk galaxies (Schweizer 1997), this does not necessarily imply *all* spheroidals formed recently.

Figure 8. The redshift-dependent comoving volume-averaged density of star formation derived from the luminosity density of ultraviolet light, H α -emission and bolometric far-infrared flux (Madau 1999). Data points with error bars refer to UV-based measures with no extinction correction (left panel) and that corrected with a constant value at 1500Å (right panel). The dotted line shows the fiducial value necessary to produce the observed background light. No star formation diagnostic is immune from difficulties and comparisons for the same sources suggest further complications arise from the timescales over which a given diagnostic applies.

5.2. Cosmic Star Formation and Mass Assembly Histories

The completion of extensive redshift surveys (Lilly et al 1995, Ellis et al 1996, Cowie et al 1999) and the identification of representative samples of Lyman break galaxies at various epochs (Steidel et al 1996, 1999) has led to analyses of the comoving density of star formation as a function of look-back time (Madau et al 1996, Madau 1999, Figure 8). Does this 'Madau plot' which indicates an extended history of star formation differentiate between hierarchical and monolithic galaxy formation as has been claimed (Baugh et al 1998)?

In quantitative detail, the star formation history is susceptible to major changes through unaccounted obscured sources, dust extinction on the detected ones, incompleteness arising from sources fainter than the survey limit, and inaccuracies or inappropriate comparisons associated with different diagnostics used, each of which has its own limitations. Although in broad agreement with the CDM predictions (Baugh et al 1998), detailed comparisons depend more on how feedback and star formation are implemented in the semi-analytical models rather than on structure formation in a cold dark matter Universe (c.f. Jimenez et al 1999). Although Figure 8 has been central in motivating optimum strategies for finding distant sources, we now seek a more fundamental way to address the physical foundations of galaxy formation theories. Ultimately this implies tracking the evolution of the assembling mass.

5.3. Field Spheroidals

The continuous formation of field spheroidals from merging spirals contrasts markedly with the traditional picture of monolithic collapse at high redshift.

This seems one of the simplest ways to differentiate two very different pictures of morphological evolution. However, in practice, a demise with redshift in the abundance of field ellipticals has been remarkably difficult to verify. Although field ellipticals can be readily discerned to $I=22-23$ in typical WFPC-2 images (and $I=24-25$ in the higher signal to noise Hubble Deep Fields), morphological number-magnitude counts (Glazebrook et al 1995, Driver et al 1996, Abraham et al 1997) give ambiguous results because of uncertainties associated with the local luminosity function. Regardless, the evolution expected in hierarchical CDM is a strong function of the cosmological model, with little evolution to $z \simeq 1-2$ expected in open or Λ -dominated versions (Kauffmann & Charlot 1998b).

Optical-infrared colours offer an alternative route to constraining the assembly history. In the monolithic case we would expect to find a sizeable population of intrinsically red systems beyond $z \simeq 1$; in fact few are found (Zepf 1997, Barger et al 1999). Menanteau et al (1999) selected a large sample of morphological spheroidals and demonstrated a marked paucity of examples with colours compatible with passively evolving systems formed at $z > 3$. The drawback is that only a modest amount of ongoing star formation occurring in objects which collapsed at higher redshift would be needed to move passively-red systems into the colour range observed (Jimenez et al 1999).

Ideally we seek a characteristic that demonstrates that field ellipticals are bluer and more inhomogeneous than their clustered counterparts from which the associated mass assembly rate could be derived. Figure 9 shows the internal colour distribution of two rather different field spheroidals located in the HDF (Abraham et al 1999a). Across a large sample of such systems, evidence is accumulating that field ellipticals are more inhomogeneous in their observed properties than those found in clusters in qualitative agreement with the predictions of hierarchical models. The patchy colour distribution of a significant fraction of field spheroidals at $z > 0.3$ may be consistent with merger-driven evolution and the challenge is now to quantify this in terms of both the observed merger fraction and the mass assembly history. Inevitably, more detail is needed on the metallicity and timescales of secondary star formation implied in the HST images.

5.4. Faint Blue Irregulars

Deep HST images confirmed that the bulk of the faint blue excess first located in faint galaxy counts arises from galaxies of irregular morphology (Glazebrook et al 1995, Ellis 1997). Such systems also contribute significantly to the increasing star formation density to $z \simeq 1$ (Brinchmann et al 1998). But what is the physical origin of these systems and why are so few found today?

Pixel-by-pixel simulations which attempt to recover the appearance of local galaxies as viewed at various redshifts taking into account instrumental and redshift-dependent effects (Abraham et al 1998, Brinchmann et al 1998) suggest that only a small fraction of the HST irregulars might be late-type spirals viewed at rest-frame UV wavelengths (a claim now supported by the NICMOS images, Figure 10).

The demise of these abundant peculiar systems with active star formation remains an outstanding question. High resolution infrared images are needed to quantify the established stellar mass (Broadhurst et al 1992, Kauffmann &

Figure 9. The internal V-I colour distribution for two spheroidals of known redshift in the Hubble Deep Field (Abraham et al 1999a). Such analyses indicate a greater degree of internal inhomogeneity, possibly consistent with secondary star formation induced by mergers, for field ellipticals than for their clustered equivalents.

Figure 10. A morphologically irregular galaxy in the HDF with the integral field unit for the Gemini CIRPASS spectrograph superimposed. Each fiber would sample a 0.15 arcsec diameter portion of the galaxy enabling detailed internal studies.

Charlot 1998a) as well as resolved dynamical data to address the possibility that many merge to form more regular systems.

This will be an enormous challenge. Merging statistics based on projected imaging data (Patton et al 1998, LeFevre et al 1999) are difficult to interpret in the context of hierarchical models because merger timescales are needed to convert pair fractions into merger rates. Will spectroscopy help in resolving this impasse? Possibly not in individual cases such as Figure 10, but the redshift-dependent mass distribution and star formation timescales of young peculiars will be an important step forward in understanding the role of feedback which apparently delays the formation of these systems.

5.5. Disks, Bars and Bulges

Numerical simulations have yet to make much headway in addressing the quantitative evolution of *internal* features in galaxies, e.g. stellar disks, bars and bulges, yet we can expect exquisite observational data on these to $z \simeq 1$ and beyond in coming years. Theorists prefer to delay disk formation to late epochs in order to avoid spiral destruction through mergers as well as an inexorable transfer of angular momentum from the stellar disk to the outer dark matter halo (Weil et al 1998).

Rotation curves are now available for almost a hundred distant spirals to $z \simeq 1$ (Vogt et al 1999) yet the associated Tully-Fisher relation shows little sign of any evolution. Structural decomposition of HST images indicate both increased disk surface brightnesses and bluer colours consistent with a higher star formation rate (Lilly et al 1998). All available data are consistent with the notion that spirals had their current space density and structural forms established at $z \simeq 1$. Is this in conflict with theory? Again, in open or Λ -dominated models, the conflict might not be so severe. Moreover, observations could be reconciled with predictions if disks were still growing whilst their star formation rates declined (Mo et al 1998). However, neither the observational data nor the model predictions are sufficiently refined to be sure what is happening at the moment.

Stellar bars and bulges offer a further clue to the growth history of spiral galaxies. CDM models predict bulges should be the very oldest components having formed at early times from mergers of the first collapsing systems and therefore have intrinsically red colours (Cole et al 1999). However, bulges selected from a sample of HDF spirals reveal a remarkable range of colours with few as red as ellipticals of the same redshift (Ellis & Abraham 1999, Figure 11) suggesting either many formed more recently or a continuous infall of gas or satellites introduced recent star formation in their cores. Bulges may also form sequentially from stellar bars, both features representing stages of instability in a well-established disk embedded in a significant stable dark matter halo (Combes 1999). A paucity of barred systems has been claimed in the HDF (Abraham et al 1999b) although the sample size remains small.

Whereas there is naturally much interest in surveying galaxies with $z > 2.5$, progress in understanding the astrophysical processes which govern galaxy formation and evolution will also require extensive studies of selected galaxies with the redshift range $0 < z < 2$ during which period it seems the Hubble sequence of types became established. HST has shown the way with exquisite optical imaging data. The challenge will be to complement this with near-infrared and

Figure 11. Representative $z > 0.5$ spirals in the Hubble Deep Field whose images have sufficient signal/noise to address questions concerning the star formation histories of bulges, bars and stellar disks (c.f. Abraham et al 1999a,b; Abraham & Ellis 1999). Resolved stellar populations from such multicolour images can be interpreted more clearly using dynamical and metallicity information obtained from integral field spectrographs.

spectroscopic data of appropriate resolution from ground-based telescopes and move away from the traditional 'number counting' approach which serves only to determine the integrated properties of the galaxy population.

6. Instrumental Themes

The above scientific overview is intended to set the scene for some of the 3-D applications discussed at this meeting. A few common instrument themes are relevant in connecting the various scientific questions:

- *Wide Field Imaging:* The Schmidt telescope revolution which influenced this field in the 1970's has been replaced by its digital equivalent - panoramic CCD cameras on 2.5-4m telescopes with $\simeq \text{deg}^2$ fields. Already such surveys are being extended into the near-IR. Multicolour surveys promise great progress in the cosmological arena through supernova searching, weak lensing and photometric redshifts aimed at tracking the growth of galaxy correlations. There are substantial challenges in each application, however. The Type Ia SN programme is hindered by the need for more detailed astrophysical information spanning a wide wavelength range (at both high and low z) rather than gross statistics. Weak lensing as a probe of large scale structure and Ω_M requires exquisite correction of instrumental effects and may, ultimately, only be possible in space⁵. Photometric redshifts are largely untested in the all-important $1 < z < 2$ range and redshift-dependent effects will have to be understood before extracting meaningful signals on correlation evolution.
- *Wide Field Spectroscopy:* Now is also an exciting time for wide field spectroscopic surveys. Although the prime spectroscopic products of the Sloan and 2dF surveys will be maps of the $B < 20$ galaxy distribution for power spectrum analyses, we can also hope to utilise the galaxy spectra to investigate the astrophysical origin of bias through luminosity or star-formation rate dependencies. The spectroscopic surveys will deliver data that is much richer than a simple redshift. Just as the Schmidt surveys formed the basis for springboard projects using other facilities so the 2dF survey is already proving useful for spinoff projects from faint radio surveys to weak lensing behind known filamentary structures. 2dF also has the potential to probe much fainter than $B \simeq 20$ and this raises the important question of an 8-m equivalent device capable of exploring to $z \simeq 1$ and beyond.
- *Resolved Spectroscopy:* HST has shown us the tantalising beauty of resolved galaxies at high redshift. At face value it seems the Hubble sequence of galaxy types was laid down by some process soon after redshift 2. Optical surveys are biased however, being sensitive only to sources with recent or ongoing star formation. We seek to extend our viewpoint of these

⁵At the moment there is no contest as the field of view of Hubble Space Telescope is very uncompetitive but a wide field modest aperture space telescope would be overwhelmingly powerful in this area.

galaxies with high resolution images in the near-infrared, where more representative stellar populations can be analysed, and through far-infrared and sub-mm detections sensitive to obscured regions. A major question is whether the optical/near-infrared should remain the fulcrum of activity in unravelling galaxy evolution. Our investment in this area is now unparalleled through HST and a new generation of 8-m telescopes equipped for high resolution work. Yet surprisingly little resolved spectroscopy has been done at moderate look-back times. Together with upcoming facilities at mid-IR and sub-mm wavelengths, we can expect to unravel the details of galaxy evolution over $0 < z < 5$ via studies of assembling mass as well as the history of star formation.

I am grateful to my collaborators and students at Cambridge and elsewhere for allowing me to present work done with their assistance. I thank Joss Bland-Hawthorn and Wil van Breugel for their encouragement and patience with this article.

References

- Abraham, R.G. et al 1997 MNRAS, 272, L47.
 Abraham, R.G. et al 1999a MNRAS, 303, 641.
 Abraham, R.G. et al 1999b MNRAS, in press (astro-ph/9811476)
 Aguirre, A. 1999 ApJ, 512, L19.
 Allen, S.W. 1998 MNRAS, 296, 392.
 Aragón-Salamanca, A., Baugh, C. & Kauffmann, G. 1998 MNRAS, 297, 427.
 Bahcall, N. et al 1999 Science, 284, 1481.
 Barger, A., et al 1999 AJ, 117, 102.
 Baugh, C., Cole, S. & Frenk, C.S. 1996 MNRAS, 283, 1361.
 Baugh, C. et al 1998 ApJ, 498, 504.
 Beckett, M. et al 1998 in *Infrared Astronomical Instrumentation*, ed. Fowler, A.M., S.P.I.E., 3354, 431.
 Blumenthal, G. et al 1984 Nature, 311, 517.
 Brinchmann, J. et al 1998 ApJ, 499, 112.
 Broadhurst, T.J., Ellis, R.S. & Glazebrook, K. 1992 Nature, 355, 55.
 Brunner, R.J., Connolly, A.J. & Szalay, A.S. 1999 in *Photometric Redshifts & High Redshift Galaxies*, eds. Weymann, R. et al, ASP Conference Series, in press (astro-ph/9907403).
 Bunn, E.F. & White, M. 1997 ApJ, 480, 6.
 Cole, S. et al 1994 MNRAS, 271, 781.
 Cole, S. et al 1998, MNRAS, 300, 945.
 Cole, S. et al 1999, MNRAS, submitted
 Colless, M. 1999 in *Large Scale Structure in the Universe*, eds. Efstathiou, G. et al, Phil. Trans. R. Soc. A., 357, 105.
 Colless et al 1999 MNRAS, 303, 813
 Combes, F. 1999 in *Building Galaxies: from the Primordial Universe to the Present*, in press (astro-ph/9904031).
 Cowie, L.L., Songaila, A. & Barger, A. 1999 AJ, in press (astro-ph/9904345).
 Djorgovski, S.G. 1996 in *New Light on Galaxy Evolution*, eds. Bender, R. & Davies, R.L., IAU Symposium 171, Kluwer, p277.

- Dressler, A. 1980 ApJ, 236, 351.
- Driver, S.P. et al 1996 ApJ, 449, L23.
- Efstathiou, G. & Silk, J. 1983 Fund. Cosmic Physics, 9, 1.
- Efstathiou, G. et al 1999 MNRAS, 303, L47
- Eggen, O., Lynden-Bell, D. & Sandage, A. 1962 ApJ, 136, 748.
- Eisenstein, D.J., Hu, W. & Tegmark, M. 1998 ApJ, 504, L57.
- Eke, V., Cole, S. & Frenk, C.S. 1996 MNRAS, 282, 263.
- Ellis, R.S. 1997 Ann. Rev. Astron. Astrophys., 35, 389.
- Ellis, R.S. & Abraham, R.G. 1999 ApJ, submitted
- Ellis, R.S. et al 1996 MNRAS, 280, 235.
- Erben, T. et al 1999 Astron. Astrophys. in press (astro-ph/9907134).
- Frenk, C.S. et al 1985 Nature, 317, 595.
- Garnavich, P. et al 1998 ApJ, 493, L53.
- Glazebrook, K. et al 1995 MNRAS, 275, L19.
- Gunn, J.E. & Oke, J.B. 1975 ApJ, 195, 255.
- Gunn, J.E. & Tinsley, B.M. 1975 Nature, 257, 454.
- Gunn, J.E. & Weinberg, D.H. 1995 in *Wide Field Spectroscopy & the Distant Universe*, eds. Maddox, S.J. & Aragón-salamanca, A., World Scientific, p3.
- Gott, J.R. et al 1976 Sci. Amer., 234, 62.
- Jain, B. & Seljak, U. 1997 ApJ, 484, 560.
- Jimenez, R. et al 1999 MNRAS, 305, L16.
- Kaiser, N. & Squires, G. 1993 ApJ, 404, 441.
- Kaiser, N., Squires, G. & Broadhurst, T.J. 1995 ApJ, 449, 460..
- Kaiser, N. et al 1999 ApJ, in press (astro-ph/9809268)
- Kauffmann, G. 1995 MNRAS, 264, 153.
- Kauffmann, G. & Charlot, S. 1998a MNRAS, 297, 23.
- Kauffmann, G. & Charlot, S. 1998b in *The Birth of Galaxies*, in press (astro-ph/9810031).
- Kauffmann, G., White, S.D.M. & Guiderdoni, B. 1993 MNRAS, 264, 201.
- Kauffmann, G. et al 1999 MNRAS, in press (astro-ph/9809168).
- Kristian, J., Sandage, A. & Westphal, J. 1978 ApJ, 221, 383.
- LeFevre, O. et al 1999 MNRAS, in press
- Lilly, S.J. et al 1995 ApJ, 455, 108.
- Lilly, S.J. et al 1998 ApJ, 500, 75.
- Lineweaver, C. 1998 in *Cosmic Microwave Background*, in press (astro-ph/9810334).
- Madau, P. 1999 in *VLT Opening Symposium*, ESO Publications in press (astro-ph/9907268)
- Madau, P. et al 1996 MNRAS, 283, 1388.
- Marzke, R. 1999 in *Photometric Redshifts & High Redshift Galaxies*, eds. Weymann, R. et al, ASP Conference Series, in press.
- Mellier, Y. 1998 Ann. Rev. Astron. Astrophys., 37 (in press, astro-ph/9812172).
- Menanteau, F. et al 1999 MNRAS, in press (astro-ph/9811463)
- Mo, H.J., Mao, S. & White, S.D.M. 1998 MNRAS, 295, 319.
- Mould, J. et al 1994 MNRAS, 271, 31.
- Narayan, R. & Bartelmann, X 1996 Lectures on Gravitational Lensing (available from astro-ph/9606001).
- Ostriker, J.P. & Steinhardt, P.J. 1995 Nature, 377, 600.
- Patton, D. et al 1998 ApJ, 475, 29.
- Peacock, J.A. 1997 MNRAS, 284, 885.
- Perlmutter, S. et al 1999 ApJ, 517, 565.
- Pettini, M. et al 1998 ApJ, 508, 539.

- Riess, A. et al 1999 AJ, submitted (astro-ph/9907038).
- Sandage, A. 1986 Astron. Astrophys., 161, 89
- Schneider, P. et al 1998 MNRAS, 296, 873.
- Schweizer, F. 1998 in *Galaxies: Interactions and Induced Star Formation*, Saas-Fee Course 26, eds. Kennicutt, R.C. et al, Springer-Verlag, p105.
- Smail, I. et al 1997 ApJ, 479, 70.
- Steidel, C. et al 1996 ApJ, 462, L17.
- Steidel, C. et al 1999 in *Large Scale Structure in the Universe*, eds. Efstathiou, G. et al, Phil. Trans. R. Soc. A, 357, 153.
- Steinhardt, P.J., Wang, L. & Zlatev, I. 1999 Phys. Rev. D., 59, 123504.
- Struck-Marcell, C. & Tinsley, B.M. 1978 ApJ, 221, 562.
- Taylor, K. 1995 in *Wide Field Surveys and the Distant Universe*, eds. Maddox, S.J. & Aragón-Salamanca, A., World Scientific, p15.
- Tegmark, M. 1999 ApJ, 519, 513.
- Thomas, P. et al 1998 MNRAS, 296, 1061.
- Turner, M.S. 1999 in *Black Holes & Relativistic Astrophysics*, ed. Sugiyama, N., 49th Yamada Conference, Universal Academy Press (Tokyo), in pres (astro-ph/9811447).
- Tytler, D. 1997 in *The Early Universe with the VLT*, ed. Bergeron, J., Springer, p3.
- de Vaucouleurs, G. 1977 in *Evolution of Stellar Populations*, eds. Larson, R. & Tinsley, B.M., Yale University Press, p1.
- Vogt, N. et al 1999 in preparation
- Weil, M.L., Eke, V.R. & Efstathiou, G. 1998 MNRAS, 300, 773.
- White, S.D.M. & Rees, M.J. 1978 MNRAS, 183, 341.
- White, S.D.M., Efstathiou, G. & Frenk, C.S. 1993 MNRAS, 262, 1023.
- Whitmore, B.C., Gilmore, D. & Jones, C. 1993 ApJ, 407, 489.
- Wright, E.L. et al 1992 ApJ, 396, L13.
- Zepf, S.E. 1997 Nature, 390, 377.

Tunable imaging filters & scientific applications

J. Bland-Hawthorn

Anglo-Australian Observatory

Abstract. Tunable imaging filters have already begun to make great strides in night time astronomy. Here we briefly review the available technologies with a specific focus on the Taurus Tunable Filter (TTF) and a low-resolution, Lyot filter customized for observational cosmology. Tunable filters – with their ability to tune precisely to a given line with a bandpass optimized for the intrinsic line width – are already producing the deepest images to date of extended emission-line nebulae. They are vastly more efficient for redshift-targetted surveys (e.g., high redshift clusters) compared with existing multi-aperture spectrographs. This is particularly true for a Lyot filter at the prime focus of a major telescope. We demonstrate the power of tunable ‘differential’ imaging when combined with CCD charge shuffling. This approach is technically superior to multi-beam cameras in that the same optical path is used for both source and reference image. Tunable imaging polarimetry should also be possible, albeit for a restricted field.

1. Introduction

It is surprising that so few observatories have moved beyond conventional imaging technology. Traditionally, nocturnal astronomy is restricted to broad spectral bands ($R \sim 5$) with the occasional forays into narrow bands ($R \sim 50$) for the purpose of isolating spectral lines. Few astronomers have shown interest in advancing the technology, believing one suspects that it has come as far as it can. This contrasts with other physical sciences – solar physics, remote sensing, underwater communications, atmospheric physics – where a wide variety of imaging technologies are routinely used. It is not obvious why nocturnal imaging should have reached the present impasse except that broadband filters are available for use at all telescopes, while narrow filters are made to order by commercial companies who guard their secrets well. Monolithic – particularly interference – filters are expensive so that only a few are purchased for use at specific wavelengths and/or redshifts.

But there are important reasons for pursuing tunable filter imaging, as we have learnt since the mid 1990’s at the Anglo-Australian 3.9m and William Herschel 4.2m telescopes. Conventional imaging has a major limitation in that images are taken sequentially which is far from ideal at even the best sites. Even small detector, instrument or atmospheric variations lead to systematic error between images. A better approach is a multi-band camera so that different bands are observed in parallel. However, even this is not ideal in that the optical

path is different for each of the filters. A tunable filter provides a powerful alternative approach since band switching can be linked to charge shuffling with the CCD (Bland-Hawthorn & Jones 1998). This is a truly differential technique which leads to much smaller systematic errors than is possible with conventional imaging.

In recent semesters, the TTF has been allocated 20% of observing nights for a diverse range of programmes. Recent published work includes: time series analysis of an x-ray binary, the first detection of weather on a brown dwarf, the HII region luminosity function in dwarf irregulars, extended ionized filaments in spiral galaxy haloes, a very extended ionized jet in a nearby Seyfert, the most extensive ionization cone in an active galaxy, very extended cooling flows, and diffuse plasma flows around radio galaxies. At higher redshifts, the TTF has determined the universal 'star formation rate' density at several discrete redshift intervals from $z \sim 0.1$ and $z \sim 2$, detected the largest ionized nebula associated with a nearby quasar, discovered a cluster of star-forming objects about a $z \sim 1$ quasar, and detected a gravitationally lensed galaxy at $z = 3.44$. Some of these results were discussed at this meeting by Baker, Cecil, Hewett, Shopbell, Tadhunter, Tinney and Veilleux.

A very different type of tunable imager, the Lyot filter, is (almost) the ultimate approach to exploiting the widest possible field with a given telescope (e.g., Palunas 2000). Spectral passbands are degraded in converging beams which is unfortunate as the widest fields are achieved at the fastest f/ratios. But, remarkably, beams as fast as f/2 can be compensated by crossed birefringent elements such that even a constant sub-Angstrom bandpass is possible over a degree-sized field! Very few telescopes offer much more than a wide-field imager at prime focus simply because it is very difficult to exploit the fast beam spectroscopically. This is unfortunate since these same telescopes are building or already using very large, very expensive CCD mosaics for the sole purpose of low-resolution, broad band imaging. The Lyot tunable filter solves this problem, as we show.

We begin with a clarification before summarizing the properties of an ideal tunable filter.

2. Tuning: a subtle distinction

We should be clear on what is meant by a 'tunable filter.' The ideal filter is a device which is able to isolate a discrete spectral band with an arbitrary bandpass $\delta\lambda$ at an arbitrary wavelength λ_0 over a very broad, continuous spectral range, preferably with a response function which is identical at all wavelengths. In this respect, the tilting interference filter is much the worst form of tunable filter. The spectral range covered almost negligible, and the filter profile varies with tilt angle. Some devices (e.g., Lyot, linearly variable filter) work only at a fixed resolving power but have the ability to tune over a wide spectral window.

The better devices allow for a wide selection of resolving powers over a wide range of wavelengths. The TTF solves this in two stages with a blue and a red device (see Fig. 1). The equation of interference is (see Fig. 2),

$$m\lambda = 2\mu\ell \cos\theta \quad (1)$$

Figure 1. The resolving powers achievable with the Taurus Tunable Filter.

Figure 2. The internal structure of the Taurus Tunable Filter.

from which it follows that

$$\frac{dR}{R} = \frac{dm}{m} = \frac{d\ell}{\ell} \quad (2)$$

where $R = mN$ for an instrumental finesse of N . The plates can be scanned over the range $1.5 - 15\mu\text{m}$, and the orders of interference span the range $4 - 40$, such that the available resolving powers are $100 - 1000$. It is important to recognize that achieving a large physical scan range is only part of the problem. The range *must* start at separations of only a few wavelengths to be useful. Several technological developments have been necessary in order to achieve both goals within the TTF (Pietraszewski 2000).

3. The ideal filter

Here are some the characteristics which make for an ideal tunable filter:

image quality, high throughput, stable;
monochromatic band, widest possible field of view;
tunable, identical, square profile;
broad spectral coverage, fast response;
large span in resolving power;
minimal reflectance phase;
minimal stray light problem;
easy to calibrate data.

As one learns in undergraduate physics, the ideal is never achieved in practice. It's hard to envisage any technique which could achieve all of the above goals, short of collecting and time-tagging all photons individually (e.g., Perryman 2000). To pinpoint a few key issues, some of the birefringent materials (e.g., LiNO_2), which have great application in solid etalon and acousto-optic filters, have microstructure which can inhibit good image quality. There are only a few filters (Sölc, Michelson) where the instrumental profile can be adjusted (e.g., squared off) straightforwardly. Fixed-element birefringent filters (e.g., Lyot, Sölc) offer only a single resolving power at a specific wavelength. The Fabry-Perot tunable filter is ultimately limited by the finite coating thickness between the mirrors, so it really only achieves the lowest interference orders ($m < 5$) at infrared wavelengths. The Michelson interferometer reaches the lowest orders even at optical wavelengths; the optical path length is adjusted between mirrors which are widely spaced.

All spectroscopic techniques rely ultimately on the interference of beams that traverse different optical paths to form a signal. The prism uses a quasi-infinite number of beams whereas the grating uses a finite number set by the number of grooves. The Fabry-Perot and the Lyot use a smaller number set by the instrumental finesse. Bell (1972) notes that as the number of beams decreases, the throughput of the spectrograph increases. Since we require two beams for interference, Bell concludes that two-beam (e.g., Michelson) interferometers are the ultimate spectrographs. In fact, the Michelson interferometer does come closest to achieving the above goals.

4. The physics of tunable imaging

A rich variety of physical phenomena can produce a finite spectral band: absorption, scattering, diffraction, evanescence, birefringence, acousto-optics, single layer and multi-layer interference, multi-path interferometry, polarizability, and so on. Most of the available optical filter technologies are listed below:

- colour film*
- coloured glass*
- hole burning device* (e.g., Keller 2000)
- Christiansen filter*
- linear / circular variable filter*
- multi-layer dielectric filter* (e.g., Cianci et al. 2000)
- Fabry-Perot interferometer* (e.g., Pietraszewski 2000)
- Michelson interferometer* (e.g., Maillard et al. 2000)
- acousto-optic filter*
- solid etalon filter*
- solid Michelson filter*
- facet optics* (e.g., Tonry 2000)
- generalized resonant grating filter*
- sub-lambda evanescent grating filter*
- volume phase holographic grating filter* (e.g., Barden 2000)
- Lyot-Öhman filter* (e.g., Palunas 2000)
- generalized Lyot filter*
- generalized Sölc filter*
- liquid crystal filter*

One needs to cast a wide net to find some of these, e.g., grating filters are largely the domain of covert research into bank note security, although there is a very interesting overlap with biologists who study the chromatic properties of insect wings, fish scales and so on.

This article can only begin to touch on what is clearly a burgeoning industry. There is a bewildering array of future possibilities, e.g., tunable crystal lattice structures (Sheng et al. 1999) with many possible variants. The physical processes involved are so diverse that it proves very difficult to describe the different technologies within a consistent mathematical framework. A reference text and an article for the *Reviews of Modern Physics* are already well advanced.

5. Summary of filter technologies

Here, we provide a brief outline of some key technologies.

Interference filter: The principle relies on a dielectric spacer sandwiched between two transmitting layers (single cavity). The substrates are commonly fused silica in the ultraviolet, glass or quartz in the optical, and water-free silica in the infrared. Between the spacer and the glass, surface coatings are deposited by evaporation which partly transmit and reflect an incident ray. Each internally reflected ray shares a fixed phase relationship to all the other internally reflected rays. For constructive interference, for a wavelength λ to be transmitted, it must

satisfy $m\lambda = 2\mu\ell \cos \theta_R$ where θ_R is the refracted angle within the optical spacer, and the optical gap is the product of the thickness ℓ and refractive index μ of the spacer. The construction of these filters has undergone a revolution through the use of dielectric, multi-layer thin film coatings (e.g., Cianci et al. 2000), and a proper description is more involved. All such filters can be tuned through a small wavelength interval ($\delta\lambda/\lambda = -\theta_R^2/2\mu^2$), which amounts to no more than 2% λ in practice. Suffice it to say, interference filters make for poor tunable devices.

Acousto-optic tunable filter (AOTF): These are electronically tunable filters that make use of acousto-optic (either collinear, or more usefully, non-collinear) diffraction in an optically anisotropic medium. The acceptance angle is generally larger than the Fabry-Perot.

AOTFs are formed by bonding piezo-electric transducers such as lithium niobate to an anisotropic birefringent medium. The medium has traditionally been a crystal, but polymers have been developed recently with variable and controllable birefringence. When the transducers are excited to 10-250 MHz (radio) frequencies, the ultrasonic waves vibrate the crystal lattice to form a moving phase pattern that acts as a diffraction grating. The diffraction angle (and therefore wavelength) can be tuned by changing the radio frequency. These devices are often water cooled to assist the thermal dissipation, although this is less important in the UV where AOTFs are particularly useful. The largest devices are 25mm in diameter as it proves to be difficult to maintain a uniform acoustic standing wave over larger areas. An additional problem is the $15\mu\text{m}$ structure in the LiNO_2 crystal which is not always optimal for good image quality.

Liquid crystal filter: Liquid crystal filters are now commercially available. The biggest we have seen has a clear aperture of 40mm, requires about 5V to scan a single order of interference, and appears to have good image quality. The tunable band for a single stage device is about $R \sim 5$ but can be tuned over the optical window. The peak transmissions are 30% or less.

Michelson filter: We forego a discussion of Fourier Transform (FT) interferometers as these are covered elsewhere in these proceedings (Maillard et al. 2000; Bennett 2000; Graham 2000). A point not made in any of these articles is that the FT device really comes into its own in space. At ground level, the atmospheric emission contributes noise at all frequencies in the Fourier domain. The TTF concentrates its effort in very dark regions between the atmospheric band-heads. It would be interesting to compare the performance of a heterodyne FT interferometer with the TTF.

Fabry-Perot filter: These interferometers are already widely used in many fields and at many observatories. They remain popular because excellent devices are commercially available and are relatively straightforward to use. The TTF is the first such device employed over a wide range of resolutions (Bland-Hawthorn & Jones 1998). The specifics of the TTF are given in Appendix A. They are by no means ideal as tunable filters (poor instrumental profile, small monochromatic field, ghost problems), but are likely to remain popular for years to come, at least for ground-based use.

Lyot filter: The Lyot filter is discussed in detail below. It consists of a chain of independent segments each composed of a birefringent plate placed between

parallel polarizers. Each stage is twice the thickness of the previous element in the optical train. A quarter-wave plate is included within each segment to allow the filter to be tuned. Each segment is rotated through twice the angle of the previous element. Lyot filters have been made up to 30 cm clear aperture for work-over a narrow spectral interval. Broadband instruments have never exceeded 10 cm due to the difficulties and expense of making achromatic retarders large enough. With the use of half-wave plates, the Lyot can be made to operate directly in beams as fast as $f/2$.

Sölc¹ filter: These highly non-intuitive filters use only two polarizers and a chain of identical retarders with varying position angles (Sölc 1959; Evans 1958). There are folded (zigzag) and fanned designs with the former having the better performance (Beckers & Dunn 1965). Title (1999, private communication) has made a tunable Sölc filter with 7 cm clear aperture. It has the extraordinary capability of tuning the spectral profile: an n -element Sölc filter can have a profile that is determined by n Fourier coefficients. The same can be achieved with polarizing filters by proper choice of crystal lengths. To our knowledge, a wide-angle Sölc filter using half-wave plates has not been attempted although it is entirely feasible.

Solid etalon filter: These are single cavity Fabry-Perot devices with a transparent piezo-electric spacer, e.g., lithium niobate. The thickness and, to a lesser extent, refractive index can be modified by a voltage applied to both faces. For low voltage systems, tilt and temperature can be used to fine-tune the bandpass. High quality spacers with thicknesses less than a few hundred microns are difficult to manufacture, so that etalon filters are normally operated at high orders of interference.

Resonant grating filter: These novel filters are inspired by the diffractive colours in many insects, and constitute dielectric gratings with three-dimensional, sub-micron microstructure (Magnusson & Wang 1993). The zeroth order reflection exhibits a broad to intermediate bandwidth ($R \sim 20$), is highly polarized and maintains useful efficiency over a $\pm 30^\circ$ tilt (or rotation) range. Gale (1998) presents one grating design that produces a roughly self-similar bandpass from 450 to 850nm over this tilt range. Grating filters, and their close relatives, evanescent gratings, show great promise but most have yet to leave the drawing board. However, since much of the research is driven by bank note security, we anticipate rapid progress. It is noteworthy that volume phase holographic gratings – in reflection – can produce a highly efficient grating filter through Bragg diffraction (Barden 2000).

6. Tunable filters: pitfalls

6.1. The phase problem

The tunable filter tries to produce a monochromatic image over a wide field but this is difficult to achieve in practice. Scratch the surface and we find that most filters obey eqn. 1 somewhere in the process. The $\cos \theta$ term constitutes the

¹The proper Czech pronunciation is ‘Sholtz’.

Figure 3. (i) An [OIII] image of CAL 83 and its LMC environs taken through a $14\text{\AA}$ FWHM filter at the CTIO 1.5m, and (ii) with Taurus-2 at the AAT 3.9m (right) using an etalon out of band ($7\text{\AA}$ bandpass). The etalon interference pattern produces an annular ring which shows unambiguously that the faintest emission in the CTIO image arises from the [OIII] $\lambda 5007$ line.

‘phase problem.’ The TTF has a phase shift of about $15\text{\AA}$ (at $H\alpha$) over its $10'$ field.

There are interesting cases of filters which have no phase effect, e.g., absorption or scattering filters (cf. Keller 2000). The Christiansen filter is made from a matrix of glass and resin beads. The dispersion curves of glass and resin cross at one wavelength. Light at this wavelength is forward-scattered regardless of incidence angle, whereas light at all other wavelengths is scattered incoherently. These filters have low transmission and very limited application.

There is one instance where the phase effect is a strength rather than a weakness. It can help in discriminating between spurious scattered light and diffuse emission intrinsic to the source. Fig. 3 presents two emission line images for the supersoft x-ray source CAL 83. Remillard et al. (1995) have argued that these sources produce hugely extended nebulae with no well defined Stromgren-like boundary. In their narrowband image, they only reach a surface brightness of a few $\times 10^{-17} \text{ erg cm}^{-2} \text{ s}^{-1} \text{ arcsec}^{-2}$; systematic uncertainties made it difficult to determine if the extended emission was real. In Fig. 3(ii), we show that the background radiation, detected at much higher sensitivity in the same exposure time, forms an annulus which can only happen in the presence of an emission line, in this case redshifted [OIII] $\lambda 5007$ emission.

But in most instances, one would far prefer a monochromatic field. Attempts to combat phase all end in disaster. One can try to counteract the $\cos \theta$ term with curved or Fresnel plates (changing ℓ), or higher or graded index

(changing μ), or slower camera or larger beam size (changing θ), but all of these compound the problem rather than solve it.

The Lyot filter offers a brilliant solution to the phase problem, as we discuss below.

6.2. Shaping the bandpass

The sharp core of almost all tunable filter profiles is a problem. A small amount of flatness in the core ensures that we avoid the problem of narrowly missing most of the line flux from the source. Since the Michelson interferometer obtains its data in the frequency domain, the profile can be partially squared off at the data reduction stage through a suitable choice of convolving function. The Sölc filter achieves this through a complex arrangement of fan-shaped or zigzagged retarder elements (§5; Evans 1958). In theory, all band-limited functions can be squared off, but in practice this is difficult. There are three experimental solutions:

(i) The use of near-identical multi-cavities used in series which requires exceedingly tolerances on matching the interference coatings. This is essentially the theoretical basis of the Butterworth function (see Jones & Bland-Hawthorn 1998).

(ii) A multi-pass arrangement whereby a tilted mirror is placed on the back face of the filter to pass the light through the interfering cavity again. This normally results in a restricted field.

(iii) The easiest practical solution which results in lower overall throughput is to oscillate the tuning about a fixed configuration (Caplan 1995): e.g., small oscillations in rotation (Lyot), small plate oscillations (Fabry-Perot).

6.3. The stray light problem

Imagers are notoriously prone to diffuse scattered light and internal ghost reflections (Bland-Hawthorn & Cecil 1997). Elaborate schemes are sometimes needed to ascertain what is real and what is an artefact. For example, some ghosts reflect about the optical axis. A shift between exposures of x arcsecs leads to ghosts moving by $-2x$ and therefore these objects are easily identified. Far more thought should go into the design of AR coatings in camera systems (Cianci et al. 2000). A particular challenge is the air-silicon detector interface which results in a diametric 10% ghost image in many systems. The most recent version of Zemax includes options to help trace the dominant scattered rays.

6.4. Fringeing in narrow bands

CCD fringeing can be a particular problem in narrow bands. We have made movies with the TTF of fringeing effects while tuning the spectral band and central wavelength. The interference pattern can undergo complex behavior like shifting sanddunes or oil slicks. There are regular spectral intervals where the fringeing goes through maxima and minima. The new generation of thick (e.g., high resistivity, deep depletion) CCDs, especially the MIT LL devices, are much less susceptible to fringeing.

6.5. The atmosphere in narrow bands

At the AAT, we have learnt the importance of accurate simulators like the TTF calculator at www.aao.gov.au/cgi-bin/ttf as there are various pitfalls to avoid. For example, in the R and I bands, there are narrow water absorption features which can produce low transmission in discrete narrow bands. The TTF calculator includes a full model for these features based on echelle data.

7. Tunable filters: advantages

7.1. Differential imaging

There are many operational advantages of the tunable filter compared to conventional imagers. The instrumental profile is usually identical for all bands over the wavelength range which greatly assists the differential imaging process. When linked to charge shuffling, two discrete bands can be obtained interleaved in time (e.g., [OIII]5007 and $H\alpha$) such that the differential errors pixel to pixel are much smaller than possible with separate filters. The differential imaging is typically so good that we can obtain excellent data even through light cirrus or with a rapidly changing sky background. In other words, photometry of the individual lines is worthless, but the pixel to pixel *comparison* is excellent. The TTF offers many charge-shuffle modes (Bland-Hawthorn & Jones 1998), including nod & shuffle strategies as discussed by Tinney (2000).

7.2. Sky background

A tunable filter which allows for a wide range of interference orders, like the TTF, can be very efficient in obtaining emission or absorption line detections. Such programmes typically require an ‘off’ (i.e. a continuum band) for comparison. The TTF allows rapid changes between narrow and broad bands, such that most of an exposure can be allocated to the narrow band, with a small overhead for the broad band. Furthermore, the broad band can oscillate to either side of the narrow band to remove a sloping spectral background.

We find today that the vast majority of TTF exposures are sky limited. The ability to choose the bandpass and tune to an exact wavelength allows us to optimize the experiment. The TTF web calculator makes this very easy to do in real time.

A curious advantage was realized only recently with the design of the 5-stage Lyot filter for the AAT. While the entrance polarizer immediately removes half the light, the overall throughput to *unpolarized* light is comparable to the TTF, as there are only half a dozen air-glass interfaces ahead of the prime focus. TTF programmes routinely exploit the very dark sky bands between atmospheric OH band heads. In bright moonlight, the sky background at roughly 90° to the moon’s direction is almost entirely scattered light in the inter-bandhead regions. For orthogonal sight lines from the moon, if we align the Lyot entrance polarizer with the moon’s direction, the sky background should be comparable to the dark sky. There are few examples in optical astronomy where dark sky astronomy can be carried out all year round (cf. echelle spectroscopy).

Figure 4. (i) In a properly matched optical system, the $\mathcal{A}_1.\Omega_1$ product of the detector measured at the telescope aperture is equal to the $\mathcal{A}_2.\Omega_2$ product of the telescope measured at the detector. The telescope diameter is D and the focal length of a spherical mirror is f ; the detector size is d . The $\mathcal{A}.\Omega$ invariant is the throughput or étendue of the system. (ii) Angle convention for an off-axis ray incident on an optical element.

7.3. Time series imaging

There are very interesting applications for tunable filters used in time series mode. In an effort to keep this article to a finite length, we refer the reader to Deutsch et al. (1998) and Tinney & Tolley (1999).

7.4. Object finding codes

Finally, it is hard to overstate the importance of object-finding codes like *Sextractor* (Bertin et al. 199x). Many TTF programmes have benefited enormously from the easy application of this code. Recent versions permit objects found in one image to be associated with objects detected in successive images. This is precisely the function required in analysing TTF scans where a series of images are taken in consecutive spectral bands, e.g., to look for emission line objects in a narrow redshift interval (e.g., Jones 1999).

8. Tunable Lyot Filter

8.1. Wide-field imaging

Traditionally, wide-field surveys have been restricted to broad photometric bands. Substantially narrower bands provide a technical challenge. Survey efficiency is often loosely stated in terms of the $\mathcal{A}\Omega$ product, where $\mathcal{A}$ is the telescope area, and Ω is the total solid angle of the sky survey. But its traditional use is to measure system throughput (see Fig. 4(i)). For a single telescope pointing,

$$\mathcal{A}_1.\Omega_1 \approx \mathcal{A}_2.\Omega_2 \approx \left(\frac{\pi^2 d^2}{4\mathcal{F}^2}\right) \quad (3)$$

where d is the detector size in mm, and $\mathcal{F}$ is the beam focal ratio ($0.5 \leq \mathcal{F} < \infty$).

For $\mathcal{A}\Omega$ to be a useful measure of survey efficiency, there are certain qualifications. We assume that the sources under study are (a) resolved by the instrument and the pixel sampling, (b) detected in a reasonable exposure time, and (c) the signal-to-noise ratio (SNR) increases as $(\text{time})^{\frac{1}{2}}$, in contrast to observations limited by detector read noise or charged-particle background. The RHS of eqn. 3 could reasonably include a natural threshold of the form $\mathcal{H}(S \geq S_L)$ where S_L is the limiting acceptable SNR, and $\mathcal{H}$ is the Heaviside operator for which $\mathcal{H}(S < S_L) = 0$, and $\mathcal{H}(S \geq S_L) = 1$.

For any telescope, fast beams are much more efficient for survey work. In eqn. 3, to sidestep subtleties involving field-expanded foci, we adopt $\mathcal{F} \geq 2$ as faster beams cannot be properly compensated by Lyot's method. But fast beams degrade the energy interval isolated through spectroscopic interference, particularly for interference filters. Interestingly, this is not aided by going to collimated beams since the focal reducer preserves the field angle θ of the incoming beam. Since the path length has an angular dependence, interference filters exhibit a phase effect (shifting central wavelength) over the field of view. An impractical solution, but not without precedent, is to produce an interference filter which covers the entrance aperture of the telescope. The UK Schmidt 35cm H α ($\mathcal{R} \sim 100$) filter (Parker & Bland-Hawthorn 1997) is the largest and most expensive interference filter made for astronomical purposes. By a simple scaling, a single monolithic filter for a 1m class telescope would cost in excess of US\$1M, rather more than the cost of the fully tunable optical system (filter + top end) proposed here.

8.2. Basic theory

There exist excellent discussions on the basic principles of the Lyot filter (Lyot 1933, 1944; Evans 1949; Beckers & Dunn 1965; Title & Rosenberg 1979, 1981). These texts assume a proper understanding of how a wavefront propagates through a birefringent medium. There are many approaches to a complete mathematical treatment, including Jones and Stokes vectors, Jones and Mueller matrices, the Poincaré sphere and impulse response analysis, to name a few. Here, we attempt an intuitive approach; a rigorous treatment is given elsewhere (in preparation). In order to arrive at the low-resolution filter proposed here, there are four key elements: (a) interference through birefringence, (b) wide-field compensation, (c) wavelength tuning, (d) retardance through differenced elements.

Figure 5. A montage of Lyot stages demonstrating the different aspects of building a Lyot filter. The elements are the polaroid (P), the retarder (R), the half-wave plate ($\frac{\lambda}{2}$), and the quarter-wave plate ($\frac{\lambda}{4}$). The black-tipped arrow shows the orientation of key elements; the white-tipped arrow indicates the direction of the fast axis; the optic axis runs horizontally. (i) The simplest Lyot filter producing the cosine behaviour in eq. 4. (ii) The wide-field compensated filter in which the retarder is split and crossed. (iii) A single-stage filter with tuning capability. (iv) A wide-field compensated filter with tuning capability. (v) A fragile, thin element filter with tuning capability. (vi) An equivalent unit to that presented in (v) with the aid of thicker elements through the differencing principle. The crossed retarders are sufficiently thick to require wide-field compensation.

The underlying principle of Lyot filters is that light originating in a single polarization state can be made to interfere with itself. The first element is a polaroid and forces unpolarized light into two orthogonal polarization components.² If the polaroid is oriented at 45° to the fast and slow axes of a birefringent crystal, one component can be delayed with respect to the other. An exit polaroid, aligned with the entrance polaroid, recombines the orthogonal components to effect interference.

Crystalline materials are usually, but not exclusively, birefringent in that the refractive index is different along two axes within the crystal. In a positive uniaxial crystal (e.g., quartz), the refractive index experienced by the ordinary ray on the 'fast' axis, n_o , is lower than for the extraordinary ray, n_e , along the 'slow' axis (see Fig. 4(ii)). The birefringence $b = n_e - n_o$ leads to a time delay $\delta t = bd/c$ where c is the speed of light in vacuo, and d is the thickness of the element. Retardance is expressed in angular units, i.e. $r = 2\pi bd/\lambda$.

For light traversing the simple system in Fig. 5(i), the transfer function is given by Malthus' law,

$$T(\lambda) = \cos^2(\pi bd/\lambda). \quad (4)$$

Lyot (1933) realized³ that a cascading series of aligned birefringent units could be used to isolate a spectral element. From equation 4, the easiest way to arrange this is to make each element twice the thickness of the preceding element (see Fig. 6). The interleaved polaroids are essential to the rejection of out-of-band emission.

Clearly, the free spectral range is determined by the thinnest element, and the spectral resolution by the thickest element. Lyot (1944) demonstrates that the instrumental response of a series of q elements as

$$T(\lambda) = \frac{1}{4^q} \frac{\sin^2(2^q \pi b d_o / \lambda)}{\sin^2(\pi b d_o / \lambda)} \quad (5)$$

It follows that the free spectral range is $\Delta\lambda = \lambda^2/bd_o$, the finesse is $\mathcal{N} = 1.13 \cdot 2^q$, and the spectral element $\delta\lambda$ is the ratio of these two quantities. These quantities can be related directly to equivalent parameters for Michelson and Fabry-Perot interferometers (Bland-Hawthorn & Cecil 1997).

The birefringence of certain crystals provides a large gain in the lever arm for isolating a spectral element compared with, say, the reflecting plates of the TTF (Bland-Hawthorn & Jones 1998). The resolving power of air-spaced Fabry-Perot plates a distance d_o apart is $\mathcal{R} = 2d_o\mathcal{N}/\lambda$, where $\mathcal{N}$ is the reflective finesse of the coated plates. This compares with $\mathcal{R} = bd_o\mathcal{N}/\lambda$, where $\mathcal{N}$ is now the Lyot finesse. The Lyot element with thickness d_o determines the order separation (fringe spacing).

For two commonly used materials, MgF_2 and crystal quartz, $b \approx 0.01$, which means that, for the same resolving power, the thinnest element of the Lyot is

²The first polaroid results in a 50% light loss, but it should be remembered that the overall throughput of a prime focus Lyot filter is 30% or better, comparable to pre-existing tunable systems which involve complex optical paths.

³The principle was independently discovered by Öhman (1938).

Figure 6. The transmission profile of a simple Lyot cascade with (top to bottom) 1, 2, 3, 4 and 5 stages. The thinnest element has retardance R and the following elements have thicknesses which are multiples of this element. The polaroids P are aligned with each other and oriented at 45° to the retarders.

two orders of magnitude thicker than the equivalent plate spacing of the TTF. Calcite ($b = -0.18$) provides an ever larger gain factor but results in overly thin optical elements. If our aim is to block neighbouring orders with conventional $\mathcal{R} \approx 5$ filters, the required TTF sub-micron spacings cannot be achieved in practice, and are subject to reflectance phase effects (Jones & Bland-Hawthorn 1998). For the Lyot, it is possible to optically bond sub-mm thick elements to a supporting substrate, or to use thicker retarding elements in a subtractive arrangement (Fig. 5(vi)) to achieve the same equivalent optical thickness.

8.3. Wide-field compensation

A narrow spectral band cannot be isolated in a fast beam using interference or etalon filters. This is because the path length through the resonating cavity cannot be equalized for all off-axis angles. Remarkably, off-axis path lengths *can* be equalized with crossed birefringent elements (Lyot 1944; Evans 1949) even for resolving powers as high as $\mathcal{R} \sim 10^5$ (Title & Rosenberg 1981). Wide-field compensation has been demonstrated to be extremely effective for narrowband

imaging by several groups, particularly in the fields of remote sensing and solar astronomy.

An intuitive explanation for why the method works is hard to come by. To begin, consider how wide-angle, zero-wave retarders are made. The retarder element in Fig. 5(i) is split perpendicular to the propagation axis into two equal thickness elements. The pieces are then bonded together after rotating the fast axis of the second element to line up with the slow axis of the first element. There is no net retardance for any angle through the retarder.

The principle of balancing fast and slow propagation is the primary method of wide-field compensation. In order to avoid the simple cancellation of the zero-wave retarder, a half-wave plate is placed between the split retarder elements (see Fig. 5(ii)) oriented at 45° to the fast and slow axes. Why this works is somewhat involved: the most accessible discussion is given by Title & Rosenberg (1979).

The half-wave plate advances the phase of the *o*-ray by 180° with respect to the *e*-ray. This rotates the plane of polarization by 90° and re-aligns the *o*-ray with the fast axis of the second element. Hence, the *on-axis* time delay of the crossed elements (Fig. 5(ii)) remains *unchanged* from the original system (Fig. 5(i)). It is the *off-axis* rays which benefit in two respects.

While the off-axis behaviour of interference filters is isotropic (circular isochromes), uniaxial crystals have a strong azimuthal dependence (hyperbolic isochromes). For rays at an arbitrary azimuthal angle ϕ (see Fig. 4), the general form of the retardance is

$$r(\theta, \alpha) = r_o \left[1 - \frac{\sin^2 \theta}{2n_o} \left(\frac{\cos^2 \phi}{n_o} - \frac{\sin^2 \phi}{n_e} \right) \right] \quad (6)$$

Note that the retardance changes sign between neighbouring quadrants, and this is the physical basis for wide-field compensation. Remarkably, with the half-wave plate in place, the complex azimuthal behaviour all but disappears. Light which enters the first split element from the direction (θ, ϕ) enters the second element from the direction $(\theta, \phi + 90^\circ)$. From eq. 6, the overall retardance is now

$$r = r_o \left[1 - \frac{\sin^2 \theta}{4n_o} \left(\frac{n_e - n_o}{n_o} \right) \right] \quad (7)$$

The equivalent form for the interference filter is

$$r = r_o \left[1 - \frac{\sin^2 \theta}{2n_o^2} \right] \quad (8)$$

The monochromatic, solid acceptance angle available to a compensated birefringent filter is a factor of $2n_o/(n_e - n_o)$, i.e. orders of magnitude, larger than possible with an interference filter. Moreover, Lyot filters have a much greater $\mathcal{R}.\Omega$ product than available to interference filters and etalon filters.

8.4. Tuning

Title & Rosenberg (1981) demonstrate several schemes for tuning a Lyot filter, none of which are particularly intuitive to follow. The output of a given birefringent element is elliptically polarized light. The trailing quarter-wave plate

transforms this to linearly polarized light whose orientation, ψ , (and therefore wavelength) is a function of the ellipticity, ε , such that

$$\psi = \tan^{-1} \varepsilon \quad (9)$$

where

$$\varepsilon = -\tan \frac{\pi db}{\lambda}. \quad (10)$$

The tuning relationship for a single stage reduces to

$$\psi = \pi db \left(\frac{1}{\lambda} - \frac{1}{\lambda_o} \right) \quad (11)$$

where the natural wavelength λ_o is selected at $\psi = 0$. In practice, tuning is achieved by rotating the exit polarizer. The exit polarizer of one stage constitutes the entrance polarizer of the next stage. Peak transmission is achieved as long as each successive stage is rotated precisely twice the rotation of the preceding thinner element. A gear system which maintains 2:1 synchrony between successive stages presents an interesting mechanical problem, particularly when anti-backlash gearing is incorporated (e.g., Beckers & Dunn 1965). The (ψ, λ) relation for the complete Lyot system over the full optical range is generally more involved than equation 11. Conveniently, the sensitivity of the wavelength tuning to angular errors decreases for the thicker elements.

The Lyot is an opto-mechanical device which makes it fairly straightforward to tune. The original design by Beckers & Dunn (1965), utilized by the GSFC Lyot (Palunas 2000), uses a succession of 2:1 gear ratios. The first element is driven by a stepper motor. The gear wheels attached to this stage rotate differentially the successive stages of the Lyot through connecting axles to gears attached to each stage. Thus, 2^m unique spectral elements requires 2^m half rotations to cover the free spectral range. Thus, wavelength tuning requires only differential rotation of the Lyot stages.

9. The AAT Lyot Filter

The proposed Lyot filter for the AAT f/3.3 prime focus is shown in Fig. 7. A full Zemax optical design shows that the pre-existing correctors are good enough to flatten out the field-expanded, hyperboloidal primary mirror. The retarder elements are to have 160mm clear aperture, and will be made from MgF_2 or sapphire which have ideal birefringent properties. The thinnest element is about 0.25mm which is well within manufacturing tolerances after bonding with neutral glass. The manufacture is actively being investigated here in Australia by the CSIRO division responsible for the LIGO optics. Quarter-wave and half-wave plates are extremely expensive and do not have ideal broadband response. We are exploring sub- λ transmissive gratings which appear to be ideal (thin, broadband, 100% transmission, large aperture). Bleached polaroids give 95% transmission over half the optical range, such that two Lyots will be required to cover the full range. The length of the filter assembly will be about 100mm; the size and aspect ratios are much like those of the TTF.

Figure 7. The basic structure of the proposed 5-stage Lyot filter. The lowest order stages serve as blocking filters to the higher orders. Stage I achieves a very thin element through differenced, thick elements which have been crossed. Stages II and III are simple Lyot elements, where the thinner element is bonded to glass. Stages IV and V use wide-field compensation. All stages are immersed in index-matched oil. Stage I and the outer polaroids are fixed; all other stages rotate differentially about the long axis. The upper rectangle shows the correct aspect ratio of the Lyot filter.

The Lyot orders are to be isolated with the revised Gunn filters – 356/60, 482/138, 626/138, 767/153, 910/137 (central wavelength/bandpass in nm) – to allow full use of the Sloan survey data. An important departure from the GSFC Lyot (Palunas 2000) is that each of the 5 segments will be tuned independently of the others using separate motors and encoders. This will allow us to explicitly correct for temperature and manufacture defects, and minimizes heating of the immersion oil through unnecessary rotations. The Lyot stages are easily aligned with a laser since each stage generates a unique cosine-squared modulation at a fixed wavelength.

10. Scientific gains of tunable filters

10.1. Observational cosmology

Wide-field imaging surveys continue to dominate many aspects of modern astrophysics. This holds true for searches of rare sources (e.g., brown dwarfs, quasars) or high-density sources (e.g., star-forming objects at cosmological redshift). Traditional all-sky surveys in broad bands have been effective in identifying several classes of astrophysical sources (e.g. photographic surveys with Schmidt telescopes). The Sloan Survey (Gunn et al. 1998) promises to identify a much wider class of sources through the use of drift-scanning with a multi-band camera mosaic, which results in much smaller differential errors between photometric bands.

In a new development, tunable filter surveys of emission-line sources at cosmological redshifts (e.g., Jones 1999; Baker et al. 1999) underline the great potential of surveys targetting narrow photometric bands, particularly those which are directed at fields with pre-existing, broadband data. To date, these studies have been confined to 10' fields. For redshift-targetted programmes, tunable filters have an enormous advantage over existing spectrographs since every pixel is used in the search for line-emitting objects. The emission-line objects detected by the TTF in 4 hours (through differential imaging) are at the sensitivity limit of LRIS on the Keck.

A tunable Lyot filter promises an even bigger advantage. It is often overlooked just how large galaxy clusters are anticipated to be at high redshift. If one considers galaxies on turn-around orbits which are just beginning to separate from the general Hubble flow, the sphere of influence of a Coma-like potential projects to $0.5a/(40h_{75}^{-1} \text{ Mpc}) \text{ deg}$ at $z \sim 1$ (e.g., Kaiser 1987), where a is the diameter of the sphere, and $h_{75} = H_0/(75 \text{ kms}^{-1} \text{ Mpc}^{-1})$. The cusps in redshift space produced by outlying galaxies are sensitive to fundamental cosmological constants. Moreover, the VIRGO Consortium simulations (e.g., Jenkins et al. 1998⁴) reveal that supergalactic structure resides on scales of several degrees. If star-forming, dwarf galaxies or Population III star clusters trace the ‘foothills’ of large-scale structure, it is not unreasonable to expect that future narrowband surveys on degree scales could directly observe the anticipated supercluster networks.

⁴See <http://star-www.dur.ac.uk/~frazerp/virgo/virgo.html>

Loeb & Rybicki (1999) have recently proposed that before the epoch of cosmological reionization, the first galaxies and quasars could possess large, polarized haloes of Ly α radiation. These photons are expected to scatter off neutral atoms expanding with the Hubble flow. This raises the question of whether the low surface brightness (~ 1 Rayleigh) haloes could even be detected in polarized light. The anticipated sources will be invisible at optical wavelengths: the escaping Ly α radiation emerges in the near-IR. In Appendix B, we propose a basic design for a high throughput, low-resolution, tunable imaging polarimeter. If photometric redshifts can reliably identify likely candidates at $z > 5$, this could be a worthwhile experiment.

10.2. General programmes

Wide-field tunable imagers open up many interesting avenues. This list is off-the-cuff and far from exhaustive:

- The detection of warm intra-cluster gas in groups of galaxies since these are thought to harbour a lot of baryons not accounted for. Nearby galaxy groups (e.g., Sculptor) subtend tens of degrees in size.
- Both low and high-redshift clusters occasionally show evidence for ‘cooling flow’ emission.
- Some radio sources subtend huge angular scales, e.g., Cen A, For A. Some radio lobes show clear evidence of patchy depolarization suggestive of intervening warm gas, e.g., Cyg A, For A.
- Starburst winds are known to extend over large scales, e.g., M82 where the soft x-rays and H α emission extend over 0.5° .
- The ionization cones in NGC 1068 have now been detected to the HI edge of the galaxy. The QSO nebula in MR 2251–178 presented at this conference suggests that some quasars may have very extended plasmas.
- Diffuse low surface brightness, stellar structures are seen in clusters and around normal galaxies and mergers. Intra-cluster and free-floating planetary nebulae offer the exciting possibility of probing the dynamics of these structures.
- A completely unexplored area is the use of tunable imagers to obtain absorption line maps of spirals, e.g., differential comparison of Mg to Fe to trace the variations in age or metallicities of the stellar population.
- The diffuse ionized gas seen in the Galaxy (‘Reynolds layer’) and in Local Group and nearby galaxies has an unidentified heat source. There is keen interest in studying this emission in diagnostic emission lines other than H α .
- The Magellanic Stream can be seen in H α as are numerous high velocity clouds. The Magellanic bridge is particularly strong in H α emission and subtends 30° .

- There are very many emission nebulae around compact sources, including photo-dissociation regions, supersoft x-ray sources, x-ray binaries, potassium shells around stars, planetary nebulae, supernova remnants, fast-moving pulsars, and expected around soft gamma-ray repeaters and gamma ray bursters.
- There is some expectation that young globular clusters may harbour warm gas.
- $H\alpha$ has been used as a first step to finding brown dwarfs combined with broadband colour data.
- Cometary tails are known to be sources of line emission.

11. Conclusion

Tunable filters provide excellent photometry of extended and compact, faint sources. When coupled to charge shuffling, they can produce fine differential measurement between bands, even in non-photometric conditions. The residual signal is often Poissonian with no sign of systematics due to, say, CCD artefacts or poorly-subtracted night-sky lines. Tunable filters can be adapted for polarimetry and have excellent applications in time series work.

We would argue that the TTF is a major step over conventional imaging. Fabry-Perot based filters are inexpensive and relatively easy to use, although it should be recognized that much of the TTF functionality has come from extensive electronic, firmware and software development at AAO/AAT. To allow for the different observing modes, the purchased controller must be heavily adapted in order to sidestep the battery of analogue switches. Ultimately, low gap Fabry-Perots are flawed tunable filters. But in practice, TTF-style devices will not be easily surpassed for ground based observations.

There is one important exception: the wide-field Lyot filter. For redshift-targetted programmes, the $A\Omega$ product afforded by the Lyot is expected to greatly exceed the performance of existing multi-object spectrographs. A low-resolution device suitable for observational cosmology, intended for the AAT, VISTA and Subaru telescopes, is currently under study at the AAO. Lyot filters offer a rare opportunity to exploit the fastest beams and the full acreage of the new generation of CCD mosaics.

First and foremost, I owe a great debt to Wil van Breugel for his warm hospitality and for the many discussions which helped to focus the ideas presented here. The cover of this book will be a living testament to his humour. Much of this work was carried out under the auspices of a Long Term Consultancy at Lawrence Livermore National Labs. I am indebted to Alice & Dickie Dunn, Ray Smartt, Bruce Gillespie and the staff at Sunspot Observatory for a memorable trip to Cloudcroft. Finally, I owe a debt to Bruce Woodgate, Povilas Palunas and Tony Danks: quite apart from their pioneering efforts to bring Lyot filters to night-time astronomy, their insights were crucial to several aspects of this work.

References

- Baker, J.C. et al. 1999, AJ, in press
- Barden, S. 2000, this volume
- Beckers, J.M. & Dunn, R.B. 1965, Internal Report, AFCRL-65-605, Airforce Cambridge Research Labs.
- Bell, R.J. 1972, *Introductory Fourier Transform Spectroscopy*, Academic Press.
- Bennett, C. 2000, this volume
- Bertin, E. & Arnouts, S. 1996, A&A, 117, 393
- Bland-Hawthorn, J. & Cecil, G. 1996, In *Experimental Methods in the Physical Sciences*, Vol. 29B, ch. 18, p. 363 (Academic Press: San Diego), eds. F.B. Dunning & R.G. Hulet
- Bland-Hawthorn, J. & Jones, D.H. 1998, SPIE, 3355, 855
- Caplan, J. 1995, In *Tridimensional Optical Spectroscopic Methods in Astrophys.*, eds. G. Comte & M. Marcellin, IAU, 149, 85 (ASP: San Francisco)
- Cianci, S. et al. 2000, this volume
- Deutsch, E.W., Margon, B. & Bland-Hawthorn, J. 1998, PASP, 110, 912
- Evans, J.W. 1949, JOSA, 39, 229
- Evans, J.W. 1958, JOSA, 48, 142
- Gale, M.T. 1998, In *Optical Document Security*, R.L. van Renesse, ed. (Artech House), ch. 12
- Graham, J. 2000, this volume
- Gunn, J.E. et al. 1998, AJ, 116, 3040
- Jenkins, A.R. et al. 1998, ApJ, 499, 20
- Jones, D.H. 1999, PhD thesis, Australian National University
- Jones, D.H. & Bland-Hawthorn, J. 1998, PASP, 110, 1059
- Kaiser, N. 1987, MNRAS, 227, 1
- Keller, C. 2000, this volume
- Loeb, A. & Rybicki, G. 1999, ApJ, 524, 527
- Lyot, B. 1933, *Compt. Rend. Acad. Sci.*, 197, 1593
- Lyot, B. 1944, *Ann. d'Ap.* 7, 31
- Magnusson, R. & Wang, S.S. 1993, *Appl. Opt.*, 32, 2606
- Maillard, J.-P. 2000, this volume
- Öhman, Y. 1938, *Nat.*, 141, 157
- Palunas, P. et al. 2000, this volume
- Parker, Q.A. & Bland-Hawthorn, J. 1997, PASA, 15, 33
- Perryman, M. 2000, this volume
- Pietraszewski, K.P. 2000, this volume
- Remillard, R.A., Rappaport, S. & Macri, L.M. 1995, ApJ, 439, 646
- Sheng, P. et al. 1999, PRL, 82, 4248
- Sölc, I. 1959, *Czech. J. Phys.*, 9, 237
- Tinney, C.G. & Tolley, A.J. 1999, MNRAS, 304, 119
- Title, A.M. & Rosenberg, W.J. 1979, *Appl. Opt.* 18:20, 3443
- Title, A.M. & Rosenberg, W.J. 1981, *Opt. Eng.* 20:6, 815
- Tonry, J. 2000, this volume

Figure 8. A schematic drawing of a fully transmissive, tunable imaging polarimeter incorporating a 5-stage Lyot filter, a three-element Wollaston prism (TEWP), a zero-wave (ZW) and a half-wave (HW) retarder. The S and P propagations are imaged in separate fields of the detector, A. The Lyot has been stretched along the optical axis: the proper aspect ratio is shown in Fig. 7.

A Taurus Tunable Filter

The Taurus Tunable Filter (TTF) is a pair of high performance, scanning air-gap etalons made by Queensgate Instruments, Ltd. A summary of the etalon properties is given below:

broad wavelength range (BTTF: 360-650nm, RTTF: 600-1000nm)
servo-controlled capacitance bridge
large aperture (75mm)
wide (1.5-15 μ m) scanning range
high finesse (40)
high throughput (85%)
compact (1.55 μ m) broadband ($\lambda/\Delta\lambda \sim 1$)
fast response time (1 kHz)
band switching linked to charge shuffling

The TTF is used with Taurus-2 at the AAT and WHT Cassegrain f/8 foci. The AAT/TTF can be used in service mode so patient souls might try it out before buying a monolithic filter. For more details, see the web site at <http://www.aao.gov.au/local/www/jbh/ttf/>.

B Tunable imaging polarimetry & application

For the Loeb-Rybicki experiment in §10, the Stokes parameters can only be measured with a dual beam system which treats both S and P states simultaneously (Fig. 8). Imaging polarimetry of faint nebular sources is notoriously difficult to undertake. The field of view should be big enough ($1'$) to include sky background since even on dark nights, the sky is polarized due to scattered solar radiation. For efficient survey follow-up, and to stand any chance of detecting these sources, the imaging polarimeter needs to tune accurately to the expected redshift. To suppress the night sky, the bandpass must be well matched to the intrinsic Ly α width (i.e. $R \sim 3000/(1+z)$).

Here, we illustrate what such a system might look like based on a low resolution, 5-stage scanning Lyot filter. (Polarimetry is not advisable using a Fabry-Perot with multilayer dielectric mirrors since the coatings impart a dominant azimuthal signature in the Stokes parameter maps.) The beam is split by the Wollaston prism before one of the states is rotated by a half-wave retarder. The zero-wave retarder is necessary to even up the optical paths along the S and P directions. With new synthetic materials, the total throughput of the polarimeter to linearly polarized (unpolarized) light is expected to be 90% (45%). Note that we do not need to rotate back the S propagation: the S and P states are imaged in separate regions of the detector. The Loeb-Rybicki haloes should be detectable in 1-2 nights on a 8-10m telescope, assuming the earliest sources can be identified from broad-band surveys.

Critical Comparison of 3-D Imaging Approaches

Charles L. Bennett

*L-43, Lawrence Livermore National Laboratory, P.O. Box 808,
Livermore, California, 94550*

Abstract.

Currently three imaging spectrometer architectures, tunable filter, dispersive, and Fourier transform, are viable for imaging the universe in three dimensions. There are domains of greatest utility for each of these architectures. The optimum choice among the various alternative architectures is dependent on the nature of the desired observations, the maturity of the relevant technology, and the character of the backgrounds. The domain appropriate for each of the alternatives is delineated; both for instruments having ideal performance as well as for instrumentation based on currently available technology. The environment and science objectives for the Next Generation Space Telescope will be used as a specific case to provide a basis for comparison of the various alternatives.

1. Introduction

It is expected that within the year, a decision will be made as to the composition of the suite of science instruments to be deployed on the Next Generation Space Telescope (NGST). It is therefore a particularly good time for a discussion of the relative merits, and appropriate domains of greatest utility for the various 3-D imaging alternatives. There has been, and no doubt will continue to be, a great deal of discussion as to which approach to 3-D imaging is "the best". There is no single correct answer, of course, since each type of instrument has its own strengths and weaknesses.

It does not seem to be widely known that, in the limiting case of photon statistical noise dominance, the performance of a 3-D imaging spectrometer based on 2-d detector arrays is the same for all architectures (Bennett et al. 1995), whether tunable filter, dispersive, or Fourier transform, provided that the same degrees of freedom are measured. In the following, I will first consider the photon statistics limited case, and show the equivalence between the various architectures. I will then generalize to the performance in the case that detector read noise, dark current, and Zodiacal background are included. I will consider specific parameters that are appropriate for the anticipated NGST environment. Finally, I will offer a suggestion for a hybrid instrument which combines the best features of all of the 3-D architectures, and offers great potential for best meeting the NGST needs.

2. Tunable Filter vs. Dispersive Spectrometer (Ideal Limit)

In comparing between the various options, it is important to assume equivalent detectors. In order to obtain 3-D data using a 2-d detector array, a series of exposures must be made. Consider an $N \times M$ pixel focal plane array, having no "gaps" between the pixel elements. Typical frames for a dispersive imaging spectrometer (DS), and a tunable filter imaging spectrometer (TF) are indicated schematically in Figure 1. In general, it is of course not necessary for the spatial samples observed by the DS to be contiguous, as implied by the arrangement displayed in Figure 1. Nor is it necessary for the spectral samples observed by the TF to be contiguous and non-overlapping, as is also implied by the configuration displayed in Figure 1. Indeed, in some cases, non-contiguous spectral sampling is desirable, and the TF system lends itself much more naturally to this mode of operation. On the other hand, for some questions, the ability to observe non-contiguous spatial samples is very important, and the DS approach, such as with a Multi-Object Spectrometer (MOS), is better suited for such measurements. For the moment, consider the case that the same spatial and spectral samples are covered by both the TF and the DS. Assume that the spectral samples represented by the various pixels along the dispersion direction in the DS correspond exactly both in terms of bandwidth and band center to the series of measurements made by the TF system, and that the spatial samples represented by the various pixels in the TF system similarly correspond exactly to the series of spatial measurements made by the DS system. In this case, if the total observation time is divided equally among the spectral samples for the TF case, and for the spatial samples in the DS case, each cell in the 3-D datacube is observed for the same exposure time, and with the same efficiency. Clearly the signal to noise performance will be the same for both of these configurations.

3. Tunable Filter vs. Fourier Transform Spectrometer (Ideal Limit)

The relation between the performance of an ideal tunable filter spectrometer with an ideal Fourier transform spectrometer is more subtle than that between the tunable filter and the dispersive spectrometer. One simplification, however, is that since the size of the image may be assumed the same for the FT and TF systems, it is only necessary to consider the information content of a single representative detector element obtained via either the TF or the FT system.

It is helpful to consider an analogy with the use of the Modulation Transfer Function (MTF) for the characterization of imaging systems. Consider an "object" spectrum having a sinusoidal intensity variation as a function of frequency. Also assume that this object spectrum is observed with a TF spectrometer having uniformly spaced filter samples, and that all of the filter samples have an equal transmission bandwidth. The "image" spectrum would also have a sinusoidal intensity variation as a function of frequency. In the case that the period of the sinusoidal intensity variation is much smaller than the characteristic width of the TF spectral channels, the "image" spectrum modulations are greatly reduced. Furthermore, if the spacing of the TF spectral samples is not sufficiently dense, the period of the modulations in the "image" spectrum may be altered by "aliasing effects".

Figure 1. A schematic comparison between a Tunable Filter Imaging Spectrometer and a Dispersive Imaging Spectrometer.

An FT spectrometer, at each of a sequence of retardance settings, directly measures the intensity of a particular sinusoidal intensity variation in the object spectrum. The set of such measurements constitutes an interferogram. In order to compare the information content of TF spectra measured in the frequency domain with FT interferograms measured in the transform domain, it is important to carefully consider the shape of the spectral response of the TF filters, their spacing, and the amount of spectral information content being measured.

In a naive approach to a TF system, it would be assumed that the transmission function for each of the TF filters had a "top hat" shape, i.e. outside the spectral bandpass of a given filter the transmission would be zero, and within a given bandpass the transmission would be unity. Viewed in terms of the response to sinusoidal modulations in the "object" spectrum, such filters have undesirable ramifications, such as contrast reversal for some modulation periods, and aliasing for others. Correspondingly, the most straightforward approach to the acquisition of interferograms by an FT spectrometer, involving equal weighting of each of the retardance measurements, produces effective spectral response functions which have undesirable negative sidelobes. It is important to consider spectral transmission functions which do not have such "sharp corners" as the "top hat" shape for the TF case, and to consider tapered weighting of the interferograms for the FT case.

Consider a sequence of measurements of the intensity of an underlying continuous spectral intensity function $S(\nu)$, dependent on the frequency ν , that is transmitted through a spectral filter $\tau(\nu)$. For a transmission filter centered at ν_0 , the observed number of photoelectrons would be given by

$$S_{\nu_0} = T_{\nu_0} \int_0^\infty S(\nu) \tau(\nu_0 - \nu) d\nu. \quad (1)$$

Here the units of spectral radiance $S(\nu)$ are photons $\text{Hz}^{-1} \text{s}^{-1}$, the exposure time for the observation is T_{ν_0} , in units of s, while the transmission function $\tau(\nu)$ is dimensionless. Also, although the integration limits extend to infinity, this is a purely formal convenience, and in this integral, as in others to follow, the integrand will always be limited to a finite range. The frequency variable, ν , is in units of Hz. (It is sometimes convenient to use the wavenumber equivalent of the frequency, defined by ν/c , and having dimensions of cycles per cm). It is assumed that the quantum efficiency is unity. The peak transmission is assumed to be unity, and the effective width of the transmission filter may be defined by

$$\Delta\nu_{eff} = \int_0^\infty \tau(\nu) d\nu. \quad (2)$$

In the case that the spectral radiance function varies slowly over the interval for which $\tau(\nu)$ is significant, the integral in Eq. (1) may be approximated by

$$S_{\nu_0} \simeq T_{\nu_0} S(\nu_0) \Delta\nu_{eff}. \quad (3)$$

The variance in the observed number of photoelectrons, in the statistical noise limit is equal to the total number of photoelectrons detected,

$$\text{var}(S_{\nu_0}) = T_{\nu_0} S(\nu_0) \Delta\nu_{eff}. \quad (4)$$

Using the relation between the observed counts and the estimate of the underlying spectral radiance function evaluated at ν_0 of Eq. (3),

$$\text{var}(S(\nu_0)) = \frac{S(\nu_0)}{T_{\nu_0} \Delta\nu_{eff}}. \quad (5)$$

For comparison with the FT spectrometer case, for which the noise spectrum $\text{var}(S(\nu_0))$ is independent of ν_0 , the dwell time T_{ν_0} is taken proportional to $S(\nu_0)$. (This assumed dwell time variation could of course only be used if the spectrum is known, and would not be applicable to multiple pixels, if they contain different spectra. The impact of varying spectral shape on the comparison between FT and TF spectrometers will be further discussed below.) The constant of proportionality may be determined by requiring that the sum over all ν_0 channels yields the total observation time,

$$T_{\nu_0} = T_{tot} \frac{S(\nu_0) \Delta\nu_0}{\int_0^\infty S(\nu) d\nu}. \quad (6)$$

Here the factor $\Delta\nu_0$ is the spacing between the TF spectral samples. For this integration time sequence the spectral variance becomes,

$$\text{var}(S(\nu_0)) = \frac{\int_0^\infty S(\nu) d\nu}{T_{tot} \Delta\nu_0 \Delta\nu_{eff}}. \quad (7)$$

With measurements made at the sample spacing $\Delta\nu_0 = \Delta\nu_{eff}$ this yields

$$\text{var}(S(\nu_0)) = \frac{\int_0^\infty S(\nu) d\nu}{T_{\text{tot}} \Delta \nu_{\text{eff}}^2}. \quad (8)$$

Measurements made at a sample spacing much finer than this produce little additional information about the continuum function $S(\nu)$, since the magnitude of $\Delta \nu_{\text{eff}}$ sets a practical limit to the fineness of the resolution recoverable, no matter how fine the sample spacing.

3.1. Derivation of the Basic Fourier Transform Relationships

The intensity of the interference pattern in a dual output port Michelson interferometer, $I(x)$, is a continuous function of the optical path difference x , i.e., the retardance, between the two mirrors, related to the continuous spectral intensity detected, $S(\nu)$, by the integral,

$$I_{\pm}(x) = \frac{1}{2} \int_0^\infty S(\nu) (1 \pm \cos(2\pi \nu x / c)) d\nu. \quad (9)$$

The two output ports correspond to the two sign values, with the “+” sign corresponding to the output port for which the two interfering beams are in phase at zero optical path difference (ZPD), and the “-” sign corresponding to the output port with out of phase beams at ZPD. As before, the product $S(\nu) d\nu$ has units of counts per second. Eq. (9) is valid for a perfectly compensated, perfectly efficient beam splitter. Real beam splitters have dispersion and are not perfectly efficient, but these complications are easily dealt with in practice. It is convenient to form the sum and difference of the signals from the two output ports of the interferometer. These two quantities are given by the integrals,

$$I_{\Sigma}(x) = \int_0^\infty S(\nu) d\nu, \quad (10)$$

and

$$I_{\Delta}(x) = \int_0^\infty S(\nu) \cos(2\pi \nu x / c) d\nu. \quad (11)$$

Note that the summed signal is independent of the optical path difference x , and is simply given by the integrated spectral intensity. Thus at each retardance setting of the interferometer the full broad band image is measured. This is because, in the absence of absorption losses, every photon entering the interferometer goes to one or the other of the exit ports. The difference signal, at the zero retardance position also becomes equal to the same full band intensity integral.

This feature of the summed signal from an FT system suggests that a desirable hybrid of FT and TF may be obtained by simply having a tunable filter placed in the optical train of an imaging FT spectrometer. In this case, the sum of the two output ports of the FT spectrometer provides the unmodulated full intensity of the light that has passed through the tunable filter. In addition, higher resolution spectral imaging may be obtained at the same time. In this hybrid approach, the summed output will be called the “panchromatic” output of

the FT, while the transform of the difference output will be called the “spectral” output of the FT instrument.

In general, it is advantageous to have the dwell time depend on retardance in order to tailor the effective spectral line shape and maximize data collection efficiency. This is typically done for radio astronomy, but is not typically done for laboratory FTIR spectroscopy. A typical interferogram would consist of a set of N discrete samples of the continuous function $I(x)$, symmetric about the point $x = 0$, each observed with dwell time T_n .

$$I_n = T_n I_\Delta(x_n), \text{ with } x_n = n\delta x, n \text{ in the range } [-N/2 + 1, N/2]. \quad (12)$$

Discrete Fourier transformation results in periodogram estimates, S_k , at integer multiples k of a fixed frequency spacing $\delta\nu$, approximately related to the continuous function $S(\nu)$ by

$$S_k \simeq T_0 S(\nu_k) \delta\nu \text{ with } \nu_k = k\delta\nu, k \text{ in the range } [-N/2 + 1, N/2]. \quad (13)$$

The approximate relation between the discrete spectral estimate and the continuous spectral function is accurate to the extent that the continuous spectral function varies sufficiently slowly in the neighborhood of the discrete sample point at $\nu = \nu_k$. This condition is similar to that used in writing expression (3) for the TF case. The spectral sample spacing $\delta\nu$ and the interferogram sample spacing δx are related by $\delta\nu = c/(N\delta x)$. The S_k values are given by the discrete Fourier transform,

$$S_k = \frac{2}{N} \sum_{n=-N/2+1}^{N/2} I_n \exp(-i2\pi kn/N). \quad (14)$$

The inverse discrete Fourier transform is

$$I_n = \frac{1}{2} \sum_{k=-N/2+1}^{N/2} S_k \exp(i2\pi kn/N). \quad (15)$$

The normalization used for the Fourier transform pair displayed in expression (14) and (15) has been chosen to most directly reflect the continuum relation of expression (11).

It follows from the convolution theorem that the spectral line shape, S_k^{res} , for a particular set of dwell times T_n is proportional to a Fourier transform,

$$S_k^{res} = \frac{1}{T_{tot}} \sum_{n=-N/2+1}^{N/2} T_n \exp(-i2\pi kn/N). \quad (16)$$

With this normalization, the peak of the resolution function at $k = 0$ is equal to unity. This resolution function plays the same role as the transmission function $\tau(\nu)$ for the TF case. Just as for the TF case, an effective width for the resolution function may be defined by summing over all k values,

$$\Delta\nu_{eff} = \delta\nu \sum_{k=-N/2+1}^{N/2} S_k^{res} = \frac{T_0}{T_{tot}} N \delta\nu. \quad (17)$$

Although the case of uniform integration times is simplest for the FT spectrometer, and indeed is the most common mode of operation of laboratory FTIR instruments, it is not the most efficient. Furthermore, for purposes of comparison with a TF spectrometer, the resolution function (a sinc function) has negative sidelobes, which cannot be realized by a physical transmission filter function $\tau(\nu)$. There are many choices for the dwell time series which produce non-negative spectral line shape functions which can be physically realized as transmission filter profiles. One of the simplest is the triangular apodization series, defined by

$$T_n = T_0 \left(1 - \frac{2|n|}{N}\right), \quad n = -N/2 + 1, \dots, 0, \dots, N/2. \quad (18)$$

The spectral line shape that results from this weighting is a sinc-squared function.

3.2. Fourier Transform Spectrometer Noise

For a real interferogram, the discrete spectrum is hermitian, i.e., $Re(S_k) = Re(S_{-k})$. While $Im(S_k) = -Im(S_{-k})$. The point $k = N/2$ corresponds to the Nyquist frequency. For a perfectly compensated beam splitter, with 100% modulation efficiency and no noise, the interferogram will also be symmetric. A real, symmetric interferogram produces a real, symmetric spectrum. Noise in the interferogram is real, and produces a hermitian contribution to the calculated spectrum. Noise in the interferogram is not necessarily symmetric, however, and thus contributes to both the real and the imaginary parts of the calculated spectrum. By virtue of the linear relation between interferogram and spectrum, and with the notation that primed quantities represent noise contributions, the spectral noise is simply the Fourier transform of the interferogram noise.

$$S'_k = \frac{2}{N} \sum_{n=-N/2+1}^{N/2} I'_n \exp(-i2\pi kn/N). \quad (19)$$

For a dual ported interferometer, with focal plane detectors having equivalent noise performance characteristics, specifically having a noise variance given by the sum of a read noise term, n_r^2 , plus a statistical noise term, the difference interferogram measurements have the noise characteristics:

$$\langle I'_n \rangle = 0. \quad \langle I'_n I'_m \rangle = \delta_{n,m} \left(2n_r^2 + \frac{T_n}{T_0} I_0 \right). \quad (20)$$

In the above expressions, the angle brackets represent an ensemble average. It is assumed that the noise is uncorrelated for different samples of the interferogram. The statistical properties of the spectral noise that follow from (19) and (20) are

$$\langle S'_k \rangle = 0, \langle \text{Re}(S'_k)^2 \rangle = \left(\frac{2}{N}\right)^2 \sum_{n=-N/2+1}^{N/2} \left(2n_r^2 + \frac{T_n}{T_0} I_0\right) \cos^2(2\pi kn/N). \quad (21)$$

Since for finite k values, the $\cos^2$ factor in expression (21) oscillates much more rapidly as a function of n than the factor T_n , it may be well approximated by $1/2$. With this approximation, the spectral noise becomes independent of k , i.e., it is “white”. The variance of the measured continuum spectrum thus is given by

$$\text{var}(S(\nu_k)) = \frac{\langle \text{Re}(S'_k)^2 \rangle}{(T_0 \delta\nu)^2} = \frac{2}{(\Delta\nu_{eff} T_{tot})^2} \left(2Nn_r^2 + T_{tot} \int_0^\infty S(\nu) d\nu\right). \quad (22)$$

In the case that of a 1-sided interferogram, with samples

$$I_n = T_n I_\Delta(x_n), \text{ with } x_n = n\delta x, \text{ } n \text{ in the range } [0, N_1 - 1]. \quad (23)$$

The variance of the measured continuum spectrum is given by

$$\text{var}(S(\nu_k)) = \frac{1}{(\Delta\nu_{eff} T_{tot})^2} \left(2N_1 n_r^2 + T_{tot} \int_0^\infty S(\nu) d\nu\right). \quad (24)$$

Although it may appear that the decrease in the variance has come “for free”, there is really no greater information content, since the density of independent spectral samples is only half as great in the spectrum derived from the 1-sided interferogram. The difference in the variance between 1-sided and 2-sided interferograms can be most easily derived (for perfectly symmetrical interferograms) by averaging each $-n$ interferogram sample with the $+n$ sample, and computing the Fourier transform of the resulting 1-sided interferogram. The statistical noise would be reduced by a factor of $1/\sqrt{2}$ for each interferogram sample, and since only half as many readouts would be required, the readout noise would be reduced a factor of two.

Expression (24), in the absence of read noise, matches expression (8) obtained for the TF case. Expression (22), similarly matches expression (7) for the TF case with a sampling interval $\Delta\nu_0 = \frac{1}{2}\Delta\nu_{eff}$, as is appropriate for the more dense sampling in frequency space. The interesting fact that the only spectral line shape parameter that enters into the spectral noise for a Fourier transform spectrometer is the effective width, $\Delta\nu_{eff}$, is novel, to this author’s knowledge (e.g., Griffiths & de Haseth, 1986). The remarkable equivalence of the noise performance over all of the various types of ideal imaging spectrometers may perhaps be interpreted in terms of an “information theory” argument.

4. Zodiacal Background and Detector Noise Terms

The Zodiacal light produces a substantial limiting background flux for NGST. For a 1 AU orbit, thermal emission from dust dominates at wavelengths longer than about $3.5 \mu\text{m}$, while for wavelengths shorter than this, scattered sunlight

Figure 2. The magnitude of the Zodiacal light background expected for NGST in a 1AU orbit.

produces the dominant background. An estimate of this background spectrum is displayed in Figure 2. The zodiacal background flux is constant, to good approximation, over the range of frequencies from 3,000 to 10,000 cycles/cm, at a level of approximately 3×10^{-4} photon $\text{cm}^{-2} \text{s}^{-1}$. Detector noise performance levels anticipated for deployment on NGST are displayed in the table below. The impact on the performance of the various 3-D imaging systems generated by these background sources are displayed in the next section.

Table 1. NGST Detector Performance Expectations

Case	Single Frame Read Noise, n_1 (e^-)	Dark Current I_d ($e^- \text{s}^{-1}$)
Current	15	0.1
Goal	3	0.02

5. The Impact of Backgrounds on 3-D Spectrometers

The signal to noise ratio, SNR, in the presence of backgrounds, (including detector dark current, I_d , read noise, n_r , and zodiacal background, $Z(\nu)$), for the TF and the DS case, is given by

$$SNR = \frac{\eta QE(\nu) S(\nu) \Delta\nu_{eff} T_\nu}{\sqrt{((QE(\nu)(S(\nu) + Z(\nu)) \Delta\nu_{eff}) + I_d) T_\nu + n_r^2}} \quad (25)$$

In this expression, the $QE(\nu)$ factor represents not only the quantum efficiency of the detector elements themselves, but also includes all other system trans-

mission losses, such as non-unity reflectance for reflecting surfaces, and non-zero absorption losses in any transmitting elements. In general this factor may vary spectrally, and in a fairly complex manner. Dispersive gratings, for example, will have reflective losses for off-blaze angles of reflection. Tunable filters have absorptive losses which tend to be greater for high resolution broadly tunable filters. The efficiency factor η in expression (25) encapsulates all effects which lower the signal level without concomitantly lowering the background level, such as surface scattering losses, for example.

The signal to noise ratio in the presence of backgrounds for the FT case is obtained from expression (24) by including the modulation efficiency factor and QE effects, adding both the zodiacal background and the detector dark current contributions to the total number of photoelectrons observed, and including readout noise,

$$SNR = \frac{\eta QE(\nu) S(\nu) \Delta\nu_{eff} T_{tot}}{\sqrt{(\int QE(\nu)(S(\nu) + Z(\nu))\delta\nu + 2I_d)T_{tot} + 2Nn_r^2}}. \quad (26)$$

This expression reduces to the expression for the spectral SNR of Graham, et al. (1998), under the simplifying assumptions made in that article. Fourier transform spectrometers have a modulation efficiency factor which enters into the η efficiency factor for the spectrally resolved SNR , but not into the pan-chromatic SNR . The magnitude of the read noise term which appears in expressions (25) and (26) depends on the number of non-destructive readouts of the detector array that are averaged to determine the estimated photo-current I for a particular spectrometer setting. This number is a compromise between integrating the photo-current for a longer time, which lowers the variance, and averaging more readouts, which decreases the amount of time available for integration. To good approximation, the optimum number of readouts is determined by I , the single frame read noise, n_1 and the time it takes to read out the array, Δt , via

$$\text{Optimal Number of Reads} = \frac{n_1}{\sqrt{I\Delta t}} \quad (27)$$

Therefore, $n_r^2 = n_1 \sqrt{I\Delta t}$.

Estimates of plausible realistic values for the system QE and η values are listed in Table 2 for operation in the K band. The NGST main telescope transmission was calculated assuming that all of the mirrors are gold coated. The dispersive spectrometer values are taken from curves computed by Satyapal (1999), for the K band at either $R=1000$ or $R=10,000$, although they are somewhat optimistic compared to the values routinely obtained at large ground based telescopes.¹ The tunable filter efficiencies are estimated on the basis of Northrop Grumman tunable Fabry-Perot performance values (Madonna & Ryan). The somewhat surprisingly low QE for the tunable filter may be attributed to the fact that the effective number of surfaces seen by the transmitted light is approximately equal to the finesse, and thus extremely high quality surfaces are

¹See, for example, the grating efficiency curves measured for the 2dF gratings, at the AAO website location: <http://www.aao.gov.au/local/www/ras/gratings/gratings.html>.

necessary to produce low losses at high finesse. The FT modulation efficiency corresponds to that of a 30° incident angle NIR-MidIR CsI Bomem beamsplitter (Villemaire 1999).

Table 2. Spectrometer Efficiency Assumptions (K band)

Case	QE detector+optics	η
TF	0.35	1
DS	0.6	0.7
FT	0.7	0.95

The Noise Equivalent Flux Density, $NEFD$, at a particular significance level is derived from the SNR equations by solving for the flux S which produces the given significance level. The $NEFD$ for observations in the K band at $2.2 \mu\text{m}$, (as one example) at the 10σ level, for a variety of imaging spectrometer options are displayed in Figure 3 as a function of spectral resolution. From these curves, for a particular problem of interest, it is easy to select the optimum instrumental configuration.

At the lowest spectral resolution, all of the 3-D instruments converge to the performance of an $R=5$, K band camera. At the highest spectral resolution, the DS has the best performance for spectroscopy, although only for the small number of objects that may be contained "within the slit". This fact is the basis for the current pre-eminence of Multi-Object Spectrometers and Integral Field Units in high resolution astronomical spectroscopy. For the purpose of imaging in a very narrow, single emission line band, the TF provides an $NEFD$ performance equivalent to that of the DS, but for every pixel in the field of view. For the purpose of obtaining complete spectra for every pixel in the field of view, the FT instrument substantially outperforms the TF or the "mapping DS". The point of equivalence between the imaging FT and the DS is when the number of settings of the DS is equal to the square of the ratio in performance between the single setting DS and the imaging FT.

For any resolution, the imaging FT instrument has the advantage that not only are spectra obtained for every pixel in the field of view, but that very deep K -band imaging (in this example, but it could be J , H , L , etc. or the entire $0.6\text{-}5.5 \mu\text{m}$ range) is simultaneously acquired. In many of the design reference missions for NGST, the data for both deep imaging and spectroscopy may be acquired simultaneously. The fact that such imaging is produced for every resolution setting of the FT instrument is indicated in Figure 3 by the lowest $NEFD$ curve labeled "panchromatic FT".

At the highest spectral resolution, the relatively strong signals required imply that for many fields of interest to NGST, the angular density of observable objects will be small enough that at most one object is expected per field of view. In this situation, it is not helpful to obtain spectra for every pixel in the field of view, and the spatial multiplexing of the imaging FT is not useful.

A very interesting hybrid approach (e.g., Beer 1992) is possible, however. This is the combination of an objective prism with an imaging FT spectrometer.

Figure 3. The Noise Equivalent Flux Density for a variety of 3d imaging approaches are displayed. The detector performance parameters are: 15 e^- read noise per readout, $0.1 \text{ e}^- \text{ s}^{-1}$ dark current. All observations extend over the K band, centered at $2.2 \mu\text{m}$, and assume a total observation time of 10^5 s . The time per readout is assumed to be 1 s . The QE and efficiency, η , values are those of Table 2 for each of the systems. TF = tunable filter, DS=dispersive spectrometer, FT=Fourier transform spectrometer. The "mapping DS" corresponds to scanning over a number of field of view settings equal to the number of spectral channels for the TF at each resolution, and thus has an NEFD equal to that of the TF spectrometer. It is only through this factor that the array size enters, as was shown in figure 1. The single setting TF corresponds to the observation of a full field of view, but through only a single filter. The panchromatic FT line indicates the NEFD sensitivity for the full K band imaging that is obtained as a function of the resolution of the imaging FT system. This sensitivity is almost independent of the resolution, since the loss of exposure time involved in the multiple readouts for the spectroscopy is a small fraction of the total observation time. The dispersed FT curve corresponds to the addition of an $R = 630$ prism in the collimated space of the imaging FT optical train, and the addition of a slit at an image plane. The dispersed FT case with slit, would not produce the panchromatic FT full field imaging. A dispersed FT operated as slit-less objective prism would produce the panchromatic FT full field imaging.

A relatively modest dispersion across one dimension of the image plane serves to reduce the spectral bandpass acceptance that is involved in the noise term for the FT spectrometer. With a slit at an image plane, the “panchromatic” output of the FT spectrometer would yield the same results as an ordinary prism spectrometer, while the Fourier transformed interferograms would enable much higher spectral resolution at much reduced NEFD. The curve labeled “dispersed FT” in Figure 3 corresponds to the assumption that a prism of dispersion equal to that of CaF_2 is placed in the collimated space of an imaging FT, and that the slit width is equal to one pixel. For objects which have much higher intensity than their surroundings, slit-less objective prism style measurements are also possible.

There are slight displacements of the curves for the various spectrometer, and imaging spectrometer configurations, depending on the choices for the detector performance parameters, and system efficiency values. Using the NGST “goal” detector performance parameters instead of the “current” performance values slightly improves the sensitivity of the TF, the DS, and the dispersed FT, but produce very little change in the imaging FT case. On the other hand, using grating efficiencies closer to those typical of ground based telescopes, lowers the DS curve, but not the TF, the imaging FT, or the dispersed FT curves. The mapping DS curve does not take into account any in-efficiencies with the precision re-pointing between observations.

In conclusion, the ability of a single instrument concept, composed of a filter wheel, programmable slit, dispersive prism, and Michelson interferometer, to deliver the performance of a wide field camera, the performance of a moderate resolution, full field imaging spectrometer, and the performance of a high resolution, limited field spectrometer seems to make this choice nearly obligatory for NGST.

This work was performed under the auspices of the U.S. Department of Energy under Contract No W-7405-Eng-48. I thank my IFIRS colleagues for many stimulating discussions and astronomical tutoring: J. R. Graham, M. Abrams, J. Carr, K. Cook, A. Dey, R. Hertel, N. Macoy, S. Morris, J. Najita, A. Villemaire, E. Wishnow, and R. Wurtz. I also thank J. Mather for the provocative suggestion to consider the dispersed FT option.

References

- C. L. Bennett, M. R. Carter and D. J. Fields, “Hyperspectral Imaging in the Infrared Using LIFTIRS”, *Proc. SPIE*, 2252, 274 (1995)
- R. Beer, “Remote Sensing by Fourier Transform Spectrometry”, J. Wiley & Sons, N.Y., (1992), p.67.
- J.R. Graham, M. Abrams, C.L. Bennett, J. Carr, K. Cook, A. Dey, J. Najita and E. Wishnow, “The Performance and Scientific Rationale for an Infrared Imaging Fourier Transform Spectrograph on a Large Space Telescope”, *PASP*, 110, 1205, (1998).
- P. R. Griffiths and J. A. de Haseth, “Fourier Transform Infrared Spectroscopy”, J. Wiley & Sons, N.Y., (1986).
- R. Madonna and R. Ryan, 1999, private communication.
- S. Satyapal, from anonymous ftp site:grisly.gsfc.nasa.gov, in directory: out.going/shobita/spectrometer_trade
- A. Villemaire, 1999, private communication.

[illegible]

One Million Galaxies in Technicolor: the Sloan Digital Sky Survey

A.S. Szalay

*Department of Physics and Astronomy,
The Johns Hopkins University, Baltimore,
for the SDSS Collaboration*

Abstract.

The Sloan Digital Sky Survey, hereafter the SDSS, is a project to digitally map about 1/2 of the Northern sky in five filter bands from UV to the near IR, and is expected to detect over 200 million objects in this area. Simultaneously, redshifts will be measured for the brightest 1 million galaxies. The SDSS will revolutionize the field of astronomy, increasing the amount of information available to researchers by several orders of magnitude. The resultant archive that will be used for scientific research will be large (exceeding several Terabytes) and complex: textual information, derived parameters, multi-band images, and spectra. The catalog will allow astronomers to study the evolution of the universe in greater detail and is intended to serve as the standard reference for the next several decades.

1. Introduction

Astronomy is about to undergo a major paradigm shift, with data sets becoming larger, and more homogeneous, for the first time designed in the top-down fashion. In a few years it may be much easier to "dial-up" a part of the sky, when we need a rapid observation than wait for several months to access a (sometimes quite small) telescope. With several projects in multiple wavelengths under way, like the SDSS, 2MASS, GSC-2, POSS2, ROSAT, FIRST and DENIS projects, each surveying a large fraction of the sky, the concept of having a "Digital Sky", with multiple, TB size databases interoperating in a seamless fashion is no longer an outlandish idea. More and more catalogs will be added and linked to the existing ones, query engines will become more sophisticated, and astronomers will have to be just as familiar with mining data as with observing on telescopes.

The Sloan Digital Sky Survey, hereafter the SDSS, is a project to digitally map about 1/2 of the Northern sky in five filter bands from UV to the near IR, and is expected to detect over 200 million objects in this area. Simultaneously, redshifts will be measured for the brightest 1 million galaxies. The SDSS will revolutionize the field of astronomy, increasing the amount of information available to researchers by several orders of magnitude. The resultant archive that will be used for scientific research will be large (exceeding several Terabytes) and complex: textual information, derived parameters, multi-band images, and

Figure 1. The view of the SDSS 2.5m telescope from the East.

spectra. The catalog will allow astronomers to study the evolution of the universe in greater detail and is intended to serve as the standard reference for the next several decades.

2. The Sloan Digital Sky Survey

The Sloan Digital Sky Survey (SDSS) is a collaboration between the University of Chicago, Princeton University, the Johns Hopkins University, the University of Washington, Fermi National Accelerator Laboratory, the Japanese Promotion Group, the United States Naval Observatory, the Max Planck Institute für Astronomie, Heidelberg, and the Institute for Advanced Study, Princeton, with additional funding provided by the Alfred P. Sloan Foundation and the National Science Foundation. In order to perform the observations, a dedicated 2.5 meter Ritchey-Chretien telescope was constructed at Apache Point, New Mexico, USA. This telescope is designed to have a large, flat focal plane which provides a 3 field of view. This design results from an attempt to balance the areal coverage of the instrument against the detector's pixel resolution.

The survey has two main components: a photometric survey, and a spectroscopic survey. The photometric survey is produced by drift scan imaging of 10,000 square degrees centered on the North Galactic Cap using five broad-band filters that range from the ultra-violet to the infra-red. The photometric imaging will use a CCD array that consists of 30 2K x 2K imaging CCDs, 22 2K x 400 astrometric CCDs, and 2 2K x 400 Focus CCDs. The data rate from this camera will exceed 8 Megabytes per second, and the total amount of raw data will exceed 40TB.

Figure 2. The SDSS system response curves. The responses are shown without atmospheric extinction. The curves represent expected total quantum efficiencies of the camera plus telescope on the sky.

A considerable care has gone into selecting the filter system. The SDSS has designed a filter system similar to *ugriz* which should transform to and from it with little difficulty, in which the *g*, *r* and *i* filters are as wide as practicable consistent with keeping the overlap small. If we wish the *u* band to be a “good” one which is almost entirely contained between the Balmer jump and the atmospheric cutoff, it cannot be significantly wider than Thuan-Gunn *u*. We call the new system *u' g' r' i' z'*. The response curves and sensitivity data with the coatings and CCDs we will use are shown in Figure 3.

The spectroscopic survey will target over a million objects chosen from the photometric survey in an attempt to produce a statistically uniform sample. This survey will utilize two multi-fiber medium resolution spectrographs, with a total of 640 optical fibers, of 3 seconds of arc in diameter, that provide spectral coverage from 3900 - 9200 Å. The telescope will gather about 5000 galaxy spectra in one night. The total number of spectra known to astronomers today is about 60,000 - only 12 days of SDSS data! Whenever the Northern Galactic cap is not accessible from the telescope site, a complementary survey will repeatedly image several areas in the Southern Galactic cap to study fainter objects and identify any variable sources.

2.1. The Data Products

The SDSS will create four main data sets: a photometric catalog, a spectroscopic catalog, images, and spectra. The photometric catalog is expected to contain one hundred million galaxies, one hundred million stars, and one million quasars, with magnitudes, profiles, and observational information recorded in the archive. The anticipated size of this product is about 250GB. Each detected object will also have an associated image cutout (“atlas image”) for each of the five filters, adding up to about 700GB. The spectroscopic catalog will contain identified emission and absorption lines, and one dimensional spectra for one mil-

Figure 3. A diagram of the data reduction software for the SDSS project. Due to the large amount of data, a fully automated reduction system is necessary. The software is organized into pipelines which communicate via the Operational Database (ODB).

lion galaxies, one hundred thousand stars, one hundred thousand quasars, and about ten thousand clusters, totaling about 50GB. In addition, derived custom catalogs may be included, such as a photometric cluster catalog, or QSO absorption line catalog. Thus the amount of tracked information in these products is about 1TB.

The collaboration will release the data to the public after a period of thorough verification. The actual distribution method is still under discussion. This public archive is expected to remain the standard reference catalog for the next several decades, presenting additional design and legacy problems. Furthermore, the design of the SDSS science archive must allow for the archive to grow beyond the actual completion of the survey. As the reference astronomical data set, each subsequent astronomical survey will want to cross-identify its objects with the SDSS catalog, requiring that the archive, or at least a part of it be dynamic.

3. The SDSS Archives

The survey archive is split into two orthogonal functionalities and the corresponding distinct components: an *operational archive*, where the raw data is reduced and mission critical information is stored; and the *science archive*, where calibrated data is available to the collaboration for analysis and is optimized for such queries. In the operational archive, data is reduced, but uncalibrated, since the calibration data is not necessarily taken at the same time as the observations. Calibrations will be provided on the fly, via method functions, and several ver-

Figure 4. A conceptual data-flow diagram of the SDSS data. The data is taken at the telescope (T), and is shipped on tapes to FNAL, where it is processed within one week, and ingested into the operational archive (ODB), protected by a firewall, accessible only the personnel working on the data processing. Within two weeks, data will be transferred into the Master Science Archive (MSA). From there data will be replicated to local archives (LA) within another two weeks. The data gets into the public domain (MPA, LA) after two years of science verification, and recalibration, if necessary. These servers will provide data for the astronomy community. We will also provide a WWW based access for the wide public, to be defined in the near future.

sions will be accessible. The Science Archive will contain only calibrated data, reorganized for efficient science use, using as much data clustering as possible. If a major revision of the calibrations is necessary, the Science Archive and its replications will have to be regenerated from the ODB. At any time the Master Science Archive will be the standard to be verified against.

High level requirements for the archive included (a) easy transfer of the binary database image from one architecture to another, (b) easy multi-platform availability and interoperability, (c) easy maintainance for future operating systems and platforms. In order to satisfy these, the Science Archive employs a three-tiered architecture: the user interface, the query support component, and the data warehouse. This distributed approach provides maximum flexibility, while maintaining portability, by isolating hardware specific features. The data warehouse, where most of the low-level I/O access happens is based upon an OODBMS (Objectivity/DB), where the porting issues depend mainly on the database vendor. Objectivity's database image is binary compatible, can be copied between different platforms, since the architecture is encoded on every page in the archive.

3.1. Accessing Terabytes of Data

Today's approaches to accessing astronomical data do not scale into the Terabyte regime — brute force does not work! Assume a hypothetical 500 GB data set. The most popular data access technique today is the World Wide Web. Most

Figure 5. A ray-traced visualization of the volume covered by the SDSS survey: the upper cone is the Northern survey, while the three narrow fans below are the Southern stripes.

universities can receive data at the bandwidth of about 15 kbytes/sec. The transfer time for this data set would be 1 year! If the data is residing locally within the building (access via Ethernet at 1 Mbytes/sec), the transfer time drops to 1 week. If the astronomer is logged on to the machine which contains the data, all of it on hard disk, then with SCSI bandwidth it still takes 1 day to scan through the data. Even faster hardware cannot support hundreds of “brute force” queries per day. With Terabyte catalogs, even small custom datasets are in the 10 GB range, thus a high level data management is needed.

4. Survey Goals

Both the photometric and the redshift survey will cover a larger fraction of the sky, a slightly elliptical region in the North, centered close to the NGP, and three strips at -10, 0 and +15 declinations. The length of these stripes is determined by extinction.

The design of the SDSS redshift survey is crucial to its effectiveness as a tool for studying large scale structure. The goal of this survey is not to measure a single statistical property of the distribution of galaxies; such as the power spectrum of density fluctuations, or the maximum size of non-linear structures. Instead, we desire an accurate description of the galaxy distribution over a wide range of physical scales, in order to address the questions listed above. To meet

Figure 6. The redshift-space distribution of galaxies. The objects lie in a slice 6deg thick along the survey equator. Galaxies are plotted at the distance indicated by their redshift, hence the appearance of clusters as “fingers-of-God.” This slice contains roughly 6% of the galaxy redshift sample. (D. Weinberg)

this goal, we plan to do a wide-angle, deep redshift survey which fully samples the galaxy distribution.

Figure 7. displays a view of a simulated SDSS spectroscopic sample, drawn from a large N-body simulation of a low-density, cold dark matter Universe. For these simulations, the spectroscopic sample consisted of all galaxies with apparent magnitude $r' < 17.55$ and apparent half-light diameter greater than 2 arcseconds, selecting just under one million galaxies over the survey area. Since the time that this simulation was done, the details of the galaxy selection criteria have changed; see below, but the qualitative nature of this figure is not sensitively dependent on these details.

This figure shows the distribution of galaxies in redshift space in a 6 deg by 130 deg slice along the survey equator, which contains about 6% of the spectroscopic sample. The median depth of the sample is $300h^{-1}\text{ Mpc}$, and there is useful information on the density field to $600h^{-1}\text{ Mpc}$, with a few galaxies as far away as $1000h^{-1}\text{ Mpc}$.

5. Large Scale Structure and Survey Strategy

A few examples from the recent history of redshift surveys illustrate how the detection of structure depends on both the survey geometry and the sampling rate. Observing in narrow pencil beams in the direction of the constellation Bootes, Kirshner et al. (1981) detected a $60h^{-1}\text{ Mpc}$ diameter void in the galaxy distribution. The geometry and sampling of the Kirshner et al. survey

Figure 7. Model power spectra. The power spectra for the SDSS Northern galaxy survey and for the luminous red galaxies (BRGs) are shown. The latter are assumed to be biased by a factor of two with respect to the galaxies. The individual points and their error bars are statistically independent. On small scales, the errors are smaller than the smallest bars. Also shown are the smallest comoving wavelength scales accessible to COBE and to the upcoming Microwave Anisotropy Probe (MAP).

was well suited to detection of a single void, but it could not answer the question of how common such structures are. De Lapparent, Geller and Huchra (1986, hereafter CfA2) measured redshifts of a magnitude-limited sample of galaxies in a narrow strip on the sky, and found that (1) large voids fill most of the volume of space and (2) these voids are surrounded by thin, dense, coherent structures. Demonstration of the first result was possible because of the geometry and depth of the survey, demonstration of the second because of the dense sampling. There were hints of this structure in the shallower survey of the same region by Huchra et al. (1983), but the deeper CfA2 sample was required to reveal the structure clearly. Sparse samples of the galaxy distribution, e.g. the QDOT (Saunders et al. 1991) survey of one in six IRAS galaxies, have provided useful statistical measures of low-order clustering on large scales, but sparse surveys have less power to detect coherent overdense and underdense regions. Perhaps because of this limitation, members of the QDOT team have elected to follow up their one-in-six survey by obtaining a complete (one-in-one) survey of IRAS galaxies to the same limiting flux, now nearing completion.

In order to encompass a fair sample of the Universe, a redshift survey must sample a very large volume. We do not know a priori how large such a volume must be, but existing observations provide the minimum constraints. Our picture of the geometry and size of existing structures became clearer as the CfA2 survey was extended over a larger solid angle (e.g. Vogeley et al. 1994). The largest of the structures found is a dense region measuring $150h^{-1}$ Mpc by $50h^{-1}$ Mpc – the “Great Wall” (Geller and Huchra 1989). Giovanelli et al. (1986) find

a similar coherent structure in a complete redshift survey of the Perseus-Pisces region, and da Costa et al. (1988) find another such “wall” in the Southern hemisphere (cf. Santiago et al. 1996). However, none of these surveys covers a large enough volume to investigate the frequency of these structures. In quantitative terms, the observed power spectrum of galaxy density fluctuations continues to rise on scales up to $> 100h^{-1}$ Mpc (Vogeley et al. 1992; Loveday et al. 1992; Fisher et al. 1993; Feldman et al. 1994; Park et al. 1994; Baugh and Efstathiou 1993; Peacock and Dodds 1994; da Costa et al. 1994; Landy et al. 1996). The observed power spectrum may be influenced by these few, largest, nearby structures. Thus a fair sample clearly must include many volumes with scale $100h^{-1}$ Mpc.

The sampling rate and geometry of a survey can strongly influence the largest structures seen (Szalay et al. 1991). A very deep pencil-beam survey by Broadhurst et al. (1990) showed remarkable clustering signatures on a scale of $128h^{-1}$ Mpc. First results from the redshift survey of Shectman et al. (1996, hereafter LC for the Las Campanas survey), has found evidence for an excess of power on similar scales (Landy et al. 1996). The LC survey covers a two narrow wedges in the Southern sky; it is deeper than CfA2 but less deep than the BEKS survey. While different survey geometries are appropriate for elucidating different features of large scale structure, a deep survey with a large opening angle and full sampling of the galaxy population is the only way to get a complete picture of the galaxy distribution. Coverage of the whole sky is extremely helpful if one wants to compare velocity and density fields in the nearby Universe, and redshift surveys of infrared-selected galaxies (Strauss et al. 1992; Fisher et al. 1995; Lawrence et al. 1995) are invaluable for this purpose. However, these surveys are much shallower than the SDSS, with almost two orders of magnitude fewer galaxies, so they are not nearly as powerful for statistical analyses of clustering.

6. Summary

We are in the middle of designing and constructing an extremely ambitious project, aiming to provide a useful tool for almost all astronomers in the world. This endeavour would not have been possible ten years ago, and even now we are pushing the limits of technology. We hope that our efforts will be successful, and the result will substantially change the way scientists do astronomy today. Having 200 million objects at our fingertips will undoubtedly lead to new major discoveries, and the spectroscopic followup of even a fraction of our fainter objects occupy astronomers for decades. The day when we have a “Digital Sky” at our desktop may be nearer than most astronomers think. Given the enormous public interest in astronomy, we hope that the resulting archive will also provide a challenge and inspiration to thousands of interested high-school students, and a lot of fun for the web-surfing public.

7. Acknowledgments

This material is based upon the Grey Book, the SDSS proposal to the National Science Foundation. The author would like to acknowledge helpful discussions

with the SDSS participants, and emphasize the heroic efforts of the collaboration on making this data set the highest quality possible.

References

- Baugh, C. M., and Efstathiou, G. 1993, MNRAS 265, 145.
- Broadhurst, T.J., Ellis, R.S., Koo, D.C., and Szalay, A.S. 1990, Nature 343, 726.
- da Costa, L. N., Pellegrini, P., Sargent, W., Tonry, J., Davis, M., Meiksin, A., Latham D., Menzies, J., and Coulson, I. 1988, ApJ 327, 544.
- da Costa, L. N., Vogeley, M. S., Geller, M. J., Huchra, J. P., and Park, C. 1994, ApJL 437, 1.
- de Lapparent, V., Geller, M.J., and Huchra, J.P. 1986, ApJL 302, 1.
- Feldman, H., Kaiser, N., and Peacock, J. 1994, ApJ 426, 23.
- Fisher, K.B., Davis, M., Strauss, M.A., Yahil, A., and Huchra, J.P. 1993, ApJ 402, 42.
- Fisher, K. B., Huchra, J. P., Davis, M., Strauss, M. A., Yahil, A., and Schlegel, D. 1995, ApJSuppl, 100, 69.
- Geller, M.J., and Huchra, J.P. 1989, Science 246, 897.
- Giovanelli, R., Haynes, M.P., and Chincarini, G. 1986, ApJ 300, 77.
- Huchra, J., Davis, M., Latham, D., and Tonry, J. 1983, ApJSuppl 52, 89.
- Kirshner, R.P., Oemler, A., Schechter, P.L., and Shethman, S.A. 1981, ApJL 248, 57.
- Landy, D.S., Shethman, S.A., Lin, H., Kirshner, R.P., Oemler, A.A., and Tucker, D. 1996, ApJL 456, 1.
- Loveday, J., Efstathiou, G., Peterson, B., A., and Maddox, S. J. 1992, ApJ 400, 43.
- Park, C., Vogeley, M.S., Geller, M.J., and Huchra, J.P. 1994, ApJ 431, 569.
- Peacock, J. A., and Dodds, S. J. 1994, MNRAS 267, 1020.
- Santiago, B. X., Strauss, M. A., Lahav, O., Davis, M., Dressler, A., and Huchra, J. P. 1996, ApJ 461, 38.
- Saunders, W., Frenk, C. S., Rowan-Robinson, M., Efstathiou, G., Lawrence, A., Kaiser, N., Ellis, R. S., Crawford, J., Xia, X.-Y., and Parry, I. 1991, Nature 349, 32.
- Shethman, S.A., Landy, S.D., Oemler, A., Tucker, D.L., Lin, H., Kirshner, R.P., and Schechter, P.L. 1996, ApJ 470, 172.
- Strauss, M.A. and Willick, J.A. 1995, Physics Reports, 261, 271.
- Szalay, A.S., Broadhurst, T.J., Ellman, N., Koo, D.C., and Ellis, R. 1991, Proc. Natl. Acad. Sci., 90, 4858.
- Vogeley, M.S., Park, C., Geller, M.J., Huchra, J.P., 1992, ApJL 391, 5.
- Vogeley, M.S., Park, C., Geller, M.J., Huchra, J.P., and Gott, J.R. 1994, ApJ 420, 525.

Design for an 8-m Telescope with a 3 Degree Field at f/1.25: The Dark Matter Telescope

Roger Angel, Michael Lesser, & Roland Sarlot

*The University of Arizona, Steward Observatory, 933 N. Cherry Ave.,
Tucson, AZ 85721*

Edward Dunham

Lowell Observatory, 1400 W. Mars Hill Rd., Flagstaff, AZ 86001

Abstract. Deep images recorded over a wide field of view are needed for astronomical projects as diverse as measuring cosmological parameters by weak gravitational lensing and predicting asteroid collisions with the Earth. Here we give the design for a three-mirror, 8.4 m telescope with an etendue of $260 (m.^{\circ})^2$, ten times greater than any other current or planned telescope. Its 3° diameter field is formed at f/1.25, and would be recorded by 55 cm diameter circular mosaics of CCDs for wavelengths $0.3 - 1\mu m$ and HgCdTe devices for $1 - 2.4\mu m$. The plate scale of $51\mu m/arcsec$ ensures that seeing-limited images will be well sampled by the $10 - 15\mu m$ pixel sizes of these detectors. When used to integrate single deep images over its 7 square degree field of view, the limiting magnitude (10σ) for a 1 night exposure will be $U=26.7$, $B=27.8$, $V=27.9$, $R=27.6$, $I=26.8$, $J=24.8$, $H=23.5$ and $Ks=22.8$. When used to find faint moving or variable objects by quickly imaging large fields, the telescope will be repositioned in 5 seconds while the detector mosaic is read out. Such agility is realized because the telescope is centrally balanced and very compact, little longer than the primary mirror diameter. By repositioning every 30 seconds, 1 steradian per night could be imaged to a limiting magnitude (10σ) of $V > 24$ or $Ks > 19$. For spectroscopic follow-up, several hundred integral field units could be used to study rare objects identified by imaging but too faint and sparsely scattered for efficient follow-up with any other large telescope.

1. Rationale and design considerations

Much of today's cutting edge astronomy makes use of just two types of optical telescope: ground based instruments having huge light grasp but with image quality limited by atmospheric blurring, and the Hubble space telescope with small aperture but exquisite image quality. In the future, larger cooled telescopes with unique sensitivity in the infrared are planned for space, while ground based telescopes will be much improved through correction of atmospheric blurring with adaptive optics. Interferometers will give still higher resolution, though at the price of proportional reduction in both field and sensitivity.

Here we consider a new type of ground based telescope, one whose strength lies not in a small, diffraction limited field of view but in having the widest possible field with the largest aperture. Equipped with a large mosaic of detectors to reach sensitivity limits set only by sky photon noise and atmospheric seeing, such a telescope has unique scientific potential. It would also be used to probe fields as large as the whole sky to unprecedented depth. By taking advantage of massive computational power only now becoming available, it would be able to identify very faint moving and variable objects. Many scientific projects are already envisaged for this Dark Matter Telescope, as described by Armandroff et al. (1999). It will also doubtless be used for projects as yet undreamed of, undertaken with the legacy of a truly enormous data base as well as with new observations.

A general measure of the power of a telescope is the figure of merit $A\Omega\eta/d\Omega$. Here A is the collecting area, Ω the solid angle of the field of view, η the detector quantum efficiency and $d\Omega$ the solid angle of the seeing limited image. In given integration time the size of field larger than Ω that can be explored to given depth is directly proportional to this figure of merit. While longer integration times can sometimes compensate for a lower figure of merit due to reduced image quality or smaller field or aperture, for some specific projects individual factors are decisive. For example, good image quality is essential to detect dark matter by slight changes in shape of distant galaxies, while long exposures cannot substitute for large aperture and good image quality in the detection of very faint moving asteroids.

To maximize the figure of merit for an imaging telescope, it should be in a site with excellent seeing and its focal length must be chosen so the detector pixels will adequately sample seeing-limited images. Experience shows that good telescopes at the best sites will deliver images of 0.5 arcsec on occasion, more frequently in the near infrared. The pixel sampling should be no worse than 0.25 arcsec (the Nyquist sampling criterion) to avoid further significant image degradation, thus each square degree on the sky must be sampled by about 200 million pixels in the detector mosaic. We would like to cover the 0.3 - 2.4 micron spectral range where the atmosphere is largely transparent and has relatively low emissivity. This will require two detector array types, most likely silicon CCDs below 1 μm wavelength and HgCdTe arrays above. Both types have quantum efficiencies, η , near unity. Individual 2048 x 2048 arrays of HgCdTe will be practical in the near future with 15 μm pixels, but not much smaller (Vural, 1999). This pixel size is consistent with future CCDs with high full well capacity, as we discuss below. Thus the optimum focal length should be around 10 - 12 m.

We have investigated optical designs for telescopes to deliver the largest possible aperture and field consistent with this focal length, with minimal degradation of 0.5 arcsec images. The largest convenient monolithic primaries are 8 m diameter, which would require imaging at $f/1.25$ to obtain 10 m focal length. Conventional telescope designs, including Schmidt cameras and other corrected systems based on one or two mirrors are incapable of wide fields at so fast a focus, but three-mirror optical systems have this capability, as we now show.

2. Three mirror telescope design

2.1. Paul optics

Three-mirror telescopes were first explored by Paul (1935). He gave a design with a parabolic primary, convex spherical secondary and a concave spherical tertiary of equal but opposite curvature. The image is formed midway between secondary and tertiary, with good correction over a wide field. One can think of the design as a reflective Schmidt telescope used as a corrector for a large afocal Cassegrain telescope. The secondary, located at the center of curvature of the tertiary, has added correction for spherical aberration in the manner of a reflecting Schmidt plate. A telescope of this type with a one degree field was built by McGraw et al., (1982), using a 1.8 m parabolic primary at $f/2.2$. The central obscuration was 22% by area, and the design gave images no more than 0.2 arcsec rms diameter at the edge of the field.

The Paul geometry is capable of still wider fields at fast focal ratio, when vignetting and central obscuration are minimized. We have explored general 3-mirror systems using ZEMAX, allowing all three mirrors to have aspheric figure optimized specifically for wide field. The wide field was not constrained to be flat, since the dedicated detector mosaics can be configured to approximate a curved field, with each flat device tangent to the focal surface.

Our first investigations were very encouraging in that we found that well corrected fields of 3° diameter or more can be formed at $f/1$ by the three mirrors alone. As the field is increased beyond 3° , so are the losses due to vignetting and central obscuration. Obscuration is caused by the large secondary and tertiary blocking the primary and by the detector obscuring the beam to the tertiary. Sky baffles needed to prevent any non-focused light from getting to the detector also increase obscuration and vignetting. An optimum design balances obscuration and field of view to maximize etendue or $A\Omega$ product for given primary aperture.

In a preprint circulated at the Austin AAS meeting (Angel, Dunham, Lesser and Millis, January, 1999), we gave a 3-mirror design with 3° diameter field and 25% central obscuration with little vignetting. While this design achieved the basic goals, it had significant limitations. The image formed at $f/1$ by a 6.5 m telescope had an uncomfortably small plate scale of $31 \mu\text{m}/\text{arcsec}$. The design also had poor achromaticity because of the introduction of a dewar window. Despite making this as thin as we dared (7 mm of sapphire), longitudinal chromatic aberration limited the breadth of filter passbands. Furthermore, to minimize vignetting there was inadequate room for thermal baffling around the detectors, or for filters to be stored in the dewar.

2.2. Revised design

Recognizing these difficulties, we have developed a more practical design with the tertiary mirror located significantly behind the primary (figure 1). The advantages of this placement for wide field imaging were recognized by Willstrop (1984). To minimize the secondary obscuration, the primary focal ratio was held at $f/1.0$, but we have used the added design flexibility to achieve a slower final focal ratio of $f/1.25$ for a better match to detector pixels. In a second major improvement, we have introduced a robust fused silica meniscus element to act as a vacuum window for the detector dewar (figure 2). This refractive element

Figure 1. Optical layout with rays at $\pm 1.5^\circ$ field angle.

together with the assigned 5 mm thick passband filter would together introduce unacceptable chromatic aberration, but this is corrected by a second, negative lens of fused silica located inside the dewar. The non-chromatic aberrations of the two lenses are balanced by reoptimization of the three reflecting elements, and by allowing the concave, inside surface of the dewar window to be aspheric. In this way the effects of longitudinal and chromatic aberration are almost eliminated. We are currently investigating the possibility that the internal element could be modified to include prisms or lenses to correct also for atmospheric dispersion.

The mirrors are arranged so the light from the secondary passes through a half-diameter hole in the primary to a near-spherical tertiary behind, and is brought to a focus near the primary vertex. Detector obscuration is minimized by making the primary and secondary together afocal. With these choices, the 3° field is completely baffled against stray sky light illumination with total obscuration and vignetting held to 26% at the field center and rising to 38% at the field edge.

The optical prescription for the new design (slightly different dewar optics are used for the infrared) are given in table 1. The dimensions are for 8.4 m primary aperture, which results in a focal length of 10.5 m, and a plate scale $51 \mu\text{m}/\text{arcsec}$. The 3° field has diameter 55 cm and is nearly flat, with a radius of curvature of 10 m (sagittal depth of 2.6 mm). The specific primary diameter of 8.4 m was chosen because similarly aspheric mirrors of this size are currently in production at the University of Arizona Mirror Lab (at $f/1.14$), and advantage could be taken of existing tooling, handling fixtures and test equipment. Such a large primary is preferred both because of its large light grasp and to obtain the

Figure 2. Detail of the dewar optics. The right hand positive meniscus lens 1.3 m in diameter forms the vacuum entrance window. The negative meniscus lens is an internal corrector. Just ahead of the focal surface is a bandpass filter.

optimum system focal length as long as 10.5 m. The telescope has an effective (unobscured) diameter of 6.9 m averaged over the field, and a collecting area of 38 m². Given detector mosaics to cover the full 3° circular field, the etendue is 260 (m.°)².

Table 1. Prescription for system configured with optical camera

Surface	Radius	Thickness	Dia.	Conic	r^6	r^{10}	r^{16}
Primary	-16.80000	-5.09996	8.40	-1.027354	-7.794474e-9	-7.959556e-13	1.808951e-18
Secondary	-6.59999	3.98969	3.43	-0.597477	-1.104251e-5	-7.162328e-8	1.858206e-10
Tertiary	-8.26926	-3.29459	4.45	-0.068663	-2.460923e-8	-6.114825e-10	1.788511e-12
window	-1.71869	-0.07875	1.28	0			
(silica)	-2.16001	-0.19262	1.28	-0.201983	0.003512024	-0.004189095	0.0214793
Corrector	-2.31283	-0.01562	1.12	0			
(silica)	-1.49723	-0.51128	1.07	0			
Filter	-2.89751	-0.00525	0.59	0			
(silica)	-2.89751	-0.04114	0.58	0			
Detector	-10.00685		0.55	-87.59419	-0.0067470		

2.3. Image quality

The telescope prescription was optimized for field angles 0, 0.7, 1.0, 1.3 and 1.5°, and for a single very broad band covering 0.45 - 0.75 μ m. Such a broad band covering the darkest sky background is of practical value to reach the faintest asteroids, whose images would be trailed in the longer exposures needed for typical photometric bands. We obtain at least 80% energy encircled within 1/3 arcsec diameter for all field positions, despite the broad chromatic range. Such achromaticity in the presence of a thick vacuum window represents a significant advance over previous Paul designs.

For the full range of optical photometric passbands, U - z, similar quality is obtained by adjusting the mirror spacings by a few mm to refocus and compensate small wavelength-dependent aberrations. Such adjustments are within the range of mirror support systems. As an example of a typical photometric band, we show in figure 3a the images that would be recorded in the R band, by combining spot diagrams for $0.6 + 0.65 + 0.7 \mu\text{m}$. All field angles show better than 80% energy into $1/3$ arcsec.

For the infrared camera, the design of the silica window and internal corrector element in the dewar was re-optimized for wavelengths $1.25 + 1.65 + 2.2 \mu\text{m}$ taken all together. Even for this very broad band, all the images are within the target of 80% into $1/3$ arcsec diameter except at the very edge. In practice, refocus allows the target to be met in any one of the J, H or K bands observed individually. The images shown in figure 3b are for the H band.

a) 0.6, 0.65, and 0.7 μm

b) 1.475, 1.65, and 1.825 μm

Figure 3. Left: spot diagrams for field angles up to 1.5° off-axis. The box size is 1 arcsec. Right: encircled energy as a function of aperture radius. The marked point is for 80% energy in $1/3$ arcsec diameter.

2.4. Sky Baffling

Through most of the telescope's spectral range, up to $1.8\mu\text{m}$, its sensitivity is limited by photon noise from optical emission by the atmosphere. To prevent additional skylight from reaching the focal surface indirectly, two black conical baffles will be used. The first, 4 m in diameter, extends 0.5 m below the secondary mirror (see figure 4). The second, 3.8 m diameter, rises 1.1 m above the primary hole. These two give complete sky baffling out to the full 1.5° field angle.

Above 2 microns wavelength black body thermal emission from telescope surfaces rises rapidly and would be the dominant background in the K band. To minimize such emission it is normal to use re-imaging optics in the dewar, so as to create an entrance pupil image where a cold stop can be located. For this telescope with very large field and fast focal ratio, pupil reimaging is prohibitively difficult and a new strategy is required. As is conventional, we will remove the upper baffle, so the secondary is seen against the colder sky background. In addition, the infrared mosaic detector is set back in the dewar, 0.8 m behind a 1.3 m diameter cold baffle just inside the entrance aperture. This reduces the solid angle accepting thermal radiation through the window to 1.4 steradians, still about 3 times the solid angle of the $f/1.25$ beam (0.45 ster). To further reduce this external radiation we take advantage of the fact that the dewar window viewed from the outside is a large surface with extremely low emission in the K band. Thus to block radiation from the lower "tube" supporting the tertiary mirror which is viewed directly by the detector, we propose to line it with gold-coated retro-reflectors. These would be curved mirror segments with their center of curvature on the dewar window. Radiation from other telescope surfaces seen reflected in the tertiary, such as the back of the primary, will be similarly blocked by placing in front of them flat retro-reflectors. In this way, we expect that the effective emissivity over the 1.4 steradians opening could be reduced to 0.2. This would cause a thermal background equivalent to that of 60% emissivity in the $f/1.25$ beam. From the estimates given by Gillett and Mountain (1998), thermal emission at this level would equal the OH background at $2.2\mu\text{m}$, would be ten times stronger at $2.4\mu\text{m}$ and 4 times weaker at $2.0\mu\text{m}$. Thus the sensitivity should be little compromised in the Ks band, from 2 to $2.2\mu\text{m}$.

3. Practical aspects

3.1. Mirrors, windows and filters

The three large aspheric mirrors needed for the telescope are manufacturable by methods already proven at the Mirror Lab (Martin et al. 1998). The primary is similar to the very fast 8.4 m $f/1.14$ mirrors for the Large Binocular Telescope, now in process. Already the $f/1.25$, 6.5 m Magellan I mirror has been finished to the diffraction limit at $0.5\mu\text{m}$. The 4.5 m concave tertiary is very fast, $f/0.8$, but is not very aspheric, and should not present any great difficulty. Perhaps the most challenging mirror is the smallest, the convex 3.5 m secondary. A similarly aspheric 1.6 m Cassegrain secondary for the MMT (Smith et al. 1997) is being made at the Mirror Lab, with stressed lap figuring and a computer-generated,

full aperture holographic test plate. The same methods would be extended to 3.5 m turntables for the figuring and hologram manufacture.

The dewar windows are 1.3 m diameter, and must take vacuum loading at sea level of about 14 tons. Their convex meniscus shape is favorable, adding shell stiffness; a preliminary finite element study kindly made by Warren Davison finds no stresses larger than 1000 psi for the shape given in the prescription, 79 mm thick at the center and 50 mm at the edge. The window would be figured with slightly more curvature than needed, to allow for the flattening under load.

We envisage intermediate waveband filters made by multi-layer coatings deposited on meniscus glass substrates 60 cm in diameter. In any optical systems with large etendue, the sharpness of interference filters is limited by the inevitable lack of collimation. Spectral broadening is proportional to the solid angle of rays passing through the filter. The smallest possible solid angle passing through a filter of area a_f is $A\Omega/a_f$. Filters 60 cm in diameter to cover the 3° field at $f/1.25$ will produce broadening $\Delta\lambda/\lambda$ of about 3%, a practical lower limit for broadening. George Jacoby has kindly calculated the profile of a 24 nm wide filter at 555 nm ($\Delta\lambda/\lambda = 4\%$), and finds the width is little affected compared to illumination at normal incidence, but the sides become less steep.

3.2. Detectors

CCD detector arrays are now a rather mature technology, and there is little doubt that a mosaic to cover the full 55 cm circular focal surface can be built for acceptable cost. One camera equipped with such a mosaic would alone be sufficient for the key dark matter and asteroid mapping projects. A similarly sized mosaic of infrared detectors represents a greater leap, but is nevertheless feasible and has enormous potential for study of the universe at high redshift and as a precursor to NGST. In both cases the individual flat detector arrays will be arranged in the form of a slightly domed mosaic to best fit the focal surface. There will remain some defocus across each flat array, but this can be held to acceptable level. Thus the 10 m radius of curvature of the focal plane results in only $\pm 8\mu\text{m}$ of longitudinal defocus for individual square arrays 25 mm on a side.

CCD mosaic A significant issue for CCDs is the control of charge blooming from bright stars. Deep wells, anti-blooming measures and the use of many smaller devices will all be important control measures. We shall suppose that detectors with $13\mu\text{m}$ pixels (0.25 arcsec) are used. These should be manufacturable with deep wells - already full well as high as 150,000 electrons has been demonstrated for $8\times 8\mu\text{m}$ pixels (Tower, 1999). In the future $13\mu\text{m}$ pixels could be optimized for still greater capacity. Anti-blooming capabilities can be incorporated in the detectors, which may reduce the full well capacity. Alternatively, anti-blooming clocking schemes may be used during integrations. To avoid uncontrolled blooming from the brightest stars, (a few really bright ones will be inevitable in a 3° field), a large number of smaller format devices may be preferred. These could be as small as 1024 pixels square, in which case 1300 devices would be needed to tile the 55 cm diameter focal plane. We find below that the read time should be no more than 5 seconds, requiring a realistic 200 kHz pixel rate to read each 1024 square device with a single amplifier.

As a way to minimize the gaps between the individual CCDs we are presently exploring detector packaging techniques which will allow the use of true, 4-side buttable devices using semiconductor industry standard packaging technologies. This development would limit the inter-device gaps to the non-imaging silicon of each detector, which is dominated by clock busses, amplifiers, and, most significantly, I/O bonding pads. With the continued industry-wide trend toward smaller I/O structures, it is not unreasonable to expect $50\mu\text{m}$ bond pads to be sufficient for future CCDs. If we therefore assume uniform gaps of $100\mu\text{m}$ around each CCD, a fill factor of 96% can be obtained.

Cooling requirements are not severe for the CCD mosaic. The criterion is that the dark rate be less than the sky photon rate in the darkest filters. We estimate that even for the U band or the narrowest 3% filter that photon rates will be $> 1.5e^-/\text{pixel}/\text{sec}$. With an MPP device, dark rates less than this can be achieved at a device temperature of about -15°C . It follows also that in the worst case of a 20 second exposure in a dark band a read noise of $\sim 4e^-$ rms will be acceptable.

While certainly a very large number of devices are required for this project, the CCDs themselves could be manufactured today. The DC shorts yield of several fabrication lines is now over 50%. If 50% of these unshorted thick CCDs are of astronomical quality, a mature lot run will yield about 25% useable devices from 6" silicon wafers at a modern facility. Assuming a 25-50% thinning and packaging yield, the final thinned device yield would be about 5-10% of the starting lot. This would then require about ~ 200 wafers to be fabricated with 100 devices per wafer, after one or two engineering and test lots. Cryogenic DC and AC wafer probing will allow rapid feedback to the fabricator on device quality and yield.

Infrared arrays The main issues for the infrared concern production time and cost for the large number of devices needed. Supposing that HgCdTe with 2048×2048 $15\mu\text{m}$ pixels are used, 250 are needed to fully populate the 3° focal surface. Production of this large a number should be feasible over a period of a few years, but techniques to lower cost need to be developed. For such a large production run, the long wave cut-off could be tailored to suit the application, and would be set at about $2.3\mu\text{m}$. Given that the fill factor already achieved for the HAWAII die is about 88% by area, we project that creative packaging combined with attention to bond pad sizes and locations on the custom multiplexer should allow for a mosaic filling 90% of the focal surface.

3.3. Mechanical considerations

The optical assembly for an 8.4 m primary is short, only 9 meters between the secondary and tertiary, with the primary and camera set midway between. This configuration is advantageous both for making an agile telescope and an inexpensive enclosure. An ideal, alt-azimuth mount would be like that of the LBT (Hill & Salinari, 1998). All three large mirrors would be stiffly mounted between two C rings, supported on a compact azimuth frame that transmits loads directly to a large diameter pier. This concept, developed by Warren Davison, is illustrated in figure 4. The stiffness of the drives gains by the square of the C ring radius, and we find that the large semicircular rings shown, 11 m

Figure 4. Concept for a rigid, fast-slewing mount. The three mirrors are held between large C rings turning on a flat azimuth platform.

diameter, coupled with the relatively small moment of inertia lead to excellent tracking and stability. Repointing by 3° in a time as short as 5 seconds is realistic. The telescope enclosure will be smaller and hence less expensive than for a standard 8 m telescope, because of the small turning radius.

4. Sensitivity and observing strategies

We have estimated imaging sensitivity in two different ways. First, we have scaled the 10σ point source magnitude limits in the Johnson photometric bands U - H ($0.35 - 1.65 \mu\text{m}$) derived ab-initio by Angel, Dunham, Lesser and Millis (1999) for a 6.5 m telescope. These included the blurring effects of atmospheric seeing and dispersion at 45° elevation. These values are listed in the second column of table 2. Second, limiting magnitudes in the infrared were derived from values computed for an adaptively corrected 8 m ground based telescope by Gillett and Mountain (1998). Corrections have been made for the following differences: 1) Collecting area was scaled for aperture and corrected for 30% vignetting/obscuration loss. 2) Sky aperture to accept 70% source flux was increased to allow for atmospheric seeing, from 0.01 to 0.50 square arcsec. 3) Results were scaled for an integration time of 20 seconds. 4) The Ks band was approximated by halving the bandwidth to $\Delta\lambda/\lambda$ to 0.1. 5) The photon background flux was doubled in Ks band to allow for imperfect telescope baffling.

The 10σ detection limits for unresolved point objects in a 20 second exposure obtained in this way for Johnson magnitudes J, H and Ks are listed in the

third column of table 2. It is encouraging to find good agreement for the J and H bands where the independent calculations overlap.

Table 2. 10σ limits for the 8.4 m telescope.

Band	$\lambda(\text{mm})$	20 second exposure		9 hour exposure
		from Angel et al	from Gillett & Mountain	
U	0.365	22.8	-	26.8
B	0.44	23.8	-	27.8
V	0.55	23.9	-	27.9
R	0.70	23.6	-	27.6
I	0.90	22.8	-	26.8
J	1.25	20.9	20.7	24.8
H	1.65	19.5	19.6	23.5
Ks	2.1	-	18.8	22.8

The depth of these 20 second exposure limits suggested to us one strategy for using the telescope, that of surveying all the observable sky in a sequence of ~ 3000 exposures. Assuming the readout is accomplished while the telescope is repointed in 5 seconds, 144 exposures an hour could be obtained in clear weather, and the survey would take about 22 hours, i.e. 3-4 nights. In another strategy, a single 3° field could be observed repeatedly during 50 hours of integration time to accumulate deep images in each of the 8 Johnson bands. With sensitivity increased as the square root of the integration time, the 10σ flux limits would be improved by 4 magnitudes to the values given in the last column. The above examples serve as a basis to calculate sensitivities for different strategies that would be used to find variable or moving objects or to obtain better spectral resolution with narrower filters.

5. Comparison with some existing and proposed imaging telescopes

The etendue at the 3° focal plane of our proposed Dark Matter Telescope is $260 (m.^{\circ})^2$, and it will be used with focal plane mosaics with a fill factor of at least 90% for a sampled etendue of $234 (m.^{\circ})^2$. The most powerful imaging telescope currently in operation is the Sloan Digital Sky Survey. It has a modified Cassegrain system with 2.5 m aperture and a 3° field at f/5. Its focal length is 12.5 m, similar to the 8.4 m telescope, and is instrumented with 30 CCDs to sample a field of 1.55 square degrees with 0.4 arcsec pixels, for a sampled etendue of $5.7 (m.^{\circ})^2$. Comparing the 8.4 m telescope with the SDSS, and allowing also for its increased pixel sampling and resolution, the advantage in figure of merit is by a factor of close to 100. The wide field optical cameras to be used with larger telescopes, such as Subaru's Suprime and MMT's Megacam have etendues which are not substantially larger than the SDSS, in the range 5 - 10 $(m.^{\circ})^2$. The recently funded British VISTA telescope is presently conceived as having a 4 m aperture with sampled etendues of 28 $(m.^{\circ})^2$ at optical and 3.1 $(m.^{\circ})^2$ at infrared wavelengths.

At infrared wavelengths, the projected point source sensitivities for the 20 second exposure of table 2 are about 4 magnitudes fainter in J and H than is

being achieved by 2MASS (Cutri, 1999), and about 3.5 magnitudes at Ks. This improvement is consistent with 40 times larger collecting area, 10 times sharper resolution (by solid angle) and 2.5 times longer integration time.

6. Spectroscopy with multiple integral field units

While the main function of the telescope will be imaging, it could be a valuable tool for spectroscopy. It has unique potential for spectroscopic follow up of rare objects with a density of a few to a few hundred in the 3° field. A good example might be the few dozen supernovae with $z \geq 1.5$ found in one field. Because of their small field, standard 8 m telescopes could obtain spectroscopic redshifts of only one or two at a time.

The spectroscopic addition discussed by Angel et al, (1999) used small, individual units to reimage objects on the fibers at longer focal ratio. These units incorporated small prisms, individually motorized for atmospheric dispersion correction. Additional mechanisms corrected continuously for atmospheric anamorphic distortion of the rotating field. The system is thus somewhat complex, and there remain concerns that sky subtraction and coupling efficiency would prevent spectroscopy of the faintest sources.

An attractive alternative is to use multiple, small-scale integral field units. For example, individual fields of 2×3 arcsec could be re-imaged onto 6×9 lenslet arrays subtending $1/3$ arcsec each, that would in turn form pupil images on an array of 54 fiber ends. In this case no mechanisms are needed to correct for either atmospheric dispersion or variable anamorphic distortion with field rotation. Images free from chromatic aberration would be reconstructed from the time sequence of dispersed spectra. The coupling losses in such systems can be kept low, and accurate sky subtraction obtained at the level needed to study the faintest galaxies. If the lenslets were sized at 1 mm, each coupler footprint need be little larger than the 6×9 mm needed for lenslet array, and several hundred units could be placed at once before crowding in the 2000 cm^2 field would become problematic.

We are grateful to Bob Millis, Peter Stritmatter, Sidney Wolff for their encouragement and interest in this concept. Reconsideration of the Paul design was stimulated by the recent Lowell workshop on Kuiper belt objects.

References

- Angel, J. R. P., Dunham, T., Lesser, M. & Millis, R. 1999, preprint
- Armandroff, T., Bernstein, G. M., Larsen, G., Millis, R. L. Pinto, P. A., Tyson, J. A., Wittman, D. M., & Zaritsky, D. F. 1999, preprint
- R. Cutri 1999, 2MASS web page
- Gillett, F.C. & Mountain, M. 1998, Science with the NGST, E.P. Smith and A. Koratkar eds, ASP Conf. Series 133, 42-51
- Hill, J. M. & Salinari, P. 1998, Proc. SPIE, 3352, 23
- Martin, H. M., Allen, R.G., Angel, J. R. P., Burge, J. H., Davison, W. B., DeRigne, S. T., Dettmann, L. R., Ketelsen, D. A., Kittrell, W. C., Miller, S. M., Stritmatter, P.A., and West, S.C. 1998, Advanced Technology Optical/IR Telescopes VI, Proc. SPIE, 3352, 194-204

- McGraw, J. T., Stockman, H. S., Angel, J. R. P., Epps, H. & Williams, J. T. 1982, Proc SPIE, 331, 137
- Paul, M. 1935, Rev d 'Optique, 14, 169
- Smith, B.K., Burge, J.H., & Martin, H.M. 1997, Optical Fabrication and Testing II, Proc. SPIE, 3134
- Tower, J. R. 1999, Proc. Inter. Conf. on Scientific Optical Imaging IV, M. B. Denton, editor, in press
- Vural, K. 1999, private communication
- Willstrop, R.V. 1984, MNRAS 210, 597 609

Not a student
of the University
of California
at Berkeley

Not a student
of the University
of California
at Berkeley

High-Redshift Gravitationally Lensed Galaxies and Tunable Filter Imaging

Paul C. Hewett

Institute of Astronomy, Madingley Road, Cambridge, CB3 0HA, UK

Stephen J. Warren

*Blackett Laboratory, Imperial College of Science Technology and
Medicine, Prince Consort Rd, London SW7 2BZ, UK*

Jon P. Willis

Institute of Astronomy, Madingley Road, Cambridge, CB3 0HA, UK

Joss Bland-Hawthorn

*Anglo-Australian Observatory, PO Box 296, Epping, NSW 1710,
Australia*

Geraint F. Lewis

*Department of Physics and Astronomy, University of Victoria, PO Box
3055, Victoria BC, V8W 3P7, Canada and Astronomy Department,
University of Washington, Box 351580, Seattle, WA 98195, USA*

Abstract. Most surveys for multiply-imaged gravitational lenses, outside of rich galaxy clusters, are based on sifting through large samples of distant sources to identify the rare examples of lensing. An alternative strategy, based on the selection of optimal lines-of-sight, offers a number of significant advantages. Utilising the multiplex capability of wide-area multifibre spectroscopy, together with tunable Fabry-Perot imaging, it is now possible to undertake such investigations. Progress in compiling a sample of high-redshift star-forming galaxies, gravitationally lensed by massive early-type galaxies at intermediate redshift, $z \sim 0.4$, is described.

1. Introduction

The potential of gravitational lenses to increase our knowledge concerning the amount and nature of dark matter, constrain key cosmological parameters and offer uniquely detailed views of faint distant objects is well known (Blandford and Narayan 1992). The development over the last decade of the field of “weak lensing” in which statistical studies of many sources, only slightly perturbed by the effects of lensing, are made has produced important new results concerning the distribution of mass in galaxy clusters and very large scale structures.

However, the assembly of examples of systems multiply imaged by individual galaxies has proved extremely hard.

Searches for individual examples of strong lensing have relied on the examination of a sample of objects, such as quasars or flat-spectrum radio-sources, where a large fraction of the sample lie at high redshift. Thus, towards each object there is a significant path-length over which an intervening deflector may interpose itself close to the line-of-sight. The lens search proceeds through the identification of sources whose morphology, multiple images or extended arcs for example, is consistent with the effects of lensing. Further imaging, at different wavelengths, and spectroscopy is then necessary to establish the source as a bona-fide lens and to obtain redshifts for the source and the deflecting galaxy. In practice, obtaining the redshifts is very difficult, particularly for radio-selected objects, and in the compilation of Kochanek et al (1999: <http://cfa-www.harvard.edu/castles>) only 19 of the 45 lensed systems possess both deflector and source redshifts.

An alternative search strategy is to examine optimal lines-of-sight by identifying a population of very effective deflectors, where it is known that any source lying behind the deflector will be significantly lensed, and then to examine the spectra of the deflectors for evidence of lensed background sources. Miralda-Escudé and Lehár (1992) pointed out that provided the surface density of faint, small, galaxies at high redshift is large, significant numbers of galaxy-galaxy lenses should exist. Subsequent observational developments have shown that the surface density of high-redshift, star-forming objects is indeed large (e.g., Steidel et al. 1996, Hu, Cowie and McMahon 1998). Provided a suitable sample of deflectors can be identified the optimal line-of-sight search strategy offers significant advantages, including i) high efficiency, the probability a lens will be seen along a line-of-sight is significant, ii) the deflector and source redshifts may be readily acquired, allowing the full lensing geometry to be defined, iii) the small, but extended, star-forming objects lead to resolved gravitational lenses, not unlike the radio-rings arising from morphologically extended radio emission, which provide much greater constraints on the deflector masses than the more familiar two- or four-image lenses of unresolved quasars.

Using APM measures of United Kingdom Schmidt Telescope B_JRI plates it is possible to identify the ideal population of deflectors – massive, bulge-dominated, galaxies at redshift $z \sim 0.4$, essentially half-way between ourselves and any high redshift source. Specifically, locating the population of relatively bright, $m_R \leq 20$, red, $B_J - R \geq 2.2$, galaxies with redshifts $0.25 \leq z \leq 0.6$ is straightforward (Warren et al. 1996). The galaxy population has a surface density of $\sim 50 \text{ deg}^{-2}$ and associated with each galaxy there is an area of sky, $\sim 1 \text{ arcsec}^2$, in which any distant source will be multiply imaged, with an associated increase in brightness of a factor $\gtrsim 10$. These early-type galaxies represent essentially optimal lines-of-sight to search for examples of strong lensing.

The presence of a lens is revealed by the detection of an anomalous emission line in the spectrum of one of the target distant early-type galaxies, so obtaining spectra of a large sample of the deflector galaxies represents the first stage in the lens survey. Examination of intermediate-resolution optical spectra of an initial sample of 160 colour-selected early-type galaxies revealed the presence of an emission line at $5589\text{\AA}$ in a galaxy with redshift $z = 0.485$. Follow-up

Figure 1. Observed-frame spectra of 403 distant early-type galaxies, redshifts $0.3 < z < 0.6$, arranged by increasing redshift. A number of prominent night-sky features are visible as vertical lines while features present in the galaxies, such as Calcium H+K and the G-band move to longer wavelength with increasing redshift.

spectroscopy (Warren et al. 1998) and imaging (Warren et al. 1999) have confirmed the B0047-2808 system as an optical Einstein ring with the source, a star-forming galaxy at $z = 3.595$, the first confirmed example of a normal galaxy lensing another normal galaxy and a demonstration of the viability of the optimal line-of-sight survey strategy.

2. The Spectroscopic Survey

With an efficient method for acquiring spectra along many optimal lines-of-sight there is the prospect of obtaining a large sample, 10–20 objects, of spatially resolved gravitationally lensed systems. The low-surface density of the galaxies on the sky means the Anglo-Australian Telescope’s 2dF multifibre instrument, with a 3 deg^2 field, is ideally suited to the initial spectroscopy. Total exposure times of $\sim 8000 \text{ s}$ produce galaxy spectra for which the completeness of redshift measurement is 95% and in which anomalous emission lines of fluxes $\sim 5 \times 10^{-17} \text{ erg s}^{-1} \text{ cm}^{-2}$ may be reliably detected i.e. fluxes comparable to those seen in high-redshift galaxy samples (e.g., Hu et al. 1998) can be reached. The unlensed fluxes are a factor ~ 10 fainter. Example spectra of galaxies from 2dF are shown in Figure 1. Two fields, producing ~ 250 galaxy spectra, can be observed per night.

Figure 2 shows the number-redshift distribution for all 485 galaxies observed in a two-night 2dF run. The median redshift is $z_{\text{med}} = 0.391$ and the sample consists of luminous, relatively high-redshift objects that barely feature in wide-angle, bright surveys (e.g. Colless 1998) or narrow-angle, deep surveys (e.g. Cohen et al. 1999; Figure 1b).

Figure 2. Number-redshift histogram for the 485 galaxies observed over two-nights at the AAT in 1998 September using the 2dF instrument.

3. Tunable Filter Imaging

Follow-up of an emission line detection in the discovery spectra would, until recently, have involved the acquisition of a second slit-spectrum to confirm the line. Subsequent narrow-band ($\sim 30\text{\AA}$) imaging, at an essentially random wavelength, to investigate the morphology of the emission, necessitates the fabrication of a custom filter, which is both time consuming and costly. However, the availability of tunable Fabry-Perot imaging, specifically the Taurus Tunable Filter (TTF) instruments at the AAT and William Herschel Telescope, enable monochromatic images to be obtained for emission lines at wavelengths covering virtually the whole optical spectrum. A resolving power of up to ~ 1000 allows the signal-to-noise ratio to be maximised by using narrow bandpasses, $\sim 10\text{\AA}$, well matched to the narrow emission features. Observing efficiency has been further improved with the recently commissioned broad-narrow shuffle mode, that allows a comparison broadband image to be obtained during the observation with only a small increase in the total exposure time. The sensitivity of the TTF system is such that confirmation of emission at the target wavelength can be achieved as quickly as with a slit spectrum. Indeed, the imaging is superior in that even with a relatively wide slit it is possible to exclude prominent emission components, which may lie anywhere on a ring up to 3 arcsec in diameter. Thus, spectroscopic follow-up may be dispensed with entirely and confirmation and narrow-band imaging achieved using the TTF.

Figure 3 illustrates the effectiveness of the TTF. The image is a composite of a broad-band J exposure, from UKIRT, and an 1800 s TTF exposure from the AAT. The candidate lens system is the double image seen at top-centre. The lower of the two components is the core of a galaxy, redshift $z = 0.519$, magnitude $m_R = 18.8$, coordinates $00^h 42^m 49.5^s -27^\circ 52' 17''$ (Equinox B1950.0). The spectrum is that of a normal early-type galaxy with the addition of a weak emission line at $5800\text{\AA}$. The upper of the components visible in Figure 3, sep-

Figure 3. Greyscale image of a 37×37 arcsec region containing the gravitational lens candidate B0042-2752 – the double image visible towards the top. North is up and East to the left.

arated by ~ 2 arcsec from the centre of the galaxy, is visible only in the TTF exposure, which covered a $\sim 10\text{\AA}$ bandpass centred at $5800\text{\AA}$. The detection of the component in the narrow-band image is unambiguous. Combining the image with two additional TTF exposures produces an image where some evidence of extended emission may be present but much deeper TTF observations are required to confirm any lower surface brightness extended structure or the presence of fainter lensed images.

The most probable identification of the emission is with $\text{Ly}\alpha$ $1216\text{\AA}$ from a star-forming galaxy at redshift $z = 3.77$. With this identification B0042-2752 is a very similar system to B0047-2808, the first lens discovered in the survey. The deflector and source in B0042-2752 are at slightly higher redshifts and the observed emission line flux is a factor ~ 5 lower, $f_\lambda \sim 1 \times 10^{-16} \text{ erg cm}^{-2} \text{ s}^{-1}$. However, without confirmation of the source redshift, via identification of a second emission line in the infrared ($[\text{O III}]$ 5007 would appear at 2.388μ), or evidence for a morphology unambiguously that of a lens, B0042-2752 remains only a lens candidate. Alternative explanations are certainly not particularly attractive. If the emission at $5800\text{\AA}$ is $[\text{O II}]$ 3727 the velocity of the line-producing region relative to the galaxy is $\sim +11,000 \text{ km s}^{-1}$, ruling out any plausible physical association, even within a rich cluster of galaxies. Chance projection of a faint H II galaxy is possible, with $z_{em} = 0.556$ if the line is $[\text{O II}]$ 3727, or $z_{em} = 0.158$ if the line is $[\text{O III}]$ 5007. There is no indication of any continuum emission from broad-band imaging data, or the TTF off-band images, and the putative H II galaxy must possess a large line equivalent-width, with an intrinsically faint continuum. The object would, for example, lie in the sparsely populated upper portion of Figure 2 of Hogg et al. (1998).

The survey spectroscopic observations are sensitive to the presence of an emission line object in an area of $\sim 5 \text{ arcsec}^2$ per galaxy, giving a surveyed area of $\sim 800 \text{ arcsec}^2$ for the sample of 160 galaxies. Assuming, rather optimistically, that [O II] or [O III] emission from a star-forming galaxy would be detected in our spectra over the redshift range $0.1 \leq z \leq 0.8$ gives a total volume of $\sim 22 h^{-3} \text{ Mpc}^3$ ($H_0 = 100 h \text{ km s}^{-1}$, $q_0 = 0.5$, $\Lambda = 0$) in which an H II galaxy may be found. Therefore, the space density of H II galaxies, with properties consistent with the constraints outlined above, must be $\gtrsim 0.01 h^3 \text{ Mpc}^{-3}$ for the *a priori* probability of finding such an object to be significant. Such a space density is substantially above that found in emission-line surveys to comparable line fluxes (e.g., Thompson, Djorgovski and Trauger 1995).

Further observations will determine whether B0042-2752 is the second lensed high-redshift galaxy to be identified in the survey. Irrespective of the outcome the combination of the AAT 2dF instrument and TTF imaging offers the prospect of compiling a well-defined sample of resolved gravitational lenses with a homogeneous deflector population. Investigation of the sample will provide unique information on the mass-to-light ratio in field elliptical galaxies and the morphology, emission line properties and masses of high-redshift star-forming objects currently too faint to be studied by any other means.

Bahram Mobasher kindly obtained the *J*-band image used to generate Figure 3. GFL acknowledges support from the Pacific Institute for Mathematical Sciences (1998-1999). The authors acknowledge the data and analysis facilities provided by the Starlink Project which is run by CCLRC on behalf of PPARC.

References

- Blandford, R.D. and Narayan, R. 1992, ARA&A, 30, 311
 Colless, M.M. 1998, in Wide Field Surveys in Cosmology, 14th IAP Meeting, (Paris: Editions Frontieres), 77
 Cohen, J.G., Blandford, R.D., Hogg, D.W., Pahre, M.A. and Shopbell, P.L. 1999, ApJ, 512, 30
 Hogg, D.W., Cohen, J.G., Blandford, R.D. and Pahre, M.A. 1998, ApJ, 504, 622
 Hu, E.M., Cowie, L. and McMahon, R.G. 1998, ApJ, 502, 99
 Miralda-Escudé J., Lehár J., 1992, MNRAS, 259, 31P
 Steidel, C.C., Giavalisco, M., Pettini, M., Dickinson, M. and Adelberger, K.L. 1996, ApJ, 462, L17
 Thompson, D., Djorgovski, S. and Trauger, J. 1995, AJ, 110, 963
 Warren, S.J., Hewett, P.C., Lewis, G.F., Møller, P., Iovino, A. and Shaver, P.A., 1996, in Astrophysical Applications of Gravitational Lensing, C.S. Kochanek and J.N. Hewitt, (Dordrecht: Kluwer), 329
 Warren, S.J., Iovino, A., Hewett, P.C. and Shaver, P.A. 1998, MNRAS, 299, 1215
 Warren, S.J., Lewis, G.F., Hewett, P.C., Møller, P., Shaver, P. and Iovino, A. 1999 A&A, 343, L35

The Power of Differential Observing

C.G. Tinney

Anglo-Australian Observatory, PO Box 296, Epping. 1710. Australia

Abstract. Although we prefer to pretend otherwise, it is generally systematic effects, rather than photon counting ones, which limit the challenging observations which can be obtained at a given facility. Recently optical astronomers have begun to “rediscover” what millimetre and infrared astronomers have known for decades – that observing differentially can obviate most systematic problems. Developments in charge shuffling have proven particularly powerful, as they allow optical astronomers to rapidly interleave reference exposures with no read noise penalty. The power of these techniques in both imaging and spectroscopic applications is highlighted.

1. Introduction

In almost all astronomical observing applications the limits to what can be achieved with a given set of equipment are set by systematic effects, rather than photon counting. Consider the fact that an 8-m telescope has four times the collecting area of a 4-m telescope. All other things being equal, therefore, obtaining a spectrum of a faint object to a given signal-to-noise ratio should only take four times as long on a 4-m telescope as an 8-m one. In which case, you’d have to wonder, (given their enormous cost), why we are building them at such an alarming rate. In practise, of course, we know that just exposing longer on a 4-m telescope doesn’t get you as faint as you’d expect. Systematic effects, like flat fielding variations and the ability to sky subtract, actually limit sensitivity. And those effects are indeed obviated by collecting more photons in a shorter period of time.

Infrared and radio astronomers have operated in a regime for many years, where instability limits sensitivity. In both disciplines, chopping plays an important role. By observing both a target and a reference (or a target and sky) rapidly interleaved, temporal changes in the instrument & sky are cancelled. Optical astronomers, on the other hand generally assume they needn’t to operate in such a mode. However, a number of observing techniques developed (and in some cases re-developed) at the Anglo-Australian Observatory (AAO) in the last few years, have shown that such techniques have enormous power in the optical as well. In this contribution I argue that optical astronomers should take much more care in planning programs. At every step they should try to answer the question “How can I make my experiment as differential as possible?”

2. CCD Astrometry

Astrometry relies on observations being repeatable over the long haul. Developments in this field have shown that CCDs make almost ideal astrometric devices. Their pixels form a regular matrix, with fairly uniform sensitivity, both between and within pixels. Because CCDs are operated in temperature controlled cryogenic environments, they are not subject to thermal variation. Parallax programs which aim to perform differential (as opposed to absolute astrometry), regularly achieve 4-5 milli-arcsecond (mas) precisions (Tinney 1999, Dahn 1997), while dedicated programs can achieve sub 1 mas precisions over several years (Harris et al. 1998, Monet et al. 1992) – that’s just $\sim 1/300 - 1/500^{\text{th}}$ of a pixel! Such programs require considerable effort to maintain the observing train (telescope, optics, filters, detectors, reduction software) constant over the course of the program, so that the only thing that *does* change is the position of the target star. As a result, while 4-5 mas precision have been obtained on common-user facilities, the best precisions have been obtained at the US Naval Observatory (USNO), where a dedicated telescope gives complete control of the observing train.

3. CCD Spectro-astrometry

The factor limiting astrometric precision in single exposures is invariably the atmosphere. A typical 5'-7' parallax field-of-view (FOV) contains many isoplanatic patches, so each line of site sees different image motions. Although these average out over time, seeing always produces residual image shifts (with the shifts becoming larger for larger fields).

One way to tackle this problem is the spectro-astrometric approach developed by Jeremy Bailey (Bailey 1998a,b). Spectroastrometry takes the differential astrometry one step further, by comparing the positions of an un-resolved target at different wavelengths. Consider an un-resolved binary with 10 mas separation, one component of which is a strong H α emitter. At the wavelength of H α the centroid of the light from the binary will be “skewed” towards the H α emitting component. At other wavelengths it won’t. So by taking a spectrum, and measuring the *spatial* position of the centroid at H α and non-H α wavelengths, one can (assuming the spectral types of the two components are known) determine the separation of the binary.

Even if the spectral types of the components are not precisely known (and so the absolute value of the separation is poorly constrained), the *position angle* and *relative separations* are well constrained. So observations of over several epochs can obtain a precise orbit, with only a single HST-resolution observation being required to calibrate the absolute separation. Because this technique works over a very small angle on the sky and has a very high degree of differentiality, it offers the possibility studying systems on scales *much* smaller than 1 mas.

4. Charge-Shuffled Narrow-band Imaging

The two techniques discussed above use differential techniques to the extent of (1) keeping an instrument constant over long periods of time, and (2) observing

Figure 1. Schematic representation of charge shuffled narrow-band imaging with TTF.

positions at one wavelength relative to positions at another over very small angles on the sky. In what further directions can one extend the differential observing model? Well, the next obvious direction is the time domain.

Narrow-band imaging is a potentially powerful astronomical observation – narrow-band imagers can have very high throughput, and for projects targetted at specific emission lines can have enormous multiplex advantages over traditional spectroscopy. Despite this, it has tended to be infrequently used, largely because of the expense of obtaining narrow-band filters, and the complexities of comparing on- and off-band exposures obtained through different filters and with different seeing & atmospheric transparency. What is needed is a tunable narrow-band filter, in which on- and off-band exposures can be so rapidly interleaved, that they see the same seeing and transparency.

This is exactly what has been developed at the AAO in the Taurus Tunable Filter (TTF, Bland-Hawthorn & Jones 1998). TTF is a pair of high-finesse, small spacing Fabry-Perot etalons used in the Taurus-2 imager. By their very nature, these etalons can be rapidly tuned to almost any wavelength in the optical, at 5-50Å resolution. Moreover, since 1995, the operation of the etalons have been integrated with the charge shuffling capabilities of the AAO-1 CCD controllers. Because the 2Kx4K CCDs used with Taurus significantly overfills the focal plane, the entire 10' FOV can be rapidly exchanged at the same time as the TTF is switched from its on- to off-band, and *vice versa* (see Fig. 1).

TTF therefore allows narrow-band imaging to be obtained where each object sees *exactly* the same optical path, wavelength band-pass, detector pixels and near-identical seeing and transparency. As a result data processing is phenomenally straight-forward – the on- and off-bands just get flatfielded and subtracted (Fig. 2). The resultant line images have near Poisson errors with minute systematic effects. Colour-term effects in subtraction can be further reduced by

Figure 2. Data reduction simplifies to subtracting one image subset from another, and removing the azimuthal sky. In this example, data are shown from a search for H α emitting brown dwarfs in IC2602.

alternating the off-band wavelength on either side of the on-band wavelength (known as “straddle mode”). Even $2-3\sigma$ events can be selected with confidence – something which is impossible with traditional narrow-band imaging.

5. Charge-Shuffled Time Series Imaging

A further wrinkle on this approach is to carry out time-series imaging. By masking the Taurus field, the CCD can be used to acquire multiple observations of a target separated by successive charge shuffles in a single direction. Once again, TTF can be used to change wavelength between each exposure, so that a highly differential sequence of on- and off-band exposures can be obtained to seek time-variability in a specific wavelength signature (see Fig. 3). This technique has been successfully used by Tinney & Tolley (1999) to search for variability in the TiO absorption in a warm brown dwarf, thought to be associated with the weather patterns expected in such a cool ($T_{\text{eff}} \sim 2000\text{K}$), rapidly rotating ($v \sin i = 30\text{km/s}$, period $\approx 3-6\text{h}$) object.

6. Nod-and-Shuffle Microslit Spectroscopy

An area demanding further differential work is optical spectroscopy. Infrared astronomers have known for years that the night sky in the $1-2.5\mu\text{m}$ region is so variable, that interleaving of object and sky exposures is a good idea. Optical astronomers working at $0.6-1\mu\text{m}$, however, have tended to think otherwise. Its incredible power, however, was brought home to me on a recent commissioning run for the upgraded LDSS⁺⁺ instrument on the AAT (Glazebrook et al: 1998). On this run, the charge shuffling techniques described above were extended to

Figure 3. Schematic representation of charge shuffled time-series. Even better interleaving can be obtained by shuffling backwards-and-forwards even more rapidly between each pair of Band 1 and Band 2 exposures.

Figure 4. Comparison of standard interpolation sky subtraction, with nod-and-shuffle sky subtraction.

a wide-field low-resolution multi-object optical spectrograph (LDSS⁺⁺). The aim was that instead of acquiring skys for each object in the traditional manner (ie. by extracting sky from the edge of a slitlet and extrapolating it across the object), skys would be obtained by noding the telescope to a blank region and shuffling a fresh region of the CCD into view (hence “nod-and-shuffle”). By rapidly (~ 15 -30s) interleaving the on-sky and on-target exposures, object and sky spectra could be obtained near-simultaneously, with only a $\sqrt{2}$ noise penalty. However, since the slitlet for each object needs to be only 1" long instead of 10" long, there is a 10 $\times$ multiplex advantage in the number of objects which can be observed (Glazebrook et al. 1999). The AAO can't claim to have invented “nod-and-shuffle” – it turns out to have been developed independently at least once before under the name “va et vient” (Cuillandre et al. 1994). And if the comments made after my presentation at Walnut Creek are any guide, probably several times before that!

However, to my mind the real power of this technique is not so much its multiplex advantage, as the fact that each object, and its sky are observed

through the same optics, with the same detector pixels and near-simultaneously. This means that (1) data reduction is an absolute doddle, since you just have to flatfield and subtract images, and (2) the resulting object-sky image has almost pure Poisson errors, with no systematic effects associated with slit unevenness, flatfielding, sky curvature, etc (as initially suggested by Bland-Hawthorn 1996). Figure 4 shows an example with LDSS⁺⁺. In this case, a single long slit has been used instead of multiple slitlets, permitting the object to be nodded along the slit, so that there is no loss of observing time on target. The left panel shows an image sky-subtracted by traditional techniques, while the right panel shows the non-and-shuffle processed image. What is noticeable is that the sky lines in the latter case don't show the systematic effects present in the former. This means that the technique can enable *extremely* deep spectra with a 4-m telescope, because you truly can integrate 20 times longer for a $\sqrt{20}$ improvement in performance. An astonishing example of the power of nod-and-shuffle was seen when data was obtained through an order sorting filter with optical power. The resulting spectra were *massively* de-focussed, with sky lines appearing as huge doughnuts. Even so, nod-and-shuffle *still* removed the sky perfectly – something which would be impossible with traditional techniques.

7. Conclusions

In the above examples I have tried to show the enormous power which differential observing techniques can offer. In each case, some aspect of the observing technique allows comparison with a reference which permits significant cancellation of systematic effects. Since it is such systematic effects which invariably limit the precision or depth of most challenging observations, their elimination should be a priority for all observers.

I am indebted to my AAO Colleagues for their work in developing innovative techniques and instruments, and for their inspiring discussions.

References

- Bailey, J.A. 1998a, In "Optical Astronomical Instrumentation", SPIE, 3355, 932
- Bailey, J.A. 1998b, MNRAS, 301, 161
- Bland-Hawthorn, J. 1996, AAO Newsletter, 79, 14
- Bland-Hawthorn, J. & Jones, D.H. 1998, SPIE, 3355, 855
- Cuillandre, J.C. et al. 1994, A&A, 281, 603
- Dahn, C.C. 1999, in "Fundamental Stellar Properties", IAU Symp.189, 9.
- Glazebrook, K. et al. 1998, AAO Newsletter, 87, 11
- Glazebrook, K. et al. 1999, in prep.
- Harris, H.C. et al. 1992, AJ, 502, 437
- Monet, D.G. et al. 1992, AJ, 103, 638
- Tinney, C.G. 1999, in "Very Low-mass Stars and Brown Dwarfs in Stellar Clusters and Associations", ed. M.-R. Zapatero Osorio & Rebolo, Cambridge University Press: Cambridge
- Tinney, C.G. & Tolley, A.J. 1999, MNRAS, 304, 119

A Four-Color CCD Imaging System for the REACT Network

A. H. Diercks², F. A. Harrison¹, S. R. Kulkarni¹, C. W. Stubbs²

1. Introduction

The REACT network is a collaboration of astronomers and institutions worldwide committed to developing a co-ordinated approach to optical follow-up of gamma-ray bursts (GRBs). In order to substantially improve the network's efficiency and provide optimum coverage for as many GRBs as possible from *HETE*, *INTEGRAL*, future Explorer missions, and *GLAST* we have designed a multi-band CCD camera (MBC) specifically optimized for GRB followup. The MBCs meet the scientific goals of the network while minimizing cost and complexity. Four goals have driven the camera design: 1) simultaneous imaging in multiple passbands, 2) high instrument stability for maximum photometric precision, 3) a compact, minimal maintenance design suitable for a variety of 1–3 m class telescopes, and 4) remote operation via the internet. The MBC uses four CCDs to simultaneously cover almost all of the wavelength range accessible with CCDs, resulting in a large multiplex advantage.

2. Optics

Light entering the camera is divided and passed to the various CCDs by a series of dichroic filters as shown in Figure 1. We have chosen to match the passbands of the Sloan Digital Sky Survey as closely as possible since this system will likely become the standard for extra-galactic work in the near future. When the survey is complete, 2% photometry of 17–20th magnitude stars will be available over a large portion of the northern sky, simplifying the task of absolute photometric calibration for targets in this region. (Gunn, 1993) This set of passbands is almost disjoint and is designed to avoid several prominent night-sky emission features such as O I $\lambda 5577$ and Hg I $\lambda 5460$. Because we will have robust four-color data, we will always have sufficient information to transform our photometry to the Sloan system with well measured color terms.

The tilted dichroics must be kept as thin as possible to minimize astigmatism, especially in the z' band which suffers two transmissions. This aberration can be eliminated by introducing tilted plates of similar thickness to the dichroics, rotated 90° about the optical axis. Ray tracing results indicate that with compensation, a 2 mm thick dichroic contributes only $0.15''$ to the overall

¹Division of physics, Mathematics and Astronomy, California Institute of Technology, 1201 East California, Mail Code 220-47, Pasadena, CA 91125

²Departments of Physics and Astronomy, University of Washington, Box 351580, Seattle, WA 98195

Figure 1. REACT multi-band camera schematic

spot size on a typical $f/10$ telescope. In the planar arrangement shown, three compensating plates are necessary, one each for the r' , i' and z' channels. We could eliminate one of the plates by rotating the g' - r' and i' - z' paths about the optical axis, however this would enlarge the size of the camera and complicate the cryogenics somewhat.

3. Detectors

The primary design choice for the detectors is the size and format of the CCDs. Ideally, the chips should be large enough so the camera can be used for transient identification (i.e., with a FOV matched to the error circles expected from *HETE*, *INTEGRAL* and the next explorer mission), as well as multicolor lightcurve photometry. *HETE* positions will range from $10''$ to $10'$ (radius), and *INTEGRAL* positions will be a few arcminutes in radius. Since detector size is a significant cost driver, we want to use as small a CCD as possible. Covering the largest *HETE* error regions would require very large detectors, at unrealistic cost. As a compromise, we have chosen an average FOV (telescope dependent) of $4' \times 4'$, large enough to completely cover a large fraction of *HETE* error boxes. In some cases, either OT identification by telescopes with larger FOVs or rastered observations with REACT telescopes will be needed prior to beginning multicolor lightcurve monitoring. The *SWIFT* mission will, if selected, supply few-arcsecond positions for GRBs with X-ray afterglows, which is more than adequate for immediate observation with the MBC. It is likely that positions from any other explorer mission would be of similar accuracy.

The second motivation for the size of the FOV is to make it wide enough to ensure that each image contains enough non-variable stars for photometric comparison. In the worst case region, the galactic pole, there are an average

of 0.35 stars brighter than 20th magnitude per square arcminute, and $4' \times 4'$ gives an average of 5.5 comparison stars per field even at this latitude (Allen 1976). To adequately sample the stellar PSF in the best seeing anticipated at some REACT telescopes, we need at least $0.25''$ per pixel.

We have chosen to use 1024×1024 , $24 \mu\text{m}$ pixel, thinned CCDs from Scientific Imaging Technologies, as they are readily available, have a proven performance record, and meet our field coverage and sampling requirements.

4. Vacuum and Cryogenics

The optics and detectors are contained in a common vacuum enclosure. In addition to simplifying the design, this arrangement keeps the optics clean and undisturbed, resulting in more stable flat-fields than are typically achieved with filter-wheel equipped cameras. We believe this feature is important, as it will not always be possible to collect the standard calibration data due to the unscheduled timing of GRBs and we want to be able to maintain good photometric precision using archival flats.

The four CCDs will be cooled with a commercially available Joule-Thompson cooler from APD Cryogenics. We have used this system on a larger CCD camera at Apache Point and found it to be highly reliable and maintenance free. Over the lifetime of the system, there is also a significant savings in liquid nitrogen cost and technician time spent filling dewars.

The vacuum will be maintained by a small 2 liter/s ion pump to improve the system stability, since without liquid nitrogen cooling there is no large tank at 77K constantly cryo-pumping the system. We have found that the much higher vacuum achievable with an ion pump is worth the design trouble as it dramatically increases thermal stability by reducing the thermal coupling to the environment.

5. Performance

By dividing incoming light into multiple passbands, we gain a significant advantage over a traditional CCD + filter wheel instrument. To ensure that the exposure times are identical in each passband, the camera has a single shutter. The CCD sensitivity and sky background differ in the various passbands, so the efficiency gain over a traditional camera depends on the signal-to-noise ratio required in each passband. Because of the high sky background and relatively low sensitivity, achieving sufficient signal-to-noise in the z' filter is often the driving factor in choosing an exposure time. With a traditional arrangement, one can choose an exposure time in each passband to maximize the signal-to-noise ratio, while we are constrained to a single exposure time. Figure 2 shows the relative gain in efficiency over a traditional design as a function of magnitude.

Almost all of astrophysically relevant information from broad-band photometry is contained in the object colors. The absolute flux is simply a measure of the distance to the target which, for objects of unknown intrinsic flux, must be measured by alternate means such as redshift. Since it has been demonstrated that extinction due to clouds is “grey” to a very high precision (Reif, 1999), si-

Figure 2. **Multi-band Camera Relative Efficiency Top Panel:** This plot shows the relative efficiency of the multi-band camera and a single CCD + filter-wheel arrangement as a function of magnitude. The solid line shows a target S/N ratio of (50,50,50,20) for g' , r' , i' , and z' respectively. The dashed line shows the same calculation for a S/N ratio of (50,50,30,5) for the same filters showing the effect of demanding high signal-to-noise in the z' band.

multaneous imaging in multiple passbands will allow us to extract the majority of the science even under non-photometric conditions.

In our experience, one of the significant difficulties with rapid-response or monitoring project observations is the lack of time available to acquire adequate calibration images such as flat-fields. Often the observations are conducted for an hour or two in the middle of the night, with no twilight time available. In addition, the optimal broadband filters are not always loaded in the instrument when an unexpected transient occurs. Our design solves this problem by putting all of the optical elements in the vacuum chamber where they will not be subject to contamination or constantly shifted during re-mounting each night.

This will allow us to generate a set of "archival" flat-field images in each filter which can be used for long periods of time without significant loss in precision. The two-color camera for the MACHO project uses this scheme to great advantage.

References

- Allen, C. W. 1982, *Astrophysical Quantities*, 3rd Ed.
- Gunn, J. R. and Knapp, G. R. 1993, in *Sky Surveys: Protostars to Protogalaxies*, Soifer, B. T., ASP:San Francisco, 267
- Reif, K. et al. 1998, in *Proc. SPIE* 3649

Anglo-Australian Tunable Littrow Articulated Spectrograph

Keith Taylor

Anglo-Australian Observatory, Epping, NSW 2121, Australia

Gordon Robertson

School of Physics, University of Sydney, NSW 2006, Australia

Sam Barden

National Optical Astronomy Observatories, Tucson, AZ 85719, USA

Abstract.

The Anglo-Australian Observatory is engaged in a program to replace its ageing Cassegrain instrumentation cluster, a work-horse intermediate dispersion spectrograph, a low dispersion multi-slit instrument and the TAURUS/TTF imaging FP system. The goal is to design a single instrument using the newest available technologies to improve significantly on the throughput, field-size and versatility of each observational mode. At the heart of this new concept is the adoption of Volume Phase Holographic gratings as dispersive elements in place of conventional surface reflection gratings. This permits a fully transmissive Littrow design conducive to both wide-field, multi-slit and interferometric spectroscopy. The instrument has been given the name **ATLAS** (the **A**nglo-**A**ustralian **T**unable **L**ittrow **A**rticulated **S**pectrograph) in recognition of the need for a variable collimator to camera angle to maintain optimal performance over a range of wavelengths ($350\text{nm} \leq \lambda \leq 1\mu\text{m}$) and resolving powers ($\mathcal{R} \lesssim 10,000$) for these tunable gratings. ATLAS on the AAT will have a rich scientific program, and be a key facility to link the output of the impending large surveys (e.g. Sloan in the optical, AXAF/XMM in X-rays) to later follow-up observations at 8-10 m telescopes. It also represents a new class of spectrograph which is well suited to further improvements in basic spectroscopy for these larger telescopes.

1. Introduction

Recognizing the need to upgrade the AAT's Cassegrain spectroscopic facilities, a concept study was commissioned to evaluate a new generation of innovative highly efficient wide-field spectrographs. By adopting Volume Phase Holographic (VPH) gratings (Barden et al. 1998) instead of conventional reflection gratings, a substantial increase in efficiency above even the best existing spectrographs is possible. A VPH grating has already been used successfully in

Figure 1. A double-beam version of ATLAS showing camera articulation at two angles characteristic of high and low dispersion modes.

the AAO's LDSS++ project (Glazebrook et al. 1998) and the use of such gratings removes the last major source of light loss in a spectrograph. We propose a fully transmitting instrument, which can entirely avoid vignetting losses.

Another advantage of VPH gratings is that the effective 'blaze' wavelength can be tuned simply by rotating the grating. Thus a single grating can be peaked in efficiency at any wavelength within a broad range. However, the camera must then be moved to the correct angle, so the collimator-grating-camera angle must be adjustable. We call this an articulated camera and the concept is illustrated in Figure 1. Such an articulated camera will introduce some engineering complication, but has the additional virtue of greatly increasing the versatility of the instrument. It will, for example, be possible to absorb both the multi-slit capabilities of LDSS (Taylor 1990) and the imaging interferometric capabilities of Taurus/TTF (Bland-Hawthorn & Jones 1998a&b) into the same instrument which carries out intermediate dispersion spectroscopy thus creating a uniquely powerful instrument by consolidating the Cassegrain instrument suite of the AAT into a single instrument.

ATLAS would deliver efficiencies substantially exceeding that of conventional spectrographs for low and intermediate dispersion spectroscopy (up to a resolution $\mathcal{R} \sim 9000$). At low resolution it would match or exceed the throughput of LDSS++. Our preliminary optical design indicates that a field size of $24'$, on a side, is feasible. Such a large field (compared with LDSS++ at $12'$) would provide a powerful multi-object survey capability, and continue the AAT's strategy of seeking to offer wide-field facilities. As a result of its efficiency and versatility, ATLAS has a strong and enormously diverse science case. Furthermore, in pioneering the application of the new VPH technology to versatile astronomical spectrographs, ATLAS would serve as a model for a new generation of spectrographs on the 8-10 metre class telescopes.

A crucial factor is the pioneering work done in collaboration with Kaiser Optical Systems Inc. (KOSI), one of only a few suppliers of VPH gratings (see Barden et al. 1998). Gratings of the size required by ATLAS (15cm collimated beam) have not yet been produced by KOSI, although there are no major technical barriers to production of this size or even greater.

2. Volume Phase Holographic (VPH) gratings - in brief

A VPH grating consists of a layer of dichromated gelatin on a suitable glass substrate. The layer is about 4 - 20 μ m thick, and is quite transparent. The grating action is achieved not by surface relief features as in a conventional grating, but by modulations of the refractive index. These impress periodic phase variations on the incoming wavefront which result in diffraction. Very high efficiency ($\sim 90\%$ at peak) can be obtained at the design wavelength, and close to this efficiency within a range around the optimum wavelength. This represents a marked improvement over the $\sim 60\%$ typical of a conventional reflection grating. The high efficiency, and high resolution obtainable, are the reasons why we expect VPH gratings to form the basis of the next generation of spectrographs. The construction and use of VPH gratings is described by Barden et al., 1998.

A key point is that the depth of the grating (4 to 20 μ m) is considerably greater than the wavelength of light. It is for this reason that the gratings are referred to as Volume Phase Holographic. The significant depth of the gratings means that the grating efficiency is high only in directions obeying the Bragg relation (effectively specular reflection from the refractive index modulations). We mention this somewhat technical point because it is this that requires a general purpose VPH spectrograph to have an additional degree of freedom: as well as rotating the grating, it must be possible to position the camera to accept a wide range of collimator-grating-camera angles. This is the concept of an articulated camera (Figure 1). The gratings are tunable in that tilting the grating with respect to the incident beam changes the wavelength which obeys the Bragg condition, and so effectively tunes the grating to peak efficiency at a different wavelength. It is this process which replaces the concept of blaze for a conventional surface grating: for VPH there is no fixed "blaze wavelength".

The VPH gratings that we are considering for use in ATLAS are transmission gratings which means that it is straightforward to achieve Littrow operation of the spectrograph (incident and diffracted beams make equal angles with the grating normal). In contrast with reflection gratings where Littrow operation

creates the difficulty that the collimator and camera must be essentially in the same place, the transmission Littrow configuration suffers no such problem while maintaining the advantage that there is no anamorphic beam expansion. The camera aperture can be substantially smaller than for a non-Littrow design, thus easing the optical design, avoiding compromises necessary to facilitate a range of beam dilations, and reducing camera complexity and cost. Furthermore the smaller pupil relief facilitates a considerably wider field of view.

Because the refractive index modulations are produced by a holographic process it is possible to have high line densities, up to ~ 6000 1/mm. This contrasts with conventional ruled reflection gratings, for which the maximum generally available is 1200 1/mm. A high line density enables high dispersion to be obtained in first order, and is another significant advantage of VPH gratings. For high line densities immersion prisms are generally required on both sides of the grating to avoid excessively large angles of incidence and diffraction, and indeed to avoid total internal reflection in extreme cases.

The application of VPH gratings does have some problems, however. Firstly, VPH gratings are not immune from the general trend that it is hard to maintain good efficiency at high spectral resolution. Thus although it would be possible to obtain resolutions greater than $\mathcal{R}=10,000$ with a collimated beam of less than 10cm by using a high line density grating, such high resolutions are inevitably associated with a large angular deviation of the beam by the grating. It is a property of VPH gratings that light in one of the polarization planes is progressively lost at larger deviation angles, thus impacting overall efficiency. At 90° deviation (within the grating medium), the p-plane light is lost entirely. In the ATLAS concept we have minimized this problem by enabling the desired intermediate dispersions ($\mathcal{R} \lesssim 9000$) to be obtained without approaching the regime in which serious p-plane loss occurs. The throughput still exceeds that of a conventional reflection grating at all design resolutions. The principal means by which we have done this is to specify a collimated beam diameter of 15cm, substantially greater than that of most 4-metre telescope transmission optics spectrographs. Resolution could also be improved by reducing the slit width, but we have maintained a design figure of $1.5''$ to match the median AAT seeing.

The second VPH property which can potentially cause trouble is the *bandwidth* of the gratings, given that the Bragg condition is strongly dependent on wavelength. Thus when a VPH grating has been tilted to tune a certain wavelength to peak efficiency (and the camera moved to receive that wavelength), the rest of the wavelength range incident on the detector is not at the optimum Bragg condition. This can result in a significant drop of efficiency towards the two ends of the wavelength range. We use the term *bandwidth* to mean the wavelength range within which the efficiency is at least half of the peak value. The *bandwidth* improves if grating thicknesses can be reduced, but in order to do this (while maintaining good peak efficiency) it is necessary to increase the amplitude of the refractive index modulations. Recent work at KOSI in collaboration with Barden suggests that gratings of thickness about $5\mu\text{m}$ and refractive index modulation ± 0.1 are feasible and this is what we have assumed in our simulations.

3. ATLAS Preliminary Design Concept

3.1. Overview of the optical system

A preliminary optical design for ATLAS has been produced by Damien Jones (of Prime Optics, Qld, Australia) and is shown in Figure 1. This figure demonstrates two articulated camera angles in the double-beam configuration.

The principal features of this design are:

- Input AAT Cassegrain f/7.9;
- All transmitting optics, thus avoiding vignetting by central obstructions and reflection losses of mirrors. Modern anti-reflection coating of lenses can reduce losses to $\sim 0.5\%$ per surface. We require good performance over at least the range 370nm to $1\mu\text{m}$. With 20 air-glass surfaces we can expect the plateau throughput of the optics to be $\sim 90\%$. (cf: GMOS on Gemini at $\sim 80\%$);
- Field of view $\sim 24'$ on a side;
- 15cm collimated beam - maximum allowable within space constraints of the Cassegrain cage without folding the instrument;
- In the collimated space there is a rotatable mount for a VPH grating with or without immersion prisms, or an etalon;
- Etalons for TAURUS mode with clear diameters of 15cm are available but extremely expensive, so we envisage that the existing 6cm ones would be used with the f/15 top end. The field size would be reduced to $\sim 12'$, but this still exceeds the current TAURUS field of $9'$;
- The camera is specified as f/2.2 providing a beneficial degree of over-sampling when used with a $1.5''$ slit (about 4 pixels per resolution FWHM) and is still able to Nyquist-sample a $0.75''$ slit or IFU input. The spatial scale at f/2.2 would be $0.36''$ per $15\mu\text{m}$ pixel;
- The camera is articulated in that it can accommodate collimator-camera angles from 0° (straight through position, for TAURUS or direct imaging modes) up to 90° or more for high dispersion gratings;
- The gratings will be used in Littrow configuration, so the exit beam has no anamorphic expansion. The camera must still accept the diverging spectrum or equivalently the full spatial field size, as in any spectrograph. Both collimator and camera designs are also eased by the fact that a transmission grating requires only a short pupil relief. This advantage would be lost if the instrument had to accommodate a collimator-camera angle of 135° (included angle 45°) for operation with reflection gratings;
- The base-line detector format is $4k^2$ with a pixel size of $15\mu\text{m}$, although a $6k \times 4k$ format is to be considered. Camera aberrations of less than $12\mu\text{m}$ rms are specified.

Figure 2. Resolution as a function of wavelength and line frequency; immersion prism angles are given in brackets.

3.2. Spectral Resolution

Figure 2 demonstrates spectral resolution, $\mathcal{R}$ -vs- wavelength for the proposed configuration (15cm beam, 1.5" slit, Littrow mount) for three different line density VPH gratings:

- 715 l/mm (with no prisms);
- 1680 l/mm (with 20° prisms) and
- 2890 l/mm (with 30° prisms).

These particular values of line density have been chosen to match the resolution of three common surface grating configurations (the last in 2nd order).

The curves for the 1680 and 2890 l/mm gratings are cut off at the long wavelength end where the angular deviation within the grating is 60°: at this value the loss of p-plane polarization reduces the average efficiency over the two polarizations to ~75% of its value for low dispersion. A resolution of nearly 10,000 is obtained at this point. These calculations allow for the dispersion of the immersion prisms (when used). Each point on these curves represents the performance if the system were set up with that wavelength as the centre wavelength.

Figure 3. An efficiency comparison for a 1200 l/mm surface grating blazed at 700 nm with respect to a 1683 l/mm immersed VPH grating.

3.3. Grating Efficiency

High efficiency is a key aspect of the ATLAS proposal. The efficiency of the optics is maximized by using a fully transmitting design and advanced anti-reflection coatings. The grating efficiency is substantially improved by using VPH technology. In this section we present quantitative data concerning the comparison of VPH with surface gratings.

Rigorous coupled-wave calculations of the diffraction efficiency of both surface and VPH gratings (as shown in Figure 3) gives a clear illustration of why we believe VPH gratings will replace conventional surface gratings in the next generation of optical/IR spectrographs. The figure shows a comparison of the efficiency as a function of wavelength of a 1200 l/mm surface grating blazed at 700 nm and a 1683 l/mm immersed VPH grating (both producing a resolution of $R=6865$ at $\lambda=700\text{ nm}$). The surface grating produces the usual rather wide blaze profile, with a peak efficiency of 68% , typical of such reflective gratings. The VPH grating efficiency is much greater, and peaks at 92% . The VPH efficiency drops off more rapidly away from the optimum wavelength - this is the bandwidth effect referred to above. But as the "scale bar" showing CCD coverage indicates, the VPH efficiency is still at or above that of the surface grating at all wavelengths which fall on the detector. (This calculation assumed a VPH grating thickness of $5\mu\text{m}$.)

Figure 4 illustrates the tunable nature of VPH gratings. It shows computed efficiency curves for a single grating (1553 l/mm , delivering $R = 6865$ at the design wavelength of 700 nm). The central curve applies when the grating tilt

Figure 4. A demonstration of the tunability of VPH gratings.

with respect to the incident beam is set to tune the wavelength centre to 700nm. By changing the grating tilt, the optimum wavelength is changed, as the other two curves show. The peak efficiency falls a little, due to the departure from the design wavelength, but the peaks as shown are still at 85%. The CCD coverage bar would slide along to follow the displaced peaks - due to the operation of the articulated camera. Thus a single VPH grating can be used at any centre wavelength over a broad range, but not over the entire visible spectrum.

Figure 5 shows a comparison of VPH and surface grating efficiency for the high resolution case, $R=8997$ at $\lambda=500$ nm. This corresponds to a 1200 l/mm immersed surface grating used in second order, or a 2894 l/mm immersed VPH grating (in first order). Again the peak efficiency of the VPH grating provides a major gain. The peak has dropped to about 80% because the loss of p-plane polarization is beginning to take effect, but the surface grating also suffers a considerable drop through being used in second order. At this high resolution the bandwidth effect of VPH gratings is more pronounced, and one might not use the entire 4k pixel wavelength range on the detector. Even so, the VPH efficiency exceeds that of the surface grating over most of the wavelength range.

Finally, figure 6 shows the results for a lower resolution configuration. With a 720 l/mm VPH grating the resolution is $R=2740$ at a λ of 700nm. Bandwidth effects are minimal at lower resolution, and at the design wavelength of the grating we have efficiency greater than 80% over almost the entire detected wavelength range. If the grating is tilted to select other centre wavelengths, efficiency greater than 65% can also be maintained from 430nm to over 1 μ m, with just the one grating.

Figure 5. A high dispersion version of Figure 3.

Figure 6. A low dispersion version of figure 5.

3.4. Overall Efficiency

The overall peak efficiency of ATLAS is estimated at 42% from above the atmosphere to detected photons. This represents a gain of about a factor of 1.5 to 2 over most modern spectrographs. It should be emphasized here that the ATLAS concept represents the basis for a whole new generation of 8-metre class spectrographs which will significantly out-class the present GMOS, DEIMOS, VIMOS instruments in its potential to deliver increases in field of view, throughput and spectral resolution coverage.

4. Design Issues

The ATLAS concept represents a considerable departure from current spectrograph designs, and as such it relies on a number of technical and engineering developments for its feasibility. In this section we summarize those issues which would need to be satisfactorily resolved during a full design study of the instrument.

4.1. VPH technology

- As noted above, KOSI have so far only produced VPH gratings in sizes up to 7.5×10 cm. In contrast ATLAS would need gratings up to 15×18 cm. It is our understanding, however, that there are no major technical impediments to the production of larger gratings; it is simply a matter of making the required commercial investment in the technology. Nevertheless, given the complete dependence of the ATLAS concept on VPH gratings, this issue must clearly be resolved before work proceeds to the construction phase.
- The restricted bandwidth of VPH gratings has the potential to limit performance. In the simulations shown above, the bandwidth effects were tolerable, in that the performance remained at least as good as that of surface gratings. The simulations assumed that gratings could be produced with a thickness of $5 \mu\text{m}$, in accordance with recent information from KOSI. However, investigation into further reductions in thickness and hence increases in bandwidth are to be highly encouraged. The uniformity of the thickness of the gratings is also an important issue – aberrations due to non-uniformity must not increase the image size more than about $15 \mu\text{m}$ rms.
- The possibility of an echellette mode for ATLAS would also be highly attractive. Barden and collaborators are investigating a 300 l/mm grating for operation in 10^{th} order. This is a higher risk development, since it requires non-sinusoidal refractive index modulations to achieve good efficiencies in higher orders. The success of ATLAS does not depend on getting a successful echellette mode, but given the large area of modern detectors it would be a very attractive facility and we believe it should be pursued. Cross dispersion could be provided by prisms mounted on the grating in a similar fashion to immersion prisms (but orthogonal to the dispersion direction) - this simply requires the slit and grating to be

rotated relative to their usual positions, and the articulated camera to be set at the appropriate angle.

4.2. Engineering aspects

1. Preliminary examination of the articulated camera concept does not indicate any over-riding difficulty. Nevertheless, it is clearly a challenge to have the camera movable while at the same time maintaining high stability of the images on the detector. We believe that a stringent limit for any flexure induced motion, such as 0.1 to 0.25 pixel in the worst case, should be aimed for. This is likely to require active compensation of some form, probably implemented via controlled motions of the detector. The error signal could be either from a lookup table or from a test beam within the spectrograph.
2. The grating mounting will also require engineering development, because of the considerable demands made of it: it must be rotatable under fine control, accommodate gratings which may have heavy prisms on them, allow interchange of gratings (and etalons) and introduce minimal flexure.

5. Other Facilities

ATLAS would be uncooled, and so would not be suitable for general infrared use. However, with a suitable detector, operation in J band and the shorter wavelength part of H could be envisaged. The design study will cover the operation of VPH gratings in the IR, and the suitability of the anti-reflection coatings.

An attractive option for use with ATLAS would be an Integral Field Unit (IFU), similar to the SPIRAL concept which is being developed by Parry, Taylor and collaborators (Parry et al., 1997) or the DensePak array developed by Barden and colleagues (Barden & Wade 1988). Such an IFU would provide two significant capabilities:

- Spatially resolved spectroscopy of objects up to about $400''$ (at $0.75''$ sampling);
- Access to higher resolution than the standard $1.5''$ slit, by providing an effective slit of $0.75''$. Resolutions, $\mathcal{R} \sim 20,000$ could then be achieved, although there would be efficiency losses in the IFU of order 20%. The f/2.2 camera would be able to provide critical sampling of the spectral profile (2 pixels per FWHM) at $0.75''$.

In the development of the ATLAS concept we must acknowledge substantial contributions made by Damien Jones (Prime Optics, Qld), guidance from KOSI and for many stimulating discussions from AAO colleagues Joss Bland-Hawthorn, Peter Gillingham, David Lee, Brian Boyle, Karl Glazebrook and many others.

References

- Barden S.C., Arns J.A. & Colburn W.S. 1998, *Proc.S.P.I.E.*, **3355**, 866
Barden S.C., & Wade R.A. 1988, *Fiber Optics in Astronomy*, A.S.P. Conf. Ser. **3**, 113
Bland-Hawthorn J. & Jones D.H. 1998a, *Publ. Astron. Soc. Australia*, **15**, 44
Bland-Hawthorn J. & Jones D.H. 1998b, *Proc.S.P.I.E.*, **3355**, 855
Glazebrook K. et al. 1998, *AAO Newsletter*, **84**
Parry, I.R., Kenworthy, M.A. & Taylor, K. 1997, *Proc.S.P.I.E.*, **2871**, 1325
Taylor K. 1990 *Modern Technology and its Influence on Astronomy*, (publ. CUP), eds. J.V.Wall & A.Boksenberg, 235.

Faint Object 2-D Spectroscopy: Promise and Limitations

M. M. Roth, T. Becker, and J. Schmoll

*Astrophysikalisches Institut Potsdam, An der Sternwarte 16,
D-14482 Potsdam, Germany*

Abstract. Since Integral Field Spectroscopy (3D) was introduced as a new astronomical observing method, it has predominantly been used for problems related to the structure and kinematics of galactic nuclei. More recently, it has been found that many other applications can take advantage of this technique. During the conceptual design phase of PMAS, an optical integral field spectrophotometer which is currently being developed at the Astrophysical Institute Potsdam, the use of 3D for faint object spectroscopy was considered to become a major task of this instrument. The issue is also interesting with regard to the 3D instruments which have been proposed for the NGST, probing deep into the universe at unprecedented faint flux levels. Several astrophysical applications are described as examples, along with a discussion of the technical problems that may be limiting 3D at faint magnitudes and a list of design features for PMAS which are intended to overcome these limitations.

1. Introduction

3D is generally an expensive technique in terms of detector space and the size of the spectrograph optics. Depending on the envisaged scientific priorities, one is forced to make a compromise between spectral resolution, spectral range, and field-of-view for a given number of available detector pixels. For PMAS, the science drivers are demanding medium spectral resolution, a large free spectral range, a moderate field of view, high throughput, and spectrophotometric capabilities. In particular the latter requirement has led to a design approach that attempts to provide an extremely stable, well-calibrated and efficient instrument. Correspondingly, PMAS will become a very useful tool for the investigation of single faint objects, limited by the background or source confusion.

Since data reduction algorithms will play an important role for obtaining high quality results, we have started observations with an existing 3D instrument in order to develop the relevant experience well before PMAS is becoming operational. Most of the ideas presented here emerged from observations with MPFS, a lens-array fiber coupled integral field spectrograph in the prime focus of the Selentchuk 6m Telescope (Afanasyev et al. 1990). Despite some limitations of the site and a relatively low throughput, our very positive experience with MPFS has illustrated convincingly that 3D in fact is an ideal tool for difficult observations of faint, background limited objects, provided one is able to control the problems which are outlined below.

Figure 1. Crowded field 2D spectroscopy of a star forming region in M33. The greyscale image represents the continuum, contours indicating the emission line surface brightness in H_{α} . Integrated spectra of the star to the upper right ($m_v \approx 18$) and a nearby H II region are plotted on the right. The example demonstrates how the subtraction of the nebular emission would have been underestimated from slit spectroscopy, taking e.g. background values from above and below the star with a vertically oriented slit. MPFS data kindly provided by S. Fabrika, V. Afanasiev, S. Dodonov (SAO).

2. Scientific Applications

Some scientific applications of faint object spectroscopy using 3-D methods are extragalactic stellar spectroscopy, extragalactic abundance determinations, spectral energy distributions and abundances in high redshift galaxies, and optical identifications of X-ray detections or follow-up to other observations with ill-defined positions (see Roth & Laux 1998).

The general argument is that for faint point sources, which are embedded in the surface brightness distribution of their host galaxy, the 3D method provides a huge advantage in terms of background subtraction, comparable to CCD photometry as opposed to aperture photometry (see Fig. 1). For extended sources 3D works as an effective image slicer, capturing essentially all the flux of the object — contrary to a slit. In each case, the larger number of sky elements around the target helps to provide a better estimate of the background level.

3. The Response Calibration Problem

For the observation of faint objects at less than a few percent of the sky brightness or any other local background surface brightness distribution, it is necessary to accurately calibrate the pixel-to-pixel non-uniformity of the detector response, otherwise flat-fielding errors will make it impossible to extract the object signal from the background noise. In 3D, in particular with instruments of the fiber

Figure 2. Response calibration residuals from the presence of 5% cross-talk between neighbouring spectra. This simulation of a 32×32 element 3D field demonstrates that even modest overlap perpendicular to the direction of dispersion, combined with a variable response from spectrum to spectrum (panels labelled 1%, 5%, ..., 20% r.m.s.), results in a systematic noise pattern, numbers of top and bottom row indicating the corresponding r.m.s errors.

coupled lens-array type, the individual spatial elements may exhibit a throughput variance that is much larger than the typical 1% non-uniformity of a good CCD. It is therefore mandatory to devise a very accurate calibration scheme, if one wants to measure faint objects with 3D. However, even if such a flat-field procedure provides a perfectly homogeneous illumination of the integral field unit, the presence of cross-talk between spectra with different throughput efficiencies results in a non-linear behaviour, giving rise to systematic response calibration errors for measurements when the illumination is *not* flat. An example of the effect is demonstrated in Fig. 2. If even through careful instrument design it is not possible to reduce these errors to an acceptable minimum, a way out may consist in observational strategies like the beam switching method. See also the interesting charge shuffling approach discussed by Tinney (2000).

4. Data Reduction Problems

The dedicated effort of data reduction algorithms also has an impact in how far the goal of faint object spectroscopy may actually be reached, or not. Leaving aside the problem of complexity when dealing with hundreds of spectra on a single CCD frame, there are a several key issues that need special attention. First of all, the extraction method is a concern. One can simply integrate the signal in a certain swath, use optimal extraction, or apply more complicated algorithms in the difficult situation when the spectra are located very close to each other and show overlap in the wings. Taking into account the response cali-

Figure 3. Variation of PSF across the detector. The greyscale image represents the slit function of MPFS on 14 different positions of the detector (black: peak intensity, white: zero), each horizontal row of pixels corresponding to the profile of a monochromatic fiber image, derived from a special etalon calibration mode (cf. Fig. 4).

bration problem, it is highly desirable to have a precise knowledge of the overlap contributions which may then be calibrated out. However, in the situation of closely packed spectra, this information is not normally available (an extremely useful exception being the etalon mode at MPFS). Fig. 3 shows that generally one should not simply make the ad-hoc assumption of a constant cross-talk contribution for each fiber. Next, the effect of CCD non-uniformity must be taken into account. If mechanical flexure of the spectrograph is present during the exposures, a compensation in software should be attempted. Finally, the removal of instrumental artefacts like straylight and ghosts must be accomplished.

5. Instrumental Effects

Like in ordinary spectrographs, the presence of instrumental effects can have a disturbing influence on the limiting accuracy, e.g. straylight and ghost images, illumination effects of calibration sources, and, in the case of fiber instruments, the effect of optical fibers on the spectrograph PSF. Concerning reflected light and the diffuse instrumental background, Wyse and Gilmore (1992) have emphasized the importance of actually measuring these contributions for faint object spectroscopy with fiber MOS, and the same argument is true for 3D instruments. Fig. 4 gives an example of the effect that could be measured in the case of MPFS. An etalon calibration frame (see caption for explanation) was conveniently tested for the presence of straylight — which is easy to detect in this case because of the large spacing of the etalon spectra. On normal science frames it is much harder to accurately measure the diffuse background because of the tightly packed arrangement of the spectra. Wyse and Gilmore have pointed out that without a measurement of the background from gaps between the spectra or otherwise blank detector space “*one would remain blissfully unaware*” of these artefacts, and the corresponding systematic errors.

Another instrumental effect occurring in fiber spectrographs with influence on the cross-talk problem is the possible dependance of the spectrograph PSF on the transmission characteristics of optical fibers. The phenomenon of focal ratio degradation which is sensitively depending on individual circumstances, in particular the onset of mechanical stress, results in random variation of the fiber output far field patterns, giving rise to different PSFs in the case of spectrographs

Figure 4. Diffuse straylight on the detector as derived from the MPFS etalon calibration mode. In this mode, only one dedicated fiber for every group of 15 is illuminated by light that has passed a Fabry Perot etalon, producing a chain of spots for each calibration spectrum, i.e. etalon orders, left image. Right image: diffuse background, etalon spots removed. Etalon frames are normally used for determining the geometry and tracing of the MPFS science spectra. They are also useful for an investigation of the behaviour of the spectrograph PSF, see Fig. 3.

that need re-imaging microlenses to accommodate slow spectrograph beams. In addition, modal noise may introduce temporal variations making it very hard to achieve accurate calibrations if the fluctuations are strong (on the order of 1%). The effects have been investigated in the lab and with real spectrographs by Schmoll et al. (1998), and Becker et al. (2000).

Again like in ordinary slit spectrographs (as well as in direct imaging cameras), the illumination from calibration light sources, which is usually not perfectly identical to the telescope illumination, is a potential source of errors. This issue should be given particular attention for the design of the NGST instruments, being intended to measure at the very faintest flux levels. Regarding this task, the presence of an intermediate pupil in the foreoptics of lens-array 3D instruments in combination with fiber scrambling is a useful feature for improving the illumination.

6. PMAS — The Potsdam Multiaperture Spectrophotometer

Taking into account the various instrumental pitfalls with influence on the limiting magnitude, we have chosen a design which should make PMAS particularly suitable for observing in the background-limited case (for a more complete description, see Roth et al. 1998).

PMAS is based on the lens-array, fiber-coupled spectrograph principle. First of all, other than with integral field units that come as an add-on module to a conventional slit spectrograph, PMAS employs a *dedicated fiber spectrograph*

with high throughput and a very good image quality. Because the fibers are coupled directly to the spectrograph with no need to reimage into the pseudo-slit, FRD and modal noise do not transform into PSF variations. The fiber interfaces at the lens-array and at the collimator input are optimized for throughput and for the least possible introduction of stress, using a technique of index matching immersion. The optical fiber train will be optimized for very small throughput variance from fiber to fiber, poorly performing fibers being removed and replaced by better ones. Starting with a prototype configuration, the total number of spatial elements, i.e. spectra on the CCD, has deliberately been chosen to be relatively small ($256 = 16 \times 16$) in comparison with the nominal capacity of the spectrograph which may accept up to 1024 fibers. This will allow us to better understand and control quantitatively the straylight properties and overlap characteristics thanks to a wide spacing of the spectra. A novel design lens-array with aspheric surfaces strongly reduces the aberrations at the critical point where the micropupils must be centered on the tiny diameter of the fibers. The re-imaging foreoptics connecting the telescope focal plane to the lens array has been optimized, providing an accessible intermediate pupil which is extremely useful for adjustment purposes. It can be connected to an internal calibration source whose optical elements have an apochromatic correction and which provides a nearly perfect, white illumination of the micropupils. The data reduction software is being developed with special care for a correct treatment of the various instrumental effects.

The PMAS project is supported by the German Verbundforschung under grants 05 3PA4141 and 05 AL9BA1, and by DFG under HA1850/10.

References

- Afanasiev, V.L., Vlasiuk, V.V., Dodonov, S.N., Silchenko, O.K. 1990, Preprint No.54, Special Astrophysical Observatory, Selentchuk
- Becker, T., Roth, M., Schmoll, J., 2000, this volume
- Roth, M.M., Seydack, M., Bauer, S.M., Laux, U. 1997, SPIE 2871, 1235
- Roth, M.M., Laux, U. 1998, in *Fiber Optics in Astronomy III*, ASP Conf. Ser. 152, eds. Santiago Arribas, Evencio Mediavilla, Fred Watson, p. 168
- Roth, M.M. et al. 1988, in *Optical Astronomical Instrumentation*, ed. S. D'Odorico, SPIE3355, 798
- Schmoll, J., Popow, E., Roth, M.M. 1998, in *Fiber Optics in Astronomy III*, ASP Conf. Ser. 152, eds. Santiago Arribas, Evencio Mediavilla, Fred Watson
- Tinney, C. 2000, this volume
- Wyse, R.F.G., Gilmore, G. 1992, MNRAS 257, 1

Probing the Metallicity Evolution of Disk Galaxies via Imaging Spectroscopy

D. R. Garnett

Steward Observatory, University of Arizona, Tucson, Arizona 85721

Abstract. Imaging spectroscopy of star-forming galaxies at intermediate to high redshifts can greatly expand our understanding of the metallicity evolution of disks, through study of nebular emission and stellar absorption lines arising from the star-forming regions themselves. In contrast, QSO absorption line systems sample only pinhole regions in the outer parts of individual galaxies. Imaging spectroscopy on large ground-based and space-based telescopes can most efficiently study star-forming disks at high spatial resolution. Simple calculations indicate that a ground-based 8-meter class telescope working at 0.2 arcsec resolution can study individual giant H II regions in galaxies at $z \approx 0.5$ with reasonable exposure times. More distant galaxies will probably require a large space-based telescope. Such observations can provide complementary data on dynamical, morphological, and chemical evolution over time that will greatly enhance our understanding of galaxy evolution.

1. Heavy Element Abundances and Galaxy Evolution

Although we now have a fairly good understanding of the distribution of heavy element abundances in the present-day interstellar medium of disk galaxies, we still have only a rudimentary idea of how the metallicity evolves with time. In the solar neighborhood, the age-metallicity relation for stars (e.g., Edvardsson et al. 1993) provides a model for local metallicity evolution, but this is restricted to the region within 100 pc of the sun. Recent studies of star clusters in the Magellanic Clouds are now providing data on metallicity evolution in our nearest neighbors (Olszewski, Suntzeff, & Mateo 1996), and with predictable improvements in light-gathering power it is reasonable to expect such work to shed light on the evolution of most of the galaxies in the Local Group. At the same time, it is unlikely in the foreseeable future that we will be able to obtain such data for galaxies beyond the Local Group.

Given that restriction, studies of metallicity evolution for a variety of galaxy types will have to focus on evolution as a function of redshift. A number of issues are ripe for further study: (1) is the metallicity evolution of galaxies consistent with the star formation rate over time (cf. Madau et al. 1996)? (2) How important is SN-driven ejection of metals into galaxy halos in the evolution of abundances? (3) Does environment affect metallicity evolution? There is evidence that the interaction of a disk galaxy with hot cluster gas affects ISM abundances, possibly as a result of truncation of cosmological infall (Skillman

et al. 1996). This needs to be confirmed with larger statistical samples. (5) In spiral galaxies we radial composition gradients as a rule. How do these gradients evolve with time? Models for metallicity evolution in spirals predict a variety of behavior: some predict that the gradients flatten with time (e.g., Carigi 1996, Mollà, Ferrini, & Diaz 1996), while others predict steepening gradients with time (e.g., Chiappini, Matteucci, & Gratton 1997). The critical factor appears to be the timescale for infall of material, with slower infall leading to steepening composition gradients; the range of infall timescales studied lies between 3 and 8 Gyr. Thus, an observational determination of the rate at which composition gradients change with time could constrain the gas infall rate in spirals. Secular evolution influenced by bar-driven gas radial inflows may also be a factor (Friedli & Benz 1995).

The first steps toward understanding metallicity evolution for disk galaxies in general have been taken in studies of damped Lyman α absorption line systems toward QSOs (Pettini et al. 1999, Lu et al. 1996). These studies have shown remarkably little metallicity evolution in galaxies with $0.7 < z < 3.0$. This may be in part due to the fact that the DLAs sample only pinprick regions within galaxies, and the connection between the absorbing gas and associated bright galaxy is tenuous. Examination of a nearby DLA system (the M81 system; Yun, Ho, & Lo 1994) provides an illustration. Within 50 kpc (projected) of M81 are a dusty starburst (M82), a low-luminosity peculiar irregular (NGC 3077), a low-surface-brightness irregular (Ho IX), and a variety of tidal streamers. Random sightlines through this system would penetrate a variety of metallicity regimes. The gaseous environments of high-redshift DLA systems may be similarly complex. Thus, measurement of abundances in the visible star-forming regions of high redshift galaxies may provide a more reliable measure of abundance evolution.

An example of what can be learned from intermediate-high redshift abundance studies is provided by Kobulnicky & Zaritsky (1999), who observed a set of compact narrow emission line galaxies (CNELGs) with $z = 0.1-0.5$, postulated to be the progenitors of local dwarf spheroidal elliptical galaxies. Kobulnicky & Zaritsky, however, demonstrated that the metallicities of these systems are greater than those of nearby dEs, and thus inconsistent with the original hypothesis. The nature of these CNELGs is therefore still unknown.

The Keck observations of the relatively luminous CNELGs by Kobulnicky & Zaritsky (1999) are near the limit of what can be reasonably done today. However, in the near future it should be possible to improve upon this with the commissioning of a new generation of 8-meter class telescopes with adaptive optics. Moreover, the possibility that imaging spectrometers will be among the instrument arsenal raises exciting possibilities for studying galaxy structural and metallicity evolution over long lookback times. I discuss some of the observational considerations below.

2. Observational Considerations

I will discuss (without prejudice as to the preferred instrument) the logistics of an observational program of spatially-resolved study of abundances from H II regions in spirals at redshifts out to $z \approx 1$. Metallicity studies of distant galaxies

would benefit greatly from imaging spectroscopy. Conventional slit spectroscopy of H II regions in nearby spirals generally obtains data for only one or two objects at a time, even with a long slit. Acquiring a sufficient number of H II region measurements (≥ 10 regions) to characterize the metallicity distribution in a galaxy thus requires an entire night or more. Imaging spectroscopy, by contrast, allows one to obtain data for an entire galaxy simultaneously, given a sufficiently large area imaged. Imaging spectroscopy can also simultaneously obtain a wealth of supporting data: stellar absorption line strengths (which trace metallicity evolution in the stellar component), rotation curves from the emission lines, and even surface photometry (with judicious choice of spatial and spectral binning). Thus, much of the chemical and structural information needed to compare distant spirals with nearby ones can be obtained with a single observation!

The primary abundance indicator for H II regions in the rest-frame visual spectrum is the emission-line intensity ratio

$$R_{23} \equiv \frac{I([OII]\lambda 3727) + I([OIII]\lambda\lambda 4959, 5007)}{I(H\beta)}.$$

The R_{23} index has an empirically calibrated relationship with oxygen abundance O/H (e.g., Edmunds & Pagel 1984). This relation becomes double-valued for O/H < 0.2 solar, but the ambiguity may be resolved in principle by measuring the [N II] $\lambda 6583$ line. Thus, a minimum observation set consists of measurements of [O II] $\lambda 3727$, [O III] $\lambda 5007$, H β , and a line to estimate reddening. $H\alpha/H\beta$ is the most robust reddening indicator, but could be difficult to measure when $H\alpha$ is redshifted beyond $1\mu m$ for $z > 0.5$; higher-order Balmer lines can be used with increased uncertainty in the reddening estimate. I take as the limiting factor a measurement of the H β line with 10% uncertainty; in a typical spiral, such as M101 (Kennicutt & Garnett 1996), all but a small fraction of H II regions have either $I([O II]) > I(H\beta)$ or $I([O III]) > I(H\beta)$, or both, so the error in H β limits the accuracy with which we can measure R_{23} .

The various chemical evolution models for spirals generally predict a factor two to three increase or decrease in the slope of the metallicity gradient at $z \approx 1$. Present-day spirals have composition gradients of about 0.2-0.3 dex per unit disk scale length, or about 1-1.7 dex over an optical radius (Zaritsky, Kennicutt, & Huchra 1994, Garnett et al. 1997). Measuring a change in the gradients should therefore be easy, given sufficient sensitivity to obtain (≥ 10) H II region measurements per galaxy.

With the H β flux as the limiting factor, we can use the H II region luminosity function to determine how faint we need to go to observe an interesting number of regions per galaxy. The work of Kennicutt, Edgar, & Hodge (1989) and Rozas, Beckman, & Knapen (1995) show that typical luminous spirals have ≥ 10 H II regions with $L(H\beta) > 10^{38.7}$ ergs s $^{-1}$, a sufficient number of regions for determining the slope of any composition gradient. H II region of this luminosity will have sizes of order 0.5-1 kpc. This sets the desired spatial resolution element: 1 kpc corresponds to approximately $1''$ at 200 Mpc, $0.2''$ at $z = 0.5$, and $0.1''$ at $z = 1$. For a galaxy similar to M101, with an optical radius of about 30 kpc, the useful spectrometer field of view would have a diameter of about $60''$ at 200 Mpc and $7''$ at $z = 1$. For illustration purposes, I assume a cosmology with $H_0 = 65$ km/sec/Mpc and $\Omega = 0.2$ and no cosmological constant.

Figure 1. Predicted S/N for $H\beta$ as a function of exposure time and redshift for the case of an 8.4-m ground-based telescope working at 0.2 arcsec resolution. The solid curves show results for $\log L(H\beta) = 39.0$ at exposure times of 1000s, 3600s and 10^4 s (left to right curves); the dotted curves are for $\log L(H\beta) = 38.7$ for the same exposure times. See text for details of the calculations.

Two cases will be discussed here. Case 1 is for an 8.4 meter ground-based telescope with adaptive optics capable of delivering 0."2 seeing for $\lambda > 5000$ Å. Case 2 is for an 8.4-meter telescope in space, with an imaging spectrometer having 0."1 spatial elements. I assume in both cases what I hope are a reasonable set of operating parameters for the spectrometer and detector: spectral coverage from 5000-10000 Å; $> 20\%$ throughput for telescope + spectrograph across the spectrum; detector read noise of $5 e^-$; dark current $< a$ few counts/hour; and spectral resolution $\lambda/\Delta\lambda$ of 5,000-10,000 (which is optimum for both emission line detectability and OH line avoidance); and of order 3-8 read-outs for long exposures. I have also assumed the following values for the night sky surface brightness in mag/arcsec²: $R = 20.9$, $I = 19.9$ (ground-based); $R = 22.3$, $I = 21.9$ (space).

Figure 1 shows the expected signal-to-noise ratios for $L(H\beta) = 10^{39.0}$, $10^{38.7}$ for Case 1 as a function of exposure time and redshift, while Figure 2 shows the same calculations for the space-based telescope of Case 2.

The figures show that a ground-based 8-m class telescope working at 0."2 resolution can study several individual star-forming regions and composition gradients in galaxies at intermediate redshifts ($z = 0.3$ - 0.5) in reasonable exposure times of 1-3 hours per galaxy. Simple metallicity evolution from observation of the most luminous H II regions could be done to higher redshifts. However, it appears that to spatial composition variations in galaxies at $z \approx 1$ may re-

Figure 2. Predicted S/N for H β as a function of exposure time for the case of a space-based 8.4-m telescope working at 0.1 arcsec resolution. The key is the same as in Figure 1. The solid lines are for exposure times of 3600s and 10⁴s, while the dotted curves are for 10⁴s and 10⁵ s.

quire exposure times of order 10⁵ s per target. This would require a space-based telescope such as NGST that will have long observing windows. A wide-field multi-object capability (multiple IFUs, for instance) would enhance the feasibility of such a project.

3. Discussion

These rough calculation show that it is feasible, with planned telescopes and instruments, to study metallicity evolution, as well as galaxy dynamics and structure, at relatively high spatial resolution at redshifts up to 1, a significant fraction of the lookback time. It should be possible to do this with large ground-based telescopes to $z \approx 0.5$ with reasonable exposures. More distant galaxies will likely require a space-based observatory that allows for long exposure times. The new prospects for imaging spectroscopy open up exciting prospects for direct study of the morphological, dynamical, and chemical evolution of disk galaxies.

DRG thanks the organizers of this conference for the opportunity to speak, and gratefully acknowledges support from NASA through Long Term Space Astrophysics grant NAG5-7734.

References

- Carigi, L., 1996, RMxAC, 4, 123
- Chiappini, C., Matteucci, F., & Gratton, R., 1997, ApJ, 477, 765
- Edvardsson, B., Andersen, J., Gustafsson, B., Lambert, D. L., Nissen, P. E., & Tomkin, J., 1993, A&A, 275, 101
- Edmunds, M. G., & Pagel, B. E. J. 1984, MNRAS, 211, 507
- Friedli, D., & Benz, W. 1995, A&A, 301, 649
- Garnett, D. R., Shields, G. A., Skillman, E. D., Sagan, S. P., & Dufour, R. J., 1997, ApJ, 489, 63
- Kennicutt, R. C., Jr., Edgar, B. K., & Hodge, P. W., 1989, ApJ, 337, 761
- Kennicutt, R. C., Jr., & Garnett, D. R., 1996, ApJ, 456, 504
- Kobulnicky, H. A., & Zaritsky, D., 1999, ApJ, 511, 118
- Lu, L., Sargent, W. L. W., Barlow, T. A., Churchill, C. W., & Vogt, S. S., 1996, ApJS, 107, 475
- Madau, P., Ferguson, H. C., Dickinson, M. E., Giavalisco, M., Steidel, C. C., & Fruchter, A., 1996, MNRAS, 283, 1388
- Mollá, M., Ferrini, F., & Díaz, A. I., 1997, ApJ, 475, 519
- Olszewski, E. W., Suntzeff, N. B., and Mateo, M. 1996, ARA&A, 34, 511
- Pettini, M., Ellison, S. L., Steidel, C. C., & Bowen, D. V. 1999, ApJ, 510, 576
- Rozas, M., Beckman, J. E., & Knapen, J. H., 1996, A&A, 307, 735
- Skillman, E. D., Kennicutt, R. C., Jr., Shields, G. A., & Zaritsky, D., 1996, ApJ, 462, 147
- Yun, M. S., Ho, P. T. P., & Lo, K.-Y., 1994, Nature, 372, 530
- Zaritsky, D., Kennicutt, R. C., Jr., & Huchra, J. P., 1994, ApJ, 420, 87

Science & Surveys with Integral Field Spectrographs

Roger L. Davies

*Physics Department, University of Durham, Rochester Building, South
Rd., Durham, DH1 3LE, UK*

Abstract.

The analysis of galaxy dynamics and stellar populations are central to advancing our understanding of how galaxies are assembled. Optical measurements were firmly established as important astrophysical diagnostics when quantitative long-slit spectroscopy came into widespread use in the 1970s. Through this work we have learned that galaxies are complex dynamical systems, with a rich and varied star formation history and extended dark halos. The limitations of long slit spectroscopy for mapping the motions of gas and stars or the stellar populations of galaxies have severely hindered further progress. Integral Field Spectroscopy, where whole galaxies can be mapped in a single integration, promises to revolutionise the study of nearby galaxies of all types. Here I describe recent work on mapping the stellar populations, the velocity fields and the physical conditions in galaxies.

1. Introduction

Our understanding of the formation and evolutionary history of the Milky Way arises in large part from the connections that have been made between the age and metallicity of stellar populations and their dynamics. Those connections illuminate the early history of the Milky Way, for example through the low rotation, large anisotropy and low abundances of the globular clusters, and carry through to ongoing events such as the assimilation of the Sagittarius dwarf spheroidal galaxy where the orbital period is found to be less than one-tenth of the age of the oldest stellar components (Ibata et. al. 1997). We are beginning to make such connections in external galaxies (eg. the Mg- σ relation for early type galaxies, Colless et al 1999) but to understand the complex dynamics and varied star formation history within galaxies we need to be able to map them efficiently. Integral field spectroscopy (IFS) promises to revolutionise this field by enabling emission line and absorption line maps of galaxies to be made in both the optical and the infrared that will match in precision and detail the HI data cubes of spiral galaxies that were crucial to establishing beyond doubt the presence of dark halos around spiral galaxies (eg. van Albada et al 1985).

In addition to the huge increase in data gathering power offered by IFS, a key aspect of the interpretation of the spectra is the ability to re-construct the physical location where each spectrum arises. Here I present a look at new data from optical and infrared IFS that highlight how area spectroscopy

Figure 1. $H\beta$ vs $[MgFe]$ for the early type galaxies in the Fornax cluster, the elliptical galaxies are filled circles, the open circles are the S0 galaxies. Overplotted is the expected relationship, for single stellar populations with a range of age and metallicity, from the models of Worthey.

provides physical diagnostic data that would be very much harder to interpret unambiguously if obtained by the, far more time consuming, long slit method. The next three sections of this paper are case studies in the use of IFS for probing the evolutionary history of early type galaxies (section 2), mapping the velocity field in the obscured nucleus of a radio galaxy (section 3) and in section 4 probing the physical conditions in the gas surrounding the nucleus of NGC 4151. I summarise in section 5.

2. Stellar populations in early-type galaxies

Determining the age of the stellar population in early type galaxies from measurements of the integrated light has been a controversial topic throughout the 1990's (Faber et al., 1995). Separating the effects of changes in the heavy element abundances from changes in age is difficult and requires careful modelling. For example in Worthey's models of the Lick indices (Worthey 1994, Trager et al. 1998) the strength of the Balmer absorption lines are successfully used as an age discriminator and combinations of magnesium and iron lines are used to measure metallicity. Figure 1 taken from Kuntschner and Davies (1998) shows that elliptical galaxies in the Fornax cluster form a sequence of roughly constant age, varying in metallicity from solar values to three times solar abundance. The S0 galaxies however populate a region lower in metallicity and have a much greater spread in age, from luminous objects as old as ellipticals to lower luminosity S0s that have luminosity weighted ages of just 2-3 billion years.

Figure 2. The relationship between the strength of the Fe3 index and Mgb index for the elliptical galaxies in the Fornax cluster. For each galaxy several points are plotted showing the radial trend from the centre (taken from Kuntschner 1998). Overplotted is the expected relationship for single stellar populations with a range of ages and metallicity from the models of Worthey (1994).

Figure 3. As figure 2 for S0 galaxies in the Fornax cluster.

Luminous elliptical galaxies exhibit an overabundance of magnesium (Peletier, 1989, Faber et al. 1992, Davies et al. 1993) which has been interpreted as evidence for an early phase of star formation (Bender 1996). This is illustrated in Figure 2 (taken from Kuntschner 1998) for the elliptical galaxies in Fornax. A similar result for some of the low luminosity elliptical galaxies in the Virgo cluster has been reported recently by Halliday (1998). Kuntschner shows, for the first time, that most Fornax S0 galaxies do *not* exhibit a magnesium overabundance (see figure 3). If this proves to be a general result it suggests a systematic difference in the star formation history of early type galaxies with and without disks. An appealing scenario is that ellipticals experience an early burst of star formation producing plentiful Mg from type-II supernovae that is locked into the first generations of stars. Subsequently a wind, established in a time less than that needed for significant iron production from type-I supernovae (1-2 Gyrs), removes the galaxy's ISM and inhibits any further star formation. The galaxy would form the bulk of its stars in less than 2 Gyrs and would have an excess of magnesium over iron. If in the disks of S0s this loss mechanism does not function, later generations of stars (still roughly 5 billion years old) will build up the solar abundance ratios as has happened in the disk of our own Milky Way.

We wish to explore whether these differences arise primarily because the disk population in the S0s is younger and has continued to form stars over an extended period or whether these are primarily differences between massive and less massive galaxies. By mapping line indices across entire galaxies we can determine the age and metallicity of the stellar population in each morphological component to discover *where* this population with strong $H\beta$ absorption and Mg-overabundance resides in the galaxy.

The SAURON consortium¹ has built a wide field, integral field spectrograph and recently commissioned it on the 4.2m William Herschel Telescope on La Palma (see Miller et al in this volume). This instrument has a $33'' \times 41''$ field-of-view and takes spectra from 4800Å to 5400Å. Using SAURON the team are mapping the $H\beta$, Mg and Fe features in elliptical, S0 and early-type spiral galaxies to determine the location of the strong $H\beta$ absorption population. Figure 3 shows the preliminary results for NGC3377, an E5 elliptical with a disk extending from about 5-40''. The figure shows the image of NGC3377 created by summing the spectra, and images in the Mgb, Fe 5270, and $H\beta$ indices of the Lick system. Analysis of these and similar spectra will allow us to determine whether the younger population resides in the disk component.

3. A stellar velocity map on the centre of NGC1316

NGC1316 is the most luminous early type galaxy in the Fornax cluster, $M_B = -21.9$, based on a distance modulus of 31.3 deduced by Silbermann et al. (1998) from the Cepheids in NGC 1365. It lies somewhat on the periphery of the cluster, is a very dusty, S0 galaxy and hosts the radio source Fornax A. Schweizer

¹R. Bacon, P. T. de Zeeuw, E. Emsellem, J. R. Allington-Smith, M. Carollo, G. Monnet, Y. Copin, R.F. Peletier, M. Bureau, B. Miller, E. Verolme & the author

Figure 4. Images of the E5 galaxy NGC3377 in the total light (top left) and in the absorption lines of magnesium, iron and $H\beta$ taken using the SAURON spectrograph on the William Herschel telescope.

Figure 5. The intensity map (top), velocity field (lower left) and velocity dispersion map (lower right) of the luminous S0 galaxy NGC 1316 (Fornax A) taken using the MPE 3D instrument at the Anglo-Australian Observatory. The shaded bars below the lower figures indicate how the grey-scale maps to velocity.

(1981) reported detailed ground based photometry and spectroscopy, Shaya et al. (1996) present the HST Planetary Camera images. Extended stellar kinematics were measured in the optical by D'Onofrio et al. (1995) and by Longhetti et al. (1998). The stellar rotation velocity shows a shallow rise to 50 km/s in the central 10'', (a radius of 0.9 kpc at a distance of 18.3 Mpc). The velocity dispersion is 250 km/s and roughly constant with radius.

Kormendy and Richstone (1995, KR) show that there is a correlation between the luminosity of the spheroidal component of early-type galaxies and the mass of the central black hole that they host. One might therefore expect a luminous elliptical hosting a radio source to harbour a massive black hole. KR find that typically the BH mass is 0.2% of the mass of the spheroid. Their relationship suggests that NGC1316 should contain a nuclear black hole in excess of $3 \times 10^9 M_{\odot}$. Shaya et al present models with a stellar mass-to-light ratio that is constant and add a central dark mass. The available kinematic measurements led them to conclude that a central mass of $2 \times 10^9 M_{\odot}$ could be present, alternatively they recognised that an increase in the radial anisotropy towards the centre could account for the measurements without a central mass. The published kinematics are based on optical measurements using Fourier Quotient techniques, it is possible that significant random or systematic motions are hidden by the extensive obscuring dust. To map the stellar kinematics free from

Figure 6. The emission in [FeII] (left) and Pa β (right) in the nucleus of NGC 4151. The box is $8 \times 6''$. The galaxy center is marked with a cross. North is in the top right, east in the top left. This figure is taken from Haynes et al. (1999)

the effect of obscuring dust we² used the MPE 3D IFS (Weitzel et al. 1996) at the Anglo-Australian Observatory to map the stellar kinematics using the first overtone CO lines at 2.2μ , over the central $6''$, with a spectral resolution ($\lambda/\Delta\lambda$) of 2100, in $0.7''$ seeing. The results are shown in figure 5. It is immediately apparent, even in this preliminary reduction of modest quality data, that there is no steep central velocity gradient, nor is there a highly peaked dispersion profile, that might be taken to be indicative of an otherwise hidden black hole. The central velocity dispersion determined using the FCQ technique (Bender 1990) from the infrared 3D data, gives a central velocity dispersion of ~ 230 km/s at all radii less than $1''$, in fact new optical data analysed in the same way also shows a low velocity dispersion. Using Shaya et al.'s models this low value of the central dispersion implies a central dark mass somewhat less than $10^9 M_\odot$.

4. The physical conditions in the gas in the centre of NGC4151

The environs of active nuclei host a complex mix of astrophysical phenomena, star formation and shocks as well as the active nucleus itself are likely to contribute to the ionising radiation field. A further complication is that galaxy nuclei are frequently very dusty (Peletier et al. 1999), requiring line ratio analyses to be carried out in the infrared if we are to be confident in the results. In-

²J. E. H. Turner, P. T. de Zeeuw, J. R. Mould, H. Kuntschner, L. E. Tacconi-Garman and the author,

frared IFS provides a powerful tool to map the source(s) of ionisation in AGNs. A Durham-Victoria team³ have used the new SMIRFS, infrared integral field unit (Haynes et al. 1999) at UKIRT to map the infrared [FeII]/Pa β ratio in NGC4151. This ratio is a sensitive indicator of whether the ionisation arises from young stars or shocks, [FeII] is 10 times stronger than Pa β in supernova remnants where shocks ionise the gas, but 100 times weaker than Pa β in Orion where the photoionising radiation arises from young stars. Figure 6 shows the distribution of light in the [FeII] and Pa β . The Pa β flux is centred on the stellar nucleus but the [FeII] emission peaks SW of the nucleus in the region that is roughly coincident with the extended narrow line [OIII] emission seen by HST and the radio jet seen by MERLIN at 6cm. The angular resolution of the SMIRFS images are considerably worse than those from HST/MERLIN and so the interpretation of these features remains uncertain. Nevertheless, the spatial coincidence of the region of high [FeII]/Pa β ratio with the radio jet and NLR suggest that the ionisation could arise from supernova remnants that result from episodes of star formation induced when the plasma jet impacts onto the circumnuclear ISM. High angular resolution, spatially resolved spectroscopy (perhaps, in the future, benefiting from adaptive correction) spanning a wide range of wavelengths is a powerful technique that can be used to unravel the mechanisms at work in complex environments.

5. Summary

I have illustrated a limited range of astronomical applications of IFS taking examples from extragalactic astronomy that involve mapping both emission and absorption lines from optical to infrared wavelengths. These techniques are particularly powerful when we wish to test hypotheses relating to the spatial structure of objects, and to identify where particular physical processes are occurring when we have no prior knowledge of the geometry. There is a rich variety of types of IFS under construction or development (many discussed at this symposium) each emphasising the different advantages of the IFS technique in specific applications. Many IFS systems will be brought into use on the new generation of 8m class telescopes. It seems very likely that these instruments will replace the traditional single object slit spectrographs over the next decade.

I am particularly grateful to my collaborators in the SAURON team : R. Bacon, P. T. de Zeeuw, E. Emsellem, J. R. Allington-Smith, M. Carollo, G. Monnet, Y. Copin, R.F. Peletier, M. Bureau, B. Miller and E. Verolme for allowing preliminary SAURON data to be used to illustrate this paper. I thank J. E. H. Turner, P. T. de Zeeuw, J. R. Mould, H. Kuntschner, L. E. Tacconi-Garman and the 3D team who are my collaborators in the investigation of NGC1316. I am also grateful to the IFS team in Durham: J. R. Allington-Smith, R. Content, G. Dodsworth, R. Haynes, J. E. H. Turner and J. Webster for their continuing innovation and dedication in implementing a challenging series of integral field spectrographs.

³C. Done, J. R. Allington-Smith, J.E.H. Turner, S. Morris & S. Chapman

References

- Bender R. 1990, A&A., 229, 441.
- Bender, R. 1996, Proc. IAU Symp 171, P. 181, eds. R. Bender & R.L. Davies, Kluwer, Dordrecht.
- Colless, M.C., Burstein, D., Davies R.L., McMahan R.K., Saglia R.P., & Wegner, G. MNRAS., 303, 813.
- Davies R.L., Sadler E.M., Peletier R.F., 1993, MNRAS., 262, 650.
- D'Onofrio, M., Zaggia, S.R., Longo, G., Caon, N., Capaccioli, M. 1995, A&A., 296, 319.
- Faber S.M., Worthey G., & González J.J., 1992, Proc. IAU Symp 149, eds. B. Barbuy & A. Renzini, Kluwer, Dordrecht.
- Faber S.M., Trager S.C., González J.J., Worthey G., 1995, Proc. IAU Symp. 164, eds. Van Der Kruit P.C., Gilmore G., Dordrecht, p. 249.
- Halliday C., 1998, PhD thesis, University of Durham, Durham.
- Haynes, R., Lee, D., Allington-Smith J.R., Content R., Dodsworth, G., Lewis, I., Sharples R., Turner J., & Webster J., 1999, accepted for publication in PASP.
- Ibata, R.A., Wyse, R.F.G., Gilmore, G., Irwin, M. J., & Suntzeff N.B., 1997, AJ., 113, 634.
- Kormendy J. & Richstone D., 1995, Annual Reviews of Astronomy and Astrophysics, 33, 581 (KR).
- Kuntschner, H., 1998, PhD thesis, University of Durham, Durham.
- Kuntschner, H., & Davies R.L., 1998, MNRAS., 295, L29.
- Longhetti M., Rampazzo, R., Bressan, A., and Chiosi C., 1998, A&A Suppl., 130, 267.
- Peletier R.F., 1989, Ph.D. thesis, Univ. of Groningen, Groningen.
- Peletier R. F., Balcells, Davies R.L., Andredakis Y., Vazdekis, A., Burkert A., & Prada F. 1999, submitted to MNRAS.,
- Schweizer, F., 1981, ApJ., 246, 722.
- Shaya, E.J., Dowling D.M., Currie, D.G., Faber, S.M., Ajhar, E.A., Lauer, T.R., Groth, E.J., Grillmair C.J., Lynds, C.R., & O'Neil, E.J., 1996, AJ., 111, 2212.
- Silbermann N. A., Harding, P., Ferrarese, L. Stetson, P.B. Madore, B. F., Kennicutt, R. C., Freedman, W. L., Mould, J. R. Bresolin, F., Ford, H., Gibson, B. K., Graham, J. A., Han, Mingsheng; Hoessel, J. G., Hill, R. J., Huchra, J., Hughes, S. M. G., Ilingworth, G. D., Kelson, D., Macri, L, Phelps, R., Rawson, D., Sakai, S., Turner, A., 1998, ApJ., 515 1.
- Trager, S.C., Worthey, G., Faber, S.M., Burstein, D., Gonzalez, J. J., 1998 ApJS., 116, 1.
- van Albada T.S., Bahcall, J.N., Begeman K., & Sancisi, R., 1985, ApJ/., 295, 305.
- Weitzel, L., Krabbe, A., Kroker, H., Thatte, N., Tacconi-Garman, L.E., Cameron, M., and Genzel, R., 1996, A&A, 119, 531.
- Worthey G., 1994, ApJS., 95, 107.

Fabry-Perot Observations of Spiral Galaxies: The Impact on Mass Distribution.

Sébastien Blais-Ouellette

*Université de Montréal and Observatoire du mont Mégantic,
C.P. 6128, Succ. centre ville,
Montréal, Québec, Canada. H3C 3J7
and IGRAP, Observatoire de Marseille, 2 Place Le Verrier, F-13248
Marseille Cedex 04, France
e-mail: blaisous@astro.umontreal.ca*

Claude Carignan

*Université de Montréal and Observatoire du mont Mégantic,
e-mail: carignan@astro.umontreal.ca*

Philippe Amram

*IGRAP, Observatoire de Marseille.
e-mail: amram@observatoire.cnrs-mrs.fr*

Abstract. New CFHT Fabry-Perot H α observations are combined with high sensitivity VLA or Westerbork H I data to accurately study the mass distribution of a sample of late-type spiral galaxies. Using the combined rotation curves and best fit models, it can be seen that the M/L ratio and the dark-to-luminous mass ratio can differ significantly from their values derived from HI rotation curves only. This shows the importance of the inner, rising part of the rotation curve for an accurate determination of the parameters of the global mass distribution and suggests that such a fine tuning of the rotation velocities using high resolution 2-D kinematics is necessary to look at correlations between the parameters of the dark matter and other properties of galaxies. These new precise rotation curves also represent an ideal tool to test different halo density profiles.

1. Introduction

It has often been argued that since H I observations probe the gravitational potential well outside the optical radius, in the dark matter dominated region, they are best suited to derive the characteristics of the mass distribution, thus of the dark mass distribution. However, this is overlooking the fact that the parameters of the mass models (and especially of the DM distribution) are very sensitive, as shown below, to the rising inner part of the rotation curve which can be derived with greater precision using 2-D H α observations. The example of NGC 5585 will show that this inner part is crucial since it strongly constrains two of the three free parameters of the mass model.

The rotation curves in Fig. 1, which are used to study the mass distribution of NGC 5585, are based on a now commonly used method (Carignan & Freeman 1985). First, the rotation curve is obtained by fitting a “tilted-ring” model to the velocity field in order to represent the warp of the H I disk which is very often present. The accuracy of the model representation is then checked by looking at the residuals map. Then, the luminosity profile, in the reddest available band in order to probe the mass dominant population, is transformed into a mass distribution for the stellar disk assuming a constant $(\mathcal{M}/L_B)_*$. For the contribution of the gaseous component, the H I radial profile is used, scaled by 1.33 to account for He. The difference between the observed rotation curve and the computed contribution of the luminous (stars & gas) component is thus the contribution of the dark component.

The Fabry-Perot observations of the H α emission line were obtained in 1994 February at the Canada-France-Hawaii Telescope (CFHT) and are described in details in Blais-Ouellette et al. 1999.

2. Mass models

The isothermal sphere (Carignan & Freeman 1985) is a functional form often used to model galactic halos. Fig.1 (top-left) shows the best fit mass model using only the H I data (Côté et al. 1991) and gives a disk mass-to-light ratio of $0.3 \mathcal{M}_\odot/L_\odot$, a core radius of 2.8 kpc and a velocity dispersion of 52.9 km/s. When the H I data is replaced in the inner 120 arcsec by the higher resolution H α observations (bottom-left), the mass distribution changes drastically to $(\mathcal{M}/L_B)_* = 1.0 \mathcal{M}_\odot/L_\odot$, $r_c = 4.3$ kpc and $\sigma = 53.6$ km/s. The result is a less centrally concentrated halo and a dark-to-luminous mass ratio reduced by $\sim 30\%$ at the Holmberg radius.

Having in hand this more precise rotation curve, it is easier to test some other halo density profiles like the parametrization described in Navarro et al. 1997 (hereafter NFW). The concentration, c , and the circular velocity at virial radius, V_{200} , are used as free parameters (Navarro 1998). This density profile is more centrally concentrated than the isothermal sphere and is known to disagree with late-type spiral and dwarf galaxies rotation curves. It is worth trying to look whether this is due to systematic errors on the inner points of the H I rotation curves or to a real discrepancy.

Fig.1 compares the best fit models for the isothermal sphere and the NFW profiles with different data sets. The absolute value of the χ^2 is very dependant of the error bars and must thus be taken with caution when comparing different data sets. The comparison between the two halo profiles for each set remains the best way to test the validity of a model.

The top figures reproduce the known disagreement between NGC 5585 H I rotation curve and a NFW density profile. The situation is quite different for the H α data (middle) where one can clearly see that the χ^2 is similar for both models. However, the composite rotation curve (bottom) seems to confirm that it is hard to reconcile a NFW halo with the NGC 5585 rotation curve, although one can see that the difference between the two profiles is less dramatic than for the H I alone case. The best fit using a NFW halo gives a $(\mathcal{M}/L_B)_* = 0.5 \mathcal{M}_\odot/L_\odot$ for the disk which is more realistic.

Figure 1. Best fit mass model of NGC 5585 for H I (top), H α (middle) and H α + H I (bottom) data using the isothermal spherical halo (left) and the NFW profile (right).

Figure 2. Best fit mass model of NGC 3109 for H I data using the isothermal spherical halo (left) and the NFW profile (right). New H α observations are also plotted.

While the two profiles gives closer results for NGC 5585 (Sd) when using higher resolution data, the discrepancy remains for NGC 3109 (Sm) where the new H α data agree very well with the VLA H I data. This is not surprising since these H I observations are of higher resolution than for NGC 5585 and the rotation curve is even shallower, minimising any beam smearing effect.

3. Conclusions

1. Combining new H α CFHT FP data with Westerbork HI observations reduced the ratio $\mathcal{M}_{dark}/\mathcal{M}_{lum}$ by $\simeq 30\%$ at the Holmberg radius via a decrease of the central density by nearly a factor of 3 for the late-type spiral NGC 5585. If such large errors are common, one can imagine that it could mask any physical correlation between the parameters of the dark and the luminous matter.
2. The fit of a NFW halo profile to the NGC 5585 rotation curve is now closer to the fit of an isothermal sphere but remains too concentrated for that Sd galaxy.
3. The new H α data of NGC 3109 agree very well with the VLA H I observations confirming the difficulty of a NFW profile to fit the mass distribution of this galaxy.

References

- Blais-Ouellette, S., Carignan, C., & Amram, P. 1999 AJ submitted
 Carignan, C., & Freeman, K. C. 1985, ApJ, 294, 494
 Côté, S., Carignan, C., & Sancisi, R. 1991, AJ, 102, 904
 Navarro, J.F., Frenk C.S., White, S.D.M. 1997, ApJ, 490, 493
 Navarro, J.F. 1998, ASP conf 136, eds. D. Zaritsky, 409

New Improved Photometric Redshifts of Galaxies in the HDF

H. Furusawa, K. Shimasaku, M. Doi and S. Okamura

*Department of Astronomy, School of Science, University of Tokyo,
Bunkyo-ku, Tokyo 113-0033, Japan*

Abstract.

We report new improved photometric redshifts of 1048 galaxies in the Hubble Deep Field (HDF). A standard χ^2 minimizing method is applied to seven-color UBVIJHK photometry by Fernández-Soto *et al.* (1998). We use 187 template spectral energy distributions (SEDs) representing a wide variety of morphology and age of observed galaxies based on a population synthesis model by Kodama & Arimoto (1997). We introduce two new recipes. First, the amount of the internal absorption is changed as a free parameter in the range of $E(B-V)=0.0$ to 0.5 at intervals of 0.1. Second, the absorption due to intergalactic HI clouds is also changed by a factor of 0.5, 1.0, and 1.5 around the opacity given by Madau (1995). The total number of template SEDs is thus $187 \times 6 \times 3 = 3,366$, except for the redshift grid. The dispersion σ_z of our photometric redshifts with respect to spectroscopic redshifts is $\sigma_z = 0.08$ and 0.24 for $z < 2$ and $z > 2$, respectively, which are smaller than the corresponding values ($\sigma_z = 0.09$ and 0.45) by Fernández-Soto *et al.* Improvement is significant, especially in $z > 2$. This is due to largely reduced systematic error, which is reported by previous authors. We discuss $N(z)$ and $E(B-V)$ value of galaxies based on our new photometric redshifts.

1. Introduction

‘Photometric redshift technique’ enables us to measure redshifts of faint galaxies. Though the accuracy of photometric redshifts is lower than that of spectroscopic redshifts, there are two big advantages to use photometric redshifts. We can measure redshifts of hundreds of galaxies without making enormous amounts of efforts for multi-object spectroscopy. Also we can go very deep, since the applicable flux limit of photometric redshift technique is very close to the limit of photometry.

In this work we use a photometry catalog of galaxies in the Hubble Deep Field (Fernández-Soto *et al.* 1998), which is the deepest catalog available with some spectroscopic redshifts. We adopt a photometric redshift technique, based on a standard χ^2 minimizing method (Yee 1998) with simulated SEDs. We treat internal absorption of galaxies and extinction due to intergalactic matter more carefully than previous work, which enables us to obtain smaller error in photometric redshifts.

2. METHOD

We use the χ^2 minimizing method which can include spectral evolution of galaxies directly. This method searches the best-fitted SED and the redshift of a galaxy by comparing an observed SED and templates.

2.1. Template SEDs

A stellar population synthesis model by Kodama & Arimoto (1997; hereafter KA97) is adopted to make template SEDs. KA97 includes the stellar evolutionary tracks after the asymptotic giant branch as well as the conventional tracks. Hence the model predicts UV flux of galaxies reasonably well, at least, as good as any previous work.

The template SEDs we use consist of the spectra of a pure disk, a pure bulge of age < 15 Gyr and a composite made by interpolating the two. Pure disk models correspond to young or active star-forming galaxies, and pure bulge models correspond to elliptical galaxies. We make intermediate SED types by adding pure disk component and pure bulge component with the same age. We also prepare very blue young SEDs of age < 1 Gyr, corresponding to blue star-forming galaxies reported by recent deep surveys. In total, our basic template consists of 187 kinds of SEDs.

2.2. The Absorption Effects

To simulate the observed SEDs, we take into account two effects. One is the internal absorption due to dust in each galaxy.

For simplicity, we assume that all galaxies have the same extinction curve, although absolute amount of extinction indicated by $E(B-V)$ varies galaxy to galaxy. We adopt the curve by Calzetti (1997), and the range of $E(B-V)$ covers 0.0(no dust) to 0.5 at intervals of 0.1.

We also include the effect of absorption due to the intergalactic medium. We adopt the opacity derived from Madau (1995). The continuum absorption below $912\text{\AA}$ as well as Lyman limit systems is also considered. It should be noted that we use 3 values of intergalactic absorption, 0.5, 1.0, and 1.5 times the value given by Madau in order to take into account the variance of Lyman absorptions statistically.

2.3. Rejection By Criteria

Galaxies which have $\chi_B^2 \geq 40$ in the χ^2 minimizing procedure are recognized as failure, and rejected. At $z > 1.5$, galaxies whose z_{photo} give unreasonable absolute magnitude, $M_B < -25$ or $M_B > -11$, are also rejected as failure. These rejection criteria hardly affect final scientific results.

3. RESULTS

3.1. Comparison with Spectroscopic Redshifts

We apply our method to 108 galaxies in the Hubble Deep Field (HDF; Williams et al. 1996) which have spectroscopic redshift. The photometric catalog by

Fernández-Soto *et al.* (1998), which also includes one of the most recent results of photometric redshifts in the HDF, is used to estimate z_{photo} .

Figure 1 shows the comparison among authors (Gwyn's and Sawicki's results are reported by Sawicki(<http://www.astro.utoronto.ca/~sawicki/>)). Fernández-Soto's and ours are newly shown here. The 'catastrophic errors' at $z \simeq 3$ is reduced. The systematic error that $z_{\text{photo}} < z_{\text{spec}}$ at $z_{\text{spec}} > 2$ is also improved.

Figure 1. The comparison of z_{spec} and z_{photo} of the HDF galaxies among various authors. Sawicki's and Gwyn's results use optical 4-band photometry for 74 galaxies. Fernández-Soto *et al.*'s and this work are based on 7-band photometry for 108 galaxies.

3.2. Properties of The HDF Galaxies

The photometric redshift distribution of 925 galaxies we obtained is shown in Figure 2. For comparison, the results by Fernández-Soto *et al.* are superposed. Our distribution has only one peak at $z \simeq 1$ and a moderate decrease at $z > 1$, while $N(z)$ by Fernández-Soto *et al.* shows a rapid decrease at $z \simeq 2$ in $26 < I < 28$, which is likely to be an artifact.

We also show in Figure 3 histograms of $E(B-V)$ value of galaxies with respect to redshift, which is one of the parameters derived from the best-fitted SED. It is found that the mean value of $E(B-V)$ remains almost constant ($E(B-V)=0.13-0.18$) from $z = 0$ to 5 with a possible peak $E(B-V)=0.18$ at $z = 3 - 4$. These $E(B-V)$ values are consistent with $E(B-V) \simeq 0.3$ of lyman-break galaxies at $z = 2.5 - 3.5$ by Sawicki *et al.* (1997). The distribution of $E(B-V)$ changes drastically between $z = 3 - 4$ and $z = 4 - 5$. This might indicate that the bulk of field galaxies are formed at $z \sim 4$, though further observation are necessary to confirm the change.

4. CONCLUSIONS

We obtained improved photometric redshifts, especially at high- z , by using a template SED set with 6 values of internal absorption and 3 values of intergalactic absorption. The redshift distribution of all the photometric sample galaxies show the single peak at $z \simeq 1$ with a moderate decrease at $z > 1$.

Figure 2. (Left) $N(z)$ of the HDF sample. Solid and dotted lines indicate our and Ferández-Soto *et al.*'s results, respectively.

Figure 3. (Right) Histogram of $E(B-V)$ for five redshift bins. $E(B-V)$ is derived from the best-fitted SED. Arrows indicate the mean of $E(B-V)$ in each redshift bin.

$E(B-V)$ value of galaxies remains almost constant in $z = 0 \sim 6$ with a possible slight increase at $z = 3 \sim 4$.

References

- Calzetti, D. 1997, AJ, 113, 162
 Fernández-Soto, A., Lanzetta, K.M., Yahil, A. 1998, astro-ph/9809126
 Kodama, T. & Arimoto, N. 1998, A&A, 334, 99; KA97
 Madau, P. 1995, ApJ, 441, 18
 Sawicki, M.J., Lin, H., Yee, H.K.C. 1997, AJ, 113, 1
 Sawicki, M.J. & Yee, H.K.C. 1998, AJ, 115, 1329
 Sawicki, M.J. <http://www.astro.utoronto.ca/~sawicki/>
 Williams, R.E. *et al.* 1996, AJ, 112, 4
 Yee, H.K.C. 1998, astro-ph/9809347

Comparison of Tunable Filter and Broadband-Selected Redshift Surveys for Star-Forming Galaxies

D. Heath Jones

*Research School of Astronomy and Astrophysics, The Australian
National University, Private Bag, Weston Creek P.O., Canberra,
A.C.T. 2611, Australia*

Abstract.

Recent evidence suggests a falling average star-formation rate over $z \lesssim 1$. However, many of the galaxy samples that delineate this decline were pre-selected through broadband imaging. Such broadband selection is insensitive to the $H\alpha$ and $[O II]$ emission-line signatures of star-forming galaxies.

In this paper a comparison is made between the detection of star-forming galaxies by a traditional broadband-selected redshift survey (the Autofib Galaxy Redshift Survey; Ellis et al. 1996), and a tunable filter-selected sample (The TTF Field Galaxy Survey; Jones & Bland-Hawthorn, in prep.). While both methods have advantages and drawbacks, the emission-line samples they furnish are sufficiently different to warrant a re-think of our current perception of the star-forming population.

1. Introduction

The TAURUS Tunable Filter (TTF; Bland-Hawthorn & Jones 1998) is a pair of tunable Fabry-Perot Interferometers covering 3700–6500 Å (blue ‘arm’) and 6500–9600 Å (red ‘arm’). Although both devices have the appearance of conventional Fabry-Perot etalons, (in that they each comprise two highly polished glass plates), they differ in two important respects. First, the spacing between the plates is adjustable over a much wider range, giving TTF its high degree of tunability. Second, the plate spacing is much narrower than a conventional etalon, giving the instrument its narrowband imaging properties. In effect, the TTF affords monochromatic imaging with an adjustable passband of between 6 and 60 Å, through plate spacings ranging 2 to 13 μm (Jones & Bland-Hawthorn 1998).

TTF is typically used to take a series of narrowband images (or *slices*) of a given field, each one offset slightly in wavelength. In this way, a narrowband scan (or image cube) is built through 10 or so images covering $\sim 200 - 300$ Å in wavelength. As the individual slices are narrowband images (FWHM ~ 20 Å), they are well-suited to the detection of objects seen through faint emission-lines. The scanning is useful for distant galaxies for which the emission has been redshifted to some unknown wavelength.

This is where TTF has a powerful advantage in the detection of faint line emission over conventional redshift surveys. Deep pencil-beam surveys typically pre-select objects down to $B \sim 24$ (Glazebrook et al. 1995) or $I \sim 22$ (Lilly et al. 1995), because these represent the practical spectroscopic limits on the telescopes used. However, these limits are set assuming the objects are dominated by continuum. Objects with continuum flux beyond the spectroscopic limit but with emission-lines above the limit would be detectable by these same instruments, albeit through the lines alone. However, because the initial selection is made from the broadband flux, such objects are excluded on the basis of having too faint a continuum level.

In order to match TTF, a broadband-selected redshift survey would need to select objects to a fainter broadband limit, knowing that for the faintest objects, only those with emission-lines would be register a detection spectroscopically. This however would represent a vastly inefficient way of finding emission-line galaxies. Between $B \sim 24$ and $B \sim 28$, the number of objects on the sky (per magnitude interval) increases by an order of magnitude. Such object densities would be impractical to search given the limitations of current multi-slit spectrographs in terms of object multiplex. In comparison, TTF is an imaging instrument and as such has access to every object in the field bright enough for detection.

However, much of the above is conjecture in the absence of a direct comparison between the two survey techniques. Such a comparison is the subject of this paper.

2. Comparison Between Surveys

The *TTF Field Galaxy Survey* (Jones & Bland-Hawthorn, in prep.) is a survey for redshifted emission-line galaxies in the field. The survey comprises 15 scans of 10 slices at random high-galactic latitude fields scattered around the sky. The scans are distributed among the spectral regions isolated by the 707/26, 814/33 and 909/40 nm TTF intermediate-band blocking filters. The survey is most sensitive to star-forming galaxies seen through $H\alpha$ at redshifts $z = 0.08, 0.24$ and 0.40 , corresponding to the 707/26, 814/33 and 909/40 windows. It comprises over 3000 emission-line candidates, the vast majority of which ($\sim 95\%$) are *line-only* objects, with no detectable continuum. The remaining objects have both line and continuum detected in the narrowbands (*line-on-continuum* candidates).

We tested the TTF detection of emission-line galaxies against a conventional broadband-selected redshift sample. The Autofib redshift survey (Ellis et al. 1996) is a composite sample of 1700 galaxies from redshift surveys probing a range of magnitudes and areas on the sky. There are 827 members of the Autofib redshift survey that can be classified as emission-line galaxies, defined as such by $[O II]$ rest-frame equivalent widths greater than $10 \text{ \AA}$. Of these, there are 18 within one of the TTF overlap fields at redshifts placing $H\alpha$ within the relevant scan window (707/26, 814/33 or 909/40). We examined the detection success of TTF on these galaxies as a function of both b_J magnitude and $[O II]$ rest-frame equivalent width.

Table 1 summarises the results. In addition to the TTF detection rate, we show how many of the TTF object detections were found as line-on-continuum

Range	TTF Detections*	Undetected	Line-Only	Line-on- Continuum
By b_J magnitude:				
$18.0 \leq b_J < 19.0$	1/1	0	0	1
$19.0 \leq b_J < 20.0$	4/4	0	1	3
$20.0 \leq b_J < 21.0$	4/5	1	1	3
$21.0 \leq b_J < 22.0$	3/7	4	2	1
$22.0 \leq b_J < 23.0$	1/1	0	1	0
By [O II] EW:				
$EW \geq 60 \text{ \AA}$	1/2	1	0	1
$50 \leq EW < 60 \text{ \AA}$	0/1	1	0	0
$40 \leq EW < 50 \text{ \AA}$	1/1	0	0	1
$30 \leq EW < 40 \text{ \AA}$	5/6	1	1	4
$20 \leq EW < 30 \text{ \AA}$	1/2	1	0	1
$10 \leq EW < 20 \text{ \AA}$	5/6	1	4	1

Table 1. TTF Detection of the Autofib sub-sample. (* number of TTF detections relative to the number of Autofib emission-line galaxies available)

and how many were line alone. Table 1 shows a trend of fainter galaxies ($b_J \gtrsim 21$) becoming H α line-only detections in increasing numbers. This we expect since the narrow TTF passband is optimised for the detection of faint line flux and not faint continuum. There is also the suggestion of a declining detection rate with fainter limiting apparent magnitude, although it is difficult to tell with such small galaxy numbers. If line flux is purely independent of continuum, we would expect the b_J magnitude to affect *how* it is detected by TTF (*i.e.* line alone or continuum), not whether it is detected. However, Table 1 could be interpreted as suggesting otherwise. When ranked according to [O II] rest-frame equivalent width, little is seen in the way of TTF detection trends. Again, the small size of the overlap sample is a limitation, particularly because the majority of objects have [O II] emission with low to moderate rest-frame equivalent widths ($10 \lesssim EW \lesssim 40 \text{ \AA}$).

Of course, we have assumed here throughout that all galaxies exhibiting [O II] emission will have equally strong H α . Although this is commonly true it is not always the case, and some of the TTF non-detections in Table 1 may have simply been because there was little or no H α . Tresse et al. (1999) find a small but significant fraction of the Stromlo-APM Survey (11%) that show [O II] but no H α . They attribute this to strong stellar absorption and do not find such galaxies to be more prevalent at a particular redshift. So it is possible that Table 1 underestimates the detection success of TTF in these cases.

Figure 1. Comparison between Autofib and TTF flux measurements, with the measurements from each plotted on the abscissa and ordinate of each panel. (a) TTF-measured (H α + [N II]) line fluxes against Autofib [O II] equivalent widths. (b) TTF-measured (H α + [N II]) equivalent widths against those of [O II] from Autofib. The row of points along the top denote galaxies detected by TTF through (H α + [N II]) line flux alone, for which the (H α + [N II]) equivalent width was indeterminable. (c) TTF-measured continuum flux versus b_J photographic magnitude. The best-fit $b_J = -2.5 \log F(\text{cont}) + 22.62$ is also shown (dotted line).

Figure 1 shows the relationship between spectral quantities measured around H α by TTF and [O II] by the Autofib survey for the galaxies in common. In panels (a) and (b), we find little correlation between either the H α and [O II] emission-line fluxes or between the corresponding equivalent width measures. We do not see the EW(H α + [N II]) versus EW([O II]) correlation of Kennicutt (1992) or Tresse et al. (1999), probably because this correlation is weak with large scatter and our overlap sample is small. The large scatter in Fig. 1(b) may also be due to a tendency of spectroscopic equivalent widths to be overestimates at the limits of a broadband-selected sample. Either way, our results lend fur-

ther weight to Kennicutt (1992) who finds [O II] flux and equivalent width to be loosely connected at best to the star-formation rate in a galaxy.

In Fig. 1(c) we show the relationship between the Autofib b_J magnitudes and the TTF narrowband continuum measurements. There is reasonable agreement between the two with the line of best-fit,

$$b_J = -2.5 \log F(\text{cont}) + 22.62, \quad (1)$$

where b_J is the photographic magnitude and $F(\text{cont})$ is the narrowband continuum in the TTF bandpass ($10^{-16} \text{ ergs}^{-1} \text{ cm}^{-2} \text{ band}^{-1}$). In making the fit we fixed the slope to -2.5 and weighted the points by the $1/\Delta F^2$ continuum photometry uncertainties. The agreement in (c) indicates that the scatter in the equivalent width measurements of (b) is due primarily to the dispersion in line flux seen in (a). Equation (1) also implies that the continuum detection limit of TTF is about the same as the $b_J = 24.0$ limit of the broadband-selected survey.

We turn our attention now to why the overlap between the two surveys (827 and 3360 galaxies) is so small (18). Since the vast majority of our sample are *line-only* candidates, they would have had insufficient continuum flux for inclusion in the Autofib survey. This is not to say that they are beyond the reach of the spectroscopic detection limits used by Autofib, merely that their low continuum saw them excluded from Autofib at the outset. Conversely, the reason TTF had so few of the Autofib galaxies to use was because the 707/26, 814/33 and 909/40 scan intervals represent quite narrow redshift windows compared to the very wide range of redshifts sampled by any broadband-selected survey.

Figure 2 is a comparison between the [O II] redshift measurements from the Autofib redshift survey and the ability of TTF to infer redshifts from ($\text{H}\alpha + [\text{N II}]$). The measurements appear consistent although the TTF measurements in the 918 set appear to be systematically higher by $\Delta z \sim 0.1$. This may not necessarily be due to TTF measurement error. Galaxies with strong [N II] will be seen as such through the dominant [N II] $\lambda 6583$ line, which will raise TTF-measured redshifts by $\Delta z \sim 0.05$ at $z = 0.4$ if taken for $\text{H}\alpha$ alone. In any case, the narrowband redshift measurements derived from TTF data are unlikely to introduce uncertainties into our flux-to-luminosity conversions that dominate over the photometry uncertainties.

3. Conclusions

The results of our comparison between the TTF and Autofib surveys can thus be summarised as follows. There is surprisingly little overlap between the emission-line selected galaxies found in the the broadband-selected Autofib redshift survey and the *TTF Field Galaxy Survey*. The vast majority of TTF-selected galaxies have continua too faint for inclusion in the broadband catalogue while the redshifts of the Autofib line-emitters are too wide-ranging for the relatively narrow scan windows of TTF. For the galaxies in common, TTF detects the brighter Autofib galaxies ($b_J \lesssim 21$) in both line and continuum while fainter line-emitters are almost exclusively detected in the line. While there is good agreement between the redshifts and continuum fluxes measured by TTF and Autofib around $\text{H}\alpha$ and [O II], there is considerable scatter between the measurements of line flux and equivalent width. Improved insight into the relative merits of TTF over

Figure 2. Comparison of TTF ($H\alpha+[N II]$) and Autofib [$O II$] redshift measurements for galaxies common to both surveys. Redshifts are grouped according to membership of the (a) 707/26, (b) 814/33 and (c) 919/40 TTF sub-samples. A $y = x$ line is also shown (dotted line).

broadband-selected emission-line samples would be obtained through a redshift survey with near-infrared spectra overlapping with the TTF wavelengths intervals directly.

It is a pleasure to thank Matthew Colless for his helpful comments on this and related work. Thanks also to the Australian Time Allocation Committee for their support of the TTF observations. DHJ acknowledges the previous assistance of a Commonwealth Australian Postgraduate Research Award.

References

- Bland-Hawthorn, J. and Jones, D. H. 1998, in *Optical Astronomical Instrumentation*, Proc. SPIE 3355, ed. S. D'Odorico, 855
- Ellis, R. S., Colless, M., Broadhurst, T., Heyl, J. and Glazebrook, K. 1996, MNRAS 280, 235
- Glazebrook, K., Ellis, R., Colless, M., Broadhurst, T., Allington-Smith, J. and Tanvir, N. 1995, MNRAS 273, 157
- Jones, D. H. and Bland-Hawthorn, J. 1998, PASP 110, 1059
- Kennicutt, R. C., Jr. 1992, ApJ 388, 310
- Lilly, S. J., Le Fèvre, O., Crampton, D., Hammer, F. and Tresse, L. 1995, ApJ 455, 50
- Tresse, L., Maddox, S. J., Loveday, J. and Singleton, C. 1999, preprint, (astro-ph/9905384)

SAURON: Integral-field Spectroscopy of Galaxies

B. W. Miller, M. Bureau, E. Verolme, P. T. de Zeeuw

Sterrewacht Leiden, Postbus 9513, 2300 RA Leiden, The Netherlands

R. Bacon, Y. Copin, E. Emsellem

*Centre de Recherche Astronomique de Lyon, 9 Avenue Charles André,
69561 Saint Genis-Laval Cedex, France*

R. L. Davies, R. F. Peletier, J. R. Allington-Smith

*Physics Department, University of Durham, South Road, Durham, DH1
3LE, UK*

C. M. Carollo

*Dept. of Physics and Astronomy, Johns Hopkins University, 3400 N.
Charles St., Baltimore, MD 21218 USA*

G. Monnet

*European Southern Observatory, Karl-Schwarzschild Str. 2, D-85748
Garching, Germany*

Abstract. We present the first results from a new and unique integral-field spectrograph, SAURON. Based upon the TIGER concept, SAURON uses a lens array to obtain two-dimensional spectroscopy with complete spatial coverage over a field of $33'' \times 41''$ in low-resolution mode ($0''.94$ lenslets) and of $9'' \times 11''$ in high-resolution mode ($0''.26$ lenslets). The spectra cover the wavelengths from $4800\text{\AA}$ to $5400\text{\AA}$ with a spectral resolution of $\sim 3\text{\AA}$ ($\sigma \approx 75 \text{ km s}^{-1}$). SAURON achieved first light during commissioning on the William Herschel Telescope on 1 February 1999. We are now commencing a systematic survey of the velocity dispersions, velocity fields, and line-strength distributions of nearby ellipticals and spiral bulges. The wide field of SAURON will be crucial for unraveling complicated velocity structures. In combination with available long-slit spectroscopy of the outer regions of the galaxies, HST spectra of the nuclei, HST imaging, and dynamical modeling, we will constrain the intrinsic shapes, mass-to-light ratios, and stellar populations in spheroids.

1. Introduction

Understanding the formation and evolution of elliptical galaxies and spiral bulges is complicated by the fact that many, if not all, of these systems are triaxial or have multiple kinematic components. Both lead to rich velocity structures

that are difficult to map using long-slit spectroscopy. In addition, the central few arcseconds of spheroids are often “kinematically decoupled”: the inner and outer regions appear to rotate around different axes (e.g., Illingworth & Franx 1989). This makes *two-dimensional* (integral-field) spectroscopy of stars and gas essential for determining the dynamical structure of these systems. Therefore, the galactic dynamics groups in Lyon, Leiden, and Durham decided to build SAURON, a unique integral-field unit (IFU) with a large field of view and high throughput that is optimized for studies of the gaseous and stellar kinematics in galaxies.

2. The SAURON Instrument

SAURON (Spectroscopic Areal Unit for Research on Optical Nebulae) is a TIGER-like integral field spectrograph (Bacon et al. 1995) currently being used at the 4.2-m William Herschel Telescope of the Isaac Newton Group on La Palma. The focal plane is sampled by an array of 1520 square lenses whose projected size on the sky can be set to $0''.94$ or $0''.26$, yielding fields of view of $33'' \times 41''$ or $9'' \times 11''$, respectively. About 70 dedicated lenses sample an area offset $1''.7$ from the main field, to allow sky subtraction. The spatial coverage is 100%. SAURON complements OASIS on the CFHT, which allows high spatial resolution sampling of the nuclear regions (e.g., one can obtain a $4''.1 \times 3''.3$ field with $0''.11$ sampling for use with the Adaptive Optics Bonnette).

The grism is optimized for the spectral domain 4800–5400 Å, which includes important absorption/emission lines such as the Mg triplet, Fe lines, $H\beta$, $[OIII]\lambda 4959, 5007$, and $[NI]\lambda 5199$. The instrumental velocity dispersions are 82 km s^{-1} ($0''.94$ sampling) or $\sim 64 \text{ km s}^{-1}$ ($0''.26$ sampling).

SAURON was commissioned on the WHT 1–5 February 1999. The instrument performed admirably. Most of the effort went to focusing the optics, aligning the CCD, grism, and lenslet array, and adjusting the baffling to remove stray light and reflections. To prevent the spectra from overlapping, the angle between the grism and the lenslet array was set to $5^\circ 19'$. Flexure was found to be less than 1 pixel in both the spatial and spectral directions in ~ 1800 seconds.

3. The SAURON Project

During the project we hope to observe about 80 bright galaxies covering all Hubble types. Initially, we are concentrating on E, S0, and Sa types. We are restricting the sample to bright ($M_B < -18$) northern galaxies with $v_\odot < 3000 \text{ km s}^{-1}$ and available HST imaging and/or spectroscopy. This will be a *representative* sample covering the full parameter space of absolute magnitude, effective radius (R_e), flattening, rotational support, and Mg_b index.

We require $S/N > 25$ per resolution element to map the velocity fields and $S/N > 40$ to measure accurate velocity profiles. Exposure times of ~ 2 hours per position are needed to reach these levels at $\mu_B \approx 20 \text{ mag arcsec}^{-2}$. We also aim to map as much of the inner $1R_e$ as possible, using multiple pointings if necessary.

Figure 1. Example SAURON spectra for NGC 3377 and NGC 4278. The NGC 3377 spectrum shows absorption lines from $H\beta$, Mg, and Fe. NGC 4278 has lines of $H\beta$, [OIII], and [NI] in emission as well as Mg and Fe in absorption.

The first science run was 14–21 February, 1999. Figure 1 shows preliminary SAURON spectra from the centers of NGC 3377 and NGC 4278. These spectra were made by summing a strip 3 pixels wide perpendicular to the dispersion direction. The sky has been subtracted but no flat-fielding has been attempted. SAURON can reach $S/N \approx 65 \text{ Å}^{-1}$ in 1800 seconds in high surface brightness regions. Emission and absorption lines from $H\beta$, [OIII], Mg_b , [NI], and FeI are visible.

The real advantage of SAURON is the ability to produce maps of velocities, velocity dispersions, and line strengths. Figure 2a shows a preliminary velocity field for NGC 3377. It reveals a rotating disk with a maximum velocity of $\sim 80 \text{ km s}^{-1}$. The velocity dispersion (Figure 2b) is lower in the outer parts of the disk and then increases towards the nucleus. Line-strength maps indicate increasing metallicity towards the center (Davies 2000).

We are currently in the process of refining the data reduction procedure. The method will be similar to that used to reduce and analyze OASIS data from the CFHT. Tight packing of the SAURON spectra means that the extraction process must be precise. To account for contamination from neighboring spectra, we employ an optimized method that models the distortion of the optics and fits a PSF at each wavelength of each spectrum. We are also developing a data reduction pipeline and a data archive.

The data will be compared with fully general numerical galaxy models built with Schwarzschild's orbit superposition method. All appropriate imaging and spectral data will be used as input for the models: high resolution HST spectra

Figure 2. Preliminary SAURON maps of NGC 3377. Each panel shows a field $26''$ on a side. a) Stellar velocity field, showing evidence for rotation. The contour interval is 20 km s^{-1} . b) Stellar velocity dispersion map, indicating an increase in σ towards the nucleus.

for constraining the mass of the central black hole; SAURON data cubes for unraveling complicated velocity structures of the stars and gas within $1R_e$; long-slit spectra for detecting the presence of dark matter halos at $(1 - 2)R_e$; and HST images and SAURON line-strength maps for estimating the structure and M/L of the starlight. The galaxy models make no assumptions about velocity anisotropy and account for PSF convolution and the finite size of the spatial element (i.e., microlens). It has been applied successfully to determine black hole masses in the E3 galaxy M32 (van der Marel et al. 1998) and in NGC 4342, an E7/S0 in Virgo (Cretton & van den Bosch 1999).

The output of the SAURON project will be a large, uniformly processed and analyzed sample. Such a dataset is essential for understanding the formation and evolution of elliptical galaxies, spiral bulges, and bars.

We would like to thank the Dutch and UK PATTs for allocation of time on the WHT and the ING staff for their support during the SAURON commissioning and science runs. B.W.M. also thanks the Leids Kerkhoven Bosscha Fonds for financial support to attend this conference.

References

- Bacon, R., Adam, G., Baranne, A., Courtès, G., Dubet, D., Dubois, J. P., Emsellem, E., Ferruit, P., Georgelin, Y., Monnet, G., Pécontal, E., Rousset, A., & Sayède, F. 1995, *A&AS* 113, 347
- Cretton, N., & van den Bosch, F.C., 1999, *ApJ*, 514, 704
- Davies, R. 2000, this volume
- Illingworth, G. D., & Franx, M. 1989, in *Dynamics of Dense Stellar Systems* (Cambridge), 13
- van der Marel R.P., Cretton, N., de Zeeuw, T., & Rix, H.W. 1998, *ApJ*, 493, 613

Multiple Deployable Integral Field Units in the Near-IR

G. S. Wright

*UK Astronomy Technology Centre, Royal Observatory, Blackford Hill,
Edinburgh, EH9 3HJ, UK*

Abstract. Integral field spectroscopy has the advantage of providing simultaneous spectral and spatial information. Technologies are now being developed to enable integral field spectroscopy of several targets over a wider field than can be covered with a single IFU. The scientific rationale for such instruments is discussed, and some instruments currently being developed for 8m class telescopes are described.

1. Introduction

The rapidly growing field of infrared spectroscopy is blossoming as array development enables increasingly sophisticated instruments. Extending the techniques of optical spectroscopy to the infrared is very important because it allows obscured regions such as the nuclei of active galaxies and young stellar objects to be probed, as well as the study of high redshift galaxies via appropriate redshifted spectral features. Optical spectrometers have been revolutionised by the development of integral field devices such as that in the TIGER instrument on CFHT (Bacon et al 1995), and there has been an increasing realisation that integral field spectroscopy is frequently a very effective approach because the often complex image morphology can be re-constructed with the same angular resolution as the spectroscopic data. The recent development of large format $1k \times 1k$ arrays has led to dramatically improved IR instrumentation capabilities and the potential for similar exciting breakthroughs in the infrared. Integral field spectroscopy in the near-IR is important for understanding the distribution, excitation and kinematics of gas and/or stars in objects ranging from obscured young star formation regions to morphologically peculiar high redshift galaxies, and several IR spectrometers with integral field capability are now available or under development.

It is clear that the superb image quality of 8m telescopes such as Gemini and the ESO-VLT, with or without adaptive optics, will make possible the study of both spatial and spectral structure in unprecedented detail. Since many important advances will come from the study of statistical samples of faint astrophysical objects, simultaneous integral field spectroscopy of many sources in the field will greatly increase the efficiency. While the projected development of IR arrays will lead to $2k \times 2k$ devices and the possibility of larger formats via mosaics, the number of such arrays in any one instrument is likely to be small. This limited number of pixels available in foreseeable infrared arrays means that fully covering the focal plane of an 8m-class telescope by an IFU with fine spa-

Figure 1. Distribution of spatial and spectral pixels

tial sampling will be impracticable. Figure 1 shows schematically the options for dividing pixels between spatial and spectral sampling for an integral field spectrograph. For a 4000×4000 pixel array and ~ 4000 spectral elements, the maximum possible field for an IFU will be $\sim 9.5'' \times 9.5''$, which is small compared to the many arcminute diameter available fields of the 8m class telescopes. Thus the ability to position a number of small field IFUs anywhere within a larger patrol field is needed to allow high spatial resolution integral field spectroscopy of several targets. Such instruments will be the powerful next generation of IR imaging spectrometers which on large telescopes will enable the detailed study of faint red-shifted or obscured sources. This multi-object capability is complementary to the "classical" multi-slit/multi-fibre approach because it will provide spatially resolved spectra for a smaller number of targets. At least 3 instruments are already under-development to provide a multiple-IFU capability at optical and IR wavelengths. Before describing these instruments and the technological challenges of including the K-band, some illustrative scientific areas on which such instruments will have a major impact are briefly reviewed.

2. Multiple Integral Field Science

Multiple integral field units, each of which is individually position-able within the field of view of an 8m class telescope, will address key scientific issues ranging from the nature of high redshift galaxies to the structure of matter around young stars on the smallest spatial scales. Many of these targets are faint, requiring long

exposures even on large telescopes, and so even a relatively small multiplexing advantage of ~ 5 -30 is an important gain. The now well known advantages of IFU spectroscopy over slit spectroscopy include the ability to take spectra of contiguous areas, the elimination of slit losses, easier acquisition and the ability to reconstruct images of each source. Thus multiple-IFU spectroscopy compared to multi-slit spectroscopy enables follow-up of survey sources unhindered by pointing errors and slit losses, as well as more complete spatial information.

Extremely deep surveys, such as the Hubble Deep Field and those planned for the future, are revolutionising the study of galaxy evolution by providing samples of high redshift galaxies for which detailed comparison of the distribution, excitation and kinematics of the gas and stars with nearby analogues may be made. Such studies ultimately require full two-dimensional imaging spectroscopy in the J H, and K bands, to cover the key rest-frame optical diagnostic lines at $z = 2 - 3$. Galaxies at high redshift are not point sources, and are often not even vaguely symmetrical, so integral field spectroscopy will be essential to determine their detailed internal properties. For such faint sources an additional advantage of the multi-IFU technique is that, when the signal-to-noise is too low, either the spatial or spectral information (or both) can be binned to yield integrated properties without slit losses. K-band surveys find ~ 50 galaxies/arcmin² to $K \sim 22$ (Davies 2000), but only a small fraction of these will fall in the selection criteria of interest (e.g. $z > 2$, etc.), so a typical programme may select a few tens of objects distributed over a 1 - 10 arcmin field. Spectroscopy of multiply lensed images in high red-shift clusters such as A2218 can provide information about the evolution of the mass distribution in clusters. For lensed galaxies spatial scales are magnified, enabling studies of the distribution of star-formation in high redshift galaxies. Since the lensed objects appear as arcs, they again require spatially resolved spectroscopy. Typical sources have ~ 10 sets of multiple images over a 1 arcminute field.

For nearby star-burst and active galaxies there is now a wealth of images showing small scale structure across regions of up to 1-2 arcmin. By mapping the kinematics of hot-spots, super star clusters, the old stellar population, and filamentary gas structures in photo-ionisation/shock/star-burst/AGN diagnostic lines, multiple integral field unit spectroscopy will determine the influences on gas and stellar populations in sufficient detail to understand the processes which induce star-formation and other activity. Since many star-bursts and AGN are heavily obscured at optical wavelengths and often occur in dusty and morphologically peculiar interacting and merging galaxies, the use of integral-field infrared spectroscopic diagnostics will be a key to unravelling the connection between AGN and star-burst activity.

Near-IR integral field spectroscopy of proto-stars and young star formation regions is needed for detailed velocity mapping of the gas and dust in proto-stellar disks and molecular outflows. Large samples are needed to search for differences in dust disks around stars of different ages and evolutionary effects due to processing and planet formation. Studies of the temperature, density and composition on scales of 45au - 500au are crucial for understanding disk and planet formation. In the Ophiuchus cloud, where 45au corresponds to 0.3'' there are typically ~ 10 sources per $10' \times 10'$, so a multiple-IFU instrument could make significant progress in this area of science. On smaller spatial scales deployable

integral field units which fully sample the diffraction limit of an 8m telescope and patrol the AO corrected field will be a powerful tool for the study of structure around low mass young stars in clusters. T-Tauri stars are low mass young stars with active accretion disks. AO-corrected IR imaging of these stars shows that there is a wealth of structure on scales as small as $0.1''$. The study of the kinematics in large samples of these stars is key to understanding the generation of YSO winds. T-Tauri stars occur in small clusters, so simultaneous study of several will be both desirable and feasible. For example in the Trapezium region there are ~ 100 T-Tauri stars in a $\sim 0.5 \text{ deg}^2$ area.

3. Instruments under development

The illustrative science goals for multiple-IFU instruments discussed above indicate that the capability to observe at most a few tens of objects distributed over about 1 - 10 arcmin is desirable. A spectral resolution of several thousand will be required both for the science and to limit the effects of the strong atmospheric OH emission. At this resolution 500 - 4000 spectral pixels will be needed to provide reasonable spectral coverage. Assuming modest developments in IR array sizes the combination of several integral field units with spectroscopy at $R > 4000$ will be possible. For example, if each IFU has a $2.3'' \times 2.3''$ field, and a $4k \times 4k$ array is available, then the data from about 5-30 deployable IFUs could be fitted onto the array. This is typical of the number of multiple IFUs in currently proposed instruments.

GIRAFFE is an instrument being developed for the ESO VLT which will have an array of 15 deployable mini-IFUs, and which is described elsewhere in these proceedings (Monnet 2000). A conceptual design study for FMOS, a fibre multi-object spectrometer for Subaru is also in progress. This instrument includes a multiple IFU mode using lenslet+fibre IFUs (Dalton 1999). FMOS will have both optical and IR fibres and so will operate over $0.4 < \lambda < 1.8\mu\text{m}$. Each IFU has 200 pixels with $0.3''/\text{fibre}$ and a $5''$ diameter field. A "fishing-rod" type of positioner will be used to deploy 15-20 such IFUs anywhere in the $30'$ accessible field at the Subaru prime focus. The Gemini telescopes have recently funded design studies aimed at developing a multi-IFU capability in the near-IR which will have good sensitivity in the K-band ($2 - 2.5\mu\text{m}$). The Gemini development studies have as a baseline the inclusion of both IFUs with $0.15''$ pixels and a $\sim 3'' \times 3''$ field deployable over at least a $5.5'$ diameter region and IFUs with $0.05''$ pixels patrolling the AO corrected field.

Although the imminent availability of large format IR detectors makes $1-1.8\mu\text{m}$ fibre fed IR instruments like FMOS possible without cryogenics, for longer wavelength IR performance the additional design challenges resulting from cryogenic cooling of both the optics and any deployment mechanism have to be solved. The capability to work in the K-band, $\lambda > 1.7\mu\text{m}$, where thermal emission dominates the background, is crucial for the much of the science described above. For example the study of $\text{H}\alpha$ emission in galaxies at $z > 1.8$, or using CO, HeI and HeII lines to probe heavily obscured young star-formation regions all require integral field spectroscopy with $\lambda > 2\mu\text{m}$. The UK-ATC and Durham University are therefore carrying out a technology development study for Gemini to produce a conceptual design for cold deployable image slicers, while the per-

Figure 2. Pyramidal mirror pick-off to feed multiple IFUs

formance of fibre technologies in the K-band is being investigated in a parallel study by the AAO.

4. Deployable Image Slicers

The use of image slicers to provide a multiple IFU capability at $\lambda > 2\mu\text{m}$ has the same rationale as their use in many single integral field unit IR spectrographs. When the instrument is cooled the slicer remains confocal and aligned since the optical surfaces and mounts may be machined from the same material, and the all mirror construction provides a high throughput and low emissivity if high reflectivity coatings are used. In addition the all reflective image slicers will work in the $2.5\text{--}5\mu\text{m}$ region, which is important for studies of highly obscured systems. Although the sensitivity in the L and M bands is poorer than in the near-IR due to the atmospheric thermal background, many star formation regions and nearby galaxy nuclei are bright enough for detailed study in these bands with Gemini. Image slicers based on the AIS type design of Content (1997) are small and similar IFUs could be used to provide the multiple IFU capability. Since the image slicers are cold any deployment mechanism will have to operate reliably at cryogenic temperatures and the design of the mechanism and optics for deployability is the key technological problem to solve. There are two basic approaches to deployable image slicers - physically move them in the focal plane or tilt/move a pickoff mirror to direct the desired field into a fixed image slicer.

Moving the IFU itself means that the position of the spectra on the detector will vary depending on the position of the IFU in the field, but since this is also the case for multi-slit spectroscopy it is not a major problem. However the design

Figure 3. Multiple IFU pick-off arm layout

of a cryogenic mechanism which does not obstruct the moving IFU modules would be very difficult. It could perhaps resemble the type of mechanism used to move slitlet arrays, but these are mechanically complex and are confined to a single slit for any given row of detector pixels. Worse, although "small", the IFUs are larger than typical slitlets and for even a minimal number of IFUs the mechanical complexity increases. A more efficient and mechanically simpler method is to use a pick-off mirror system to steer the desired field area into a fixed array of image slicers. This also has the advantage of producing a fixed spectrum at the detector.

One method of selecting small areas of the field for the IFUs is to place a multi-sided "pyramidal" mirror in the telescope focal plane to divide the field into large segments, as shown on the left of Figure 2. A small tilting steering mirror is then used to select the region of interest and steer the beam through path-length compensation optics in the module shown at the right of Figure 2 to the IFUs which lie behind the extraction segments. There are two disadvantages to this method. Firstly the path length compensation optics will need to be controlled as well as the tilting mirror, and secondly the pyramidal extraction mirror becomes very large if big fields at the focal plane of an $f/16$ telescope are to be divided. The mirror shown in Figure 2 has a base of ~ 10 cm diameter for only a $2'$ field, so this system is not considered suitable for the Gemini IR Mos

Figure 4. Instrument layout showing IFU deployment

where positioning the IFU field within a patrol field of 5-10 arcmin is required. Such a system may prove to be more practical for small patrol field applications.

A system for directly placing small pick-off mirrors in the telescope focal plane is probably the simplest option from a mechanical and thermal point of view. Figure 3 shows a conceptual design for using articulated arms to position an attached pickoff mirror near the focal plane. Each arm can move its mirror in one quarter of the 10arcmin total field, and there is a region in the centre which all the arms can reach. In this design, since a collimated beam is passed round the "elbow" and "shoulder" of the mechanical arm, (as shown on the right of the figure) there are no other moving parts to control. The common patrol area in the middle of the field means that the multiple IFUs may be used to study contiguous areas, effectively acting as a larger field single IFU. The four arms in the figure place the pick-off mirrors in a plane. More pick-off mirrors, and hence IFUs, can be accommodated by using a similar set of arms to place the pick-off mirrors slightly above or below the ones shown in Figure 3. This is illustrated in Figure 4 which shows a cross-section of the layout of an instrument with three sets of four pick-off arms, packaged in a cryostat which fits the Gemini instrument envelope. The development of this concept via a detailed design study of the mechanisms and optics, is now underway.

5. Conclusion

A near-IR multiple-IFU instrument on an 8m class telescope will address a wide range of scientific issues. Two instruments using positionable multiple fibre IFUs are now in development. A technology development study for deployable image slicers to provide a multi-IFU capability with good K-band performance for Gemini is underway. The development of cold multiple-IFU instruments will take infrared spectroscopy into a new astrophysical regime.

As always the development of a new instrument concept is the work of many people. I'd especially like to thank my collaborators on the Gemini IR MOS study, particularly Peter Hastings, Martyn Wells, Ray Sharples and Jeremy Allington-Smith, for their significant contributions to the concepts discussed in this paper. Peter Hastings also produced the drawings of IFU deployment mechanisms. I also thank Gavin Dalton, John Peacock and Antonio Chrysostomou, who provided as yet unpublished images that were used to illustrate my talk.

References

- Bacon, R., Adam, G., Baranne, A., Courtes, G., Dubet, D., Dubois, J.P., Emsellem, E., Ferruit, P., Georgelin, Y., Monnet, G., Pecontal, E., Rousset, A., & Say, F. 1995 *A&AS*, 113, 347
- Content, R., 1997 *SPIE* 2871, 1295
- Dalton, G. 1999, private communication
- Davies, R., 2000, this volume
- Monnet, G., 2000, this volume

Supermassive Black Hole Searches with 3-D Spectroscopy

R. Bacon, E. Emsellem, Y. Copin

*Centre de Recherche Astronomique de Lyon, Observatoire de Lyon,
69560 Saint-Genis-Laval Cedex, France*

G. Monnet

*European Southern Observatory, Karl-Schwarzschild Strasse 2,
D-85784 Garching bei Munchen, Germany*

Abstract. Galactic nuclei are complex dynamical systems, showing evidences for non-axisymmetry, multiple structures, and dynamical decoupling. Detailed interpretation of one or even a few long-slit data sets can thus be strongly impaired by the prejudices consisting on placing them e.g. along an a priori defined principal axis. Three-dimensional spectroscopy at high spatial resolution is an ideal tool to study the central regions of nearby galaxies. Some illustrations of the progress made in that field using the TIGER and OASIS integral field spectrographs are given.

1. Introduction: the need for high spatial resolution and 2D coverage

Demography of quasars and active nuclei shows that they have been numerous in the past, with a peak of activity around $z \sim 2$ (Shaver et al. 1996). The standard hypothesis is that the engine of this activity is a supermassive black hole (hereafter SBH), which implies that most of today's galaxies classified as normal have been active in the past and should therefore contain a SBH. Un-fueled SBH may then sit quietly in the centre of galaxies, in which case the only expected detectable signature may well be their gravitational influence on stars (as of now). In the past ten years, rapid progress in high-spatial resolution imaging and spectrographic capabilities have allowed in some cases to nearly fully discard alternatives to SBH, e.g. clusters of dark objects. The strongest case is our own Galaxy, for which Eckart and Genzel (e.g. Genzel et al. 1997, and these proceedings) provided conclusive evidences for the existence of a $2.6 \cdot 10^6 M_{\odot}$ SBH using stellar proper motions. We are unfortunately very far from being able to use the same techniques for external galaxies. Most of the other strong cases of SBH detection have been established via stellar or gas kinematics using spectrography in the visible, with the remarked exception of NGC 4258 for which the presence of water masers allowed the detection of keplerian motion at a very small, intrinsic scale (Miyoshi et al. 1995).

Most of the SBH detections rely on measurements of stellar kinematics and light distribution. To infer the spatial mass density from observations is not a straightforward task. The procedure involves deprojection of the light

Figure 1. Left and right panels: the OASIS integral field spectrograph at CFHT. Central panel: OASIS lens array.

surface brightness, complex kinematical modeling and comparison of the derived projected velocity distributions with the observed ones. Single axis velocity profiles do not fully constrain the orbital characteristics of stars within a given potential. The proportion of radial orbits to tangential ones, known as the anisotropy parameter, is generally a free parameter. For low spatial resolution data, radial anisotropy can mimic the effect of a SBH (Binney & Mamon 1982). High spatial resolution kinematics is necessary to remove the degeneracy between anisotropy and central dark mass (van der Marel 1994). It is also required to nail down the radius in which the dark mass is concentrated. A proper spatial coverage is then critical to determine the internal dynamics of a stellar system, as most stellar orbits in galactic nuclei are spatially three-dimensional. Integral field spectrography (hereafter IFS) at high spatial resolution is then the ideal observational tool, providing *both* the high spatial resolution and a 2D spatial coverage.

Using gas kinematics to derive the mass distribution in galaxy nuclei can be, in some sense, more straightforward than using stellar kinematics. Gas being a dissipative component, the relevant orbits in a stable system cannot intersect, and observed gas motions can in principle be more easily converted into a mass profile. However non-gravitational perturbations (e.g. shocks, outflows) are commonly observed in galactic nuclei. 2D spatial information is then required (but may not be enough) to disentangle their relative contributions.

2. The integral field spectrographs TIGER and OASIS

SBH detection, and more generally, kinematics of nuclei of nearby galaxies, have been one of the main scientific objective of the TIGER project. TIGER

Figure 2. Mean stellar velocity and stellar velocity dispersion field of NGC 3115 (TIGER/CFHT).

was the first realization of a lenslet based integral field spectrograph. A detailed description of the concept is given in Bacon et al (1995). TIGER has been regularly operated at the CFH 3.6 m telescope since 1987. TIGER has now been replaced by OASIS (Bacon 98) which saw its first light in 1997. OASIS is a TIGER type IFS but with enhanced spatial and spectral resolutions and field of view. It is also a common-user instrument with full remote control operation and user-friendly control software. It comes with a dedicated data reduction package including a graphical user interface. OASIS superior performance comes in part from the spatial resolution attainable with the CFH/PUEO adaptive optics bonnette (Rigaut et al. 1998). OASIS has been optimized to be used in combination with the adaptive optics system in the $0.4 - 1 \mu m$ range, and can acquire 1200 spectra simultaneously with a spatial sampling ranging from 0.04 to 0.4 arcsec. Although performances of adaptive optics at optical wavelength are modest in terms of Strehl ratio, a significant gain in FWHM and contrast is achieved with PUEO in typical Mauna Kea seeing conditions. Depending on the guiding source, FWHM from 0.2 to 0.5 arcsec have been measured.

3. Stellar dynamics of SBH candidates

A good illustration of the advantage of IFS in stellar dynamics is given by the dynamical study of the central part of NGC 3115. Observations were carried out using the TIGER IFS. The mean stellar velocity and velocity dispersion maps are presented in Figure 2. A raise in velocity and velocity dispersion were already observed using classical long-slit spectrography. The central velocity dispersion, as observed with the HST FOS spectrograph, peaks at about 440 km.s^{-1} (Kormendy et al. 1996). Such observations do constrain the SBH mass, which was then estimated to be $2 \cdot 10^9 M_{\odot}$. Subsequent detailed kinematical modeling (Emsellem et al. 1999) discarded this value at a 3σ level and have shown that the SBH mass was ill-constrained when excluding the 2D kinematical TIGER maps. Two and three-integral models finally provided a value of $8 \pm 2 \cdot 10^8 M_{\odot}$ for the SBH in NGC 3115.

Figure 3. Reconstructed continuum image, mean stellar velocity and stellar velocity dispersion field of M 31 (OASIS, December 98, CFHT).

In 1993, HST WFPC1 images of the nucleus of M 31 resolved the light asymmetry in a double peaked structure (Lauer et al 1993). Spectrographic TIGER observations (Bacon et al 1994) strongly emphasized additional asymmetries in the velocity and velocity dispersion fields: the centre of rotation being roughly coincident with the fainter light peak and the maximum of velocity dispersion being significantly offset from that centre. The question of the nature of the secondary (brighter) light peak was left open. The spatial resolution obtained with TIGER for the M 31 data set was unfortunately not high enough (~ 1 arcsec FWHM) to clearly resolve the two peaks observed with HST.

Various hypotheses have then been invoked to explain the observed properties: an eccentric stellar disk with stars aligned on coplanar keplerian orbits (Tremaine 1995), or a stellar cluster on the verge of tidal disruption by SBH gravitational forces (Emsellem and Combes 1997). Since then, observations at better resolutions have been obtained using the long-slit spectrograph SIS at CFHT (Kormendy & Bender 1999), HST/FOS apertures (Pei et al. in prep.) and the long-slit mode of HST/FOC (Statler et al. 1999). We have recently obtained OASIS data with improved spatial resolution (0.45 arcsec FWHM). The velocity and velocity dispersion maps are presented in Figure 3. The various morphological quantities (rotation centre and principal axis, maximum of velocity dispersion) can be located *a posteriori* with respect to the light distribution, using the continuum image reconstructed from the data cube. Relative astrometry can be done up to a small fraction of the lens diameter, depending of the signal-to-noise ratio of the observations and the morphology of the object. Typical accuracy from a few hundredths to a tenth of the lens diameter can be achieved: about 0.01 arcsec in the case of the M 31 observations. With long-slit observations it is generally much more difficult to reach equivalent performances taking into account the lack of information outside the slit, all the more since the "transverse" position of the slit may not be accurately known. This is one of the fundamental advantage of 3D spectroscopy which will be essential to understand the complex structure observed in the nucleus of M 31. Analysis of the OASIS data set is presently under way and will be presented in a forthcoming paper.

Figure 4. TIGER observations of NGC 4278. Left side. Upper panel: continuum map, lower panel: stellar velocity map. Right side. Upper panel: [NII] emission map. Lower panel: [NII] velocity map.

4. Gas kinematics

In the classical scheme, gas fuels the engine of AGNs. As a dissipative system whose angular momentum can change with time, it plays a key role in the evolution of the dynamical structure and, as outlined in the introduction, can be a tracer of the gravitational potential. Ionized gas kinematics has thus been used to infer SBH masses. The most convincing results come from the HST WFPC2 and FOS observations of NGC 4374 (M 84; Bower et al. 1998) and M87 (Harms et al. 1994) where SBH masses of $1.5 \times 10^9 M_\odot$ and $2.4 \times 10^9 M_\odot$ were inferred from simple keplerian disk models. Using TIGER we have now completed a study of a few early-type galaxies meant to examine the morphological and dynamical coupling between the stellar and gas components (Emsellem et al., in preparation). Kinematical decoupling of gas and stars is frequently observed. An example is given in Figure 4 which presents the kinematical maps in the central region of NGC 4278. Note the decoupling of the gas with a rotation axis tilted by $\sim 30^\circ$ with respect to the star rotation axis. Such property strongly suggests an external origin for the gas which is being accreted onto a triaxial potential. In the case of NGC 128 (Emsellem & Arsenault 1997), the tilt of the gas rotation axis (and negative spin) can be explained assuming the potential is triaxial and tumbling, showing evidence for the existence of the so-called

anomalous retrograde orbits. Such measurements would have been difficult to obtain with a long-slit spectrograph, the slit being generally positioned along the major-axis measured on the broad band image. No *a priori* assumptions regarding the morphology are required with an IFS.

5. Prospects

The small sample of solid SBH detections outlines a correlation between SBH mass and the bulge luminosity of the parent galaxy (Magorrian et al. 98). However measurements are significantly biased towards large SBH mass particularly as the distance of the object increases.

Understanding the relationship between the global properties of the galaxy and the SBH requires both high spatial resolution IFS to investigate their nuclear properties and IFSs with a larger field of view to uncover their large scale dynamical state. OASIS is able to deliver high spatial resolution data on a limited number of nearby galaxies. At larger scale, the SAURON IFS (see Miller et al in this conference), with its relatively large field of view ($35'' \times 44''$), is being used for a systematic survey of kinematical properties of nearby elliptical and spiral bulges.

To extend these studies to higher redshifts will require progress in both spatial resolution and collecting power. On the VLT, the SINFONI instrument under development (see Thatte, this conference) will bring near-IR high spatial resolution IFS capabilities. The NGST with its exquisite spatial resolution, 8m collecting power, dark sky and infrared wavelength coverage could be a decisive tool, provided it is also equipped with IFS capabilities (see IFMOS presentation, Lefevre, this conference).

References

- Bacon, R., et al., 1995, A&AS, 113, 347
- Bacon, R., Emsellem, E., Monnet, G., Nieto, J.-L., 1994, A&A, 281, 691
- Bacon, R., 1998, in CFHT users meeting, eds P. Martin & S. Rucinsky, P.25
- Binney, J., Mamon, G. A., 1982, MNRAS, 200, 361
- Bower, G. A., et al, 1998, ApJ, 492, L111
- Emsellem, E., Dejonghe, H., Bacon, R., 1999, MNRAS, 303, 495
- Emsellem, E., Arsenault, R., 1997, A&A, 318, 39
- Genzel, R., Eckart, A., Ott, T., Eisenhauer, F., 1997, MNRAS, 291, 219
- Harms, R. J., 1994, ApJ, 435, 35
- Kormendy, J., et al., 1996, ApJ, 459, 57
- Kormendy, J., Bender, R., 1999, ApJ in press
- Lauer, T. R., et al., 1993, AJ, 106, 1436
- Magorrian, J., et al., 1998, AJ, 115, 2285
- Miyoshi, M., et al. 1995, Nature, 373, 127
- Rigaut, F., et al. 1998, PASP, 110, 152
- Shaver, et al., 1996, Nature, 384, 439
- Statler, T. S., et al., 1999, AJ, 117, 894
- Tremaine, S., 1995, AJ, 110, 628
- van der Marel, R. P., 1994, ApJ, 432, 91

The Nature of Obscured Nuclear Starbursts: M 82

N. M. Förster Schreiber

*CEA/DAPNIA/SAP, Centre d'Etudes de Saclay, F-91191
Gif-Sur-Yvette, France*

Abstract.

Results of near-infrared integral field spectroscopy of the starburst galaxy M 82, obtained with the MPE 3D instrument, are presented. These data provide the spatial distribution of key diagnostics (*e.g.* H I and He I nebular emission lines, molecular/atomic stellar absorption features) for the composition of the populations of hot massive stars and cool evolved stars, on scales of 20 – 25 pc. Together with mid-infrared spectroscopy from the SWS on board *ISO*¹ and data from the literature, the 3D data are analyzed with a starburst model to constrain quantitatively the cutoffs of the IMF, and the spatial and temporal evolution of starburst activity.

1. Introduction

In the past two decades, “starburst” galaxies, experiencing vigorous star formation activity in their central regions, have attracted considerable attention. However, little quantitative information exists on the properties of star formation in starburst environments and on the evolution of starburst activity. Detailed investigations have been hindered by the lack of spatially resolved data, and by dust obscuration often preventing studies at optical/UV wavelengths.

Near-infrared (NIR) integral field spectroscopy is ideal for investigating obscured starbursts. At these wavelengths, starburst activity is conspicuous and extinction is lower than in the optical/UV regimes ($A_K \approx 1/10 A_V$). The results of such observations of M 82 are presented here. Complemented with mid-infrared spectroscopy obtained with *ISO* and observations at various wavelengths from the literature, these data provide detailed constraints notably for the composition of the stellar population. With the application of a starburst model, the data are used to constrain *quantitatively* the cutoffs of the IMF as well as the burst age, timescale and intensity across the starburst core of M 82.

M 82 is one of the nearest starburst galaxies ($D = 3.3$ Mpc, implying $1'' \approx 15$ pc). The triggering of its starburst activity is generally attributed to gravitational interaction with M 81 $\sim 10^8$ yr ago or, alternatively, to a stellar bar possibly induced by the interaction. The starburst core of M 82 is enclosed by a molecular ring ≈ 500 pc in diameter, with inner ionized edge and roughly

¹*ISO* is an ESA project with instruments funded by ESA member States (especially the PI countries: France, Germany, the Netherlands and the United Kingdom) with the participation of ISAS and NASA. The SWS is a joint project of SRON and MPE.

centered on the nucleus. An important complex of supernova remnants extends over 600 pc along the plane of the galaxy, and material is flowing along the minor axis out to several kpc. For a more comprehensive description, see *e.g.* Kronberg *et al.* (1985), Larkin *et al.* (1994), Satyapal *et al.* (1995), Shen & Lo (1995), Shopbell & Bland-Hawthorn (1998), and references therein.

A representative region $250 \text{ pc} \times 150 \text{ pc}$ was observed, including the nucleus and extending to the west along the major axis up to the inner edge of the molecular ring. Various authors have modeled the global properties of M82 in the past, in particular Rieke *et al.* (1980; 1993). Satyapal *et al.* (1997) first presented models of selected stellar clusters seen at NIR wavelengths. In this work, the *entire* region mapped is modeled on scales of 20 – 25 pc.

2. Observations and results

The NIR data were taken with the MPE integral field spectrometer 3D (Weitzel *et al.* 1996; see also Genzel 1999, this volume), in January 1995 and 1996. 3D uses an image slicer and a grism to image a square field of view and to provide simultaneously the spectrum of each spatial pixel across the entire H - (1.45 – 1.85 μm) or K -band (1.9 – 2.5 μm). The data were acquired at $0.5''/\text{pixel}$ and spectral resolution $R \sim 1000$, in $\approx 1.5''$ seeing. 3 – 40 μm spectroscopy at $R \sim 1000$ was obtained with the Short Wavelength Spectrometer (SWS; de Graauw *et al.* 1996) on board *ISO* (Kessler *et al.* 1996) in March 1996. The SWS field of view coincides approximately with that for the 3D raster. Details of the data acquisition and reduction are given in Förster Schreiber *et al.* (1999a).

The 1 – 40 μm range is rich in signatures of starburst activity; the following are discussed below. Recombination lines of H I and He I, and collisionally-excited fine-structure lines of most abundant atoms (*e.g.* Ne, Ar, S) originate in nebulae ionized by OB stars. The He I 2.058 μm /Br γ 2.166 μm , He I 1.701 μm /Br 10 1.736 μm and [Ne III] 15.6 μm /[Ne II] 12.8 μm flux ratios probe the spectral shape of the ionizing radiation field, sensitive to the effective temperature of the OB stars ($T_{\text{eff}}^{\text{OB}}$). Thus, they constitute diagnostics for the most recent starburst activity (up to ~ 10 Myr ago). The NIR continuum emission of starbursts provides diagnostics for earlier star formation. Usually dominated by the light from cool evolved stars, it exhibits absorption features from various atoms and molecules present in the atmosphere of these stars. The feature strengths vary with stellar temperature and luminosity, and allow one to distinguish between red supergiants (formed $\sim 10 - 50$ Myr ago) and red giants (formed much longer ago). Particularly useful are the equivalent widths of the ^{12}CO (2,0) and ^{12}CO (6,3) bandheads at 2.29 μm and 1.62 μm , and of the Si II feature at 1.59 μm ($W_{2.29}$, $W_{1.62}$ and $W_{1.59}$; Origlia *et al.* 1993).

Maps of the K -band emission, of the Br γ and He I 2.058 μm line emission, and of $W_{2.29}$ extracted from the 3D data cube are shown in Fig. 1. The NIR spectrum of two representative regions are plotted in Fig. 2: the nucleus and the Br γ source “B1” (see Fig. 1). The tracers of the young and of the evolved stars show distinct burst sites as small as a few tens of parsecs, typical of giant H II regions/molecular clouds or complexes of such systems. Each stellar population is concentrated at different locations, as emphasized by the $\log(L_K/L_{\text{Ly}\alpha})$ map (Fig. 1d; the Lyman continuum luminosity $L_{\text{Ly}\alpha}$ is $\propto L_{\text{Br}\gamma}$). Along the outward

Figure 1. Images of M 82 obtained with 3D. The crossed circle indicates the position of the nucleus. Spectra extracted within the boxes are shown in Fig. 2. (a) K -band image from 2.0 to 12.0 ($\times 10^{-15} \text{ W m}^{-2} \mu\text{m}^{-1} \text{ arcsec}^{-2}$). (b) $\text{Br}\gamma$ linemap, from 5.0 to 25.0, and contours of the $\text{HeI } 2.058 \mu\text{m}$ emission, starting at 5.0 in steps of 1.0 ($\times 10^{-18} \text{ W m}^{-2} \text{ arcsec}^{-2}$). (c) Map of the $W_{2.29}$, from 8 Å to 16 Å. (d) Map of the $\log(L_K/L_{\text{Lyc}})$, from -1.2 to 0.0. (Grayscale intensities are given from white to black).

sequence $\text{N} \rightarrow \text{B1}$, the strength of the features relative to the continuum vary significantly, with shallower absorption features and stronger emission lines.

3. Population and evolutionary synthesis modeling

The HeI/HeII ratio maps and the global $[\text{NeIII}]/[\text{NeII}]$ line ratio of M 82 are interpreted using the photoionization code CLOUDY (Ferland 1996) and the model atmospheres for hot stars of Pauldrach *et al.* (1998). The most massive stars present have $T_{\text{eff}}^{\text{OB}} \approx 37400 \text{ K}$, corresponding to masses $\approx 30 M_{\odot}$. Throughout the regions mapped, little spatial variations in $T_{\text{eff}}^{\text{OB}}$ are found, indicating a similar evolutionary state for the youngest stellar populations.

The $W_{2.29}$, $W_{1.62}$ and $W_{1.59}$ are analyzed using the stellar spectral atlases obtained with 3D (Förster Schreiber 1999b) and from Kleinmann & Hall (1986), and Origlia *et al.* (1993). K5 supergiants dominate the continuum emission in the central few tens of parsecs of M 82. At larger radii, the shallower absorption features do not allow to discriminate between red giants and supergiants. However, the slight decrease in stellar mass to K -band luminosity ratio and the presence of supernova remnants favour red supergiants outside of the nucleus as well. The spatial variations in the equivalent widths are due to intrinsic variations in the evolved stellar population (with *e.g.* warmer K2 supergiants at B1) as well as to dilution by hot dust emission (up to $\approx 30\%$ at B1).

Figure 2. 3D spectra of the nucleus and of B1, taken in the apertures shown in Fig. 1, and normalized to unity in the interval 2.2875 – 2.2910 μm .

The starburst models combine the evolutionary synthesis code STARS (Kovo & Sternberg 1999) and CLOUDY. The star formation rate is $\propto \exp(-t_b/t_{sc})$, where t_b and t_{sc} are the burst age and timescale. A Salpeter IMF ($dN/dm \propto m^{-2.35}$) is adopted, between 1 $M_\odot$ and upper mass cutoff m_{up} . Model curves are plotted in Fig. 3 for $[\text{NeIII}]/[\text{NeII}]$, L_{bol}/L_{Lyc} , L_K/L_{Lyc} and $W_{2.29}$. For individual regions, an “equivalent $[\text{NeIII}]/[\text{NeII}]$ ratio” is obtained from the T_{eff}^{OB} inferred from the NIR HeI/HI ratios. The bolometric luminosity L_{bol} includes the contribution from OB stars and evolved stars (derived from the infrared and K-band luminosities; see Telesco *et al.* 1993; McLeod *et al.* 1993). The models are applied to the global properties of M 82, and to each resolution element in the regions mapped; properties for M 82, the nucleus and B1 are shown in Fig. 3.

From the diagnostics dominated by OB stars ($[\text{NeIII}]/[\text{NeII}]$ and L_{bol}/L_{Lyc}), burst timescales of a few Myr only are inferred. The absence of very massive stars ($> 50 M_\odot$) can be attributed to aging and rapidly declining star formation activity. Longer timescales are only possible for $m_{up} \sim 35 M_\odot$. However, this does not seem likely as in local high-mass star-forming regions, very massive stars form independently of the metallicity or of other global properties of the host system (*e.g.* Massey 1998). On scales of a few tens of parsecs, short timescales are easily understood since stellar winds and supernova explosions quickly disrupt the local ISM. The global ratios for M 82 require short burst durations on large scales as well. These results indicate *strong negative feedback effects* of starburst activity, both locally and globally.

Figure 3. Variation of selected properties with burst age, for $t_{sc} = 1$ Myr and 5 Myr to show the trend for increasing timescales, and for $m_{up} = 100 M_{\odot}$ (solid), $50 M_{\odot}$ (dashed), $35 M_{\odot}$ (dash-dotted), and $25 M_{\odot}$ (dotted). The global properties for M82, and those for the nucleus and B1 are indicated.

L_K/L_{Lyc} and $W_{2.29}$, with increasing contribution from the evolved stars, imply older ages than $[NeIII]/[NeII]$ and L_{bol}/L_{Lyc} . The various diagnostics can be reconciled if two populations formed in *successive short bursts* contribute to the properties, globally and along each line of sight. The two bursts inferred locally likely result from projection effects, since the red supergiants are distributed in the disk and peak at the nucleus while the OB stars are concentrated along a circumnuclear ring and along the bar (*e.g.* Larkin *et al.* 1994).

The results from the spatially detailed modeling suggest the following scenario: (1) Following tidal interaction with M81, $\sim 10^8$ yr ago, the ISM in M82 experienced large-scale torques, loss of angular momentum, and infall towards the nuclear regions possibly channeled by the bar. In the absence of dynamical resonances, the ISM can reach the nucleus. (2) A first starburst episode took place 8 – 15 Myr ago throughout the central 500 pc of M82, declined rapidly, and was particularly intense near the nucleus. (3) A second episode occurred 4 – 6 Myr ago predominantly in a circumnuclear ring and along the bar, was likely triggered by bar-induced dynamical resonances, and also decayed rapidly. (4) A supernova-driven starburst wind has now broken out of the galactic plane, the dramatic aftermath of the powerful nuclear starburst.

The two, short episodes are well revealed by the global temporal evolution of the starburst activity, reconstructed in detail by integrating the burst intensity as a function of age over the regions modeled, shown in Fig. 4. Similar results for the nature and evolution of starburst activity have been obtained for IC 342, based on 3D data (Böker *et al.* 1997). A full account of this study of M82 will be presented in a forthcoming paper (Förster Schreiber *et al.* 1999a).

Figure 4. Global starburst history of M82 derived from spatially detailed starburst modeling (R_0 is the star formation rate).

I wish to thank R. Genzel and L. Tacconi for precious guidance, the 3D and SWS teams at MPE, A. Sternberg, M. Rieke, G. Rieke for help and discussions, and C. Telesco and D. Gezari for their mid-infrared image.

References

- Böker, T., Förster Schreiber, N.M. & Genzel, R. 1997, *AJ*, 114, 1883
 de Graauw, T. *et al.* 1996, *A&A*, 315, L49
 Ferland, G.J. 1996, *Hazy, a Brief Introduction to Cloudy*, University of Kentucky Department of Physics and Astronomy Internal Report
 Förster Schreiber, N.M. *et al.* 1999a,b, in preparation
 Kessler, M.F. *et al.* 1996, *A&A*, 315, L27
 Kleinmann, S.G. & Hall, D.N.B. 1986, *ApJS*, 62, 501
 Kovo, O. & Sternberg, A. 1999, in preparation
 Kronberg, P.P., Biermann, P. & Schwab, F.R. 1985, *ApJ*, 291, 693
 Larkin, J.E. *et al.* 1994, *ApJ*, 420, 159
 Massey, P. 1998, in *The Stellar Initial Mass Function*, ASP Conf. Series vol. 142, eds. G. Gilmore and D. Howell (San Francisco: ASP), p. 17
 McLeod, K.K., Rieke, G.H., Rieke, M.J. & Kelly, D.M. 1993, *ApJ*, 412, 111
 Origlia, L., Moorwood, A.F.M. & Oliva, E. 1993, *A&A*, 280, 536
 Pauldrach, A.W.A., *et al.* 1998, in *Boulder-Munich Workshop II: Properties of Hot Luminous Stars*, ASP Conf. Series vol. 131, ed. I. Howarth (San Francisco: ASP), p. 258
 Rieke, G.H., Lebofsky, M.J., Thompson, R.I., Low, F.J. & Tokunaga, A.T. 1980, *ApJ*, 238, 24
 Rieke, G.H., Loken, K., Rieke, M.J. & Tamblyn, P. 1993, *ApJ*, 412, 99
 Satyapal, S. *et al.* 1995, *ApJ*, 448, 611
 Satyapal, S. *et al.* 1997, *ApJ*, 483, 148
 Shen, J. & Lo, K.Y. 1995, *ApJ*, 445, L99
 Shopbell, P.L. & Bland-Hawthorn, J. 1998, *ApJ*, 493, 129
 Telesco, C.M., Dressel, L.L. & Wolstencroft, R.D. 1993, *ApJ*, 414, 120
 Weitzel, L. *et al.* 1996, *A&AS*, 119, 531

BEAR Imaging FTS: High Resolution Spectroscopy in Infrared Emission Lines

J.P. Maillard

Institut d'Astrophysique de Paris - CNRS, Paris, France

Abstract. An imaging FTS makes 3D spectroscopy possible at high resolution in the infrared. The coupling of the step-scan FTS of the CFH Telescope, to a NICMOS3-type array has been realized to exploit this concept in the 1 to 2.5 μm domain. Any spectral resolution up to $\sim 30,000$, in a spectral range within these limits, is possible. The most appropriate applications are the studies of the physical conditions and the kinematics of various gaseous media through their emission lines (like Br γ , H $_2$ 2.12 μm , HeI 2.06 μm ...). Hence, young stellar objects interacting with their molecular clouds, planetary nebulae at different phases of evolution, as well as the inner region of the Galactic Center, are suitable targets.

1. Properties of an IFTS

The main characteristics of an astronomical Imaging FTS (IFTS) can be summarized as follows (Maillard 1995):

- the entrance field is imaged on a 2-D array through a Michelson-type interferometer. Each pixel acts as a detector matched to the corresponding point on the sky. The pixel size must be adapted to the seeing disk (or the diffraction pattern) to correctly sample it. An IFTS makes it possible Integral Field Spectroscopy (IFS) over a field, the size of which is limited only by the pixels number and the pixel scale.
- data acquisition is only compatible with a step-by-step scanning of the path difference. An image is recorded at each step.
- spectral resolution is continuously adaptable, up to a maximum value which depends on the beam diameter, the field and the pixel size. However, the overhead time associated with the recording of each image, which must be as short as possible, imposes a practical limit.
- the spectral coverage can in theory be as wide as wanted. However, since the noise increases with the bandpass, filters selecting spectral domains of interest are preferable.
- by applying a FT to the interferogram for each spatial element, the spectral data cube is directly produced. Each plane is a monochromatic image of the field. Imaging through the equivalent of an ideal square filter is obtained, by summing on slices of the cube.

Despite these constraints, when compared to any other spectro-imaging devices, an imaging FTS offers the maximum flexibility, in the choice of spectral resolution and spectral coverage, and over a sizeable field.

2. Presentation of the BEAR instrument

The facility FTS installed at the infrared f/35 focus of the CFH Telescope, currently used for high-resolution spectroscopy (60 cm max path difference), with InSb mono-detectors in the 1 – 5.4 μm range (Maillard and Michel 1982) has been coupled to a 256×256 HgCdTe near-infrared camera (Simons *et al.* 1994, Maillard 1995). Since it is a step-scan instrument, the data acquisition was directly adapted to work in imaging mode. The interferometer uses cat's eye retroreflectors in each arm to provide the symmetric and asymmetric outputs. The field of view of 24" is determined by the size of the secondary mirror of the cat's eyes. An interface box provides a means of bringing the two images of the telescope's focal plane produced by the FTS onto a single array. The final scale on the detector is 0.35 arcsec/pixel (4 pixels in the core of a stellar image). The dead time to read and store a frame is currently of 2.5 s. To limit the total overhead time of a scan to no more than $\simeq 40$ min, the number of frames in a raw data cube must be limited to ~ 1000 . Thus, if a high resolution is needed (large path difference), the spectral coverage must be limited, which permits an increase of the step size and hence, a total number of frames not higher. These considerations determine the optimum use of the instrument. In practice, a minimum of two hours and ~ 1000 frames devoted to one scan is needed, to obtain an efficiency better than 50%. Allowing for source acquisition, four cubes can be recorded in one night. Resolutions of ~ 3000 for a $\Delta\lambda/\lambda \simeq 15\%$ filter (e.g. a K' filter), and of $\sim 30,000$ for a 1.5% filter (e.g. a 2.12 μm filter for H_2) are typical values for data cubes acquired with BEAR. At $\sim 10 \text{ km s}^{-1}$ resolution, through a narrow filter, line fluxes in H_2 1-0 S(1) as faint as $10^{-19} \text{ W m}^{-2}/\text{arcsec}^2$ are detected.

2.1. BEAR software

Specialized software, (all written in IDL), is also part of the BEAR development effort. The available procedures are :

- *bearprocess* : generates the spectral cube from the raw data cube (~ 1000 frames of 128×256 pixels). This two-part package, first, conducts standard image processing (flat-fielding, dark subtraction, correction for bad pixels), flipping of one image about the Y-axis and spatial registration with respect to the other image to insure the complementarity of all the points in image pairs, registration of all the frames with respect to the first one to correct for guiding errors and flexure drifts. Secondly, the prepared cube is scanned in X and Y to extract the interferogram from each pair of complementary points. The zero path difference is then determined in order to obtain the spectrum directly as the real part of the interferogram FT, and restore the optimum spectral resolution of the data.
- *psubcube* : extracts a sub-cube from the main cube, limited to a spectral region of interest (e.g. a line profile), in which the original channel width (determined from the Shannon's criterion) is divided into an equal number of intervals. It results in a local oversampling, making it possible, without generating too large

a cube, to apply the correction of the natural phase error present throughout the field and, to dispose of the full line position precision.

- *merge_cube* : merges two cubes (original cube or sub-cube) recorded to increase the spatial coverage of an extended object. It assumes a stellar-like object in the common area (e.g. Fig. 1).
- *bear_calib* : provides the absolute calibration of spectral cubes, by using a spectrum extracted from a similar cube recorded on a reference star.
- *cubeview* : permits inspection of any spectral cube. Monochromatic images can be displayed and the spectrum of any point of an image extracted through an adjustable square aperture, to be shown and stored with the scale in wavenumber, wavelength or velocity (e.g. Fig. 1 and 2).

2.2. Scientific programs

IFS on emission-line objects at high spectral resolution in the near infrared, are the most appropriate and unique programs for BEAR. This choice is dictated by the desire to obtain the optimal sensitivity, with the highest possible spectral resolution. As the CFHT-FTS is a room-temperature instrument, spectro-imaging beyond $2.5\ \mu\text{m}$ is excluded. With narrow-band filters, the background emission can be limited and for emission-line objects the mean energy within a filter is low, all of which contribute optimizing the S/N ratio. The 1 to $2.5\ \mu\text{m}$ range provides access to many common spectroscopic indicators including, atomic hydrogen (in particular $\text{Br}\gamma$ $4616.6\ \text{cm}^{-1}$), molecular hydrogen ($1 - 0\ \text{S}(1)$ $4712.9\ \text{cm}^{-1}$), HeI ($4857.5\ \text{cm}^{-1}$) and also FeII and CO ($\Delta v = 2$) lines. Young stellar objects, late type stars in the planetary nebula phase, star clusters and starburst galaxies, less than 1 arcmin in size, are suitable targets for BEAR. Brief reviews of some programs are given below.

3. The inner region of the Galactic Center

Observations of the central parsec of the Galaxy have been undertaken with BEAR, with a filter isolating the $\text{Br}\gamma$ line, first by Simon and Maillard (1996), and more recently by Morris and Maillard (2000, this volume). With a spectral resolution of $32\ \text{km s}^{-1}$, the kinematics of the main components of the minispiral traced by this line have been modelled in the context of the gravity field created by the $2.6 \times 10^6\ M_{\odot}$ black hole. Observations have also been conducted with a $2.058\ \mu\text{m}$ filter to reanalyze the HeI star cluster (Krabbe *et al.* 1995) detected in the IRS16 region. The profile of this line in the HeI stars has been obtained with a resolution of $40\ \text{km s}^{-1}$, which is superfluous for profiles $1000\ \text{km s}^{-1}$ wide. However, with this resolution it has been possible to detect, for the first time, helium streamers, with a linewidth of $\sim 50\ \text{km s}^{-1}$.

This gas must originate from the mass loss of the HeI stars. It was possible to separate this contribution from the emission of the stars and to eliminate those which were improperly identified as HeI stars. The comparison of the velocity field of the ionized gas in these two tracers ($\text{Br}\gamma$ and $2.058\ \mu\text{m}\ \text{HeI}$), which extends over $700\ \text{km s}^{-1}$ will improve our understanding of the complex kinematics in the vicinity of the central black hole.

Figure 1. Image in the $\text{Vel} = 0 \text{ km s}^{-1}$ plane, showing the patchy structure of the flows and, line profiles in different points (Maillard *et al.* 2000). The bright spot at $[0,0]$ is the young source S106-IR.

4. The environment of young stellar objects

The massive young stellar objects (YSOs) are hot stars (O type or early B). In their early phases they are still embedded in their parent molecular cloud, and are thus only accessible in the infrared. They contribute efficiently to the dispersion of their dusty and gaseous environment. Mapping the velocity of the gas, in order to study the turbulent conditions in the vicinity of the hot central star is a unique capability provided by BEAR. In this case, the maximum resolution (up to 10 km s^{-1}) is attempted as the lines are narrow, which requires the use of a narrow-band filter isolating a single line. Br γ was used to study the ionized regions, and the $1 - 0 \text{ S}(1)$ and $2 - 1 \text{ S}(1)$ H $_2$ lines to derive line ratios to test the excitation conditions of the molecular cloud. A selection of high-mass YSOs has been observed, which represents a sequence of evolution from a deeply embedded YSO to its visible stage. Complementary data have also been obtained in the mid-IR with the ISO satellite. The current ground-based observations provide a ten times better spatial resolution but on a smaller field. Typical objects are :

- GL2591 : an object in a very early phase. A HII region has not yet developed. The surrounding H $_2$ is relatively undisturbed by the central object.
- S106 : in contrast presents a very developed HII region with a broad bipolar structure (Fig. 1). Information on the 3-D structure of the flows in H and H $_2$ can be extracted from the data, with a study of their interaction and also with PAH grains, by comparison with 10-micron maps (Maillard *et al.* 2000).

- LkH α 234 : an YSO in a more evolved state with deflected jets.
- NGC 7023 : a prototypical reflection nebula. The hot star is no longer enshrouded. Remnants of the parent molecular cloud remain in filaments excited by the UV field of the star. Excitation conditions, densities, velocity fields are deduced from the data (Lemaire et al. 1999).

Figure 2. Velocity profile of the 1–0 S(1) H_2 line in the PN BD+30°3639, extracted with *cubeview* (see text) from the 15 km s^{-1} resolution cube with a square aperture covering the full field. Several velocity components are seen, representative of distinct ejection events.

5. Study of planetary nebulae

The planetary nebulae (PN) represent the final stages of evolution of 1 to $8M_{\odot}$ stars. Generally, the remnant of the red giant circumstellar envelope forms an outer molecular shell surrounding the ionized region excited by the central white dwarf. The integrated spectrum in the infrared shows a rich variety of emission lines from neutral and ionized atoms, and molecular hydrogen as well. A K' filter covering simultaneously several H_2 lines, $B\gamma$, HeI and HeII lines was used, making it possible a clear separation of the regions of excitation of each species. Complementary studies were made at high resolution (up to 10 km s^{-1}) through narrow-band filters to study the velocity field in the prominent emission lines. Different PNs which form a sequence of evolution, were observed. Depending on their stage of evolution and the UV-field intensity, these lines are more or less intense :

- CRL 2688 and CRL 618 : two examples of pre-PN. The H_2 envelope is well developed with a bipolar structure while $B\gamma$ is not detected (Cox et al. 1997).
- NGC 7027 : the prototype of a young PN with an almost spherical, hollow, ionized region, inside a symmetrical H_2 envelope (Cox et al. 1997).
- BD+30°3639 : a young PN slightly more evolved. With the K' filter, the ionized region (with $B\gamma$ and $2.058 \mu\text{m}$ HeI) of same shape than in NGC 7027, and the H_2 envelope, still with a general bipolar shape but very irregular, were

detected. From a high resolution study with a filter isolating the $2.12\ \mu\text{m}$ line, the molecular envelope appears with a complex velocity field (Fig. 2).

6. Conclusion

The results reviewed above show that the combination of high spectral resolution with 3D spectroscopy provides a unique diagnosis for complex sources. This capability is provided by spectro-imaging made with an IFTS. However, the BEAR instrument does not represent the last word for this method. Further developments are possible. It would be tempting to extend the same technique toward the mid-infrared. Because of the increasing thermal background, the sensitivity is degraded on ground-based telescopes. On the contrary, this extension is promising in space and is proposed for an IFTS on the NGST (Graham 1998, Maillard & Renault 2000) with the extremely low background environment it offers. The combination of high spectral resolution with large spectral coverage is also a feature of an IFTS. In the imaging mode, the quantity of data to store would rapidly become a limiting factor. In the NGST application, a rationale completely different than for an instrument like BEAR is adopted : very low spectral resolution (~ 100), on a very large spectral domain (1 to $5\ \mu\text{m}$) and a field of several arcmins. Spectro-imaging at the diffraction limit of an 8m telescope would become possible. Note that could be already feasible on a small field with an IFTS behind an adaptive optics system. Finally, the optical simplicity of an IFTS, even with a large field, is a property which should be exploited for a spaceborn instrument.

References

- Cox, P., Maillard, J.P., Forveille, T., Simons, D., Guilloteau, S., Huggins, P.J., Omont, A., and Bachiller, R. 1997, A&A, 321, 907
- Graham, J.R., Abrams, M., Bennett, C., Carr, J., Cook, K., Dey, A., Najita, J., Wishnow, E. 1998, PASP, 110, 1205 and this volume
- Krabbe, A. *et al.* 1995, ApJ, 447 L95
- Lemaire, L. *et al.* 1999, A&A, in press
- Maillard, J.P. 1995, in *Tridimensional Optical Spectroscopic Methods in Astrophysics*, IAU Col. 149, G. Comte, M. Marcelin (ed), ASP Conf. Series, 71, 316
- Maillard, J.P. & Renault, J.C. 2000, this volume
- Maillard, J.P. & Michel, G. 1982, in *Instrumentation for Astronomy with Large Optical Telescopes*, IAU Col. 67, C.M. Humphries (ed), D. Reidel (Pub), 92, 213
- Maillard, J.P., Joblin, C., Mitchell, G.F., Vauglin, I. and Cox, P. 1999, in *The Universe as seen by ISO*, Paris (20-23 Oct 1998), P. Cox (ed), in press
- Morris, M. & Maillard, J.P. 2000, this volume
- Simons, D.A., Clark, C.C., Smith, S., Kerr, J., Massey, S., Maillard, J.P. 1994, in *Instrumentation in Astronomy VIII*, SPIE, 2198, 2198
- Simons, D.A. & Maillard, J.P. 1996, in *The Galactic Center*, 4th ESO/CTIO Workshop, R. Gredel (ed), ASP Conf. Series, 102, 233

CIRPASS: A Near-IR Integral Field Spectrograph

I. R. Parry, A. J. Dean, R. S. Ellis, D. King, C. D. Mackay, R. G. McMahon, S. R. Medlen, J. M. Pritchard, A. N. Ramaprabh

Institute of Astronomy, Madingley Rd, Cambridge, CB3 0HA, UK

Abstract. We are in the final stages of constructing a near-IR integral field spectrograph for use on the Gemini 8m telescopes. CIRPASS, as this instrument is known, will have a fibre-based 499 element integral field unit (IFU). Several sampling scales and fields of view will be available by interchanging the fore-optics that feed the IFU. CIRPASS operates in the wavelength range 0.9 – 1.8 microns and offers spectral resolutions from 3,000 to 26,000. Other features of CIRPASS are a special masking technique which removes the OH airglow well upstream of the detector and a multi-fibre feed for multi-object spectroscopy.

1. Introduction

High resolution images of distant galaxies obtained with HST and the promise of excellent image quality on the new generation of ground-based 8m telescopes have encouraged many groups to devise instruments that can perform spatially resolved spectroscopy to determine internal dynamics, and study stellar populations etc. And because these systems are at high redshift and of low surface brightness this work requires a very sensitive near-IR integral field spectrograph.

This paper briefly describes such a spectrograph that we are building primarily for use on the two Gemini 8m telescopes. CIRPASS (the Cambridge IR PANoramic Survey Spectrograph) uses a Rockwell 2048 Hawaii-II array giving a large combination of FOV and wavelength coverage in a single exposure. Other ingredients of the design include the versatility to be readily used on many telescopes and with AO systems; very faint limiting magnitudes through a combination of hardware and software suppression; and a custom-made IFU using a very high quality macro-lens array coupled with very high precision to a fibre-optic bundle.

2. The Integral Field Unit

The CIRPASS IFU is a larger version of the 100 element IFU (Kenworthy et al, 1998a, 1998b, Kenworthy 1998) that we built for the COHSI spectrograph (Piché et al 1997; Ennico et al 1998, Ennico 1998). It uses a macro lens array of 499 lenses with each lens forming a micro-pupil image on to an optical fibre. The fibres are then fed to the spectrograph which is located on the observing floor. As for the COHSI IFU each lens in the array is hexagonal and is 3mm

across the corners. To reimage the telescope focal plane on to this array at a useful scale, two lenses are used. By changing the first of these two lenses the scale on the sky can be varied. Three sampling scales will be available when used with the Gemini f/16 focal plane: 0.36 arcsec/lens with a FOV of 13×5 arcsec; 0.25 arcsec/lens with a FOV of 9.2×3.5 arcsec and 0.12 arcsec/lens with a FOV of 4.4×1.7 arcsec. A fourth scale option will be available with the University of Hawaii's adaptive optics system (Hokupa'a, Graves et al, 1998) giving 0.05 arcsec per lens and a FOV of 1.8×0.7 arcsec.

Figure 1 illustrates the layout of the lens array on the sky for an image scale of 0.25 arcsec/lens. The mapping of the fibres from the two-dimensional lens array to the one-dimensional spectrograph slit is such that fibres that are adjacent on the slit are also adjacent on the sky.

3. The Spectrograph

The CIRPASS spectrograph design is of the Baranne type and features an intermediate spectral image where it is possible to place a suppression mask to remove light from the strong OH airglow emission. The optics from the fibre slit to the cryogenic camera are the same optics as were previously used in the COHSI spectrograph. In the CIRPASS case however the light coming from the mask is not undispersed by a second grating as it was in COHSI but is simply fed directly into the cryogenic camera which reimages the mask surface on to the detector preserving the high spectral resolution but scaling the image down so that the physically large format at the mask fits on to the Rockwell array.

By using a mask rather than a plane mirror the strongest source of light in to the spectrograph (i.e. the OH) is removed at the earliest possible position. Thus the amount of scattered light in the spectrograph optics (and in particular in the camera) will be reduced and the lowest possible background between the OH lines will be achieved. We calculate that a sensitivity gain of 0.25 magnitudes will be obtained by removing the OH light at the mask surface if the optics between the mask and the detector scatter 1% of the propagating light. We use the term *scattering* here to include any effects that cause light to arrive at detector pixels well away from those predicted by the ray-traced PSF. A major contributor is likely to be out of focus ghost images caused by the relatively high reflectivity of the detector surface and the filters. The actual performance however has not been measured yet because at the time of writing the camera is still under construction.

The optics of the camera operate at 77K and the rest of the spectrograph operates at 218K. The camera works at f/1.35 and uses 5 single lenses made of flint glass. The camera also has a fixed filter for blocking light longward of 1.8 microns and two filter wheels to optionally provide further blocking and order sorting. All the filters operate at 77K. The detector is tilted slightly to allow for the chromatic focal shift of the camera design. The beam size of the spectrograph is 150mm diameter and the largest lens in the camera is 224mm in diameter.

Figure 2 shows the optical layout. The light enters the spectrograph via the optical fibre slit which is at the focal plane of a large classical schmidt system. After striking the primary mirror and passing through an off-axis sub-aperture

of the aspheric plate the light is collimated in to a 150mm diameter beam and this falls on the reflection grating. A grating of 400 lines per mm is used to give a spectral resolution of ~ 3000 in first order. The grating is used in a near-Littrow configuration and the light passes back through the schmidt system forming a spectrum of ~ 300 mm length on the mask surface for the whole spectral range 1.0-1.8 microns. In practice, because the fibre slit has to share the same focal plane, two mask mirrors (one each for J and H) are used to selectively reflect the light present in the intermediate spectral image. In figure 2 only the J mask is shown for simplicity.

The mask surface has a pattern on it which very accurately matches the OH spectrum. In the regions where light from OH lines hits the mirror the surface coating is a multi-layer anti-reflection coating with a reflectivity of less than 0.5%. In the other regions which are unaffected by the OH airglow the coating is gold with a reflectivity of $\sim 98\%$. The gold areas add up to more than 90% of the total mask surface area. We have developed special techniques for making these mask mirrors. For details of these techniques see Ennico, 1998.

Light that reflects off the mask is then collimated by the other half of the large schmidt optics. The light therefore passes through the schmidt system (one mirror and one aspheric plate) three times in total. The collimated light is folded by the steering mirror and imaged on to the detector. By adjusting the tilt of the steering mirror any region on the mask surface can be brought to the centre of the detector's FOV. The 2048×2048 array can image more than 0.4 microns of wavelength coverage so it is possible to observe all of H or all of J in one shot.

As well as allowing the use of suppression masks the Barranne design with its near-Littrow configuration also allows steeply angled gratings to be used which provide very high spectral resolution. Grating options with $R \sim 26,000$ and $R \sim 10,000$ will be provided. These will be used with plane mirrors rather than mask mirrors and the OH suppression will be done entirely in software. There will be a set of order sorting filters in the filter wheel in the camera dewar.

After six to nine months of operation with the IFU we will also provide a plug-plate system with ~ 100 -150 fibres for multi-object spectroscopy. Each fibre will be 1 arcsec in diameter on the sky and the FOV on Gemini will be ~ 6 arcmin in diameter.

Our thanks go to The Raymond and Beverley Sackler Foundation, PPARC and the Gemini project for the support they have given to this project. Thanks are also due to Alfonso Aragon-Salamanca, Karl Glazebrook, Matthew Kenworthy, Kimberly Ennico and the staff at UKIRT for their help with COHSI.

References

- Ennico, K. A., Parry, I. R., Kenworthy, M. A., Ellis, R. S., Mackay, C. D., Beckett, M. G., Aragon-Salamanca, A., Glazebrook, K., Brinchmann, J., Pritchard, J. M., Medlen, S. R., Piché, F., McMahon, R. G., Cortecchia, F., 1998, "The Cambridge OH Suppression Instrument (COHSI): Status After First Commissioning Run", *Proc of SPIE*, 3354, 668.
- Ennico, K. A., 1998, "Near Infrared Faint Object Spectroscopy", Ph.D. Thesis, University of Cambridge.

Figure 1. Integral Field Spectroscopy of a galaxy merger using CIRPASS: the spectral resolution will facilitate internal dynamical studies. The galaxies shown here are CFRS 14.1129 in HST F814W. At $z=0.831$ H-alpha is at 1.202 microns. In this illustration the fore-optics are arranged to give 0.25 arcsec across each lenslet.

- Graves, J. E.; Northcott, M. J.; Roddier, F. J.; Roddier, C. A.; Close, L. M., 1998, "First light for Hokupa'a: 36-element curvature AO system at UH" Proc. SPIE Vol. 3353, p. 34-43, Adaptive Optical System Technologies, Domenico Bonaccini; Robert K. Tyson; Eds.
- Kenworthy, M. A., 1998, "The Development of New Techniques for Integral Field Spectroscopy in Astronomy", Ph.D. Thesis, University of Cambridge.
- Kenworthy, M. A., Parry, I. R., Ennico K. A., 1998a "COHSI: a lens array and fibre feed for the near infra-red", proc. of the conference "Fibre Optics in Astronomy - III", ASP conf. series vol 152, 300, eds. Arribas, Mediavilla and Watson.
- Kenworthy, M. A., Parry, I. R., Taylor, K., 1998b, "Integral Field Units for SPIRAL and COHSI", Proc. SPIE Vol. 3355, p. 926-931, Optical Astronomical Instrumentation, Sandro D'Odorico; Ed.
- Piché, F., Parry I. R., Ennico, K. A., Ellis, R. S., Pritchard, J. M., Mackay, C. D., McMahon, R. G., 1997, "COHSI: The Cambridge OH Suppression Instrument", Proc of SPIE vol 2871, 1332.

Figure 2. The optical layout of CIRPASS. Light from the telescope is injected into the spectrograph from the fibre slit which is located in the centre of the system. The chief ray from the central fibre at the central wavelength (1.2 microns) is shown. In this example the steering mirror is looking at the J mask. The H mask, which sits above the fibre slit, is not shown for clarity.

Kinematics of Ionized Gas in the Central Parsec of the Galaxy from High-Resolution Spectroscopy of the Brackett- γ Line

Mark Morris¹

Dept. of Physics & Astronomy, UCLA, Los Angeles, CA 90095, USA

Jean-Pierre Maillard

Institut d'Astrophysique, 98 bis Bd. Arago, 75014 Paris, France

Abstract. With the BEAR Fourier Transform Spectrometer on the CFHT, spectro-imaging of the Br γ emission line was carried out toward the central portions of the SgrA West HII region. Here, we focus on the kinematics of three structures in this region: the mini-cavity and the Northern and Eastern Arms. Both arms appear to be continuous streams of gas accelerating toward the $2.6 \times 10^6 M_{\odot}$ black hole at the center, and both culminate in a bright, narrow tip at their leading edges, presumably resulting from compression by the converging gravitational field of the black hole. The mini-cavity, located just 2 arcsec from the central mass, has the apparent morphology of a hollow, spherical shell, but the data indicate that its kinematics are substantially more complex than simple spherical expansion, in part because of the strong gravitational field of the nearby black hole, and in part because the source of the kinetic energy in the walls of the cavity may lie outside the cavity.

1. Introduction

The HII region Sgr A West, lying within the Galaxy's central parsec, is composed of at least three gas streams that appear to converge within several arcseconds of the central black hole. Much of the gas in this region has been ionized by radiation from the central cluster of hot, young stars. The kinematics of Sgr A West have been studied using 12.8- μ m NeII emission (Lacy et al. 1980; Serabyn et al. 1988), radio recombination lines (Schwarz, Bregman & van Gorkom 1989; Roberts & Goss 1993; Roberts, Yusef-Zadeh & Goss 1996), and near-IR recombination lines, notably Br γ (Geballe et al. 1987; Herbst et al. 1993). The picture that has emerged is that the ionized arms of Sgr A West represent either accretion streams (the "Northern Arm" and the "Eastern Arm"), the inner edge of the circumnuclear disk (the "Western Arc"), or the kinematically complex "bar" of emission passing near SgrA*. Within the bar lies the Mini-Cavity, an apparently evacuated region near the central black hole.

¹CNRS Chercheur Associé, Institut d'Astrophysique de Paris

Figure 1. *Figure 1:* Br γ intensity integrated between -400 and $+400$ km s^{-1} toward the two mosaicked fields of Sgr A West. The prominent features are indicated. Also, several hot stars show up as point sources in this image because of their Br γ emission. In this figure, North is oriented in position angle 20° .

Here, we report new Br γ observations of Sgr A West, made with the BEAR FTS spectrometer on the CFHT (Simons et al. 1994; Maillard 1995). These observations were made with an unprecedented combination of spatial and spectral resolution (0.6 arcsec & 33 km s^{-1}), allowing us to follow the gas streams into and through complex emitting regions, and to distinguish the characteristic broad-line profiles of the spatially unresolved stellar sources from the extended gas.

2. Observations

Two overlapping circular fields of diameter 24 arcsec (1 pc) were observed toward the Galactic center on 1997 July 24-26 with a spectral resolution of 33 km s^{-1} . Continuum maps were made from portions of the spectrum lying beyond the range of velocities in which Br γ emission is present, and these were subtracted from the entire data set to produce the resulting line cube. Figure 1 shows the integrated line image of the two merged fields.

Figure 2. *Figure 2:* The Northern Arm of SgrA W, plotted in a position-velocity diagram. The horizontal axis is distance along the ridge of the Arm, starting at its northernmost point in figure 1. The model locus of the N Arm is superimposed (an ellipse with eccentricity 0.74, inclination= 63° , and a periaapse distance of 2.5×10^{17} cm (0.1 pc)).

2.1. The Northern Arm

The Northern Arm, a well-defined, narrow gas stream with a monotonically increasing velocity gradient along its length, apparently results from the sheared infall of gas toward the central black hole on a highly eccentric orbit. In our data, the main ridge of the Northern Arm can be followed to where it abuts the mini-cavity, ending abruptly at that point in a narrow tip. This is the point of closest approach to SgrA*, and the point at which the radial velocity gradient and magnitude (-400 km s^{-1}) both have maximum values (Fig. 2). This indicates that the Northern Arm is accelerating toward its point of closest approach to SgrA*, and that it is being compressed at its tip as it moves into the converging gravitational field of the central point mass. The Northern Arm appears to have a weaker, more diffuse secondary branch which passes to the South of the mini-cavity, but it, too, ends in the vicinity of the mini-cavity.

The Br γ emission depicted here reveals only the ionized component of the gas. There is probably a dominant neutral component as well (Jackson et al. 1993; Davidson et al. 1992). Our images show diffuse Br γ emission at velocities similar to that of the Northern Arm extending across the entire area between the N and E Arms. We suggest that this emission arises from the extended surface layers, or from the partially ionized interior, of the neutral portion of the structure underlying the Northern Arm.

2.2. The Mini-Cavity

Because of its bright rim with well-defined inner edges, and its relatively hollow interior, the mini-cavity has long been regarded as a structure created by a wind or perhaps by a broadly-directed jet. If so, the kinematics of the gas in the rim should be indicative of the mechanism which evacuated the cavity. Our results indicate that the situation is much more complex than simple spherical expansion by itself. While the mini-cavity may basically have resulted from a wind, its kinematics are probably also affected by at least two other factors: 1) the fact that it lies at the apparent confluence of various gas streams, including the N Arm and the gas in the poorly understood bar, and 2) the exceptionally strong gravity of the nearby black hole. One possible indication of a relatively high-velocity expansion motion is the presence of relatively weak emission at $-400 \pm 50 \text{ km s}^{-1}$ across the face of the cavity. This emission, which appears to be continuous in velocity and space with the bright emission at the tip of the N Arm, could be explained in terms of gas expanding toward the observer on the near side of the mini-cavity.

2.3. The Eastern Arm

Like the N Arm, the E Arm is apparently an infalling gas stream. It is kinematically distinct from the rest of SgrA W, although it is continuous in projection with the emission in the N Arm and the bar. The radial velocity in the E Arm is remarkably constant at $+150\text{--}160 \text{ km s}^{-1}$ over most of its 14-arcsec E-W extent (the gradient of $\sim 30 \text{ km s}^{-1}$ across its 3-4 arcsec width is substantially larger than the gradient along most of its length). At its western tip, however, it ends in a feature showing a strong velocity gradient of a few hundred km s^{-1} in just a few arcseconds, reaching a maximum velocity of $+350 \text{ km s}^{-1}$ at its westernmost tip. The E Arm is analogous to the N Arm in three respects: 1) both are relatively narrow streams of gas that curve toward the central mass and are therefore likely to be infalling structures, 2) both end in bright, narrowing tips of emission at their points of closest approach to SgrA*, and 3) both show a strong velocity gradient at their tips. However, both the radial velocity and velocity gradient of the E Arm have the opposite sign to those of the N Arm.

References

- Davidson, J.A. et al. 1992, *ApJ*, **387**, 189.
 Geballe, T.R., Wade, R., Krisciunas, K., Gatley, I. & Bird, M.C. 1987, *ApJ*, **320**, 562.
 Herbst, T.M., Beckwith, S.W.V. & Forrest, W.J. 1993, *AJ*, **105**, 956.
 Jackson, J.M. et al. 1993, *ApJ*, **402**, 173.
 Lacy, J.H., Townes, C., Geballe, T., Hollenbach, D. 1980, *ApJ*, **241**, 132.
 Maillard, J.-P. 1995, in *Tridimensional Optical Spectroscopic Methods in Astrophysics*, IAU Coll. 149, eds: G. Comte & M. Marcelin, ASP Conf. Ser. No. 71, p. 316.
 Roberts, D.A. & Goss, W.M. 1993, *ApJS*, **86**, 133.
 Roberts, D.A., Yusef-Zadeh, F. & Goss, W.M. 1996, *ApJ*, **459**, 627.
 Schwarz, U.J., Bregman, J.D. & van Gorkom, J.H. 1989, *A&A*, **215**, 33.
 Serabyn, E., Lacy, J.H., Townes, C.H. & Bharat, R. 1988, *ApJ*, **326**, 171.
 Simons, D.A., Clark, C.C., Smith, S., Kerr, J., Massey, S. & Maillard, J.-P. 1994, in *Instrumentation in Astronomy VIII*, SPIE, **2198**, 185.

The New Palomar Integral Field Spectrograph

T. W. Murphy, Jr., K. Matthews, and B. T. Soifer

*Palomar Observatory, California Institute of Technology, 320-47,
Pasadena, California 91125*

Abstract. A new near-infrared integral field spectrograph employing image slicer optics has been built and operated on the Palomar 200-inch Telescope. The all-cryogenic spectrograph design acts as a front end to either of two 256×256 infrared array cameras, covering the 1–2.5 micron or 1–5 micron wavelength ranges. Standard machining techniques, including diamond fly-cutting of the image slicer optical surfaces, are used in the production of the all-aluminum image slicer, allowing for an inexpensive and rapid construction project. The $5.4'' \times 9.6''$ field of view is split into eight slits, and feeds two independent spectrographs within the same dewar. Two resolution modes are available, as well as a broadband imaging mode for photometry and target acquisition. Spectral resolutions offered are $R \approx 1300$ and $R \approx 600$. The PIFS was first used on the 200-inch Telescope in 1998 September. The total system throughput, from the top of the atmosphere to photo-electron counts, is measured at $\sim 22\%$, enabling detection of a $K = 18$ continuum source at 5σ in one hour at the high resolution setting. Example data are presented showing velocity-resolved images of the $\text{Pa}\alpha$ emission line in an ultraluminous infrared galaxy.

1. Introduction

Large-format infrared detector arrays have enabled great advances in infrared astronomical instrumentation in the past decade. In particular, three-dimensional integral field spectrographs have become practical, permitting very efficient packing of data on the detector array. This technique was first exploited in the near-infrared by Weitzel et al. (1996), with an instrument called 3D. In this paper we describe a new instrument, the Palomar Integral Field Spectrograph (PIFS), a cryogenic, image slicing integral field spectrograph for the 200-inch Telescope.

The PIFS is designed to study gas kinematics in ultraluminous infrared galaxies, though it is also well suited for observations of high-redshift radio galaxies, circumnuclear environments of galaxies, and quasar host environments. A more detailed description of the PIFS is contained Murphy et al. (1999).

Figure 1. Schematic representation of the image slicing concept. The rectangular field is sliced into eight strips, or sub-slits, which are rearranged in an end-to-end configuration such that light from each slit may be dispersed without overlapping. The PIFS image slicer produces two effective long slits, for reasons described in the text.

2. The Image Slicer

The image slicer performs the task of reformatting a contiguous two-dimensional field into a one-dimensional pattern such that spectra may be produced without convolving spatial and spectral information. Figure 1 illustrates this concept.

An all-reflective image slicer is used in the PIFS rather than a fiber bundle to take advantage of the higher throughput of the reflective slicer, and to avoid the technical challenge of maintaining lenslet-to-fiber alignment in the cryogenic environment. The large plate scale and slowly diverging beam at the $f/70$ focus of the 200-inch Telescope facilitated the image slicer approach.

The double-slit design of the PIFS image slicer is adopted to double the effective field of view on the sky, at the cost of reduced spectral coverage. The PIFS design is constrained by the decision to use existing infrared cameras to record the spectra produced by the spectrograph. Matching the focal plane and pupil in the camera sets the pixel scale at $0''.167$ per pixel, leading to a $38''$ field on the array. With a slit width of $0''.67$ (4 pixels), the total area of a slit spanning the array is about 25 arcsecond^2 . Creating a second slit, as in Figure 1, doubles the field to 50 arcsecond^2 , with only half of the array available for spectral dispersion of each slit. A rectangular field pattern offers greater observing flexibility (provided the field may be rotated), and requires fewer sub-slits in the image slicer. As such, the PIFS field of view is $5''.4 \times 9''.6$, divided into eight slits.

The image slicer optical elements and mounting fixture are made entirely out of aluminum alloy. The optical elements have diamond fly-cut reflective surfaces coated with protected silver. The first set of image slicer optics are eight planar facets each measuring $1.17\text{mm} \times 17.8\text{mm}$ in size. These facets sit at the $f/70$ focal plane, directing the incident light into eight different directions.

Figure 2. Arrangement of the image slicer optical elements. The optical elements are aluminum plates, diamond fly-cut on the ends to form optically flat reflective surfaces. The stack of thin mirrors at the focal plane redirects the field onto eight collector mirrors, at left, which send light towards the spectrograph in the end-to-end pattern depicted in Figure 1.

Sitting roughly 100mm away from the focal plane are a second set of planar facets, these being $2.92\text{mm} \times 17.2\text{mm}$ in size. Upon reaching these mirrors, the eight slits have separated into two groups of end-to-end patterns, as depicted in Figure 1. Each of the second set of mirrors intercepts the light from one of the flat mirrors in the focal plane and sends it toward the spectrograph as if originating from a common pupil at the distance of the telescope secondary mirror. Figure 2 depicts the physical arrangement of the end-polished optical elements in the image slicer, and shows the optical paths through the system.

Producing an image slicer with eight separate optical paths yet preserving the $f/70$ pupil to a few percent of the pupil diameter requires angular alignment of each mirror to ~ 0.1 milliradian. The technique for fabricating the image slicer optics is such that random alignment errors are minimized. The image slicer optics are mounted onto a monolithic aluminum fixture, which helps to control random alignment error and provides a homogeneous material environment such that thermal contraction is the same for both fixture and optics. Stainless steel pins align the optical elements on the fixture, and screws with Belleville washers clamp the elements onto the fixture.

Each slicer optical element is a thin plate roughly $17\text{mm} \times 60\text{mm}$ in size, with a flat optical surface on one end. The focal plane elements are 1.17mm thick, and the second set of mirrors are 2.92mm thick. Each plate has a common set of two pinholes establishing the final position and alignment of the optic on the mounting fixture. One of these holes is centered on each plate near the front—or reflecting—end of the optic. The second pinhole is at the back end, in one corner of each plate. A third pinhole at the back end of the plate, in combination with the front, or pivot, pinhole establishes an alternate configuration such that the

Figure 3. The image slicer optics are diamond fly-cut simultaneously in the stack configuration illustrated at left. The final configuration, with a variety of reflector orientations, is shown at right. Precision pinholes used in the fly-cut and final configurations completely determine the relative angles of the reflective surfaces. See text for details.

line between these two pinholes is perpendicular to the optical surface. When all the plates are stacked together with a pair of pins running through this alternate set of pinholes, a single, coplanar surface is presented. A flat surface is diamond fly-cut on the end of this stack to simultaneously produce all of the optical surfaces. The final configuration, established by the pivot pinhole and the common corner pinhole, will then fan the reflecting surfaces out in a variety of angles. The relative angle between surfaces is solely determined by the angle between the three pinholes in any given plate. Thus controlling alignment to ~ 0.1 milliradian with pinholes separated by ~ 50 mm requires a pinhole positional accuracy of $5\mu\text{m}$, which is within the reach of an accurate milling machine. Figure 3 illustrates the principle involved. All of the image slicer optics are produced in this way, though the final arrangement of the second set of reflectors is side-by-side rather than in a stack, as with the focal plane elements.

The figure accuracy of the thin focal plane optical elements is measured to be flat to $\sim 200\text{nm}$, peak-to-valley, while the thicker, more robust second set of elements has a peak-to-valley figure accuracy of $\sim 60\text{nm}$.

3. The Spectrograph

The PIFS dewar contains the image slicer, gratings and various reflective optics. The sliced and dispersed output is formatted to match the focal plane of an existing infrared camera, which is externally coupled to the PIFS dewar. The primary design camera contains a 256×256 HgCdTe array, which operates from $1-2.5\mu\text{m}$. Also available for use with the PIFS is a 256×256 InSb array camera, extending the PIFS wavelength coverage out to $5\mu\text{m}$. The PIFS optical layout is shown in Figure 4.

Figure 4. The PIFS dewar and optical path. The infrared camera is a separate dewar attached to the PIFS dewar at left.

Within the PIFS dewar is a pair of identical spectrographs, one for each of the two slits produced by the image slicer. The two spectrographs in effect separate the spectra of the two slits. Following the optical path in Figure 4, the $f/70$ beam from the telescope deflects off of the tertiary mirror into the PIFS dewar through an anti-reflection coated sapphire window. After another fold, the beam comes to focus on the thin slices of the image slicer, and then is split into eight beams arranged in two effective slits. The light from each slit then encounters a spherical collimator, the grating, a spherical camera mirror with the same focal length as the collimator, then reflects off a fold mirror and through the CaF_2 exit window. An $f/70$ focal plane is produced at the focus of the external infrared camera. Liquid nitrogen in the upper canister of the PIFS dewar cools the entire PIFS optical assembly to temperatures near 80°K .

The replica grating is tunable by rotating the aluminum substrate upon which it is affixed. The grating substrate has two grating surfaces replicated onto it, providing two different spectral resolutions. The primary grating surface delivers a resolution of $R \approx 1300$, while the alternate grating provides a lower resolution of $R \approx 600$. The spectral coverage in each of these modes is about 25 resolution elements, or $\sim 5000 \text{ km s}^{-1}$, and $\sim 12,000 \text{ km s}^{-1}$, respectively. Order sorting is accomplished via broad-band filters, and a 5% circular variable filter (CVF) in the $1.5\text{--}3.0\mu\text{m}$ range. A flip-in mirror may be rotated into the beam in front of the grating, which results in an image of the slit, and thus of the PIFS field. The imaging mode is quite useful for target acquisition and broad-band photometric calibration.

Figure 5. PIFS data of the double-nucleus ultraluminous infrared galaxy, IRAS 1019+13, shown here in pure continuum and continuum-free Pa α . Note the isolated clump of line emission in the red Pa α image, with no continuum counterpart. Note also the very diffuse velocity-differentiated gas surrounding the source at right. Crosses mark the continuum peak positions. The velocity-binned images are each 350 km s^{-1} apart and 350 km s^{-1} wide.

Total throughput of the PIFS dewar is 55–60%, with the image slicer itself contributing only a $\sim 5\%$ light loss. Overall, 22% of the photons incident upon the atmosphere are detected with the PIFS in high resolution spectral mode.

4. Example Data

Figure 5 shows PIFS data of the ultraluminous infrared galaxy, IRAS 1019+13. The images shown are collapsed out of a single datacube, representing 40 minutes of on-source integration time. The 350 km s^{-1} velocity bins in Figure 5 are coarser than the 230 km s^{-1} resolution of the PIFS. The comparison of the continuum image with the images of Pa α in different velocity channels shows substantial differences. This rapidly changing structure with wavelength is an illustration of the importance of integral field spectroscopy in unraveling the complex kinematics of gas in the nuclear regions of ultraluminous infrared galaxies.

References

- Murphy, T. W., Matthews, K., and Soifer, B. T. 1999, PASP, to appear September 1999
- Weitzel, L., Krabbe, A., Kroker, H., Thatte, N., Tacconi-Garman, L. E., Cameron, M., and Genzel, R. 1996, A&AS, 119, 531

Near-IR Integral Field Spectroscopy with Adaptive Optics

N. Thatte, S. Anders, F. Eisenhauer, M. Tecza, S. Mengel, A. Eckart,
and R. Genzel

*Max-Planck-Institut für extraterrestrische Physik, Giessenbachstraße,
D-85748 Garching, Germany*

G. Monnet, and D. Bonaccini

*European Southern Observatory, Karl-Schwarzschild-straße 2, D-85748
Garching, Germany*

Abstract. Integral field spectroscopy, in conjunction with adaptive optics systems, has the unique potential of providing spectra at spatial resolutions close to the diffraction limit of the telescope. We present first results from integral field spectroscopy in conjunction with an adaptive optics system, achieving diffraction limited images and spectra at the Calar Alto 3.5 meter telescope. SINFONI, an adaptive optics assisted integral field spectrometer currently being built for the VLT, is one of the few adaptive optics optimized integral field instruments. We present a brief summary of its features, together with a brief description of the image slicer employed in SINFONI.

1. Introduction

Adaptive optics systems hold the promise of delivering diffraction limited images, allowing almost an order of magnitude gain in spatial resolution over that delivered by seeing. However, adaptive optics systems have mostly been used in conjunction with imaging devices (Hofmann et al. 1995, Rigaut et al. 1998, etc), although narrow band Fabry-Perot imagers have been employed with some systems (Chalabaev et al. 1999). Spectroscopy at resolutions close to the diffraction limit of the telescope holds immense promise of yielding fresh insights for a large range of astronomical problems. However, the difficulty of coupling spectrographs with adaptive optics systems has precluded their widespread use so far.

2. Spectroscopy at the diffraction limit

Three major difficulties stand in the way of adapting existing long slit spectrographs for use with adaptive optics systems.

- If the slit size employed by a long slit spectrograph is comparable to the diffraction limit of the telescope, then slit diffraction plays a dominant role in determining the point spread function (PSF) of the spectrograph.

Diffraction at the slit, located in a focal plane, results in a diffuse elliptical image of the telescope pupil. Any stop placed at this pupil image must either exclude a significant fraction of the light passing through the slit, or let in excess background emission (thermal or night sky) which would reduce the sensitivity of the instrument significantly.

- The light entering the spectrograph suffers diffraction at the slit, resulting in a smaller numerical value for the focal ratio than determined by geometrical optics. A spectrograph designed for large slit widths (when slit diffraction is unimportant) would suffer additional light loss due to vignetting by spectrograph optics and baffles.
- Adaptive optics systems typically deliver a PSF which consists of a diffraction limited core surrounded by a seeing limited halo. The relative amount of energy contained in the two components depends on a large number of factors, which include atmospheric seeing, guide star magnitude, angular separation of guide star and object, number of modes corrected, frequency of operation etc.. It is a common occurrence that the PSF varies on time scales of a few minutes, mostly due to variations in seeing and mechanical flexure of telescope components. Spectroscopy of extended objects needs information about the PSF, so that the various spatial contributions to the observed spectra can be identified and accurately estimated, even if no deconvolution is done. Since a long slit spectrograph only samples the PSF along one spatial direction (along the slit), it is impossible to recover sufficient information about the PSF to allow such analysis.

The first of these difficulties can be surmounted by including a pupil stop before the spectrograph slit. The stop (and succeeding optics) would need to be cooled if the instrument was cryogenic. The second problem may be overcome by over-designing the spectrograph optics, so as to include almost all the light diffracted by the slit, minimizing light loss. The third problem is more fundamental and may be overcome only by using a spectrograph which can recover two dimensional information of the PSF *during* the observation, namely, an integral field spectrometer. A Fabry-Perot spectroscopic imager may also be used, but the observer must then rely on the adaptive optics system to deliver a stable PSF during an entire scan covering the wavelength range of interest, which is rarely true in practice.

The OASIS spectrograph built for use with the PUEO adaptive optics bonnette of the CFHT (Bacon et al. 1999) is the first integral field spectrograph to be used in conjunction with an adaptive optics system. However, OASIS works at visible wavelengths, whereas adaptive optics systems deliver best performance in the near infrared. To exploit this unique area, we have mounted the MPE 3D near infrared integral field spectrometer at the focus of the ALFA adaptive optics system of the Calar Alto 3.5 meter telescope.

3. The AIM module

The MPE 3D integral field spectrometer (Weitzel et al. 1996) provides *simultaneous* spectra of 256 spatial pixels, arranged in a 16×16 square field of view on

the sky. Spectral resolutions of 1000 or 2000 are available, covering the H and K atmospheric windows. For further details of the 3D spectrograph, please refer to Weitzel et al. 1996. The ALFA laser star adaptive optics system (Hippler et al. 1998, Davies et al. 1998), is a collaborative effort of MPE and MPIA, operational since 1997 at the Cassegrain focus of the Calar Alto 3.5 meter telescope. ALFA's adaptive optics system is based on a Shack-Hartmann type wavefront sensor driving a deformable mirror with 97 actuators. Operating frequencies of upto 1200 Hz are possible, yielding closed loop bandwidths of ~ 120 Hz. A separate tip-tilt sensor allows operation with a laser guide star. The laser itself is a 3 Watt continuous wave laser operating at the frequency of the Na D line, producing an artificial star of $\sim 11^{\text{th}}$ magnitude at a height of about 90 km.

The MPE 3D integral field spectrometer was built to provide $0''.5$ pixels for the nominal f/45 focus of the Calar Alto telescope. To adapt the pixel scales of 3D for operation in combination with the ALFA adaptive optics system, the AIM (Aperture Interchange Module) interface has been designed (Anders et al. 1998). AIM performs four important functions.

- The pixel scale needed to correctly sample the diffraction limited images produced by ALFA is $0''.07$. However 3D's field of view (FOV) is always 16 times the pixel size in each linear dimension, so that choosing $0''.07$ per pixel yields a FOV of only $1''.1 \times 1''.1$. AIM allows on-the-fly scale changes between $0''.07$ and $0''.25$ per pixel, the latter corresponding to a $4'' \times 4''$ FOV.
- The AIM scale change optics include a small dewar with a Peltier cooled cold stop, which is inserted in the beam before the image slicer when the $0''.07$ pixel scale is chosen, thus avoiding any light loss or excess background due to slit diffraction effects.
- The small FOV implies that no pixels cover blank sky. In order to provide correct background subtraction, especially at near-IR wavelengths where the OH sky emission varies on time scales of a few minutes, AIM incorporates an ON-OFF mirror mechanism which allows 3D to observe two offset pointings within the ALFA FOV without moving the telescope. Using the ON-OFF feature, the time between successive 'nods' of the telescope can be increased from 2 minutes to 20 minutes, thus dramatically reducing the overhead needed to unlock and lock the adaptive optics system each time the telescope is pointed at blank sky. The 'chops' (on-offs) and 'nods' can be combined to yield a data set completely free of background flux from telescope and sky.
- AIM provides a set of pupil imaging optics which re-images the pupil of the adaptive optics system on the spectrograph image slicer. A reconstructed image of the 3D field may then be used to align the spectrograph pupil with that of the telescope and adaptive optics system, ensuring maximum throughput to the spectrograph.

Details of the construction of AIM can be found in Anders et al. 1998.

4. First results

We present first results from the commissioning run in August 1998, which used MPE 3D together with the ALFA adaptive optics system for the first time. Figure 1 shows a reconstructed image of the close binary HEI 7, which was observed with 3D + ALFA at a spectroscopic resolution of 2000 over the wavelength range 2.15 to 2.45 μm . The orbit of HEI 7 is well measured, with a separation between the primary and secondary components at the epoch of observations of $0''.26$ (Hershey 1975). The primary component has a visual magnitude $m_V = 7.1$, while the secondary component has $m_V = 10.0$. The spectral type of the primary is known to be K0 V, that of the secondary could only be estimated from colours to be M0 V. The reconstructed image of the PSF reference HD 183051, also $m_V = 7.1$, taken a few minutes prior to the observations of HEI 7, is shown on the left hand side of figure 1.

Figure 1. Reconstructed K band images of the binary star HEI 7 (central panel), the PSF reference star for HEI 7 (left panel) and the deconvolved data cube of HEI 7 (right panel). The images were made by collapsing the data cube in the spectral dimension. The PSF reference clearly shows a diffraction limited core ($0''.14$ FWHM), as well as part of the first Airy ring. The imperfect PSF of the adaptive optics system causes residual aberrations which distort the Airy pattern. The deconvolved image in the right hand panel shows the two components of the binary. The data cube also contains spectral information about each spatial pixel.

The image of the PSF reference clearly shows the need for two dimensional information of the PSF. Part of the first Airy ring around the diffraction limited core can be made out, but low order aberrations which cause flux to spill over into the first Airy ring can be easily seen. The PSF of the primary causes a $\sim 8.6\%$ contamination to the flux of the secondary.

The spectra extracted from a few pixels centered on the core of the primary and secondary components are presented in figure 2. The bottom part of the figure shows two spectra of template stars, taken from a library of near-IR stellar spectra, which have the same spectral types as those of the primary and secondary, the latter estimated from colours. As expected, the ^{12}CO bandheads (at 2.29, 2.32, 2.35 and 2.38 μm) are almost twice as deep in the secondary

compared to those in the primary, as are the Na ($2.21\ \mu\text{m}$) and Ca ($2.26\ \mu\text{m}$) absorption features. In contrast, the Mg feature (at $2.27\ \mu\text{m}$) has the same depth in both the observed spectra as well as the templates. It should be noted the the signal-to-noise ratio of the spectra decreases dramatically toward the long wavelength edge of the K band, where thermal background emission from warm parts of the telescope and optics increases sharply.

The right hand panel of figure 1 shows a deconvolved reconstructed image of the binary star, created by deconvolving each plane of the data cube using the reconstructed image of the PSF reference star as the reference image. The deconvolution, though good, is not perfect, showing that even within a time span of a few minutes, and telescope inclination variations of a few degrees, the PSF delivered by the adaptive optics system varies substantially. In the case of the ALFA system, this dramatic variation in PSF is most probably caused by changes in flexure. In any case, these variations underpin the need to sample the PSF (in two dimensions) relatively frequently.

5. Future prospects

Several integral field instruments are planned for the 8 meter class telescopes which have recently been or will soon be put into operation. However, only a fraction of these integral field spectrographs plan to operate at near-IR wavelengths, and of those, only a fraction will be optimized for use with an adaptive optics system. In the following section, we report on SINFONI, the adaptive optics assisted integral field spectrometer for the ESO VLT.

5.1. SINFONI

SINFONI (Thatte et al 1998) is conceived from the start in the framework of a scientific collaboration between the Max-Planck-Institut für extraterrestrische Physik (MPE) and the European Southern Observatory (ESO) covering the field of high spatial resolution studies of compact objects (e.g. star forming regions, centers of galaxies, cosmologically distant galaxies). SINFONI is a combination of two instruments, a near-infrared integral field spectrometer, SPIFFI II (Tecza et al. 1998), and a curvature sensor based adaptive optics system MACAO (Bonaccini et al. 1998).

The near infrared integral field spectrometer, SPIFFI II, which is the responsibility of the MPE group, obtains 1024 simultaneous spectra covering a 32×32 contiguous square field of view, with a filling factor close to 100%. The pixel size may be chosen to be $0''.25$ per pixel for seeing limited observations, or $0''.03$ per pixel for diffraction limited observations, or $0''.1$ for intermediate observing conditions. The corresponding fields of view are $8'' \times 8''$, $1'' \times 1''$, and $3''.2 \times 3''.2$ respectively. Four different gratings are available, providing spectral resolutions of ~ 4500 covering one of the J, H or K bands, or both H and K bands simultaneously at a spectral resolution of 2000. The high spectral resolution settings would allow an *OH avoidance* technique to be implemented. Since 98% of the night sky emission flux in the wavelength range from 1.0 to $2.2\ \mu\text{m}$ is contained in OH lines, a substantial reduction in sky background flux may be achieved by simply blanking out those pixels in the spectrum which are contaminated by OH line emission.

Figure 2. Spectra of the two components of HEI 7 are shown in the top panel for the long wavelength half of the K band. Two spectra of template stars from a stellar library are illustrated in the bottom panel for comparison. The spectral types of the template stars are chosen to match those of HEI 7. The spectrum of HEI 7 suffers from enhanced noise longward of 2.3 μm , due to thermal background from telescope and warm optics. The CO bandhead at 2.29 μm is twice as deep in the secondary as in the primary, while the Mg feature at 2.27 μm has the same depth in the two stars, as expected.

In addition, SPIFFI II incorporates a *Himmelspinne*, which allows light from one of four pre-selected offset positions to be redirected to the image slicer, and then onto the detector. This precludes the need to nod the telescope every couple of minutes in order to measure the variable sky background accurately. Combined with the OH avoidance capability, this would allow integrations of roughly 1 hour duration. An imaging channel with $25'' \times 25''$ FOV is available, in order to ease acquisition of faint targets. A pupil imaging system is also included, so as to ensure accurate alignment of the telescope and adaptive optics system pupils with that of the spectrometer.

The curvature sensor based adaptive optics system MACAO (Bonaccini et al. 1998), which is the responsibility of ESO, incorporates a deformable mirror with 67 actuators. It uses photon counting Avalanche Photodiodes (APDs) for wavefront sensing, providing extremely high sensitivity. The system is specifically optimized to provide low order correction using very faint guide stars. Its performance is well adapted to spectroscopic applications, where energy concentration is more important than achieving diffraction limited image quality. MACAO can use a guide star within a field of $2'$ diameter centered on the scientific target. Its novel design implies only three warm reflections (apart from the primary and secondary) before the light enters the cryogenic spectrometer. The dewar window itself serves as the dichroic, reflecting the visible light to the wavefront sensor. The extra thermal background added by the relay optics is thus kept to an absolute minimum, dramatically improving the sensitivity of SINFONI in the K band, where thermal emission from warm telescope optics and surroundings can dominate the background flux seen by the spectrometer. MACAO can also be used with a laser guide star, which is foreseen to be a standard facility for the VLT ANTU (formerly UT1) telescope.

5.2. The SINFONI Image Slicer

The heart of SPIFFI II is the image slicer, which rearranges the light from a two dimensional field on the sky to a (pseudo) long slit feeding the spectrograph collimator. Figure 3 shows the conceptual design of the SPIFFI II image slicer, which is based on the image slicer employed in the MPE 3D spectrograph. The slicer consists of two stacks of plane mirrors, dubbed the *small slicer* and *big slicer*. The stack of small slicer mirrors resembles a spiral staircase. It is composed of gold coated mirrors, each 9.6 mm long, but only 0.3 mm thick. Each mirror has a different tilt to the incoming light path, and redirects the light to a particular mirror segment of the big slicer.

The big slicer consists of 32 segments, one for each of the small slicer mirrors, whose midpoints are located on a parabola. The small slicer sits at the focus of the parabola. In addition, the big slicer mirrors are arranged in two layers, forming a brick-wall pattern, with 16 segments in each layer. Each big slicer mirror redirects the light from the corresponding small slicer mirror so that it is parallel to the incoming beam. The image slicer thus preserves the pupil position, which is set to infinity for the entrance pupil. The brick-wall arrangement of big slicer mirrors allows each mirror to be almost twice as long as the nominal slitlet length it re-images, thus avoiding any light loss due to defocus effects:

The total re-imaged slit length is 307.2 mm, and is composed of 32 slitlets, each 9.6 mm long. Each slitlet is in turn imaged onto 32 detector pixels, thus

Figure 3. Conceptual design of the SPIFFI II image slicer. The incoming beam (with entrance pupil at infinity) is incident on a stack of plane mirrors (small slicer) which redirect the light in different directions depending on its height in the image plane. The second stack of plane mirrors (big slicer) then redirects the light from each slice (or slitlet) so that it is parallel to the incoming beam (pupil preserving). The big slicer mirrors form a two level brick-wall pattern. Their centers are located on a parabola, whose directrix is the location of the virtual slit.

providing 1024 spectra covering a 32×32 square field of view. The re-imaged slit is virtual, and is located behind the big slicer mirrors, along the directrix of the parabola which defines the positions of the big slicer mirrors. The computed throughput of the SPIFFI II slicer has an average value of 89%, degrading slowly from the center to the edge of the field along each slitlet. The entire slicer is cooled to the instrument's operating temperature of 77 K. The individual slices of the small slicer are optically contacted to each other, thus avoiding any glue (which could lead to misalignments and non-reflecting zones) and extremely high image quality over the whole field.

5.3. Alternative Slicer Concept

We are also pursuing an alternative concept for the SPIFFI II image slicer, using OH free silica-silica fibers to do the image slicing. These fibers show adequate transmission ($\geq 90\%$) out to the long wavelength edge of the K band ($2.5 \mu\text{m}$), if short lengths ($\sim 15 \text{ cm}$) are used. They are much more robust than "infrared fibers" (e.g. ZrF) and standard handling and polishing techniques may be used. Optical fibers have been used for image slicing for a long time at visible wavelengths (Arribas et al. 1998). There are typically combined with an array of lenslets so as to provide good filling factor in the image plane (Kenworthy, Parry and Taylor 1998). The lenslets form tiny images of the telescope pupil matched to the diameter of the fiber core. However, the difficulty of maintaining the alignment between lenslets and fibers at cryogenic temperatures has precluded their use in infrared instruments.

The *flared fiber* concept (Tecza and Thatte 1998) circumvents the cryogenic alignment problem by combining the microlens and the fiber into a single monolithic unit. The fiber end is heated and "flared", in a process akin to the reverse of drawing the fiber. The flared tip is then ground, so as to remove the cladding material, leaving bare core material in the shape of a "taper". The front surface of this taper is then polished to form a lenslet with the required radius of curvature, and the sides are ground to form a regular hexagon. One such flared fiber is shown in the left hand panel of figure 4. These individual units can then be arranged in a close packed geometry to form the image slicer, as illustrated in the right hand panel of figure 4. Since each fiber and its microlens are part of the same monolithic unit, relative alignment is maintained even when the whole unit is cooled to cryogenic temperatures.

The other ends of the fibers are all arranged to lie along a straight line, forming the entrance slit of the cryogenic spectrograph. This concept has the advantage that it is easily expandable to a large number of image points (fibers). Small bundles may also be used to form *deployable IFUs*, which could be placed by a cryogenic robot within a large field of view. Our preliminary tests of this concept have been very encouraging, yielding fibers with $>80\%$ coupling efficiency. We will pursue this concept further, for use in future instruments.

We would like to acknowledge the tremendous effort of the ALFA + 3D team, in particular, R. Davies, W. Hackenberg, S. Hippler, M. Kasper, T. Ott, S. Rabien and R. Sosa-Brito during the ALFA + 3D run in July-August 1998. We also thank the MPE staff for their help in building instruments and interfaces. We are grateful to the Calar Alto staff for their help at the telescope.

Figure 4. Photograph of a single flared fiber is shown in the left hand panel. The hexagonal lenslet has a size of $600\ \mu\text{m}$. The unflared part of the fiber is held in a metal ferrule to protect it during the polishing process. The right hand panel illustrates the concept of an image slicer built using flared fibers. Each lenslet forms an image of the telescope pupil on to the fiber core. The numerical aperture of the fiber is chosen to match the focal ratio of light from the lenslet.

References

- Anders, S., Maiolino, R., Thatte, N., and Genzel, R. 1998, *Proc. SPIE*, 3354, 222
- Arribas, S. et al. 1998, *Proc. SPIE*, 3355, 821
- Bacon, R. et al. 1999, in press
- Bonaccini, D., Rigaut, F., Dudziak, G., and Monnet, G. 1998, *Proc. SPIE*, 3353, 553
- Chalabaev, A., le Coarer, E., Rabou, P., Magnart, Y., Petmezakis, P., and Le Mignant, D. 1999, *Proc. of ESO/OSA meeting "Astronomy with Adaptive Optics"*, Sonthofen, Germany, in press
- Davies, R. et al. 1998, *Proc. SPIE*, 3353, 116
- Hershey, J. 1975, *AJ*, 80, 662
- Hippler, S. et al. 1998, *Proc. SPIE*, 3353, 44
- Hofmann, R., Brandl, B., Eckart, A., Eisenhauer, F., and Tacconi-Garman, L. E. 1995, *Proc. SPIE*, 2475, 192
- Kenworthy, M., Parry, I., and Taylor, K. 1998, *Proc. SPIE*, 3355, 926
- Rigaut, F. et al. 1998, *PASP*, 110, 152
- Tecza, M. and Thatte, N. 1998, in "Fiber Optics in Astronomy III", *ASP Conference Series*, 152, 271
- Tecza, M., Thatte, N., Krabbe, A., and Tacconi-Garman, L. E. 1998, *Proc. SPIE*, 3354, 394
- Thatte, N. et al. 1998, *Proc. SPIE*, 3353, 704
- Weitzel, L., Krabbe, A., Kroker, H., Thatte, N., Tacconi-Garman, L. E., Cameron, M., and Genzel, R. 1996, *A&AS*, 116, 531

NGC 4388: Imaging and Kinematics of the Nuclear Region through Integral Field Spectroscopy

S. Ciroi¹, V. Afanasiev², S. Dodonov², M. Radovich³, P. Rafanelli^{1,3}, G. Richter¹, S. Temporin⁴

¹*Astrophysikalisches Institut Potsdam, An der Sternwarte 16, D-14482
Potsdam, Germany*

²*Special Astrophysical Observatory, Nizhnij Arkhyz, Zelenchukskaya,
Karachaevo-Cherkesia, Russia, 357147*

³*Dipartimento di Astronomia, Università di Padova, Vicolo
dell'Osservatorio 5, I-35122 Padova, Italy*

⁴*Institut für Astronomie Innsbruck, Technikerstr. 25, A-6020
Innsbruck, Austria*

1. Introduction

The Integral Field Spectroscopy (IFS hereafter) is a powerful method for investigating the extended emission-line regions of the AGNs. In this poster we present an application of IFS to the circumnuclear and extranuclear regions of the Seyfert-2 galaxy NGC 4388. NGC 4388 (Fig. 1) is a highly inclined spiral galaxy ($i=75^\circ$), located in the core of the Virgo cluster. The presence of a bar with associated dust lanes obscures the nucleus of the galaxy (Phillips and Malin, 1982). Narrow band images have revealed the existence of excitation gas appearing distributed into two opposite cones (Pogge, 1988). Furthermore, the galaxy contains a double-peaked radio source close to the optical nucleus (Stone et al., 1988) aligned with the extended emission-line regions. Weak broad H α emission components (FWZI ≈ 4000 km s⁻¹), attributed to scattered light originating in a hidden Seyfert-1 broad line region, have been detected at several locations far from the nucleus (Shields and Filippenko, 1988).

2. The Observations

The observations were performed in March 1998 at the Special Astrophysical Observatory of the Russian Academy of Sciences with the 6-m telescope, equipped with the MPFS, the Multi Pupil Fiber Spectrograph, designed and developed by the Spectroscopy and Photometry Laboratory of the observatory. The MPFS is a combination of a 16 $\times$ 16 lenses array joined to a bundle of 256 optical fibers, which carry the signal to the spectrograph. A group of 8 fibers, with an offset of 4', is dedicated to the sky observation. As detector, the instrument is provided with a CCD TK 1024 $\times$ 1024 (pixel size = 24 μ).

The nuclear region of NGC 4388 has been observed in the spectral range 4400–7100Å, with a spectral resolution of about 10Å, combining two consecutive 600^s exposures. A binning mode applied during the observations, produced a spectral dispersion of 5.2Å/px. Each of the 256 fibers covers a region of 1'' of

size (0.16 kpc, assuming $H=75 \text{ km sec}^{-1} \text{ Mpc}^{-1}$), so that the total field covered by the array is $16'' \times 16''$ (Fig. 1). The typical seeing was about $4''$.

Figure 1. Broad band R image of NGC 4388 (left) and the position of the array over the galaxy (right)

3. Data Analysis and Results

The data have been reduced and analysed with the IRAF software. A special package, developed within our group in IRAF environment, allowed us to identify and extract each single spectrum, performing tracing with a good precision. The spectra were calibrated in wavelength, using a He-Ne-Ar lamp, with a mean accuracy of about $1\text{\AA}$. After cosmic rays removal and sky subtraction, they have been calibrated in flux, using the observations of the spectrophotometric standard star HZ 44. The $[\text{OIII}]\lambda 5007$ and $\text{H}\alpha$ emission line images of the galaxy have been reconstructed, after continuum subtraction and reddening correction of the measured fluxes (Fig.2).

The ionization map, resulting from the ratio between these images, revealed part of the South-West cone, already found by Pogge (1988). The cone has an axis at P.A. $\approx 200^\circ$ and an aperture angle of about 70° , less than what was indicated by Pogge, probably due to our smaller field of view. A map of the $E(B-V)$ extinction, calculated using the case B recombination value $\text{H}\alpha/\text{H}\beta=2.86$, shows values ranging from about 0.4 in the nuclear regions to less than 0.1. The $[\text{OIII}]\lambda 5007$ and $\text{H}\alpha$ heliocentric velocity fields, obtained by measuring the position of these emission lines in each spectrum, are in good agreement. The axis of the nodes appears to have a P.A. $\sim 55^\circ$, consistent with the value $60^\circ \pm 15^\circ$ found by Petitjean and Durret (1993).

Figure 2. The emission line images overlaid with their velocity fields (*top*), the ionization map (*bottom left*) and the extinction map (*bottom right*)

A more detailed analysis of the physical properties of each $1'' \times 1''$ region, using the Veilleux–Osterbrock diagnostic diagrams, points out the AGN-like characteristics of all spectra (Fig. 3). Actually a different ionization degree can be observed, distinguishing between spectra coming from the cone and spectra from the surrounding regions: the first ones have been modelled with CLOUDY, using a power law continuum $F_\nu = \nu^\alpha$ with spectral index $\alpha = -1.2$, electron density N_e ranging from 10^1 to 10^4 cm^{-3} and a sulphur abundance $S = 1.5S_\odot$; the others have been modelled again using a power law, but with $\alpha = -1.5$, consistent with the hypothesis that outside the ionization cone the nuclear radiation is absorbed, and applying $N = 0.6N_\odot$ and $S = 0.7S_\odot$.

Finally, even if our spectral resolution was not so high, we attempted to verify the detection by Shields and Filippenko (1988) of a weak broad $H\alpha$ emission line component at several distances from the nucleus. By fixing the $[\text{NII}]\lambda 6583/[\text{NII}]\lambda 6548$ ratio to 3 and applying the same FWHM for both lines, a multicomponent gaussian fitting of the $H\alpha + [\text{NII}]\lambda\lambda 6548, 6583$ blend was performed in randomly selected spectra: a broad $H\alpha$ component (FWHM $\approx 2000 \text{ km s}^{-1}$) could be detected in different regions of the ionization cone (Fig. 4).

Figure 3. Veilleux-Osterbrock diagnostic diagrams and photoionization models for spectra from the ionization cone (*top*) and from its surroundings (*bottom*); the ionization parameter U ranges from 10^{-4} to 10^{-2} and the electron density N_e from 10^1 (+) to 10^4 (Δ) cm^{-3}

Figure 4. Examples of the presence of a broad $H\alpha$ component inside the ionization cone (*top*) and of its absence outside the cone (*bottom*)

References

- Petitjean, P. & Durret, F. 1993, A&A, 277, 365
 Phillips, M.M. & Malin, D.F. 1982, MNRAS, 199, 905
 Pogge, R.W. 1988, ApJ, 332, 702
 Shields, J.C. & Filippenko, A.V. 1988, ApJ, 332, L55
 Stone, J.L. Jr., Wilson, A.S., Ward, M.J. 1988, ApJ, 330, 105

Integral Field Spectroscopy of Ultraluminous Infrared Galaxies

L. Colina¹, S. Arribas², K. Borne³

(1) *Instituto de Física de Cantabria, CSIC-UC, 39005 Santander, Spain*

(2) *Instituto de Astrofísica de Canarias, CSIC-IAC, 38200 La Laguna, Tenerife, Spain*

(3) *Raytheon Information Technology and Scientific Services, NASA
Goddard Space Flight Center, Astrophysics Data Facility, Greenbelt, MD
20771, USA*

Abstract.

We have initiated a program with the aim of studying the stellar/gas kinematics and ionization sources in ultraluminous infrared galaxies (ULIRGs). Integral field optical spectroscopy of a small sample of ULIRGs covering different stages of the merger sequence are being obtained with the INTEGRAL system (Arribas et al. 1998). INTEGRAL is a fiber based 3D spectrograph that generates about 200 spectra over a region which can vary from $8'' \times 7''$ to $34'' \times 30''$, at different spatial resolutions. Preliminary results for the ultraluminous infrared prototype Mrk 273 fully illustrate the capabilities and potential of 3D spectrographs such as INTEGRAL for the study of ULIRGs.

1. The INTEGRAL System

INTEGRAL is a common user instrument mounted in the Nasmyth No. 1 platform (GHRIL) of the 4.2 m William Herschel Telescope at La Palma Observatory. A cylindrical structure (Acquisition, Guiding, and Calibration box structure) is bolted to the Nasmyth mechanical rotator. As this structure rotates synchronously with the focal plane when the telescope tracks, the sky position angle is fixed during an observation.

The Swing Plate (SP) is located in the telescope focal plane area, and this is the connection point for the acquisition, and scientific fibre bundles. By rotating the SP any installed bundle can be placed at the centre of the focal plane, while the other bundles are light blocked. Hence, only one bundle at any one time can feed the spectrograph. For acquisition a coherent fibre bundle transmits the light from the focal plane (via the SP) to a standard TV camera.

Up to six scientific bundles can be simultaneously mounted in the SP, although in the standard configuration only three are used (called sb1, sb2, and sb3). These bundles are 5.5 m long, while their fibre core diameter (in sky units) is: $0.45''$ (sb1), $0.9''$ (sb2), and $2.7''$ (sb3), respectively. The bundles are simultaneously connected at the entrance (*pseudo-slit*) of the WYFFOS spectrograph.

Figure 1. The figure presents the roughly 200 calibrated $H\alpha - [SII]$ spectra region located at their corresponding position in astronomical coordinates relative to the central fiber (110). The numbers on the upper right corner of each spectrum indicate the number of the fiber. These spectra are used to generate the $H\alpha$ emission line distribution and its velocity field.

They can be interchanged very easily, with an overhead of a few seconds. Hence, depending on the prevailing seeing conditions the instrument can be easily optimized for the scientific program.

Guide stars may be acquired by moving an XY slide containing two coherent fibre bundles. This unit moves in an off-axis region and the bundles feed a TV based autoguider. A liquid fiber is used to feed a calibration source (currently Copper-Argon, Copper-Neon, or Tungsten lamps) to the focal plane.

A detailed description of the general layout of INTEGRAL and its various operational modes can be found in Arribas et al. (1998)

2. 3D INTEGRAL Spectroscopy of Mrk 273

2.1. Relevant Observing Parameters

The ultraluminous infrared galaxy Mrk 273 was observed using a total of 219 fibers of $0.9''$ in diameter (i.e. fiber bundle sb2) and a field-of-view of $16.5'' \times 12.3''$. The spectra covered the $5000-7900 \text{ \AA}$ range with an effective resolution of $4.4 \text{ \AA}$. The seeing was $1.0''$ (FWHM) during the observations. A full account of the reduction and calibration process will be presented elsewhere.

2.2. Preliminary Results

Mrk 273 is one of the closest (redshift of 0.037) prototypes of ULIRGs characterized by a long (about 40 kpc) tidal tail, a Seyfert 2 type optical spectrum (Sanders et al. 1988), and possibly the presence of circumnuclear starbursts (Goldader et al. 1995).

The roughly 200 calibrated spectra obtained with INTEGRAL for the central region of Mrk 273 (Figure 1) are used to generate maps of the warm ionized gas distribution, its velocity field, and the true (i.e. free of emission lines) continuum (Figure 2).

The stellar light distribution (traced by the continuum images) and the warm emitting line gas (traced by the $H\beta$ and $H\alpha$ images) show a rather different morphology. The continuum images show a single peak dominating the visual (4900Å) and red (7600Å) light distribution, and most likely associated with the true nucleus of the galaxy. The stellar component of Mrk 273 is elongated along the direction of the known long tidal tail.

The emitting line gas has two well differentiated emission peaks separated by about 4" (i.e. ~ 3 kpc at a distance of 148 Mpc). The dominant emission peak (northeast) coincides (within the 1" resolution) with the nucleus of the galaxy but the second peak (southwest) has no obvious continuum counterpart. This secondary emission peak, which is the dominant [OIII] emission line peak, has the optical emission line ratios of a high excitation Seyfert 2 while the northeast peak is classified as a low excitation LINER (Colina, Arribas & Borne, in preparation).

The $H\beta$ and $H\alpha$ velocity fields (Figure 2; central panels) show the complexity of the gas motions in Mrk 273. While the gas associated with the secondary, southwestern, peak has an almost constant velocity, a tongue of high velocity gas (velocity difference of ~ 300 km s $^{-1}$) is detected in the region between the two dominant emitting peaks. Also, regions of compact (i.e. smaller than 1" or 0.7 kpc) high velocity (up to 500 km s $^{-1}$) gas components (shown as white spots in the velocity panels) are detected at various locations within the field of view. A complete analysis of all these results and their implications will be presented elsewhere.

References

- Arribas, S., Carter, D., Cavaller, L., del Burgo, C., Edwards, R., Fuentes, J., García, A., Gentles, B., Herreros, J. M., Jones, L., Mediavilla, E., Pi, M., Pollacco, D., Rees, P., & Sosa, N. 1998 SPIE, 3355, 821.
- Goldader, J.D., Joseph, R.D., Doyon, R., Sanders, D.B., 1995 ApJ, 444, 97
- Sanders, D., Soifer, B.T., Elias, J.H., Madore, B.F., Matthews, K., Neugebauer, G., & Scoville, N.Z. 1988 ApJ, 325, 74.

Figure 2. Figure showing the $H\beta$ and $H\alpha$ emitting gas distribution (bottom left and right panels), its velocity field (central panels; velocity range of 500 km s^{-1} is represented), and the stellar light distribution at $4900\text{\AA}$ and $7600\text{\AA}$ (top left and right panels). Each box is $17'' \times 17''$ and position (0,0) corresponds to the central fiber (110) of INTEGRAL. North is left and East down.

Imaging Spectropolarimetry: The Nature of Hyperluminous Infrared Galaxies

Dean C. Hines

Steward Observatory, The University of Arizona, Tucson, AZ 85721

Abstract. In recent years the realization has grown that our best, and in some cases sole, optical view of many active galactic nuclei is offered by light scattered off surrounding material (see Antonucci 1993 for review). This spectral component becomes strongly linearly polarized when scattered through large angles, permitting the nuclear spectrum to be separated from the surrounding (unpolarized) emission using modern spectropolarimetry techniques. Imaging polarimetry can resolve the spatial distribution of this scattered component, thus enabling the construction of detailed models of the central structure of these objects. By incorporating polarimetry information, we add a fourth dimension to 3D imaging studies. I review recent imaging spectropolarimetry results which show conclusively that many Hyperluminous Infrared Galaxies harbor energetic QSOs surrounded by an optically and geometrically thick dusty torus. Viewed from the vantage point of the scattering material, these galaxies would be indistinguishable from typical optically selected QSOs.

1. Hyperluminous Infrared Galaxies

Hyperluminous Infrared Galaxies (HIGs) are characterized by “warm” far-infrared colors [$0.25 \leq F_\nu(25\mu\text{m})/F_\nu(60\mu\text{m}) < 3$], indicating thermal emission from dust heated by a QSO-like UV continuum source (Low et al. 1988). Formally, $\log(L_{\text{IR}}/L_\odot) \geq 13^1$ (Cutri et al. 1994), but there are clearly objects with very similar properties which have $11.5 \leq \log(L_{\text{IR}}/L_\odot) \leq 13$ (e.g. Young et al. 1996; Wills & Hines 1997). The UV/optical rest-frame spectra are dominated by high excitation narrow emission lines and galaxy starlight. Optical and near-infrared images from WFPC2 and NICMOS often show disturbed morphology and associated companion galaxies suggesting strong interactions (Surace et al. 1998; Hines et al. 1999a). Like “warm” Ultraluminous Infrared Galaxies (Scoville et al. 1999), central unresolved sources are revealed at longer wavelengths.

2. Polarimetry of HIGs

HIGs typically show high, non-time-variable polarization ($p \approx 5\% - 20\%$) caused by scattering (Hines & Wills 1993; Hines et al. 1995; Wills & Hines 1997).

¹ $H_0 = 75 \text{ km s}^{-1} \text{ Mpc}^{-1}$ and $q_0 = 0.1$.

Extended ionized and polarized “bicones” are seen in at least two HIGs (Hines et al. 1999b,c; Armus et al. 1999; Fig.1). This morphology is similar to highly polarized bipolar reflection nebulae such as CRL 2688 (Cohen & Schmidt 1982; Sahai et al. 1998), and implies that a bright nucleus is surrounded by a dusty torus which blocks our direct view, but allows light to escape at the poles to ionize gas and be scattered (thus polarized) into our line of sight.

The spectrum of scattered light in the bicones shows a power-law UV/optical continuum and broad permitted emission lines, both typical of luminous QSOs (Figure 1). However, the narrow emission lines in the ionization bicones are nearly unpolarized, confirming that the scattering material is mixed within the narrow line gas. It is evident that the scattering material and ionized gas have a direct view of a powerful QSO nucleus, and that the observed narrow-line equivalent widths are overemphasized. By counting the number of ionizing photons necessary to produce the observed narrow-line flux, we find that the intrinsic UV/optical luminosities are 10 – 50 times larger than observed directly. Thus, the QSO nuclei are sufficient to power a large fraction of the bolometric luminosities without appealing to enhanced star formation. Therefore, viewed along the symmetry axis (i.e. the vantage point of the scattering material), HIGs would be indistinguishable from typical QSOs, and would have been easily identified in optical surveys such as the Palomar-Green sample.

These results imply an intimate connection between QSO activity and the HIG phenomenon. Furthermore, large numbers of QSOs may have been missed by optical selection, because these surveys are biased against obscured objects (e.g. Wills & Hines 1997). The existence of such buried luminous objects implies that QSOs may contribute significantly to the energy budget of the universe, and in particular to the Infrared and X-ray Cosmic Backgrounds.

3. The Fourth Dimension

As the specific example above illustrates, future combined imaging spectropolarimetry, especially using large ground-based telescopes, will provide a powerful tool for investigating the geometry and structure of AGN in more detail. By capturing the Stokes parameters that characterize the polarization, as well as the intensity and wavelength as a function of position, we make use of all the information provided by the electromagnetic radiation from astronomical objects.

This work has involved a large supporting cast, and it is a pleasure to thank Bev Wills, Gary Schmidt, Frank Low, Paul Smith and Roc Cutri. I acknowledge support from NASA grant NAG5-3042 to the NICMOS instrument definition team.

References

- Antonucci, R.R.J., 1993, *ARA&A*, 31, 473
- Armus et al. 1999, *ApJ*, submitted
- Cohen, M. & Schmidt, G.D. 1982, *ApJ*, 259, 693
- Cutri, R.M., Huchra, J.P., Low, F.J., Brown, R.L., & Vanden Bout, P.A. 1994, *ApJ*, 424, L65
- Hines, D.C., & Wills, B.J. 1993, *ApJ*, 415, 82
- Hines, D.C., Schmidt, G.D., Smith, P.S., Cutri, R.M. & Low, F.J., 1995, *ApJ*, 450, L1
- Hines, D.C., Low, F.J., Thompson, R.I., Weymann, R.J. & Storrie-Lombardi, L.J. 1999a, *ApJ*, 512, 140
- Hines, D.C., Schmidt, G.D., Wills, B.J., Smith, P.S. & Sowinski, L.G. 1999b, *ApJ*, 512, 145
- Hines, D.C., Low, F.J., Evans, A.S., Scoville, N.Z., Rieke, M., Thompson, R.I. & Schneider, G., 1999c, in prep.
- Low, F.J., Huchra, J.P., Kleinmann, S.G. & Cutri, R.M., 1988, *ApJ*, 327, L41
- Sahai, R., Hines, D.C., Kastner, J.H., Weintraub, D.A., Trauger, J.T., Rieke, M.J., Thompson, R.I. & Schneider, G. 1998, *ApJ*, 492, L163
- Scoville et al. 1999, *ApJ*, submitted
- Surace, J.A. Sanders, D.B., Vacca, W.D., Veilleux, S., & Mazzarella, J., 1998, *ApJ*, 492, 116
- Wills, B.J., & Hines, D.C. 1997, in *Mass Ejection from AGN*, ASP Conf. Series, Vol. 128, Arav, Shlosman, & Weymann, eds., p. 99
- Young, S., Hough, J.H., Efstathiou, A., Wills, B.J., Bailey, J.A., Ward, M.J., & Axon, D.J. 1996b, *MNRAS*, 281, 1206

Figure 1. Imaging polarimetry and spectropolarimetry of IRAS 09104+4109 (after Hines et al. 1999b): (a & b) The rotated, normalized Stokes fluxes $q' \times F$, $u' \times F$ from WFPC2 imaging polarimetry ($q' = q \cos 2\theta + u \sin 2\theta$ and $u' = u \cos 2\theta - q \sin 2\theta$). The light gray lines indicate the average polarization position angle $\theta = 98^\circ$. The majority of the flux is contained in $q' \times F$ confirming that the extended structure is composed almost entirely of light polarized by scattering. (c) Comparison between the rotated Stokes (polarized) flux density spectrum ($q' \times F_\lambda$) with a mean quasar spectrum. The dashed curve is a power law fit to the polarized flux density continuum ($\alpha_\nu = -0.5$; $F_\nu \propto \nu^{\alpha_\nu}$). The dotted line indicates the location and FWHM of the broad H β emission line.

3-D Spectroscopic Observations of the IRLG Mrk266

T. Ishigaki¹, M. Yoshida, and K. Aoki

National Astronomical Observatory of Japan, Tokyo 181-8588, Japan

H. Ohtani, H. Sugai, T. Hayashi², S. Ozaki, T. Hattori, and M. Ishii

*Department of Astronomy, Faculty of Science, Kyoto University, Kyoto
606-8502, Japan*

1. Introduction

The Infrared Astronomical Satellite (IRAS) established the ubiquity of infrared-luminous galaxies (IRLGs) with an infrared luminosity $L_{\text{IR}} > 10^{11} L_{\odot}$. Most of IRLGs show the signatures of starbursts and/or AGN activities (Sanders & Mirabel 1996) and have large optical emission-line nebulae with a galactic-scale extent (Armus et al. 1990). The spectroscopy of the emission-line nebulae will provide the good informations on the activities in IRLGs and might give clues to investigate the energy source of enormous IR emissions.

As a case study, we made 3D optical spectrophotometric observations of Mrk266 ($L_{\text{IR}} = 3 \times 10^{11} L_{\odot}$), one of the archetypes of the IRLGs with AGNs (see Wang et al. 1997 and references therein). Mrk266 is apparently an ongoing merger system with two nuclei (the NE LINER nucleus and the SW Seyfert nucleus). Mrk266 has a large (> 20 kpc) and luminous ($L_{\text{H}\alpha + [\text{N II}]} > 10^{42} \text{ erg s}^{-1}$) emission-line nebula (Armus et al. 1990). However, the ionization mechanisms for a large part of the extended emission-line regions still remain unclear. The purpose of this study is (1) to identify the ionization mechanisms of the extended emission-line regions and (2) to probe heating sources for the infrared emission.

The redshift of Mrk266 is 0.028 and we will adopt a distance of 115 Mpc ($H_0 = 75 \text{ km s}^{-1} \text{ Mpc}^{-1}$) according to Wang et al. (1997). This implies a scale of $1'' = 540 \text{ pc}$.

2. Observational Data

Narrow-band images of Mrk266 were obtained at the f/18 Cassegrain focus of the Okayama Astrophysical Observatory (OAO) 188cm telescope. The Kyoto tridimensional spectrograph I (Ohtani et al. 1998) was used in the mode of imaging Fabry-Perot interferometer with a $5 \mu\text{m}$ gap Queensgate ET-50 etalon. The etalon provides bandpasses of $25 \text{ \AA}$ at $6929 \text{ \AA}$ and $16 \text{ \AA}$ at $5015 \text{ \AA}$. Observed emission lines were $\text{H}\alpha$, $[\text{S II}]\lambda 6716$, $[\text{S II}]\lambda 6731$, and $[\text{O III}]\lambda 5007$.

¹Present address: Department of Applied Physics, Graduate School of Engineering, Hokkaido University, Sapporo 060-8628, Japan

²Present address: Toyama Science Museum, Toyama 934-8084, Japan

Figure 1. (a) $H\alpha$ image, (b) HST image with the contour plots of $[O III]$ image, and (c) $[O III]/H\alpha$ ratio map. The sizes of panels are (a,c) $60'' \times 60''$ and (b) $30'' \times 30''$. East is toward the left and north is toward the top. (c) The gray scale indicates that the darker regions have the larger ratios. (d) The sketch of Mrk266 is shown to designate conspicuous morphological features.

Long-slit scan observations of Mrk266 were also carried out by using the OAO Cassegrain spectrograph in the Spectro-NebulaGraph (SNG) mode (Kosugi et al. 1995). Two spectral regions were observed with a spectral resolution of $2.6 \text{ \AA}$: a region from $6300 \text{ \AA}$ to $7100 \text{ \AA}$ and a region from $4600 \text{ \AA}$ to $5400 \text{ \AA}$. The slit was oriented in PA 0° . The bright part of the extended emission-line regions of Mrk266 was scanned with 10 slit positions.

In addition, an archival HST F606W image of Mrk266 was also obtained.

3. Results

The emission-line image and the broad-band HST image are shown in Figure 1. In the HST image, it is conspicuous that the vicinity of the SW nucleus is very dusty. Comparisons with the ground-based images show that the appearance of the ground-based images around the SW nucleus are largely affected by the dust lanes. In particular, the distribution of $[O III]$ emissions shows close anti-correlation with that of dust lanes. It may mean that the $[O III]$ emissions around the SW nucleus are absorbed by dusts. In the HST image, filaments

Figure 2. Decomposition of the line profiles of the NE nucleus. (a) $H\beta$ and $[O\ III]\lambda\lambda 4959,5007$. (b) $H\alpha$ and $[N\ II]\lambda\lambda 6548,6583$.

corresponding to the northern tail-like region and the eastern arm-like region are also found. It seems that the northern tail-like region is accompanied by dusts and these dusts connect with the dust lanes around the SW nucleus.

The $[O\ III]/H\alpha$ map (Figure 1c) and the long-slit spectroscopic data show that the northern tail-like region, the eastern arm-like region, and the west knot are in high excitation states with large $[O\ III]/H\beta$ and $He\ II/H\beta$ ratios.

In Figure 2, the continuum-subtracted spectrum of the NE nucleus is shown. The line profile of the NE nucleus shows a blue asymmetry and the decomposition of the line profiles reveals that the blue component has larger $[O\ III]/H\beta$ and $[N\ II]/H\alpha$ ratios than the main component.

4. Discussion

In Figure 3, the results of the long-slit spectroscopic observations are plotted in the manner of line-ratio diagnostic diagram and compared with the predicted ratios of the various models. In this diagram, almost all regions of Mrk266 show the ‘‘Seyfert-like’’ or ‘‘LINER-like’’ line ratios.

The observed spectral properties of the high excitation regions (the northern tail-like region, the eastern arm-like region, and the west knot) can be well explained by the photoionization model by a power-law continuum. If we assume that the northern tail-like region is photoionized by the SW Seyfert nucleus, the required rate of ionizing-photon emission by the SW Seyfert nucleus is 10^{55-56} photons s^{-1} in order to explain the excitation and the $H\alpha$ luminosity of this region. It leads to a fact that the SW Seyfert nucleus of Mrk266 is one of the most luminous Seyfert nuclei (cf. Padovani & Rafanelli 1988).

Figure 3 shows that the line ratios of the main component of the NE nucleus can be explained by an HII-region model and those of the blue component are plotted near the positions of the shock models of Shull & McKee (1979). It can be interpreted as follows; the LINER nucleus has a nuclear starburst corresponding to the main component and the gas corresponding to the blue component is ionized by shock due to a superwind from the starburst region. The emission-line luminosities of each component can be accounted for by a central starburst with a bolometric luminosity of about $2 \times 10^{10} L_{\odot}$.

Figure 3. Line-ratio diagnostic diagram, $[N II]/H\alpha$ vs. $[O III]/H\beta$. Regions are indicated by symbols shown at lower left. The various models are also shown: photoionization models by a power-law continuum calculated using CLOUDY (Ferland 1996) (dotted lines), shock models of Dopita & Sutherland (1996) (solid line) and Shull & McKee (1979) (capital letters), and HII-region models of Evans & Dopita (1985) (dashed lines).

The line ratios of the SW nucleus and its north circumnuclear region can be interpreted as those of a composite of the gasses photoionized by a power-law continuum and those by starbursts.

Estimation of contributions from the above activities to the energy source of the enormous infrared emission of Mrk266 shows that the Seyfert nucleus can be the dominant source. However, the contributions of the starburst activities around the nuclei may be also not negligible and be 20-30%.

We thank the staffs of the Okayama Astrophysical Observatory.

References

- Armus, L., Heckman, T. M., Miley, G. K. 1990, *ApJ*, 364, 471
- Dopita, M. A. & Sutherland, R. S. 1996, *ApJS*, 102, 161
- Evans, I. N. & Dopita, M. A. 1985, *ApJS*, 58, 125
- Ferland, G. J. 1996, Univ. of Kentucky Internal Report
- Kosugi, G. et al. 1995, *PASP*, 107, 474
- Ohtani, H. et al. 1998, *Proc. of SPIE*, 3355, 750
- Padovani, P. & Rafanelli, P. 1988, *A&A*, 205, 53
- Sanders, D. B. & Mirabel, I. F. 1996, *ARA&A*, 34, 749
- Shull, J. M. & McKee, C. F. 1979, *ApJ*, 227, 131
- Wang, J., Heckman, T. M., Weaver, K. A., Armus, L. 1997, *ApJ*, 474, 659

Merging Signatures in the Core of Mkn938

P. Rafanelli^{1,2,3}, A. Rifatto², V. Afanasiev⁴, S. Dodonov⁴, K. Birkle⁵, P. Böhm³, C. Cannavacciuolo², G. Richter³, A. and J. Vennik⁶

¹*Dipartimento di Astronomia, Università di Padova, Padova, Italy*

²*Osservatorio Astronomico di Capodimonte, Napoli, Italy*

³*Astrophysikalisches Institut Potsdam, Potsdam, Germany*

⁴*Special Astrophysical Observatory, Zelenchukskaya, Russia*

⁵*Max-Planck-Institut für Astronomie, Heidelberg, Germany*

⁶*Tartu Observatory, Tõravere, Estonia*

Abstract. Spectra of the central region ($16'' \times 16''$) of the Seyfert-2 galaxy NGC34 = Mkn938 have been obtained using the Multi Pupil Fiber Spectrograph (MPFS) at the 6 m telescope of the Special Astrophysical Observatory of the Russian Academy of Sciences at Zelenchukskaya, Caucasus. Gas and stars show a complex morphology and kinematics. The stellar radial velocity field in the lines of NaID doublet shows the existence of two counterrotating dynamical centers, the position of which corresponds to two distinct infrared sources detected at 8.8μ and 12.5μ . The presence of emitting gaseous knots ionized by stars on the south and on the west of the active nucleus support the idea that we are observing a dynamic merging of two galaxies.

1. Introduction

Mkn938 = NGC34 = IRAS00085-1223 ($z_{hel}(3K) = 0.0186$, scale ~ 360 pc/arcsec for $H = 75 \text{ km s}^{-1} \text{ Mpc}^{-1}$) is an irregular galaxy the classification of which is controversial in literature. Osterbrock & Dahari (1983) classified it as an emission line galaxy, dominated by star forming regions, while Veron-Cetty & Veron (1986) classified the galaxy as a Seyfert-2. Mulchaey et al. (1996) on the basis of narrow band images found that Mkn938 is a rather weak emitter of $[OIII] \lambda 5007$ compared to most of the Seyfert galaxies observed in their survey. In addition they found that the $H\alpha$ emission is strong and extended over the entire galaxy, further suggesting that the ionization of the gas is not related to any Seyfert activity. Morphologically the galaxy is apparently undergoing a merger as evidenced by the presence of optical tidal tails (Figure 1). This is confirmed by high resolution mid-infrared images at 8.8μ and at 12.5μ , obtained by Miles et al. (1996) in the frame of a work addressed to investigate the nature of luminous ($L_{FIR} \geq 10^{11} L_{\odot}$) far-infrared galaxies $L_{FIR}(\text{Mkn938}) \sim 10^{11.31} L_{\odot}$. Both images show in the nuclear region a double source separated by ~ 1.2 arcsec, with the fainter source at a position angle of $\sim 180^\circ$ from the brighter one.

Figure 1. Laplace adaptive filtered R image with overlaid B contours at levels $\mu=25.0$ and $\mu=26.5$ mag/arcsec² (length of bar 30''). North at the top, East to the left.

2. Optical Imaging

B and R images obtained at the Cassegrain focus of the ESO-MPIA 2.2 m telescope, equipped with EFOSC2, were reduced using adaptive filtering procedures developed at the Astrophysical Institute Potsdam (Richter et al. 1991). The contours of the B image at levels $\mu=25.0$, 26.5 are overlaid to the adaptive Laplace filtered R image in Figure 1. Tidal tails in the NE-SW direction, which resemble the typical shapes produced in simulations of close encounters between galaxies, suggest that a merging is going on. Fine structures are revealed by the Laplace filtered R image on the upper tail and in the core of the galaxy (central 16×16 arcsec), where many knots are present and the nucleus itself shows on the SE a close point-like emitting region at an angular distance of ~ 2 arcsec. The concavities of two emitting caps are facing the nucleus from symmetrically opposite positions along the North-South direction. Their separation from the nucleus is ~ 6 arcsec and their opening angle (as seen from the nucleus) is $\sim 70^\circ$.

3. 3D-Spectroscopy

Spectra of the nuclear region of Mkn938 were obtained using the Multi Pupil Fiber Spectrograph (MPFS) at the 6 m telescope of the Special Astrophysical Observatory (SAO). The array, consisting of 16×16 fibers, covered a field of view of $\sim 16 \times 16$ arcsec centered on the nucleus and was oriented along the North-South direction. Two different grating inclinations were needed to cover the spectral ranges 4250–5640 Å and 5660–7050 Å with a typical spectral resolution of ~ 4 Å.

Figure 2. a(left): Slices of the data cube in the $H\alpha$ line (contours) and corresponding nearby continuum. Tracings show the NaID profile. b(right): Slices of the data cube in the $[OIII] \lambda 5007$ line (contours) and corresponding nearby continuum. Tracings show the $H\beta$ and $[OIII]$ line profiles.

Contour maps of $H\alpha$ emission and $[OIII] \lambda 5007$ emission are overplotted to the images derived from their nearby continuum in Figure 2a and Figure 2b respectively. The distribution of the emitting gas is characterized by two emitting regions, one on the West and one on the South of the nucleus. It is confirmed that the emission in $H\alpha$ is spatially more extended than in $[OIII] \lambda 5007$ (Mulchaey et al. 1996), but the diagnostic line ratios $[SII]/H\alpha$, $[NII]/H\alpha$, $[OI]/H\alpha$ versus $[OIII]/H\beta$ (Veilleux & Osterbrock 1987) derived from the single fiber spectra, show that the main ionization mechanism in the nucleus is photoionization by a non-thermal source. Since the profiles of both forbidden and permitted lines have the same width ($FWHM \sim 400 \text{ km s}^{-1}$) Mkn938 can be classified as a Seyfert-2. The effects of the nuclear source are present also in the spectra of the western and southern regions, which show features typical of emitting gas photoionized by a mixture of nonthermal and thermal sources.

The NaID $\lambda\lambda 5896, 5890$ absorption doublet, which is a tracer of the stellar component, shows line profiles characterized by a different shape in the single fiber spectra (Figure 2a). A cross-correlation analysis shows that both NaID lines have two components characterized by different wavelengths in each fiber spectrum. A 3-dimensional plot of the radial velocities derived in this way (Figure 3) shows the presence of two distinct stellar concentrations, separated by $\sim 150 \text{ km s}^{-1}$ in the velocity space and each characterized by a radial velocity dispersion of $\sim 100 \text{ km s}^{-1}$ around its dynamic center. On the sky projection

Figure 3. Line of sight velocity distribution (LSVD) of the NaID absorption line, as derived from cross correlation analysis.

(X-Y plane) the centers of the two stellar concentrations are located on the North-South direction, spaced by ~ 1 arcsec. Their size and relative position fit very well the brightness centers identified in the infrared by Miles et al. (1996). Analysis of the radial velocity distribution along directions at different position angles on the X-Y plane indicates that stars move in opposite directions around the two dynamic centers.

4. Conclusions

The analysis of the data clearly indicates that the nuclear region of the galaxy is harbouring an active nucleus surrounded by star forming regions. The two distinct stellar components, clearly separated both in velocity and in position, can be interpreted as the nuclei of two galaxies in the act of merging. This confirms what is suggested by the irregular morphology of the galaxy and is an example of a dynamically revealed actual merging of galaxies.

References

- Miles, J.W., Houck, J.R., Hayward, T.L., Ashby, M.L.N., 1996, *ApJ*, 465, 191
- Mulchaey, J.S., Wilson, A.S., Tsvetanov, Z., 1996, *ApJS*, 102, 309
- Osterbrock, D.E., Dahari, O., 1983, *ApJ*, 273, 478
- Richter, G.M., Böhm, P., Lorenz, H., Priebe, A., Capaccioli, M., 1991, *Astr. Nac.*, 312, 345
- Veilleux, S., Osterbrock, D.E., 1987, *ApJS*, 63, 295
- Véron-Cetty, M.-P., Véron, P., 1986, *A&AS*, 65, 241

Imaging Spectroscopy at mm-Wavelengths

N. Z. Scoville and A. I. Sargent

*Astronomy Department, California Institute of Technology, 105-24,
Pasadena, California 91125*

Abstract.

Millimeter-wave imaging with the existing arrays provides arcsec angular resolution at a few km s^{-1} kinematic resolution in the molecular rotational lines (eg. CO, HCN, and CS). The environments range from planetary atmospheres, proto-stellar and proto-planetary nebulae, stellar mass-loss envelopes, the star forming regions of nearby galaxies, and the starburst nuclei of ultraluminous galaxies at both low and high redshift (out to $z=4.7$). In this review, we highlight the spectacular results from these areas and discuss advances expected in the next decade.

1. Introduction

At present there exist four mm-wave interferometers : $6\times 10\text{m}$ at Nobeyama (NRO), $6\times 10\text{m}$ at Caltech (OVRO), $10\times 6\text{m}$ at BIMA, and $5\times 15\text{m}$ at IRAM. Despite the modest sizes and costs of these arrays, they yield images at arcsecond to sub-arcsecond resolution for the thermal emission in molecular rotational lines and the dust continuum; in contrast, typical resolutions using the VLA on the 21-cm HI line are $10\text{--}30''$. The results from the mm-arrays have been spectacular in their own right and underscore the tremendous potential of future, even larger arrays.

Aperture synthesis imaging in atmospheric windows at $\lambda \sim 3$, 1.3 and 0.8 mm is an incredibly powerful technique for observations of the dense interstellar clouds responsible for virtually all star formation and probably also the source of fuel to active galactic nuclei. The mm-wave arrays have yielded aperture synthesis images of molecular emission in domains ranging from cometary and planetary atmospheres and the circumstellar envelopes of evolved stars, to both low and high mass star-forming regions in the Galaxy and protoplanetary disks. In extragalactic observations, the arrays have provided high resolution images of the central bars and spiral-arm molecular gas concentrations in nearby galaxies in addition to extensive studies of starburst and AGN nuclei in more distant ultraluminous galaxies. CO has been detected and imaged at arcsecond resolution in eight high redshift quasars and ultraluminous galaxies at $z = 2.3\text{--}4.7$. In the following sections, we present a sample of the results from studies of a proto-planetary disk around the young star MWC 480, spiral structure imaging in M51, the interacting system NGC 4038/39, the ultra-luminous IR galaxy Arp 220 and two high redshift submm continuum sources.

MWC480 CO(2→1) Emission

Figure 1. MWC 480: Comparison between the observed CO(2-1) emission and emission predicted by a model of a flattened, Keplerian rotating disk around the central star (Mannings, Koerner & Sargent 1997). The bottom panel shows spectra for the model and observations made by integrating each of the separate velocity images. Tic marks on the images are spaced by 1".

2. Proto-planetary Disks

It is now well established that a large fraction of solar-mass pre-main sequence stars are surrounded by disks of gas and dust during the first 3–30 Myr of their existence (c.f. Beckwith & Sargent 1996; Sargent 1996). Disk masses range from ~ 0.001 and $0.1 M_{\odot}$, and outer radii are between 100 and 1000 AU, comparable to those inferred for the primitive solar nebula during the epoch in which the planets formed (c.f. Lissauer 1993).

The OVRO array has recently been used for a survey of Herbig Ae/Be stars. Most of the eleven Ae stars in the sample have unresolved regions of thermal continuum emission (Mannings & Sargent 1996; Mannings et al. 1997). Typically, upper limits to radii are between 200 and 300 AU, and masses range from 0.005 to $0.234 M_{\odot}$. Six H Ae stars also display molecular line emission with elongated morphologies and double-peaked spectra. The ordered velocity gradients along the major axes of the observed emission strongly suggest that gas is orbiting these stars in inclined disks. For MWC 480, synthetic spectral line maps for a disk in Keplerian rotation and observed CO (2–1) channel maps are compared in Figure 1; the excellent agreement between the synthetic and observed images provides unambiguous evidence for a rotating disk of gas and dust bound to the star (Mannings et al. 1997). MWC 480 is a higher-mass counterpart of a T Tauri system and a possible precursor to a β Pic-like debris disk. This disk should survive long enough for planetesimals to grow – its age greatly exceeds (by a factor of almost 100) current estimates of the formation time for km-size planetesimals (Weidenschilling 1997).

3. Spiral Structure – M51

The relationships of the HI, H₂ and HII gas *vis-a-vis* the galactic spiral arms; the nature of spiral arm streaming motions; and the possible existence of higher order resonance patterns are all major outstanding issues in spiral structure. Although these questions might logically be addressed at higher resolution in our Milky Way galaxy, the inherent line-of-sight confusion in the Galactic plane favors nearby galaxies for such studies. In Figure 2, the CO (1–0) emission in the central disk of M51 is overlaid on a visible wavelength image of M51 from HST (Aalto et al. 1999). This image obtained with the OVRO array shows remarkably clear definition of the molecular gas associated with the major spiral arms and the positional coincidence between this gas and the visible dust lanes and offset H α emission. The spiral-arm molecular cloud associations of mass 10^7 – $10^8 M_{\odot}$ are clearly resolved with a radial width of approximately 150 pc. Across the arms, streaming motions (20 – 50 km s^{-1}) due to the potential well of the spiral arm are seen in CO.

In addition to the large scale patterns of molecular emission in the M51 disk, a CO emission peak is seen on the nucleus (This is hard to see in Figure 2 due to the bright nuclear background). This molecular gas is possibly an “accretion” disk with $10^7 M_{\odot}$ of H₂ at 50 pc radius (Scoville et al. 1997).

Figure 2. CO emission contours are shown superposed on an HST (B,V,R, $H\alpha$) color composite image for M51. The CO image is at $2.3''$ resolution, corresponding to 100 pc; it was made from a mosaic of 19 pointing centers (Aalto, Huettemeister, Scoville, & Thaddeus 1999). The HST image is missing complete coverage due to the higher resolution pixels in the PC.

4. Interacting Starburst Galaxies – NGC 4038/39

As the closest example of a merger of two massive spiral galaxies, the Antennae, NGC 4038/39, offers a unique opportunity to study physical and dynamical properties of the interstellar medium in an early merger system. The Antennae has one of the largest populations of luminous star clusters known (Whitmore & Schweizer 1995) and infrared (ISO) observations have revealed optically-invisible regions of massive star formation that are likely to be the early stages of globular clusters (Vigroux et al. 1996).

Images of the CO emission in NGC 4038/39 have been made using both the BIMA (7", Gruendl *et al.* 1999) and OVRO arrays (3", Wilson et al. 1999). In Figure 3 the BIMA CO contours are overlaid on the HST image of the Antennae (Whitmore & Schweizer 1995). The image reveals a wealth of fine detail as well as extended fainter emission throughout the region. The excellent spatial correlation between the dust lanes in the optical image and the CO emission is striking. The total molecular gas mass is $1.610^{10} M_{\odot}$, half of which is concentrated in the overlap region. This region consists of two main parts : a vertical streamer stretching from the northern to the southern galaxy made up of gas at velocities from 0 to -120 km s^{-1} with respect to the systemic velocity and a highly concentrated clump at the bottom of the streamer (Gruendl *et al.* 1999). The second component is partly made up of gas moving at high blueshifted velocities (up to -200 km s^{-1}). The brightest CO concentrations appear to be gravitationally bound, with radii of 600-900 pc and masses of $4 - 8 \times 10^8 M_{\odot}$ – a factor of 10 greater than the most massive molecular structures in the nearby spiral, M51 (Wilson *et al.* 1999).

Most intriguing is the correspondence between the ISO $15 \mu\text{m}$ continuum emission and CO (Mirabel et al. 1998) in the region where the two galaxies overlap. Here the $15\mu\text{m}/\text{CO}$ ratio is 3-4 times larger than elsewhere, consistent with greatly enhanced star formation efficiency, but the mass, size, and velocity dispersion of the molecular gas are similar to those of the other bright clumps.

5. Nearby Galaxy Surveys

Two major interferometric surveys of nearby galaxies have been conducted recently : the joint NRO/OVRO survey of 20 nearby spirals in CO(1-0) at $2-4''$ resolution (Sakamoto et al. 1999a) and the BIMA survey of 44 galaxies at $\sim 7''$ resolution (Helfer *et al.* 1998). We mention these surveys to underscore the speed of mm-wave aperture synthesis which can yield mosaiced images of nearby galaxies even with the existing small, 6-10 telescope arrays.

A major result of the NRO/OVRO survey comparing equal numbers of barred and nbarred spirals is that barred spirals exhibit a higher degree of central gas concentration than normal spirals. In Figure 4, the surface density of H_2 averaged over the central kpc is plotted versus the H_2 surface density averaged over the entire optical disk. The barred spirals (filled squares) have generally lower overall surface density but, more significantly, *systematically higher central concentrations* (i.e. they lie in the upper left of the figure) – typical central gas fractions are three times higher than in the normal spirals. This constitutes the first clear evidence of the oft-quoted prediction of more rapid radial-transport of

Figure 3. HST image of the cores of the colliding galaxies: NGC4038 and NGC4039. Contours show concentrations of CO emission obtained by combining BIMA array at 7" resolution with NRAO 12m single dish maps (Gruendl *et al.* 1999).

Figure 4. The surface density of molecular emission averaged over the central kpc plotted against the surface density of molecular emission averaged over the optical disk out to R_{25} for the 20 galaxies in the NRO-OVRO galaxy survey (Sakamoto *et al.* 1999a). There is a clear trend for the barred spirals to have significantly greater central concentrations, supporting models of barred-induced radial accretion of gas.

gas in barred spiral galaxies. The trend for lower overall H_2 surface densities in the barred galaxies suggests that the concentration of gas in their centers has led to more rapid expenditure of the galactic ISM, presumably in star formation.

6. Ultra-Luminous IR Galaxies – Arp 220

With infrared ($8\text{--}1000\mu\text{m}$) luminosities in excess of $10^{12}L_{\odot}$, ultraluminous infrared galaxies rival QSOs in their total energy output (Sanders & Mirabel 1996, Soifer 1987). These systems result from interactions and merging of gas-rich, spiral galaxies. Their extreme luminosities result from dynamically triggered starbursts and active galactic nuclei (AGN) fueled by molecular gas driven inward by gravitational torques and dissipative accretion. For both the nuclear starbursts and AGN, the power originates at UV-to-optical wavelengths but is absorbed by dust and re-emitted in the infrared. In these merging systems, the dust extinction in the active nuclear regions can be extraordinarily high (mean $A_V \geq 2000$ mag, see Scoville *et al.* 1997) and mm-wave imaging provides a

Figure 5. The merging nuclei of Arp 220 are shown in $0.5''$ resolution imaging of the CO(2-1) and dust continuum emission. These data clearly resolve the two nuclei and reveal for the first time counter-rotating disks in each nucleus. The panels show : a) continuum-subtracted CO(2-1) (using only high resolution data), b) the CO mean velocities, c) the 1.3 mm dust continuum, and d) the total CO emission (including both low and high resolution interferometry). Crosses indicate the 1.3 mm continuum positions of the nuclei.

unique capability to probe the ISM distribution and kinematics. More than 20 luminous ($\geq 10^{11} L_{\odot}$) infrared galaxies have now been imaged, primarily at OVRO and IRAM (Scoville *et al.* 1991, Scoville *et al.* 1997, Downes & Solomon 1998, Bryant & Scoville 1999). Virtually all display massive concentrations of molecular gas in the central few kpc.

Arp 220, at 77 Mpc, is one of the nearest and perhaps the best known ultra-luminous merging system. It has been the subject of a number of OVRO and IRAM interferometer studies imaging in the 2.6 mm CO line (Scoville *et al.* 1991), 3 mm HCN (Radford *et al.* 1991), and 1.3 mm CO (Scoville, Yun & Bryant 1996, Downes & Solomon 1998, Sakamoto *et al.* 1999b). The CO (2-1) line emission was first mapped at $1''$ resolution at Owens Valley and showed two peaks separated by, $0.9''$ and an inclined disk of molecular gas (Scoville *et al.* 1996); somewhat higher resolution obtained by Downes & Solomon (1998) clearly resolved the two peaks. These peaks correspond well with the double 'nuclei' seen

in near-infrared and radio continuum images. The $0.5''$ resolution CO and 1.3 mm continuum maps obtained recently by Sakamoto et al. (1999b) using OVRO are displayed in Figure 5. These reveal for the first time **counter-rotating** disks of gas in each of the nuclei. The kinematic data clearly require very high mass concentrations in each nuclei, consistent with their being individual galactic nuclei. The fact that they are counter-rotating is consistent with the concept that more complete merging may be associated with counter-rotating precursor galaxies in which there can be greater angular momentum cancellation. The masses in each nucleus are apparently dominated by the molecular gas – a common finding of the the ULIG galaxy studies.

7. High Redshift Galaxies – Sub-mm Continuum Sources

In recent years there has been an explosive growth in observations related to the high redshift era of galaxy formation. Particularly intriguing is the evidence for the presence of large quantities of heavy-element-enriched gas at $z \geq 2-5$. Massive quantities of both dust and molecular gas (CO) have now been detected out to $z = 4.7$ – pushing the epoch of early metal enrichment *in these objects* back to within a few billion years of the Big Bang. In Table 1, CO detections in distant galaxies are listed.

CO imaging provides critical constraints on the early metal enrichment, in addition to estimates for the gas and, most critically, the dynamical masses. These cannot be derived at other wavelengths. The gas mass-fraction is a clear signature of the evolutionary state of the galaxy, be it protogalactic or more like present-day galaxies with just a small fraction of the mass in the ISM. The mm/submm observations therefore constrain the cosmic epoch for initial star and galaxy formation and should establish whether galaxies formed hierarchically (from the bottom up) or as pre-assembled protogalaxies with their present-day masses. They should also clarify whether the first generation star-formation occurred in *superluminous*, dynamically-triggered bursts or over longer periods at lower luminosity.

Recent sub-mm surveys using SCUBA have uncovered a population of dust-rich galaxies at high redshift (Smail et al. 1997; Barger et al. 1998; Hughes et al. 1998; Eales et al. 1998). mm-Interferometric observations are critical for unambiguous identification of any optical counterparts to the sub-mm sources; the $15''$ SCUBA beam often leaves room for uncertainty. OVRO has been used for searches for CO emission in the three brightest sub-mm sources which have well-determined optical redshifts. Two were detected : SMM J02399 at $z = 2.8$ and SMM J14011 at $z = 2.6$ (Frayser et al. 1998, 1999). Contour maps of the CO(3-2) emission are overlaid on optical images in Figure 6. The extremely high molecular gas masses and the sub-mm/radio flux ratios for both objects are consistent with the infrared luminosities being generated by star-formation. Indeed, the properties of SMM J02399 and SMM J14011 are similar to those of local ultra-luminous infrared galaxies (ULIGs) but the co-moving number density of the sub-mm population is more than a factor of 100 times higher than that of low-redshift ULIGs.

Figure 6. CO (3-2) contours for SMM J02399 (top) and SMM J14011 (bottom) are overlaid on optical images (B-band for SMM J02399 and I-band for SMM J14011). The contour levels are -3,3,4,5,6,7,8 σ .

Table 1
Distant CO Detections

Name	Redshift	Transition	Telescope	Reference
F10214+4724	2.286	CO(3-2)	12m	Brown & V'B 91
		CO(3-2)	30m	Solomon et al. 92
3C 48	0.37	CO(1-0)	OVRO	Scoville et al. 93
H1413+117	2.558	CO(3-2)	PdB	Barvainis et al. 94
BR1202-0725	4.69	CO(5-4)	NRO	Ohta et al. 95
		CO(5-4)	PdB	Omont et al. 96
BRI 1335-0415	4.41	CO(5-4)	PdB	Guilloteau et al. 97
53W002	2.39	CO(3-2)	OVRO	Scoville et al. 97
MG 0414+0534	2.64	CO(3-2)	PdB	Barvainis et al. 98
SMM J02399-0136	2.808	CO(3-2)	OVRO	Frayser et al. 98
APM 0827+5255	3.911	CO(4-3)	PdB	Downes et al. 99
SMM J14011+0252	2.565	CO(3-2)	OVRO	Frayser et al. 99

8. Future Advances

Within the next decade, we expect several major advances in instrumentation and techniques for high frequency aperture synthesis imaging. Most phenomenal, will be the ALMA array planned as more than 64 12-m telescopes sited at 16500 ft in Chile. This array should be capable of operation down to $\lambda = 350\mu\text{m}$ with resolution $\leq 0.1''$. Assuming optimistic schedules and funding it could be operating within ~ 10 yrs. On timescales of a few years, one can expect the sub-mm array (SMA) in Hawaii with ~ 8 6-m dishes also operating down to $350\mu\text{m}$. The OVRO and BIMA have also recently proposed to join their mm-arrays with 15 telescopes at a 9000 ft site in California with 3-4 yrs – an instrument called CARMA. Thus, we can look forward to arrays with vastly larger collecting area and number of elements, resulting in much higher sensitivity and imaging speed.

There are also very exciting technical developments – extremely broadband correlator systems using high speed programmable array chips (Hawkins & Woody 1999) and active atmospheric phase compensation. The latter uses radiometers on each dish to monitor the atmospheric water vapor emission along the sight-line from telescope to the astronomical source (Woody, Carpenter & Scoville 1999). Based on the derived H_2O column, one can then remove the short-term phase fluctuations and obtain higher coherence imaging and improved image fidelity. This is the mm-wave equivalent of adaptive optics and the successful implementation will be just as fundamental at mm-wavelengths as in the optical. Both of these developments are highly desirable in order to push the interferometry into the submm regime.

It is a pleasure to thank S. Vogel and S-W Lee for communicating the BIMA results shown here.

References

- Aalto, S., Huttemeister, S., Scoville, N., & Thaddeus, P. 1999, *Ap.J.*, (in press)
- Barger, A.J., Cowie, L.L., Sanders, D.B., Fulton, E., Taniguchi, Y., et al. 1998, *Nature*, 394, 248
- Barvainis, R., Alloin, D., Guilloteau, S., & Antonucci, R. 1998, *ApJL*, 492, L13
- Barvainis, R., Tacconi, L., Antonucci, R., Alloin, D., & Coleman, P. 1994, *Nature*, 371, 586
- Beckwith, S.V.W., & Sargent, A.I. 1996, *Nature*, 383, 139
- Bryant, P. M. & Scoville, N. Z. 1999, *A.J.*, 117, 000.
- Brown, R.L., & Vanden Bout, P.A. 1991, *AJ*, 102, 1956
- Djorgovski, S.G., DeCarvalho, R.R., Gal, R. R. et al. 1998, *IAU Symp.* 179, 424.
- Downes, D., Neri, R., Wiklind, T., Wilner, D.J., & Shaver, P.A. 1999, *ApJL*, 513, 1
- Downes, D., & Solomon, P.M.. 1998, *ApJ*, 507, 615
- Eales, S.L., Lilly, S., Gear, W., Dunne, L., Bond, J.R., et al. 1999, *ApJ*, 515, 518
- Frayser, D.T., Ivison, R.J., Scoville, N.Z., Evans, A.S., Yun, M.S., et al. 1998, *ApJ*, 506, L7
- Frayser, D.T., Ivison, R.J., Scoville, N.Z., Evans, A.S., Yun, M.S. et al. 1999, *ApJ*, 514, L13
- Gruendl, R. A., Lo, K. Y., Lee, S.-W., Lee, T.-H and Gao, Y 1999, *Ap.J.*, (in preparation)
- Guilloteau, S., Omont, A., McMahon, R.G., Cox, P., & Petitjean, P. 1997, *A&A*, 328, 1
- Helfer, T.T., Regan, M. W., Thornley, M. D., Wong, T., Sheth, K., Vogel, S. N., Bock, D. C.-J., Blitz, L., Harris, A., 1998, *The Evolution of Galaxies on Cosmological Timescales*, Tenerife
- Hawkins, D. & Woody, D. P. 1999 (in preparation)
- Hughes, D.H., Serjeant, S., Dunlop, J., Rowan-Robinson, M., Blain, A., et al. 1998, *Nature*, 394, 241
- Lissauer, J.J. 1993, *ARA&A*, 31, 129
- Mannings, V., Koerner, D.W., & Sargent, A.I. 1997, *Nature*, 388, 555
- Mannings, V., & Sargent, A.I. 1997, *ApJ*, 490, 792
- Ohta, K., Yamada, T., Nakanishi, K., Kohno, K., Akiyama. M., et al. 1996, *Nature*, 382, 426
- Omont, A. , Petitjean, P., Guilloteau, S., McMahon, R. G., Solomon, P. M. & Pecontal, E. 1996, *Nature*, 382, 428
- Sakamoto, K., Okumura, S. K., Ishizuki, S. & Scoville, N.Z. 1999a, *ApJ*, (in press)
- Sakamoto, K., Scoville, N.Z., Yun, M.S., Crosas, M., Genzel, R., et al. 1999b, *ApJ*, 514, 68
- Sanders, D.B., & Mirabel, I.F. 1996, *ARA&A*, 34, 749
- Sargent, A.I. 1996, in *Disks & Outflows around Young Stars: Lecture Notes in Physics*, Vol 465, eds. S.V.W. Beckwith, H. Staude, A. Quetz, & A. Natta (Berlin:Springer-Verlag), 1
- Scoville, N.Z., Sargent, A.I., Sanders, D.B., & Soifer, B.T., 1991, *ApJ*, 366, L5
- Scoville, N.Z., Yun, M.S., Armus, L., & Ford, H. 1998, *ApJL*, 493, L63
- Scoville, N.Z., Yun, M.S., & Bryant, P.M. 1997, *Ap.J.*, 484, 702
- Scoville, N.Z., Yun, M.S., Windhorst, R.A., Keel, W. C., & Armus, L. 1997, *ApJL*, 485, L21
- Soifer, B.T., Sanders, D.B., Madore, B. F., Neugebauer, G., Danielson, G.E., et al. 1987, *ApJ*, 320, 238
- Weidenschilling, S. 1997, *Icarus*, 127, 290
- Whitmore, B.C., & Schweizer, F. 1995, *AJ*, 109, 960
- Wilson, C.D., Madden, S., & Scoville, N.Z. 1999, in preparation
- Woody, D. P., Carpenter, J. M. & Scoville, N. Z. 1999, in preparation

Submm Continuum Surveys for Obscured Galaxies

Ian Smail

*Department of Physics, University of Durham, South Road, Durham
DH1 3LE*

Rob Ivison

*Department of Physics and Astronomy, University College London,
Gower Street, London WC1E 6BT*

Andrew Blain

Cavendish Laboratory, Madingley Road, Cambridge CB3 0HE

Jean-Paul Kneib

*Observatoire Midi-Pyrénées, CNRS-UMR5572, 14 Avenue E. Belin,
31400 Toulouse*

Frazer Owen

NRAO, P.O. Box 0, 1003 Lopezville Road, Socorro, NM 87801

Abstract. We discuss deep surveys in the submm regime using the SCUBA bolometer array on the JCMT. At $850\mu\text{m}$ SCUBA has resolved the bulk of the submm background (SMBR) detected by *COBE* into discrete sources brighter than 0.5mJy . The on-going identification and characterisation of this population at other wavelengths suggests that the bulk of the submm sources brighter than $\sim 1\text{mJy}$ lie at $z \gtrsim 1$, with a median redshift for the population of $\langle z \rangle \sim 2-3$. The optical/near-infrared properties of the counterparts to the submm sources breakdown as follows: roughly two-thirds have reliable identifications with the others being more ambiguous. Of those with identifications about half are optically bright ($I \lesssim 23$) mergers or interactions with a high proportion showing signatures of AGN activity and the other half are optically faint ($I \gtrsim 25$) including both blank fields and Extremely Red Objects (EROs). We conclude that a population of distant, highly obscured ultraluminous infrared galaxies (ULIRGs) dominates the SMBR. The wide range in the characteristics of the optical counterparts is consistent with the dispersion in the restframe UV properties of local ULIRGs. We suggest that the faint submm population comprises a class of high redshift dusty, mergers associated with the formation of present day luminous elliptical galaxies.

1. Introduction

The energy density in the extragalactic background radiation at optical/UV wavelengths is roughly equal to that seen in the far-infrared/submm (e.g. Bernstein et al. 1999; Puget et al. 1996; Fixsen et al. 1998). The simplest interpretation of this observation (ignoring contributions from dust-enshrouded AGN) is that averaged over all epochs, around half of all the star formation in the Universe has occurred in highly obscured regions. Clearly if we wish to obtain a complete and unbiased view of the star formation history of the Universe, necessary to constrain models of galaxy formation and evolution (Baugh et al. 1998), we need to understand in detail what fraction of star formation is obscured by dust and how this varies with epoch and environment. To achieve this we must investigate the nature and origin of the far-infrared/submm background and that requires the resolution of the background into discrete sources and the study of their individual and collective properties.

This review begins by discussing observational programmes in the submm with the Sub-millimeter Common User Bolometer Array (SCUBA, Holland et al. 1999) on the 15-m JCMT which have achieved the aim of resolving the bulk of the extragalactic background at $850\,\mu\text{m}$ into discrete sources. We then discuss follow-up observations of these sources and use these to explore the broad characteristics of the populations contributing to the submm background.

As a benchmark for the following discussion we note that a ULIRG similar to Arp 220 with a far-infrared luminosity of $L_{\text{FIR}} \sim 3 \times 10^{12} L_{\odot}$ and a star-formation rate (SFR) of $\sim 300\,\text{M}_{\odot}\,\text{yr}^{-1}$ would have a $850\text{-}\mu\text{m}$ flux density of $\gtrsim 3\,\text{mJy}$ out to $z \sim 10$ in a spatially flat Universe and $\gtrsim 0.3\,\text{mJy}$ for an open Universe with $q_0 = 0.05$ (Hughes & Dunlop 1997).¹ In two nights of observing in good conditions with SCUBA it is possible to achieve a 3σ flux limit of $3\,\text{mJy}$ across a $160''$ -diameter field, probing a volume of $10^6\,\text{Mpc}^3$ out to $z \sim 10$ for dusty galaxies as luminous as Arp 220.

2. The Number Counts of the Faint Submm Population

The advent of sensitive submm imaging with SCUBA has allowed a number of groups to undertake ‘blind’ surveys for faint submm galaxies. Results on the number density of sources in blank fields as a function of limiting $850\text{-}\mu\text{m}$ flux density have been published by three groups: Hughes et al. (1998) worked with a single deep map centered on the *Hubble Deep Field* (HDF); while Barger et al. (1998, 1999b) employed a combination of deep/narrow and wide/shallow observations of fields in the Lockman Hole and Hawaii Survey Field regions, finally there is a on-going survey of areas included in the Canada-France Redshift Survey (first results given in Eales et al. 1999). The surface densities of sources measured by the different groups are shown in Fig. 1. Due to the modest resolution of these maps, $15''$ FWHM, they are confusion limited at $\sim 2\,\text{mJy}$.

Our collaboration has taken a complimentary approach to these ‘blank’ field surveys by using massive gravitational cluster lenses to increase the sensitivity

¹We assume $q_0 = 0.5$ and $h_{100} = 0.5$ unless otherwise stated. In addition, unless identified as ‘observed’, all magnitudes/fluxes are corrected for lens amplification.

Figure 1. The cumulative 850- μ m counts from published SCUBA surveys. The latest counts from the Blain et al. (1999a) analysis of the SCUBA Cluster Lens Survey are marked by solid circles. Counts from Barger et al. (1998; B98, 1999b; B99), Eales et al. (1999; E99), Holland et al. (1998; H98) Hughes et al. (1998; HDF) and Smail, Ivison & Blain (1997; S97) are also shown. $P(D)$ indicates the limit from Hughes et al.'s confusion analysis in the HDF. The solid line shows a crude parameterisation of the counts, $N(>S) = 7.9 \times 10^3 S^{-1.1}$, while the counts from a non-evolving model based on the local *IRAS* 60- μ m luminosity function are given by the dotted.

and resolution of SCUBA. The first deep submm counts were based on maps of two clusters (Smail, Ivison & Blain 1997) and the survey was subsequently expanded to cover seven lensing clusters at $z = 0.19$ – 0.41 (Smail et al. 1998; Blain et al. 1999a). The complete sample comprises a total of 17 galaxies detected at 3σ significance or above, with 10 detected above 4σ , from a total surveyed area of ~ 40 sq. arcmin in the image plane down to a 1σ flux limit of ~ 1.5 mJy at 850 μ m. Two of these sources are identified with the central cluster galaxies in the clusters A 1835 and A 2390 and as such are removed from our analysis, although see Edge et al. (1999) for more discussion of these systems. The analysis of our catalog makes use of well-constrained lens models for all the clusters to accurately correct the observed source fluxes for lens amplification (Blain et al. 1999a). For the median source amplification, $\sim 2.5\times$, our survey covers an area of the source plane equivalent to 15 sq. arcmin to a 3σ flux limit of ~ 2 mJy. The lens amplification also results in a factor of two finer beam size at this depth so that these counts have a fainter confusion limit than the blank field observations. At higher amplifications, the survey covers a smaller region, but at a correspondingly higher sensitivity (e.g. ~ 1 sq. arcmin at $\sigma_{850} \sim 0.1$ mJy) and resolution. The uncertainties associated with our lensing analysis are included in the final error quoted on the derived counts (Blain et al. 1999a). The total uncertainty in the lensing correction is at most comparable to the typical error in the absolute SCUBA calibration.

The 850- μ m counts from the analysis of Blain et al. (1999a) are shown in Fig. 1 and are in agreement with the results from the other surveys at ≥ 2 mJy. However, the magnification produced by the massive cluster lenses allows us to

also constrain the source counts down to 0.5 mJy, four times fainter than the deepest blank-field counts published and free from confusion noise.

The cumulative 850- μ m counts down to 4 mJy from Smail, Ivison & Blain (1997) accounted for roughly 30% of the SMBR detected by *COBE* (e.g. Puget et al. 1996; Fixsen et al. 1998). The counts from later surveys down to the blank field confusion limit of JCMT at 2 mJy account for close to 50% of the SMBR, while the deepest counts from the lens fields indicate that the bulk of the SMBR is resolved by 0.5 mJy (Blain et al. 1999a). The majority of the SMBR is thus produced by sources with 850- μ m fluxes of 1–10 mJy and as we discuss in the next section these galaxies are likely to lie at $z \gtrsim 1$ and hence they have intrinsic bolometric luminosities of 10^{12} – $10^{13} L_{\odot}$ and densities of around 10^{-5} Mpc^{-3} .

In addition to the 850- μ m maps discussed above, SCUBA also provides simultaneous 450- μ m imaging of the same fields. However, the combination of modest atmospheric transmission at 450 μ m in normal conditions on Mauna Kea and the lower efficiency of the JCMT dish surface at shorter wavelengths has restricted the results appearing in this waveband. Based upon a similar analysis to Blain et al. (1999a), using those 450- μ m maps from the lens survey where useful sensitivity was obtained we derive an approximate cumulative source density of 1000 deg^{-2} brighter than 20 mJy at 450 μ m. This surface density is consistent with the number of 450- μ m detections in the Eales et al. (1999) survey and the reported lack of detections in the HDF by Hughes et al. (1998). As we discuss in the next section the relative paucity of 450- μ m sources suggests a fairly high redshift for the bulk of the submm population, $z \gg 1$.

Having resolved the background we can now study the nature of the populations contributing to the SMBR and so determine at what epoch the background was emitted. Here again our survey has the advantage of lens amplification, this time in the radio and optical/near-IR where the identification and spectroscopic follow-up are undertaken. Typically the counterparts of our submm sources will appear ~ 1 magnitude brighter than the equivalent galaxy in a blank field.

3. The Redshift Distribution of the Faint Submm Population

Photometric techniques for estimating the redshifts of candidate counterparts to submm sources have been employed by Hughes et al. (1998) and Lilly et al. (1999) leading to the suggestion that the bulk of the population lay at $z = 2$ –4 or $z = 0.1$ –3 respectively. However, these analyses are based upon the spectral energy distributions (SED) of local optically-selected galaxies and the diversity of the restframe UV/optical properties of local ULIRGs (Trentham et al. 1999) and their differences from ‘normal’ galaxies suggests that such analyses are fraught with complications. More general constraints on the maximum possible redshift of the submm population come from the search for the signature of Lyman- α absorption due to the intergalactic medium in the broad-band photometry of candidate counterparts (Smail et al. 1998). The detection of the bulk of the proposed counterparts in our survey in *B* or *V* imaging indicated that at least three-quarters were at $z \lesssim 5.5$ and more than half were likely to have $z \lesssim 4.5$. While admittedly weak, these constraints are free from concerns over the SEDs adopted for the distant ULIRGs and suggest that proportion of the faint submm population at very high redshifts, $z > 5$, is small (see Eales et al. 1999).

Figure 2. The cumulative redshift distribution for the SCUBA lens survey. We have used the spectroscopic redshifts of those galaxies known to be reliable counterparts from Barger et al. (1999a) and combined these with the probable redshift ranges of the remaining sources derived from their observed $\alpha_{1.4}^{850}$ indices or limits. The solid line shows the cumulative distribution if we assume a flat probability distribution for the sources within their allowed z_α ranges from the Carilli & Yun (1999) models and a maximum redshift of $z = 8$ for those sources where the $\alpha_{1.4}^{850}$ indices only provide a lower redshift limit. In contrast, the dashed line gives the conservative redshift distribution which is obtained if *all* sources are assumed to lie at their lower z_α limit. The dotted line is the cumulative redshift distribution for all the counterparts from Barger et al. (1999a) with two of the source identifications corrected as in Smail et al. (1999a) and the blank-field/ERO candidates placed at $z = 4$.

First results from spectroscopic surveys of the submm population are beginning to appear. In particular a Keck II spectroscopic survey of possible counterparts to the submm sources in our survey has recently been published (Barger et al. 1999a). Identifications were attempted for all the galaxies bright enough for reliable spectroscopy within the SCUBA error-boxes. This resulted in spectroscopic redshifts or limits for 24 possible counterparts to 14 SCUBA sources. The median I -band magnitude of the counterparts is $I = 22.4$, the equivalent depth for identifying candidates in a blank field submm survey would be closer to $I \sim 23$, stretching the capabilities of even the largest telescopes.

In a number of cases the spectral properties of the candidate counterparts suggested that they were likely to be the source of the submm emission (Barger et al. 1999a), in others the identification of a radio counterpart (e.g. Ivison et al. 1999), or the unusual optical-NIR colors of a candidate (see §4) or morphologies add support to the identification of the submm emission as arising from a particular galaxy. For two sources we have been able to confirm the proposed galaxy as the submm source through the detection of redshifted CO emission in the millimeter at the redshift of the optical counterpart (Frayer et al. 1998; 1999). The reliable spectroscopic identifications include a $z = 2.8$ dusty type-2 AGN/starburst (Ivison et al. 1998); a $z = 2.6$ starburst (Barger et al. 1999a; Ivison et al. 1999); a $z = 3.2$ type-1 AGN (Ivison et al. 1999), the first example of the sub-mJy submm population identified; and the lowest redshift confirmed

source, a $z = 1.06$ ring galaxy (Soucail et al. 1999). The spectroscopic observations are thus consistent with the photometrically-derived redshift limits for the bulk of the population, and also provide information about the dominant emission processes in individual galaxies. In particular, the spectra give a useful indication of the relative fractions of AGN and starbursts in the submm population (Ivison et al. 1999).

But for over half the submm sources the results are more ambiguous with none of the galaxies with spectroscopic identifications within the submm error-box showing unusual spectral features, colors or morphologies. This leaves open the possibilities that either the submm emitting region is so highly obscured that it is invisible in the restframe optical/UV, a not unreasonable suggestion, or that the submm source has a fainter optical counterpart and remains unidentified in the spectroscopic survey. To distinguish between these alternatives and attempt to determine the redshift distribution of a *representative* sample of the faint submm population we have to resort to other spectroscopic indicators, which while admittedly cruder have the advantage of not relying on the identification of an optical counterpart for the submm source. In particular this is the only way to tackle those sources with no visible counterparts in the optical (e.g. Hughes et al. 1998; Smail et al. 1998) as well as providing useful information on the small proportion of sources with very red counterparts seen only in the near-infrared (Smail et al. 1999a). General constraints on the likely redshift distribution of faint submm sources come from the spectral shape of dust emission in the restframe far-infrared (Hughes et al. 1998) and the combination of this with radio information (Carilli & Yun 1999; Blain 1999).

As discussed by Hughes et al. (1998), the ratio of 450- and 850- μ m fluxes can be used as a crude redshift indicator. Their analysis of the information provided by the 450- μ m non-detections of the five 850- μ m sources in the HDF suggested that the galaxies all lay at $z > 1$. A similar constraint arises from assuming that the same population of sources are being detected at 450 and 850 μ m (an assumption which is supported by Eales et al. 1999) and determining the flux ratio between the two wavelengths at a fixed source surface density. The 450- μ m counts are 1000 deg^{-2} at a flux limit of 20 mJy, the equivalent surface density is achieved at 850- μ m at ~ 6.5 mJy. Thus we obtain a typical S_{450}/S_{850} ratio of $S_{450}/S_{850} \sim 3$ suggesting that the median redshift for the population is $\langle z \rangle \sim 2$.

Another other method has been recently developed to estimate redshifts for faint submm sources using the 850 μ m to 1.4 GHz spectral index, $\alpha_{1.4}^{850}$ (Carilli & Yun 1999). This technique relies upon the good correlation between the strength of the far-infrared emission (reprocessed UV/optical radiation from massive stars) and radio (synchrotron emission from electrons accelerated in the supernovae from massive stars) in local star-forming galaxies (Condon 1992). The decline in emission from dust at longer wavelengths is eventually overtaken by the rising synchrotron emission to produce an upturn between the submm and radio regimes around 3 mm. This feature is seen in the SEDs of both AGN and starburst galaxies (see examples in Ivison et al. 1998, 1999) and as proposed by Carilli & Yun (1999) the spectral index in this region can provide a crude redshift estimate. The spectral index has the behaviour that it is larger for higher redshift sources and Carilli & Yun were able to show that the redshift predictions from $\alpha_{1.4}^{850}$ based on local templates spectra and model SEDs were

in good agreement with the observed redshifts for a small sample of distant submm sources. The index also has the useful property that contamination by radio emission from an obscured AGN will tend to reduce the value of $\alpha_{1.4}^{850}$ giving a low redshift estimate. Thus $\alpha_{1.4}^{850}$ can be used to place robust *lower* limits on the redshifts of the submm population.

We have used deep VLA 1.4-GHz maps of the seven clusters in our survey to measure or place limits on the radio flux from the submm sources. To make these limits as conservative as possible we have used the radio flux for the *brightest* radio counterpart within each submm error-box, irrespective of whether there are other candidates which are preferred for other reasons. This means that we obtain a strong lower limit on $\alpha_{1.4}^{850}$, and hence on the source redshift, independent of the exact source identification. The radio maps have a typical 1σ sensitivity of $\lesssim 10\mu\text{Jy}$ in the source plane and we identify radio counterparts to around half of the submm sources, with useful limits on the remainder. Using the various spectral models from Carilli & Yun (1999) we can transform these $\alpha_{1.4}^{850}$ measurements and limits into redshift ranges, z_α , for each source.

We plot in Fig. 2 the cumulative redshift distribution derived from the radio/submm spectral analysis of the sources in our survey (Smail et al. 1999b). We show two curves, the first makes the most conservative assumption that each source without a reliable spectroscopic identification in Barger et al. (1999a) lies at the minimum redshift allowed by the observed spectral index or 3σ lower limit (typically assuming an SED similar to Arp 220). The other curve assumes that the sources have uniform probability of lying anywhere in the redshift range allowed by their $\alpha_{1.4}^{850}$ index (with a maximum redshift for the population of $z = 8$). These distributions are compared to that obtained by Barger et al. (1999a) in their spectroscopic follow-up of these sources (where the two blank field sources and the recently identified ERO counterparts of Smail et al. (1999a) are all placed at $z = 4$). The median redshift for the complete spectroscopic sample, including both the reliable and possible counterparts, is $\langle z \rangle \sim 2.5$ equivalent to the lower limit determined from our most conservative spectral index analysis. Distributing the sources in a more reasonable manner within their allowed range for z_α leads to median redshifts closer to $\langle z \rangle \sim 3-4$.

Our conclusions about the source redshifts based upon both the S_{450}/S_{850} ratio and $\alpha_{1.4}^{850}$ spectral index are sensitive to the assumed dust temperature, T_d , in the sources. For both estimators reducing T_d will allow lower median redshifts for the population (see Blain 1999). However, only if we force the entire submm population to have a characteristic dust temperature less than 30 K can we start to push the median redshift much below $\langle z \rangle = 2$.

We conclude that a variety of constraints from the observed spectral properties of the submm population, as well as detailed spectroscopic observations of a small number of robustly identified sources (Ivison et al. 1998, 1999; Barger et al. 1999a), suggest that the median redshift of the submm population lies in the range $\langle z \rangle \sim 2-3$, with few if any luminous submm galaxies at $z \lesssim 1$. The submm fluxes of all the sources detected in published surveys lie in the range $S_{850} \sim 0.5-10 \text{ mJy}$, assuming they lie at $z \gtrsim 1$ then their luminosities are $\log_{10} L_{\text{FIR}} \sim 12-13$ and so they all class as ultraluminous infrared galaxies.

Figure 3. ERO counterparts to two submm sources in our survey (Smail et al. 1999a). The two panels on the left show a deep, $0.6''$ -resolution Keck II *I*-band image and a UKIRT *K*-band image of the field of SMM J09429+4658, overlaid on these are the $850\text{-}\mu\text{m}$ SCUBA map and a deep 1.4-GHz VLA map respectively. The two panels on the right show the equivalent data for the field of SMM J04433+0210. The faintest sources visible in the *I*-band exposure have $I \sim 25.5\text{--}26.0$, while the *K*-band images reach to $K \sim 20.5$. The original candidate counterparts for the submm sources are marked on the *K* images, as well as the new ERO candidates, H5 for SMM J09429+4658 and N4 for SMM J04433+0210. Each panel is $30''$ square and is centred on the nominal position of the $850\text{-}\mu\text{m}$ peak (absolute accuracy of $\lesssim 3''$). The relative radio-optical astrometry is better than $0.4''$ and hence the radio source close to the bright galaxy at the top of the SMM J09429+4658 frame is not coincident with it.

4. The Nature of the Faint Submm Population

We are still at an early stage in the study of the faint submm population and so the following discussion will concentrate on the results from the SCUBA lens survey for which a wide range of follow-up has been published. To briefly summarise the status of this survey. There are 15 non-cluster submm sources in the survey, of these around five have spectroscopic identifications which we believe reliably identifying these galaxies as the submm sources (two of these have been subsequently confirmed via detections in CO). A further three sources have either counterparts with unambiguous radio identifications or extreme optical-near-IR colors which allow us to identify them, and two more are in optically blank fields. This leaves five sources which have ambiguous identifications.

Starting with those galaxies with spectroscopic identifications, these all have relatively bright optical counterparts and a high proportion show multiple components in optical/near-IR – on separations of $\lesssim 2\text{--}3''$ (e.g. Ivison et al. 1998, 1999), at the galaxy redshifts this scale is equivalent to ~ 10 kpc. This lends weight to the suggestion that mergers and interactions are a crucial trigger of activity in distant ULIRGs (Smail et al. 1998; Lilly et al. 1999) as they are for more local examples. The identification of submm sources with merging systems suggests that we should properly view them as ‘events’ rather than galaxies.

To search for any extremely red counterparts, $(I - K) > 6$, which could have been missed in the optical identifications we have used UKIRT to obtain near-IR imaging of our fields down to $K \gtrsim 20$. We have so far identified two possible ERO counterparts to submm sources in our survey (Fig. 3, Smail et al. 1999a), both previously identified with bright $z \sim 0.5$ spiral galaxies, with the bulk of the submm error-boxes containing galaxies with optical-near-IR colors more typical of the general field, $(I - K) \sim 2\text{--}4$. Deeper *K*-band imaging of one submm source with a reliable radio position has also provided an identification

of a $K \sim 22$ counterpart, the optical limit on this galaxy is $I \gtrsim 24$. The two submm error-boxes which were ‘blank’ ($I \gtrsim 25$) in the optical search undertaken by Smail et al. (1998) show no near-infrared candidates to $K \sim 21$.

The ERO counterparts account for 15% of the submm population. If the optically ‘blank’ fields also contain EROs the proportion of submm sources with highly reddened counterparts will rise to 30%. The implied surface densities of submm-bright EROs are consistent with the number of EROs detected in targeted SCUBA observations (e.g. Dey et al. 1999) and suggest that around half of the ERO population are dusty, star forming galaxies at moderate/high redshifts (Smail et al. 1999a).

Although currently incomplete, we are slowly building up a picture of the population of distant, luminous submm galaxies which dominate the SMBR at wavelengths around 1 mm. The fact that the submm population detected by SCUBA can account for all of the *COBE* background indicates that a large fraction of the stars in local galaxies could be formed in these systems. The characteristics of the submm galaxies are similar to those of local ULIRGs, except that they contribute a submm luminosity density at early epochs that is at least an order of magnitude greater than the corresponding local population. One goal is to use this population to trace the amount of high-redshift star-formation activity that is obscured from view in the optical by dust, and so is missing from existing inventories of star-formation activity at high redshift (Smail, Ivison & Blain 1997; Hughes et al. 1998; Blain et al. 1999b). In this way a complete history of star formation in the Universe can be constructed.

From the analysis discussed in §3 we can state that the luminous submm population is roughly coeval with the more modestly star-forming galaxies selected by UV/optical surveys of the distant Universe (e.g. Steidel et al. 1999). However, the individual SCUBA galaxies have SFRs which are typically an order of magnitude higher than those of the optically-selected galaxies, as well as being dustier and probably therefore more chemically enriched. Converting the submm luminosities into an equivalent star formation density is a very uncertain process (e.g. Blain et al. 1999b), nevertheless, most reasonable conversions result in a star formation density in the submm population which is a factor of several higher than that seen in UV/optically selected samples at high redshift (Fig. 4), even after applying similarly uncertain corrections to the latter for the effects of dust obscuration on the UV luminosities. Although this comparison remains uncertain, it does show that the submm population needs to be explained by any models that claim to describe the star formation history of the Universe.

As with local ULIRGs, there is uncertainty over the exact contributions from AGN and starbursts to the far-infrared luminosity density of the SCUBA population. This adds another possible source of contamination to the comparison shown in Fig. 4. The identification and removal of obscured AGN from the submm sample can be achieved through searches for hard X-ray emission with *Chandra*. For all but the most heavily enshrouded systems ($\log N(\text{H I}) \gtrsim 24$) the hard X-ray emission should still be detectable from the central AGN (Gunn 1999). These searches will also provide an estimate of the total contribution from the dust-obscured AGN to the X-ray background (Almaini et al. 1998; Gunn 1999). Calculations which use the X-ray background to constrain the number of obscured AGN indicate that at most 20–30% of the SCUBA sources

Figure 4. The estimated star formation densities at different epochs from UV/optically-selected samples (nominally corrected for extinction) compared to that estimated from the SCUBA population assuming the conservative redshift distribution from Fig. 2 (Blain et al. 1999b). The lower-bound on the SCUBA point shows the maximum correction for AGN contamination in the sample as discussed in the text. The symbols follow those in Blain et al. (1999b) with the exception that the filled circles show the latest results by Steidel et al. (1999) and the cross that from Yan et al. (1999).

could harbor an obscured AGN. The incomplete observations available for the submm population suggest that $\geq 20\%$ of the submm population show obvious spectral signatures of an AGN. We stress, however, that this does not mean that the AGN dominates the emission in the submm and so the correction indicated in Fig. 4 remains uncertain.

One interesting issue which remains is what are the descendants of the submm population? They have high bolometric luminosities, large dust masses ($10^8 M_\odot$, Ivison et al. 1998, 1999) and large gas reservoirs ($10^{11} M_\odot$, Frayer et al. 1998, 1999). These properties are consistent with on-going massive star formation in these systems at rates $\geq 1000 M_\odot \text{yr}^{-1}$, which given the available molecular gas supply could continue for 10^8 yrs and result in the formation of an entire L^* galaxy at $z \geq 2$. The submm population also show morphological similarities to local merging ULIRGs which are thought to evolve into elliptical galaxies. The circumstantial evidence thus points towards the submm sources being precursors of elliptical galaxies. Assuming that the luminous submm phase lasts for a few dynamical times of the remanent halo, ≥ 1 Gyr, the volume density of these galaxies at high redshifts is $\sim 2\text{--}4 \times 10^{-4} \text{Mpc}^{-3}$, high enough to allow all massive ellipticals to be formed in this manner. Arguments have been advanced for a number of years for the prompt and synchronised formation of a large fraction of the luminous elliptical galaxy populations in the richest clusters at $z \geq 3$ (Bower, Lucey & Ellis 1992; Ellis et al. 1997) due to their extreme homogeneity both within individual clusters and between clusters. If the SCUBA sources are identified with the earliest, obscured phases of this activity then we would expect the submm sources to be clustered on scales comparable to that of the putative proto-clusters at these early epochs, ~ 10 Mpc or 20 arcminutes (Governato et al. 1998). Thus strong clustering of SCUBA sources

is a clear prediction of the identification of these galaxies with proto-ellipticals, searching for such structures should be a high priority for future submm surveys.

An equally interesting issue is the relationship between the submm population and the Lyman-break galaxies (Steidel et al. 1999). These objects appear to have typically lower SFRs and less dust than the submm galaxies, although they have a substantially higher number density. We would therefore identify the submm galaxies with the most energetic mergers which form massive, young ellipticals and more quiescent star formation due to secular evolution in disk systems with the Lyman-break objects. Detailed observations of the submm population should thus provide much needed observational input to models of the formation and evolution of massive galaxies (Blain et al. 1999b). In particular we anticipate CO detections of more submm galaxies to study the kinematics of these systems and hence determine their masses.

5. The Future

On-going and planned upgrades to SCUBA and the JCMT will improve the sensitivity and effectiveness of this world-class facility. Looking slightly further ahead there is a proposal for a wide-field submm imager, SCUBA-2, based upon new detector technology and capable of providing statistically reliable samples of submm galaxies. Such samples will enable us to break new ground in crucial areas of study such as clustering of submm sources, necessary to understand the evolutionary status of this population relative to other classes of high redshift source. However, at $850\ \mu\text{m}$ the large JCMT beam will remain the main restriction to probing deeper into the submm counts to identify the turn-over which should occur around $\sim 0.5\text{--}1\text{ mJy}$. This turn-over has important implications for models of the formation and evolution of obscured galaxies and the approach we have taken of employing massive gravitational lenses to increase the sensitivity and resolution of the SCUBA maps is well suited to tackling this problem. An equivalent length exposure to that obtained with SCUBA on the HDF (Hughes et al. 1998) but instead targetted on a well-constrained cluster lens such as Abell 370 would constrain the form of the $850\text{-}\mu\text{m}$ counts down to 0.3 mJy (3σ), probing the region where the counts should turn over if they are to remain consistent with the *COBE* measurements of the SMBR.

In the long term the proposed Atacama Large Millimeter Array (ALMA) will mark an enormous leap forward in the capabilities of ground-based submm mapping and imaging. The current optical/near-infrared identification programs for the minute samples of relatively bright submm sources provided by SCUBA are stretching the capabilities of 4- and 10-m telescopes. The numbers and characteristics of the sources likely to be uncovered with ALMA will probably exceed the follow-up capabilities of the available facilities (including *NGST*), even if we restrict ourselves to merely near-infrared imaging and limited spectroscopy. It may be that we have to abandon entirely short-wavelength observations of this population and rely on what we can observe in the submm and radio. The other option is to tailor the submm surveys for easier follow-up. In this regard we are pursuing submm imaging of fields around bright, $V \sim 12$, high-latitude stars (which are invisible in the submm). These can then be used as natural guidestars for high-order adaptive optics systems on 4- and 8-m telescopes, facilitating deep high-resolution imaging and spectroscopy of possible counterparts.

We are grateful to Amy Barger, Len Cowie, Katherine Gunn, Neil Trentham for useful discussions. IRS thanks the organisers for support to attend the conference and acknowledges a Royal Society Fellowship.

References

- Almaini, O., Lawrence, A., Boyle, B.J., 1999, MNRAS, 305, L59
 Barger, A.J. et al., 1998, Nature, 394, 248
 Barger, A.J., Cowie, L.L., Smail, I., Ivison, R.J., Blain, A.W., Kneib, J.-P., 1999a, AJ, in press
 Barger, A.J., Cowie, L.L., Sanders, D.B., 1999b, ApJ, 518, L5
 Baugh, C.M., Cole, S.M., Frenk, C.S., Lacey, C.G., 1998, ApJ, 498, 504
 Bernstein, R.A., Freedman, W., Madore, B., 1999, ApJ, submitted
 Blain, A.W., 1999, MNRAS, submitted
 Blain, A.W., Kneib, J.-P., Ivison, R.J., Smail, I., 1999a, ApJ, 512, L87
 Blain, A.W., Smail, I., Ivison, R.J., Kneib, J.-P., 1999b, MNRAS, 302, 632
 Bower, R.G., Lucey, J.R., Ellis, R.S., 1992, MNRAS, 254, 601
 Carilli, C.L., Yun, M.S., 1999, ApJ, 513, L13
 Condon, J.J., 1992, ARAA, 30, 575
 Dey, A., Graham, J.R., Ivison, R.J., Smail, I., Wright, G.S., Liu, M., 1999, ApJ, in press
 Eales, S.A., et al., 1999, ApJ, 515, 518
 Edge, A.C., Ivison, R.J., Smail, I., Blain, A.W., Kneib, J.-P., 1999, MNRAS, in press
 Ellis, R.S., et al., 1997, ApJ, 483, 582
 Frayer, D.T., Ivison, R.J., Scoville, N.Z., Evans, A.S., Yun, M., Smail, I., Blain, A.W., Kneib, J.-P., 1998, ApJ, 506, L7
 Frayer, D.T., Ivison, R.J., Scoville, N.Z., Evans, A.S., Yun, M., Smail, I., Barger, A.J., Blain, A.W., Kneib, J.-P., 1999, ApJ, 514, L13
 Fixsen, Dwek, E., Mather, J.C., Bennett, C.L., Shafer, R.A., 1998, 508, 123
 Governato, F., Baugh, C.M., Frenk, C.S., et al., 1998, Nature, 392, 359
 Gunn, K.F., 1999, PhD Thesis, University of Durham
 Holland, W.S., et al., 1998, Nature, 392, 788
 Holland, W.S., et al., 1999, MNRAS, 303, 659
 Hughes, D.H., Dunlop, J.S., 1997, in 'Highly Redshifted Radio Lines', eds. Carilli, C., et al., PASP
 Hughes, D.H., et al., 1998, Nature, 394, 241
 Ivison, R.J., Smail, I., Le Borgne, J.-F., Blain, A.W., Kneib, J.-P., Bézecourt, J., Kerr, T.H., Davies, J.K., 1998, MNRAS, 298, 583
 Ivison, R.J., Smail, I., Barger, A.J., Kneib, J.-P., Blain, A.W., Owen, F.N., Kerr, T.H., Cowie, L.L., 1999, MNRAS, submitted
 Lilly, S.J., et al., 1999, ApJ, in press
 Puget, J.-L., Abergel, A., Bernard, J.-P., et al., 1996, A&A, 308, L5
 Smail, I., Ivison, R.J., Blain, A.W., 1997, ApJ, 490, L5
 Smail, I., Ivison, R.J., Blain, A.W., Kneib, J.-P., 1998, ApJ, 507, L21
 Smail, I., Ivison, R.J., Kneib, J.-P., Cowie, L.L., Blain, A.W., Barger, A.J., Owen, F.N., Morrison, G., 1999a, MNRAS, in press
 Smail, I., Ivison, R.J., Owen, F.N., Blain, A.W., Kneib, J.-P., 1999b, in prep.
 Soucail, G., Kneib, J.-P., Bézecourt, J., et al., 1999, A&A, 343, L70
 Steidel, C.C., Adelberger, K.L., Giavalisco, M., Dickinson, M., Pettini, M., 1999, ApJ, in press
 Trentham, N., Kormendy, J., Sanders, D.B., 1999, AJ, 117, 2152
 Yan, L., McCarthy, P.J., Freudling, W., et al., 1999, ApJL, in press

Ionized Outflows in 3-D: Insights from Herbig-Haro Objects & Applications to Nearby AGN

Gerald Cecil

*Dept. Physics & Astronomy, U. North Carolina at Chapel Hill, NC
27599; SOAR Telescope Project, NOAO, Tucson AZ 85726*

Abstract. HST shows that the gas distributions of these objects are complex and clump at the limit of resolution. HST spectra have lumpy emission-line profiles, indicating unresolved sub-structure. The advantages of 3D over slits on gas so distributed are: robust flux estimates of various dynamical systems projected along lines of sight, sensitivity to fainter spectral lines that are physical diagnostics (reddening; gas density, T, excitation mechanisms, abundances), and improved prospects for recovery of unobserved dimensions of phase-space. These advantages allow more confident modeling for more profound inquiry into underlying dynamics. The main complication is the effort required to link multi-frequency datasets that optimally track the energy flow through various phases of the ISM. This tedium has limited the number of objects that have been thoroughly analyzed to the *a priori* most spectacular systems.

For HHO's, proper-motions constrain the ambient B-field, shock velocity, gas abundances, mass-loss rates, source duty-cycle, and tie-ins with molecular flows. If the shock speed, hence ionization fraction, is indeed small then the ionized gas is a significant part of the flow energetics.

For AGN's, nuclear beaming is a source of ionization ambiguity. Establishing the energetics of the outflow is critical to determining how the accretion disk loses its energy. CXO will provide new constraints (especially spectral) on AGN outflows, and STIS UV-spectroscopy is also constraining cloud properties (although limited by extinction.) HHO's show some of the things that we will find around AGN's.

I illustrate these points with results from ground-based and HST programs being pursued with collaborators.

1. Introduction

This paper addresses two large fields, linked by the common technique of 3D-spectroscopy and by similar data modeling to highlight underlying physics. Outflows are either hot, diffuse fans or confined to collimated particle streams (jets). Galactic Herbig-Haro objects (HHO's) were understood from the outset to be shock-excited structures. Their large scale and dense ISM allowed single-dish antennas to map the associated molecular flow. These outflows are our premiere shock laboratories. HST has resolved HHO "jets" into streams of bow shocks around "bullets" that are heavier than their surroundings (Reipurth *et al* 1997.)

In contrast, much of an AGN outflow proceeds outside the purview of optical (I include near-IR) astronomy. Unlike Herbig-Haro objects, outflows are secondary in the minds of most to the processes that act within a couple of Schwarzschild radii of the nuclear SMBH. Radio continuum and X-ray images are important flow tracers, the latter on large scales until CXO flies. AGN jets are found to have lower density than their surroundings, so buoyancy must be considered when there is a gravity gradient.

In this paper I will list the advantages of 3D-spectra on outflows, outline the expected dynamics, introduce some kinematic and shock models, highlight insights derived from Galactic outflows, show a few results on AGN outflows, and end with prospects on AGN outflows for 3D in the UV and with CXO. The version at www.thececils.org/science/3D99/3D.pdf has color figures, which help.

1.1. A Brief History

Before 3D it was known that most Seyferts have broad [O III] emission-line profiles (e.g. Whittle 1985.) If not gravitational, it was likely that these largely blue-wing asymmetric profiles indicated shocks somewhere. The literature was ripe with speculation as to whether flows were in or out, according to the dust distribution. There was discussion about radiative acceleration of photoionized clouds, largely because models of high-velocity shocks were rudimentary, and because even multiple slits sampled the few spatially extended gas distributions in an incoherent fashion that failed to reveal underlying symmetries. Subsequent radio surveys (largely by Ulvestad & Wilson e.g. 1984) showed that a subset of these galaxies had radio jets/plasmons whose interaction with the ISM would produce strong shocks. At last, an outflow axis! After a few heroic slit-surveys, 3D instruments supplanted slits and found their AGN niche. Work over the past 15 years has concentrated on the most spectacular of the jet/ISM interactions. Fabry-Perot (FP's) have isolated the most prominent shocked regions. Now these are targets for STIS (e.g. Crenshaw & Kraemer, and collaborators.)

The theory of fast shocks was developed in the series of papers by Dopita and collaborators, starting with Binette, Dopita, & Tuohy (1985). Allen *et al* (1998) have extended this work into the UV to support STIS and COS on HST. Ferruit *et al* (1997) have developed models of AGN bow shocks.

2. Advantages of 3D on Outflows

The most obvious advantage arises when variations of extinction or gaseous excitation distort the appearance of the flow. Slits presuppose an axial symmetry, but Nature may not cooperate with this laziness. 3D also allows unambiguous rejection of things that are *not* outflows. I did this recently for gas along the ionization cones of NGC 5252 (Fig. 1.) Here the breakthrough occurred when 3D allowed us to link our model velocity field to that mapped in H I outside the ionized cones. Other advantages are:

Untangle Projections Shocks are most prominent in dense media, so are often seen projected on other dynamical systems. "Gaussian deblending" is more successful on 3D- than on slit-spectra because it can be checked for consistency from many different directions (e.g. Shopbell & Bland-Hawthorn 1998.)

Figure 1. NGC 5252, with HST line and continuum images (former show the inner “ionization cones” at left (Morse *et al* 1998.) Panels (a) & (b) at right show data superimposed on rotation and outflow models, respectively. Extensive slit-spectra could not exclude outflow. But our 3D spectra showed that some gas orbits on a disk, inclined to the S0 disk. The core counter-rotates (compare model (d) to data (a)). All these features are occasionally superimposed along our line of sight, and all are expected from a merger whose debris train is illuminated as it orbits through the nuclear ionization cones. Our model velocity field (c to the right) is contiguous with H I velocities measured outside the cones where motions cannot be traced by ionized gas.

Reliable Multifrequency Comparisons Other wavebands are invariably 3D. So optical work on spatially extended targets *must* dispense with slits to avoid sampling-artifacts as the energy flow is tracked through various ISM phases.

Radial & Space Velocities HHO’s expand enough in the interval between HST TAC meetings to ensure that even shock accelerations are measured. AGN’s require models to bound the missing dimensions. This is most successful when there are symmetries to hang the model on.

3. Parameterizing Outflows

To find important data symmetries, we must parameterize the sea of emission-line profiles. The usual Gaussian fitting scheme must address multiple components because ionized gas associated with an outflow often projects on the rest of the galaxy. We often compare model and data line-profiles, not just parameter maps. Spectra must be smoothed to the resolutions of data in other wavebands.

4. What Can We Constrain with 3D Spectra?

Source Disk Opening Angle This follows from the illumination pattern of high-ionization lines. There is evidence for opening-angle dependence on photon energy. The TTF allows these maps to be obtained in low-luminosity but high excitation lines in an efficient manner, with excellent cancellation of the underlying starlight. Of course geometry may be set on unresolved scales.

Spin Misalignments Between Kinematic Subsystems I have mentioned NGC 5252. As another example, NGC 4258 has different spin axes for the small-scale accretion disk and galaxy disk (Miyoshi *et al* 1995.) 3D spectra show that its outflow *projects* along the accretion disk axis (§8) but actually lies 65° away.

Fueling Efficiency & Duty Cycle With the full flow-field mapped, one can establish if there has been a continuous or episodic outflow (HHO's), or if events have been interleaved on opposite sides of the source (flip/flop ejection).

Flow Collimation & Entrainment Oblique shocks generate the thermal cocoon. Simulations show that these can stabilize a jet against disruption (Steffen *et al* 1997.) HST sees such shocks in several HHO's (Fig. 5; Heathcote *et al* 1996.)

Galaxy Halo Properties Whether the wind-blown superbubble ruptures at its top, and details of its velocity structure as a function of distance from the nucleus, are both influenced by halo properties (Suchkov *et al* 1994.)

Chemical Enrichment 3D has not yet provided constraints because Fabry-Perot (FP) spectrometers have until recently been too inflexible in their wavelength coverage to obtain accurate photometry of a large number of lines of different excitation. The TTF changes this. Now one can exploit spatial binning and 3D subtraction of the galaxy continuum at straddling wavelengths to develop high S/N flux estimates (or limits) on faint, nebular lines (diagnostic of T.) By separating dynamical systems, 3D can lead to more robust models which highlight abundance anomalies, and perhaps correlate these anomalies with kinematic changes (e.g. grain destruction in shocks.)

Ambient B-Field This inhibits compression in high-velocity shocks. In some regimes emission-line flux ratios are altered significantly, yielding the *only probe of B-field strength in ionized gas filaments*.

Nature of the Power Source The clarifications afforded by 3D spectra outlined above improve the power of diagnostic models to constrain the ultimate energy source. Gas abundances and ionization parameter(s) have the strongest influences. The TTF and IFU's are able to measure faint, high ionization lines in the optical (e.g. [Ne V], [Fe X] $\lambda 6374$) to pin down the ionizing continuum.

5. Realities

Stellar spectra constrain the gravitational potential and provide a robust dynamical axis that allows orbiting gas to be removed from the picture. These

measurements will require IFU's not FP's. Overall, IFU's are more immune to seeing variations and are a better match to the scale of an AGN than are FP's. However, FP's and the TTF have the advantage of placing the AGN in the context of the larger galaxy. E. g., Bland-Hawthorn shows in his contribution that the influence of the AGN in NGC 1068 extends across the galaxy.

AGN To date, fewer than 20 AGN circumnuclear regions have been studied with 3D. These objects were selected because most contain radio jets, so general insights are elusive. It is time for a systematic 3D survey that does not depend on source properties (including the presence of a compact light source on which to lock the AO system!) The route to insight *via* flow energetics is straightforward:

- Deredden spectra across the outflow. Normally the Balmer decrement is used, but it will soon be efficient with the TTF and deep depletion CCD's to use the blue/IR [S II] flux ratio for the best optical baseline.
- Use velocity symmetries to deproject the flow into the "rest frame" of the galaxy or star-forming region. Be sure to account for a stellar bar.
- Based on flux/kinematical correlations, guess the dynamical state of the gas. Assume shocks in the rest of this list.
- Compare observed and model fluxes of as many line ratios as possible, including faint T diagnostics in the cooling zone, to pin down the shock velocity V_s and magnetic compression $B/\sqrt{n_0}$ (Fig. 2). The Dopita & Sutherland (1995, DS hereafter) models generate $L(H\alpha) \approx 1.5 \times 10^{36} f n_0 A_4 V_{s100}^{2.32}$ erg s⁻¹ for shock area A_4 (units 10⁴ pc²) and pre-shock gas density n_0 (cm⁻³.) Scaling $f = 0.3 - 0.8$ for $V_s = 90 - 1000$ km s⁻¹.
- But first we need to estimate pre-shock density n_0 and filling factor, because ionized mass, KE, and momentum all scale by $1/n_0$. Good luck, the [S II] lines have limited dynamic range. HHO's can be in a dense enough medium to use this doublet, but the low shock-speed means that there is a large correction to the ionization fraction (Hartigan *et al* 1994.) AGN circumnuclear regions are frequently in the low-density limit.
- If velocity discontinuities are a few hundred km s⁻¹, search for highly ionized coronal Fe-lines and apply for HST STIS time to confirm shocks with UV lines. If you can't find the Fe lines, perhaps what you see is just fluff that was compressed by the flow but is not energized entirely by X-rays from hot, post-shock gas (Wilson & Raymond 1999.)

At nearly the end of this tortuous path, for ionized gas in the outflows of the "narrow-line regions" (NLR's) of M 51, NGC 1068, NGC 3079, & M 82 we obtain KE's of $10^{54-56}/n_0$ erg. Bicknell *et al* (1998) interpret some Seyfert NLR's as structures powered by the energy and momentum of jets. They argue that for NGC 1068, a confining wind (Cecil, Bland, & Tully 1990) is not competitive with the jet in exciting the NLR clouds.

HHO's Most of the classical systems have been done with 3D. The data support more sophisticated models than for AGN because the shocks are localized and small enough to be oblique and almost mono-velocity.

Figure 2. Emission-line ratios that are robust to reddening uncertainties (DS.) Models are gridded by shock velocity and B-field, for cases with/out a radiative precursor to the shock. Mostly vertical tracks are naive photoionization models. Clustered points are the "anomalous arms" of NGC 4258 (§8) near the nucleus, where lines have widths comparable to models.

HST Followup Underway Shocks are far more prominent to HST than to earthlings because they are largely unresolved even from space, so have higher surface brightness. We need to accumulate more examples before a UV spectroscopy campaign is attractive (HST orbit politics), including robust reddening estimates. Kraemer and Crenshaw are successfully using the STIS "slitless" mode to reconnoiter several extended NLR's. If the spatio-kinematic distribution of emitting filaments is favorable, then this mode is so successful that it is difficult to get time to do the job "properly" with slits!

6. General Results on Outflows

Fountain bubbles These wind-blown structures have large opening angles that can be mistaken for ionization cones (Fig. 3-right.) There is a well-defined theoretical framework for energy input by an impulse or a continuous wind, based on robust self-similar solutions (Weaver *et al* 1977.) Complications include B-

Figure 3. Left: Structures in a high-speed shock, scaled to 400 km s^{-1} outflow into a typical spiral galaxy ISM (adapted from DS.) Right: 3D is necessary because the bow shape is usually resolved (Hartigan *et al* 1987.) Planar shock models are “painted” across the bow shape according to their perpendicular velocity component. The layer of shocked gas is assumed to be thin and locally ionized, not always valid for AGN shocks (Ferruit *et al* 1997.)

field, fluid mixing layers at dynamical instabilities, source dimming, reionization of fossil features by AGN photons, unobserved energy bands, and buoyancy.

Bow Shocks – 3D Lessons for AGN’s from Herbig Haro Objects Ionization is entirely local in the HHO shock, see Fig. 4. HHO’s provide many lines on clean backgrounds. HST sees oblique shocks along the jets (Heathcote *et al* 1996), coupled with proper motions over several epochs. This bears on entrainment of the ISM. Raymond *et al*’s (1994) suggestion of prompt entrainment was based on patterns in the 3D spectra of HH 47 (Fig. 5), and was subsequently supported by HST images. Shock velocities are low enough that no “cross-ionization” occurs from the highest velocity region near the shock apex into the wings, not the case in AGN’s (Ferruit *et al* 1997.) There are often good abundance constraints because there are many lines and many ionization stages (in the past from long-slits, now from TTF with spatial binning), as well as spectra of the host H II region. This redundancy improves the predictive power of shock models.

3D FP spectra have isolated the Mach disk (jet shock) by its excitation, kinematics, and geometry. Models have good microphysics (no need to include the source star), schematic geometry, but poor hydro. Models give the ionization fraction hence flow mass, and constrain the ambient B-field (which alters line ratios but is not strong enough to alter hydrodynamics.) Comparison with models shows a mismatch between the inferred shock speed and the observed kinematics. Fainter bow shocks at larger radii (*much* larger in results by Bally 1998) confirm this result. The interpretation is that Fu-Orionis-type ejections have occurred every few hundred years over a long timespan.

AGN Bow Shocks Radio maps from the Ulvestad & Wilson surveys were important motivators. Spirals are preferred environments because of their dense ISM. Some AGN bow-shocks do appear to be moving into a more rarified medium (e.g. TIGER IFU data on the N bow-shock in NGC 1068, Pecontal *et al* 1997.)

Shock velocities can be high enough to generate a radiative precursor, the optical spectra of which can resemble an NLR photoionized by the AGN (see Fig. 2.) The good news is that 3D coverage allows spatial binning to build up the S/N for faint lines (exploited by TTF.) At present there are limited

Figure 4. Data and best-fit models of HHO bow shocks are compared spatially (top) and spectrally (bottom). From Morse *et al* (1992.) Differences arise because the model ignores both the gas cooling distance behind the shock and emission from the Mach disk and cloulet in the flow.

Figure 5. Top left: maps of high- and low-velocity material around the jet of HH 47 (Hartigan *et al* 1993). High-velocity material is enveloped by low-, the pattern expected for "flow in a pipe". Cecil, Bland, & Tully (1990) found a similar pattern along the jet in the NLR of NGC 1068. Right: HST confirmation of "prompt entrainment" along the boundaries of the jet, at various oblique shocks (Heathcote *et al* 1996.)

Figure 6. Left: VLA image (from S. Caganoff) and our ground and HST images of NGC 3079, an edge-on, starburst galaxy with $>2000 \text{ km s}^{-1}$ outflow.

abundance constraints, and few molecular-IR fluxes. In HHO and AGN shocks, gas densities are high enough for prompt radiative cooling. For AGN shocks, $L_{\text{cool}} = 11.3 V_{300}^4 / n_0 \text{ pc}$ (units of 300 km s^{-1} and cm^{-3} .)

AGN 3D spectra are often limited to a few lines on a messy background (with poorly constrained temperatures), while radio jets with uncertain thrust may play a significant dynamical role (Bicknell *op cit.*) Radio data have advantages: no extinction so the jet can be tracked, while synchrotron spectra track reacceleration sites. The geometry delineated by masers is a tremendous constraint if it can be linked to structures on the HST/VLA scale. UV images and spectra track the shock with high contrast, and constrain the internal properties of clouds (column length, dust content.) Post-shock gas is prominent in X-rays, but its abundances and kinematics (§9) cannot be measured until CXO flies.

7. NGC 3079 – Fountain Flow

NGC 3079 is an almost edge-on spiral with energetic star formation in its disk (Fig. 6.) The most prominent feature is a 1 kpc-diameter, ovoidal superbubble with an apex on the nucleus (Fig. 7.) It is composed of dense gas that apparently depolarizes embedded relativistic particles, so the bubble is not traced in the radio. Instead, several rings at larger nuclear distance but of similar size suggest

Figure 7. HST highlights of the nuclear superbubble of NGC 3079 in $[\text{N II}]+\text{H}\alpha$, together with simulations by Suchkov *et al* (1994.) Patterns probably from Rayleigh-Taylor instabilities are evident at the top of the bubble. Gas concentrates into the walls. These are often unresolved at HST dithered WFC resolution, tightening the constraint on the gaseous filling factor that we derived from our ground-based FP spectra (Veilleux *et al* 1994.) (Because my space is limited, Veilleux discusses these results in his contribution.)

an episodic process. We used an FP to sort out the geometry (Veilleux *et al* 1994), and, based on robust energy arguments, found a possible role for an AGN.

HST Tightens the Constraint on the filling factor, hence KE and momenta, $\approx 100\times$ over our FP spectra. It is now less likely that the AGN contributes. HST also shows that the whole disk is ejecting narrow gas and dust filaments.

8. NGC 4258 – Finding the Active Jet

Efforts on NGC 4258 (M 106) have focused on understanding the nature of the large-scale (5') "anomalous arms" in this nearby spiral galaxy (Fig. 8.) Recently, the spatio-kinematic distribution of masers has provided unique constraints on the accretion disk around the SMBH (Miyoshi *et al* 1995.) Jets and accretion disks are certainly linked in the minds of theorists, but observational connections are tenuous. NGC 4258 can provide unique constraints if the active flow can be traced in a continuous fashion to link with the VLBA results at small radii.

Previous Work The anomalous arms emit radio synchrotron, optical lines, and most of the X-rays of the galaxy. Synchrotron traces active and fossil particle flows, optical lines have flux ratios consistent with shock excitation (Fig. 2,

Figure 8. $\lambda 20$ cm radio images of NGC 4258. The main image spans a region $3 \times 3'$. The insert shows $3 \times$ the area, and the full outflow across much of the galaxy at lower resolution and with more sensitivity to diffuse emission. The active jet flows almost NS and ends at two prominent hotspots.

Cecil, Morse, & Veilleux 1995), and the large spatial extent of the thermal (Cecil, Wilson, & DePree 1995) X-rays along the anomalous arms also indicate fast shocks consistent with the observed kinematics. But where is the active jet?

New Results Our VLA image (Fig. 8) has sufficient sensitivity to isolate the active jet within the anomalous arms. We find that it projects along the spin axis of the accretion disk. The active flow ends in bow-shocks (Fig. 9), both well modeled by kinematics and excitation as laying near the galaxy disk (Fig. 10.) The observed gas *backflows* from the working surfaces, around the stalled jet. The shock velocity from our models is exactly that expected from our measured line ratios. The Balmer decrement yields $A_V = 1$ for both bow shocks. There are “fossil” versions of these features that lag galaxy rotation at similar radii.

9. Coming Attractions

Promise and Perils of the UV High-speed (several hundred km s^{-1}) shocks cool by UV emission. It is straightforward to distinguish between shocks and photionization in the theoretical fantasy world of a single shock velocity or a restricted (or “optimally emitting”) population of photoionized clouds. Having to contend with arbitrary distributions of either of these blurs the distinction. Moreover, the UV lines are very sensitive to extinction (left panel Fig. 11.)

Figure 9. Left: HIFI FP results on NGC 4258 (UH 2.2m; Cecil, Wilson, & Tully 1992, CWT) Right: HST F658N-791W mosaic of the anomalous arms in $[N II]\lambda 6583 + H\alpha$, same region plotted in Fig. 8, log scaled. Jet/ISM interactions are circled, with dark circles around the N & S bow shocks at the ends of the active jet. The fastest gas is at the bow shocks, but there are $50\text{--}100\text{ km s}^{-1}$ discontinuities across each strand of the anomalous arms (Fig. 10 left.) The CO forms a channel around the active jet (Neininger 1999.)

Figure 10. (From Cecil *et al* 1999.) Data (left panel) and model (right top) long-slit spectra (3D because we placed the slit at two adjacent points!) that intersect the S and N (insert) bow shocks at the ends of the active jets in NGC 4258. In the main figure, emission lines are $H\beta$, $[O III]$, $[N II] + H\alpha$, and $[S II]$; note the different appearance of the $[O III]$ profiles, a characteristic of shocks. The model (based on structures shown in Fig. 4) reproduces the spatio-kinematic variations in the data if the flow is a backwash at $\approx 350\text{ km s}^{-1}$ around a stalled jet. Coronal Fe lines expected behind a fast shock (e.g. Wilson & Raymond 1999) are not in the data. HST images of the bow shocks are inserts at the same scale and orientation as our long slits. Note that the Mach disk (the barely resolved knot enveloped by the bow-shock in the HST insert of the S bow shock) is isolated by these WHT ISIS spectra. As for HHO's, my model includes no emission from the Mach disk. Nor have I bothered to reproduce local flux variations in the photoionized precursor.

Figure 11. Simulated UV & CXO spectra of shocks in NGC 1068 (left) & NGC 4258 (right). Leftmost panel: average extinction expected across the NLR of NGC 1068 that inhibits STIS observations. The second column shows average line spectra expected from shocks of various velocities and magnetic parameters $B/\sqrt{n}$, with the power-law photoionized spectrum from a family of matter-bounded clouds at top. Right: model spectra showing the resolution obtained with the CXO CCD (left panels) and grating (right) spectrometers and a 150-ksec exposure after combining counts along the 5'-long jets.

However, the top spectrum in the middle panel of Fig. 11 is a composite of matter-bounded clouds, and is clearly distinguishable in the UV from shocks. Our program on NGC 1068 is time-consuming yet shallow in flux limit because the STIS slit is narrow and must be stepped across the complex NLR. 3D please!

CXO: measuring the T & N_H of hot, post-shock gas Of course extinction is not a problem for energies above 0.2 keV, an energy regime accessed by gas as it cools behind a fast ($> 300 \text{ km s}^{-1}$) shock (Fig. 3 left panel.) CXO will finally combine spatial and spectral resolution with enough sensitivity to map this hot gas in the nearest, largest bow shocks. Two imaging spectrometers on CXO will play roles (Fig. 11 right panels.) The ACIS CCD will obtain $R \approx 100$ spectra (from Fano effect) across spatially complex regions. Where there are narrow jets and compact shocks, the slitless grating with the HRI offers slightly better spatial resolution and far greater spectral resolution, at the expense of much lower throughput and more limited areal coverage.

For example, our program on the anomalous arms of NGC 4258 (total $L_X \approx 2 \times 10^{40} \text{ erg s}^{-1}$) will be one of the longest in Cycle 1, yet will still require that counts be binned along 15"-long segments of the arms (see Fig. 11.) Our 3D FP data (CWT) show kinematical gyrations in each such interval. The pay-off is that we should be able to measure for the first time the chemical abundances, N_H , and perhaps spatially averaged kinematics of the hot, entrained gas.

The larger effective area of XMM will allow spectra of extended winds such as that which ruptures the superbubble of NGC 3079 (Pietsch *et al* 1998). The walls of the superbubble remain good targets for the CCD on CXO, and will provide unique constraints on gas mixing along dynamically unstable layers.

10. Summary

Despite obvious source points, AGN and HHO flows still need the serendipity lacking with slits. Reconnaissance with FP's has found kinematic discontinuities characteristic of shocks. The TTF has increased efficiency several-fold, and is now reaching deep enough to map the widespread influence of an AGN on its host and to flux faint nebular diagnostic lines. Using such instrumentation, a 4m telescope at a superb site can learn much on the host/AGN connection. Facility IFU spectrometers to appear soon on 8+m telescopes – coupled with AO on nearby objects – will finally compel most spectroscopists to become three-dimensional folk, and incidentally accelerate the study of AGN outflows.

This work was supported by NSF and NASA grants to UNC-CH. My collaborators are referenced below, with thanks for their encouragement and support.

References

- Allen, M. *et al* 1998, ApJ, 493, 571
 Bally, J. 1998, NOAO Newsletter, December issue
 Bicknell, G. V., *et al* 1998, ApJ, 495, 680
 Binette, J., Dopita, M.A., & Tuohy, I.R. 1985, ApJ, 297, 476
 Cecil, G., Bland, J., & Tully, R.B. 1990, ApJ, 355, 70
 Cecil, G., Morse, J. A., & Veilleux, S. 1995, ApJ, 452, 613
 Cecil, G., Wilson, A.S., & De Pree, C. 1995, ApJ, 440, 181
 Cecil, G., Wilson, A.S., & Tully, R. B. 1992, ApJ, 390, 365
 Cecil, G. *et al* 1999, submitted to ApJ
 Dopita, M.A., & Sutherland, R. S. 1995, ApJ, 455, 468 (DS)
 Ferruit, P., Binette, L., Sutherland, R.S., & Pecontal, E. 1997, A&A, 322, 73
 Ferruit, P. *et al* 1999, MNRAS, in press
 Hartigan, P. *et al* 1993, ApJ, 414, 121
 Hartigan, P., Morse, J.A., & Raymond, J.C. 1994, ApJ, 436, 125
 Hartigan, P., Raymond, J., & Hartmann, L. 1987, ApJ, 316, 323
 Heathcote, S. *et al* 1996, AJ, 112, 1141
 Miyoshi, M. *et al* 1995, Nature, 373, 127
 Morse, J.A., Cecil, G., Wilson, A.S., & Tsvetanov, Z.I. 1998, ApJ, 505, 159
 Morse, J.A. *et al* 1992, ApJ, 399, 231
 Neninger, N. 1999, astro-ph/9903230
 Pecontal, E. *et al* 1997, Ap&SS, 248, 167
 Pietsch, W. *et al* 1998, A&A, 340, 351
 Raymond, J.C. *et al* 1994, ApJ, 434, 232
 Reipurth, B. *et al* 1997, AJ, 114, 757
 Shopbell, P.L., & Bland-Hawthorn, J. 1998, ApJ, 493, 129 A&A, 322, 73
 Steffen, W. *et al* 1997, ApJ, 491, 73
 Suchkov, A. *et al* 1994, ApJ, 430, 511
 Ulvestad, J. & Wilson, A. S. 1984, ApJ, 278, 544
 Veilleux, S. *et al* 1994, ApJ, 433, 48
 Weaver, R., *et al* 1977, ApJ, 218, 377
 Whittle, M. 1985, MNRAS, 213, 189
 Wilson, A.S. & Raymond, J.C. 1999, ApJ, 513, 115

Fabry-Perot Imaging Spectroscopy of Starburst and AGN Winds

S. Veilleux

*Department of Astronomy, University of Maryland, College Park, MD
20742*

Abstract.

To date, the most detailed studies of galactic winds have come from 3-D spectrophotometric observations with radio and Fabry-Perot interferometers. Here, we report the latest results from a long-term optical survey of nearby active and starburst galaxies with the Hawaii Imaging Fabry-Perot Interferometer (HIFI) at Mauna Kea and the TAURUS-2 system in Australia. These data reveal that the outflows are highly complex, highly energetic ($> 10^{53}$ ergs in most cases), and the brightest emission often appears to be associated with strong shocks. The outflowing material in the starburst galaxies generally lies on the surface of bubbles or along the walls of funnel-shaped winds rapidly accelerating out of the galactic plane. These winds are sometimes lop-sided and tilted with respect to the polar axis of the host galaxy. Evidence for entrainment of (rotating) disk material is seen in some objects. Our results are combined with HST, radio and X-ray data and discussed in the context of future surveys of distant galaxies on 8-meter class telescopes.

1. Introduction

Active galactic nuclei (AGN) and nuclear starbursts may severely disrupt the gas phase of galaxies through deposition of a large amount of mechanical energy in the centers of galaxies. As a result, a large-scale galactic wind ("superwind") that encompasses much of the central regions of these galaxies may be created (e.g., Chevalier & Clegg 1985; Schiano 1985). Depending upon the extent of the halo and its density and upon the wind's mechanical luminosity and duration, the wind may ultimately blow out through the halo and into the intergalactic medium. The effects of these superwinds may be far-reaching. Bregman (1978) has suggested that the Hubble sequence can be understood in terms of a galaxy's greater ability to sustain winds with increasing bulge-to-disk ratio. Superwinds may affect the thermal and chemical evolution of galaxies by depositing large quantities of hot, metal-enriched material on the outskirts of galaxies. They also offer a natural way to create a cosmically evolving population of large, metal-enriched, kinematically-complex gaseous halos, in many ways resembling the sharp metal lines and Lyman-limit systems observed in quasar spectra.

Strong evidence for spatially-resolved superwinds now exists in several nearby starburst, Seyfert, and dwarf galaxies (e.g., Bland & Tully 1988; Cecil et al. 1990; Heckman et al. 1990; Lehnert & Heckman 1996; Marlowe et al. 1995;

Meurer et al. 1992; Veilleux et al. 1994; Colbert et al. 1996). Our group is combining Fabry-Perot imaging spectrophotometry with radio and X-ray data to track the energy flow through various gas phases. The complete spatial and kinematic sampling of the Fabry-Perot (FP) data is ideally suited to study the complex and extended morphology of the warm line-emitting material that is entrained in the wind flow. The radio and X-ray data complement the FP data by probing the relativistic and hot gas components, respectively. The high level of sophistication of recent hydrodynamical simulations (e.g., Tomisaka 1990; Slavin & Cox 1992; Mineshige et al. 1993; Suchkov et al. 1994) has provided the theoretical basis to interpret our data and to predict the evolution and eventual resting place (disk, halo, or intergalactic medium) of the outflowing material. In the present paper, we summarize the results on three representative objects from our Fabry-Perot survey and discuss possible implications.

2. The Fabry-Perot Survey

Using the gap-scanning mode of the Hawaii Imaging Fabry-Perot Interferometer (HIFI) on Mauna Kea Observatories and of the TAURUS-2 Fabry-Perot system on the Anglo-Australian Telescope, our group is carrying out a survey of twenty nearby ($z < 0.01$) starburst and Seyfert disk galaxies. Deep tunable-filter images and stare-mode FP spectra supplement some of the data cubes. The objects in our sample were selected on the basis of *a priori* evidence for large-scale nuclear outflows.

So far, high-quality data cubes have been obtained and analyzed for about a dozen galaxies, and the results have been published for ten of them. Our FP data set allows us to perform spectrophotometric analyses of the line-emitting gas at typically 10,000 – 100,000 positions across the extent of our sample galaxies. These data therefore provide very stringent constraints on the general flow pattern of the line-emitting gas entrained in these outflows.

3. Results

Because of space limitations, we focus our discussion of the results on three of our sample galaxies. These objects were chosen to illustrate the broad diversity of morphologies, kinematics, and energetics associated with galactic winds.

3.1. M82

This prototypical starburst galaxy has long been suspected to host a galactic-scale outflow (e.g., Lynds & Sandage 1963; Burbidge, Burbidge, & Rubin 1964). A recent HST image of this galaxy (Fig. 1) shows the well-known filamentary complex that extends several kpc above and below the disk of M82. The FP data from Shopbell & Bland-Hawthorn (1998) reveal a bipolar outflow of material that originates from the bright starburst regions in the galaxy's inner disk but is misaligned with respect to the galaxy spin axis. The deprojected outflow velocity indicated by the optical filaments increases with radius from 525 to 655 km s^{-1} . Double components are detected in the centers of the outflowing lobes,

Figure 1. Continuum-subtracted $H\alpha + [N II]$ HST/WFPC2 image of M82 showing dramatic filamentation and bow shock structures in the outflowing lobes. The size of the smallest resolved features in this image is about 2 pc (Shopbell et al. 1999)

with line splitting by $\sim 300 \text{ km s}^{-1}$ over a region almost 1 kpc in size. The lobes lie along an axis tilted by 15° with respect to the spin axis of the galaxy.

The filaments are not simple surfaces of revolution, nor is the emission distributed evenly over the surfaces. These lobes are best modeled as a composite of cylindrical and conical structures, collimated in the inner ~ 500 pc but expanding at a larger opening angle of $\sim 25^\circ$ beyond that radius. The wind in M82 therefore seems to be freely flowing into the galaxy halo (“free-wind” phase in the nomenclature of Weaver et al. 1977). Using this outflow geometry and assuming a filling factor of 0.1, Shopbell & Bland-Hawthorn (1998) finds that a kinetic energy of $\sim 2 \times 10^{55}$ is involved in the outflow. There is also some evidence for rotation of the wind filaments about the outflow axis in support of entrainment.

The observed filamentation probably arises from large-scale shocks from the high-speed wind plowing into the gaseous halo and entrained disk material. The line ratios suggest that photoionization by the nuclear starburst play a significant role in the excitation of the optical filament gas, but that shock ionization becomes increasingly important at large radii.

Figure 2. (left panel) Continuum-subtracted $H\alpha + [N II]$ HST/WFPC2 image of the nuclear bubble in NGC 3079. The spatial resolution is about 8 pc. (upper right panels) Results of numerical simulations from Suchkov et al. (1994). Note the resemblance with the fine structures observed in the HST image. (lower right panel) A view of the star-forming disk in NGC 3079 derived from the HST images. Elevated line-emitting chimneys and filaments are seen near the brightest $H II$ regions. (Cecil et al. 1999)

3.2. NGC 3079

In no other galaxy is the impact of superwinds more evident than in NGC 3079, a nearby edge-on spiral galaxy which is host to a spectacular kpc-scale bubble (see Fig. 2). Violent gas motions that range over $2,000 \text{ km s}^{-1}$ are detected across the bubble and diametrically opposite on the other side of the nucleus. The unusual gaseous excitation (e.g., $[N II] \lambda 6583/H\alpha > 1$) of the line-emitting gas in the bubble region confirms that shocks are important.

This is the most powerful example known of a windblown bubble ($\sim 2 \times 10^{56} \text{ N}_e^{-1}$), and an excellent laboratory to study wind dynamics. An ovoidal bubble, inflated from the nucleus with monotonically increasing velocities ($V \propto R^n$ with $2 < n < 3$) and inclined $\sim 3^\circ$ from the plane of the sky, provides a good first-order fit to the observed velocity field. The dimensions and energies of the bubble imply that it is in the blowout phase and partially ruptured. A detailed dynamical analysis of this outflow indicates that the wind alone can contribute up to $5t_{\text{outflow},8}\%$ of the total metal content of the host galaxy, where $t_{\text{outflow},8}$ is the outflow lifetime in units of 10^8 yr (Veilleux et al. 1994). The core of NGC 3079 harbors both a nuclear starburst and an AGN. The nature of the energy source that drives the outflow is not clear at present. The poorly constrained electron density and filling factor of the warm line-emitting

material in the bubble are the main sources of uncertainty in the calculation of the energetics. To shed new light on this issue, we recently obtained a WFPC2 HST image of NGC 3079 with resolution of $\sim 0''.1$ (~ 8 pc; Fig 2). The HST image reveals intricate patterns in the line-emitting material near the top of the bubble. These features share a remarkable resemblance with those observed in hydrodynamical simulations (e.g., Suchkov et al. 1994) – they are probably the signatures of Rayleigh-Taylor instabilities in the entrained material. Similar filamentary structures are observed near the brightest star-forming regions in the disk of NGC 3079, bringing further support to the idea of a dynamically active disk in this galaxy (Veilleux et al. 1995). A preliminary quantitative analysis of the HST data suggests that the AGN in NGC 3079 is not contributing significantly to the nuclear outflow.

3.3. Circinus

The FP data on Circinus show a complex of ionized filaments extending radially from the nucleus out to distances of 1 kpc (Fig. 3). The velocity field of the filaments confirms that they represent material expelled from the nucleus or entrained in a wide-angle wind roughly aligned with the polar axis of the galaxy. Extrapolation of these filaments to smaller radii comes to within $1''$ of the active galactic nucleus, therefore suggesting a AGN or nuclear starburst origin to these features. The outflow involves a fairly modest kinetic energy ($\sim 10^{53} N_{e,2}^{-1}$ ergs) and therefore appears to lie at the low energy end of the distribution for wide-angle events observed in nearby galaxies.

The complex of radial filaments and bow shocks detected in the Circinus galaxy is unique among active galaxies and does not fit within the standard evolutionary picture of windblown bubbles (e.g., Weaver et al. 1977). It is not clear at present why that is the case. The discovery of these features in the Circinus galaxy, a spiral galaxy with an abnormal richness of gas (Freeman et al. 1977), brings up the possibility that we may be witnessing a common evolutionary phase in the lives of gas-rich active galaxies during which the dusty cocoon surrounding the nucleus is expelled by the combined action of jet and wind phenomena.

4. Future Avenues of Research

The overall agreement between the simulations of Suchkov et al. (1994) and current observations of local galaxies with galactic winds augurs well for the future. However, substantial quantitative differences still remain: e.g., current models severely underestimate the outflow velocities of the entrained line-emitting material (by more than an order of magnitude in NGC 3079). A more realistic treatment of the multi-phase ISM in the host galaxy may help solve some of these problems and may help explain the morphological peculiarities of the outflow in the gas-rich galaxy Circinus.

The results from these observational and theoretical studies will eventually serve as a critical local baseline for future deeper surveys with IFUs on 8-meter class telescopes from the ground and in space. Current surveys suggest that galaxies have experienced a very active phase of star formation and nuclear activity around redshifts of 1-3 (e.g., Madau, Pozzetti, & Dickinson 1998;

Figure 3. Line flux images of Circinus, the nearest Seyfert galaxy: *a*, [O III] $\lambda 5007$ and *b*, blueshifted $H\alpha$. The position of the infrared nucleus is indicated in each image by a cross. The spatial scale, indicated by a horizontal bar at the bottom of the [O III] image, corresponds to ~ 25 arcsec or 500 pc. Note the unusual complex of radial filaments emerging from the nucleus. $PA_{\text{major}}(\text{disk}) = 30^\circ$, $i(\text{disk}) = 65^\circ$ (Veilleux & Bland-Hawthorn 1997)

Schmidt, Schneider, & Gunn 1995). Detailed comparisons of high-redshift galaxies with local superwind hosts and with state-of-the-art numerical simulations of windblown bubbles will help us quantify the impact starburst- and AGN-driven winds may have had on the chemical and thermal evolution of the galactic and intergalactic environments.

The survey described in this paper is done in collaboration with J. Bland-Hawthorn, G. Cecil, P. Shopbell, and R. B. Tully. The author is grateful for support of this research by a Cottrell Scholarship awarded by the Research Corporation, NASA/LTSA grant NAG 56547, and NSF/CAREER grant AST-9874973.

References

- Bland, J., & Tully, R. B. 1988, *Nature*, 334, 43
- Bregman, J. N. 1978, *ApJ*, 224, 768
- Burbidge, E. M., Burbidge, G. R., & Rubin, V. C. 1964, *ApJ*, 140, 942
- Cecil, G. N., Bland, J., & Tully, R. B. 1990, *ApJ*, 355, 70
- Cecil, G. N., et al. 1999, in preparation.
- Chevalier, R. A., & Clegg, A. W. 1985, *Nature*, 317, 44
- Colbert, E., et al. 1996, *ApJS*, 105, 75
- Freeman, K. C., et al. 1977, *A&A*, 55, 445

- Heckman, T. M., Armus, L., & Miley, G. K. 1990, *ApJS*, 74, 833
Lehnert, M. D., & Heckman, T. M. 1996, *ApJ*, 462, 651
Lynds, C. R., & Sandage, A. R. 1963, *ApJ*, 137, 1005
Madau, P., Pozzetti, L., & Dickinson, M. 1998, *ApJ*, 498, 106
Marlowe, A. T., et al. 1995, *ApJ*, 438, 563
Meurer, G. R., et al. 1992, *AJ*, 103, 60
Mineshige, S., Shibata, K., & Shapiro, P. R. 1993, *ApJ*, 409, 663
Schiano, A. V. R. 1985, *ApJ*, 299, 24
Schmidt, M., Schneider, D. P., & Gunn, J. E. 1995, *AJ*, 110, 68
Shopbell, P. L., & Bland-Hawthorn, J. 1998, *ApJ*, 493, 129
Shopbell, P. L., et al. 1999, in preparation
Slavin, J. D., & Cox, D. P. 1992, *ApJ*, 392, 131
Suchkov, A. A., Balsara, D. S., Heckman, T. M., & Leitherer, C. 1994, *ApJ*, 430, 511
Tomisaka, K. 1990, *ApJ*, 361, L5
Veilleux, S., & Bland-Hawthorn, J. 1997, *ApJ*, 479, L105
Veilleux, S., Cecil, G., Bland-Hawthorn, J. 1995, *ApJ*, 445, 152
Veilleux, S., et al. 1994, *ApJ*, 433, 48
Weaver, R., et al. 1977, *ApJ*, 218, 377

Deep TTF Imaging of Powerful Radio Galaxies

C. Tadhunter¹, M. Villar-Martin², R. Morganti³

¹*Department of Physics and Astronomy, University of Sheffield,
Sheffield S3 7RH, UK*

²*Institute d'Astrophysique, 98bis Boulevard Arago, F-75014, Paris,
France*

³*Instituto di Radioastronomia, CNR, via Gobetti 101, Bologna, Italy*

Abstract. Deep emission line imaging observations taken with the TAU-RUS Tunable Filter (TTF) on the 4.2m WHT telescope are presented for two nearby radio galaxies with strong jet-cloud interactions. The images show a broad distribution of emission line gas and challenge the conventional view that the extended ionized haloes of powerful radio galaxies are photoionized by the radiation cones of quasars hidden in the cores of the galaxies. Rather, the close morphological association of the emission line filaments with the radio jets, lobes and hotspots suggests that these may be objects in which the ionization of the extended gas is entirely dominated by shocks induced by interactions with the radio-emitting components. We expect future Tunable Filter observations of high redshift ($z > 0.5$) radio galaxies to reveal similar low-surface brightness emission line structures well away from the high-surface-brightness structures aligned along the radio axis. Such observations have the potential to reveal the true distributions of warm gas in the host galaxies.

1. Introduction

The extended emission line nebulae around powerful radio galaxies provide potentially valuable information about the origins of the radio activity and the evolution of the host galaxies. However, despite decades of the study, the ionization mechanism for the nebulae remains controversial. On the one hand it has been proposed that the nebulae are photoionized by radiation cones of powerful quasars hidden in the cores of the galaxies (e.g. Fosbury 1989), while on the other it has been proposed that the nebulae have been ionized by shocks generated in interactions between the radio-emitting components and the warm ISM (e.g. Sutherland et al. 1993). Evidence exists to support both models: the polarimetric detection of scattered quasar features in many powerful radio galaxies demonstrates that anisotropic quasar illumination *must* be important at some level (e.g. Cimatti et al. 1997; Ogle et al. 1997), yet jet-cloud interactions are required to explain the highly-collimated structures (e.g. Best et al. 1996) and extreme emission line kinematics (e.g. McCarthy, Baum & Spinrad 1996) of the emission line regions aligned along the radio axes of many high redshift sources

(McCarthy et al. 1987). It is clear that both AGN illumination and jet-cloud interactions are required to explain the phenomenology of powerful radio galaxies, but the balance between the two mechanisms remains to be determined.

A particular problem concerns the fact that, despite the polarimetric evidence for quasar illumination, the emission line structures in many high redshift sources show jet-like structures which are more highly collimated than the broad — $45\text{--}60^\circ$ opening half-angle — quasar radiation cones predicted by the unified schemes for powerful radio galaxies (Barthel 1989, Lawrence 1991). This raises the possibility that the warm/cool ISM in the host galaxies may be intrinsically aligned along the axes of the radio jets. Alternatively, there may be objects in which the ionization of the extended ionized gas is entirely dominated by jet-induced shocks.

Unfortunately, the published images of the high- z sources are either relatively shallow and have a low spatial resolution or, in the case of HST images, have a high spatial resolution, but are insensitive to low-surface-brightness structures. Thus, the existing images are biased towards the high-surface-brightness structures aligned along the radio axes and may give a misleading impression of the true distribution of the warm ISM in the host galaxies. In order to investigate this possibility we have carried out a pilot imaging study of two intermediate redshift radio galaxies — 3C171 ($z = 0.2381$) and 3C277.3 ($z = 0.08579$) — which are nearby prototypes of the high redshift radio galaxies in the sense that they show high-surface-brightness emission line structures aligned along their radio axes (Heckman et al. 1984, van Breugel et al. 1985). The images were taken using the TAURUS Tunable Filter (TTF) in subarcsecond seeing conditions on the 4.2m WHT telescope on La Palma (Bland-Hawthorn & Jones 1998). In each case we made emission line observations with the TTF etalon tuned to the redshifted wavelength of $H\alpha$, and continuum observations with the etalon removed and an intermediate-band filter in the beam. The exposure total times for the $H\alpha$ and continuum images were 1800s and 900s respectively, and the images reach an $H\alpha$ surface brightness of $\sim 1 \times 10^{-17} \text{ erg s}^{-1} \text{ cm}^{-2} \text{ sq. arcsec}^{-1}$. The preliminary results reported here demonstrate the potential of TTF observations for studies of active galaxies and challenge some commonly held assumptions about the ionization of the extended ionized haloes around powerful radio galaxies. A fuller discussion of the results of this study is presented in Tadhunter et al. (1999).

2. TTF Imaging Results and Discussion

The TTF images of 3C277.3 and 3C171 are shown in Figures 1 and 2. In both cases the $H\alpha$ images reveal spectacular systems of emission line filaments well away from the radio axes of the host galaxies. From the fact that these structures are not visible at the same level in the continuum images, we conclude that they are predominantly gaseous structures.

The main results can be summarized as follows:

- **Distribution of warm ISM.** The images reveal the presence of low-surface-brightness structures at large distances from the radio jet axes. Clearly, the warm/cool ISM in the host galaxies is not intrinsically aligned along the axes of the inner radio jets.

Figure 1. TTF images of 3C277.3 in the light of $H\alpha$ +continuum (left: $\lambda_c = 7124\text{\AA}$, $\Delta\lambda = 19\text{\AA}$), and line-free continuum (right: $\lambda_c = 7580\text{\AA}$, $\Delta\lambda = 280\text{\AA}$). In the images the “G” symbols indicate galaxies, while the “S” symbol indicates a faint star (unresolved in HST images). The line segment in the $H\alpha$ +continuum image indicates the inner radio axis. The spatial scale in kpc has been calculated for $H_0 = 50 \text{ km s}^{-1} \text{ Mpc}^{-1}$ and $q_0 = 0.0$.

- **AGN Illumination.** The emission line images cannot be reconciled with a simple AGN illumination model. Not only do some of the low-surface brightness emission line structures in both 3C171 and 3C277.3 lay well outside the $45\text{--}60^\circ$ opening half-angle cones predicted by the unified schemes, but there are no sharp edges that would suggest the boundaries of ionization cones, and at least one of the filaments in 3C277.3 wraps a full 180° in PA around the nucleus. It is possible to produce broader distribution of gas than the nominal $45 - 60^\circ$ ionization cones by tilting the cone axes out of the plane of the sky. However, in order to explain the observations in this way the cones would have to be tilted to such an extent that our line of sight would fall within the cones and we would see the AGN directly. Clearly this is not the case.
- **Relationship to the radio structures.** Optical/radio overlays presented in Tadhunter et al. (1999) show that the extended low-surface-

Figure 2. TTF images of 3C171 in the light of $H\alpha$ +continuum (left: $\lambda_c = 8124\text{\AA}$, $\Delta\lambda = 28\text{\AA}$), and line-free continuum (right: $\lambda_c = 8570\text{\AA}$, $\Delta\lambda = 400\text{\AA}$). The symbols are the same as for Figure 1

brightness filaments in both 3C171 and 3C277.3 are associated with the outer edges of the extended radio lobes. This suggests that the ionization of the gas is in some way linked to the radio-emitting plasma.

Given that there is no independent evidence for powerful AGN hidden in the nuclei of these two sources, the most plausible explanation for the emission line structures is that we are seeing the results of interactions between the expanding radio lobes and a pre-existing distribution of warm ISM in the host galaxies. These may be objects in which the ionization of the gas is entirely dominated by shocks induced by interactions with the radio-emission, with little contribution from direct AGN photoionization.

3. Conclusions and future prospects

The imaging observations presented here demonstrate the potential of Tunable Filter observations for both revealing the true distribution of warm ISM in the host galaxies and investigating the ionization mechanisms. The results challenge the conventional view that the distributions of emission line gas observed in

many low redshift radio galaxies are a consequence of broad-cone illumination by quasars hidden in the cores of the galaxies. They also suggest that the high-surface-brightness structures aligned along the radio axes of many high redshift ($z > 0.5$) radio galaxies give a misleading impression of the true distribution of gas in the host galaxies. We expect future deep Tunable Filter observations of high redshift sources to reveal extensive haloes of warm ISM, similar to those observed in their low redshift counterparts.

We would like to thank Joss Bland-Hawthorn for help with the observations.

References

- Barthel, P.D., 1989, ApJ, 336, 606
Best, P.N., Longair, M.S., Rottgering, H.J.A., 1996, MNRAS, 280, L9.
Bland-Hawthorn, J. & Jones, D.H. 1998, Publ. Astron. Soc. Austr., 15, 44
Cimatti, A., Dey, A., van Breugel, W., Hurt, T., Antonucci, R., 1997, ApJ, 476, 677
Fosbury, R.A.E., 1989: In: *ESO Workshop on Extranuclear Activity in Galaxies*, Meurs & Fosbury (eds), p169
Heckman, T.M., van Breugel, W.J.M., Miley, G.K., 1984, ApJ, 286, 509
Lawrence, A., 1991, MNRAS, 252, 586
McCarthy, P.J., van Breugel, W., Spinrad, H., Djorgovski, S., 1987, ApJ, 321, L29
McCarthy, P.J., Baum, S., Spinrad, H., 1996, ApJS, 106, 281
Ogle, P.M., Cohen, M.H., Miller, J.S., Tran, H.S., Fosbury, R.A.E., Goodrich, R.W., 1997, ApJ, 482, 370
Sutherland, R.S., Bicknell, G.V., Dopita, M.A., 1993, ApJ, 414, 510
Tadhunter, C.N., Villar-Martin, M., Morganti, R., Axon, D., 1999, MNRAS, submitted
van Breugel, W., Miley, G., Heckman, T., Butcher, H., Bridle, A., 1985, ApJ, 290, 496

3-D Observations of Nearby Active Galaxies with the IFU TIGER

P. Ferruit¹, E. Pécontal², and A. S. Wilson¹

¹ *Astronomy Department, University of Maryland,
College Park, MD 20742, USA*

² *CRAL - Observatoire de Lyon, 6 avenue Charles André,
69561 Saint-Genis-Laval CEDEX, France*

Abstract. We present the results of 3D (x,y, λ) observations of nearby Seyfert galaxies with the integral field unit TIGER at CFHT. After briefly reviewing the various nearby active galaxies previously observed with TIGER, we highlight the results obtained on one particular Seyfert galaxy, Mrk 573.

1. Introduction

Despite intensive studies, many key questions on the nature and structure of active galaxies remain unanswered. Within the zoo of AGN types, the nearby Seyfert galaxies are perhaps the best suited to try to answer some of these questions. Although less powerful than more distant AGN like quasars and radio-galaxies, their proximity allows their central regions to be probed with a relatively high absolute spatial resolution, even from the ground. In addition, they are morphologically complex and require multiple slit positions when studied with long-slit spectroscopy, but are ideal objects for integral field units (IFUs) like TIGER (Bacon et al. 1995).

Therefore, we have undertaken a program of 3D (x,y, λ) spectroscopy of nearby Seyfert galaxies in the visible at the Canada France Hawaii Telescope (CFHT), program initiated with TIGER and now continuing with its successor OASIS, developed by the CRAL - Observatoire de Lyon (PI : R. Bacon). The primary scientific goal of this program is to investigate the energy budget of the [extended] narrow line regions ([E]NLR) of Seyfert galaxies. In the following, we present a brief review of previous observations of Seyfert galaxies with TIGER (Sect. 2.), followed by description of the results obtained on Mrk 573 (Sect. 3.). These latter is intended to be an example of what can be done on this kind of object using an IFU like TIGER.

2. Previous work with TIGER

Our work on Seyfert galaxies with TIGER has consisted of detailed studies of individual objects. In the following, we list the various observed active galaxies, the observations of which have been published.

Figure 1. *Left panel:* HST/WFPC2 [O III] $\lambda 5007$ image of Mrk 573 (Falcke et al. 1998). *Right panel:* [O III] $\lambda 5007$ image of Mrk 573 reconstructed from the TIGER/CFHT 3D (x,y, λ) data cube.

- **NGC 5728** (Pécontal et al. 1990): This is one of the first observations obtained with TIGER, which was still a prototype at this time. These observations have been used to map the kinematics of the ionized gas in the central 3'' of the galaxy. A broad H α line was detected for the first time in this object.
- **NGC 1275** (Ferruit & Pécontal 1994; Ferruit et al. 1997b): In this extremely complex object, the TIGER observations have been used to disentangle the emission from the low- and high-velocity components and to study their spectral and kinematical properties. The map of the spectral characteristics of the low-velocity component is found to be consistent with photoionization by a varying mixture of nuclear radiation and local emission from young stars. Ionized gas clouds with velocities intermediate between those of the low- and high-velocity systems are detected (serendipitously) for the first time in these observations. This is the first evidence for a direct interaction between the two systems.
- **NGC 5929** (Ferruit et al. 1997a): The morphology and the spectral properties of the emission-line clouds in this object have been compared to the predictions of various models of the interaction between a jet and the interstellar medium of the host galaxy. None of the models is found to be completely satisfactory.
- **NGC 1068** (Pécontal et al. 1997): The kinematic properties of individual nuclear clouds were isolated for the first time in this object, after deconvolution of the TIGER data cubes. We have obtained the first 3D observations of a Seyfert galaxy in the [Fe VII] line. It shows the presence of a strong, blueshifted [Fe VII] component north of the nucleus. Comparison with the deconvolved [O III] data cube shows that it is associated with cloud G, which appears highly excited.

Figure 2. *Top panels:* $[\text{O III}]\lambda 5007$ image (left) and velocity field (right) of Mrk 573 obtained by fitting a single Gaussian profile on the $[\text{O III}]\lambda 5007$ line of each individual spectrum of the deconvolved TIGER/CFHT 3D data cube. The spatial resolution achieved in this deconvolved data cube is $0''.35$ (110 pc). Contours of the 3.6 cm radio map (Falcke et al. 1998) have been superimposed to both images. *Bottom panel - Central image:* $[\text{O III}]\lambda 5007$ map of Mrk 573 (see above) with the $[\text{O III}]$ isovelocity contours superimposed. *Bottom panel - Profiles:* $[\text{O III}]\lambda 5007$ velocity profiles from eight selected locations. The origin corresponds to the stellar systemic velocity of Mrk 573 (5150 km s^{-1} , Nelson & Whittle 1995).

3. The example of the Seyfert 2 galaxy Markarian 573

Mrk 573 is a Seyfert 2 galaxy located at a distance of 65 Mpc (yielding a scale of $315 \text{ pc arcsec}^{-1}$). It has been observed with TIGER at CFHT in November 1993 in two spectral configurations, the first one including $H\beta$ and $[\text{O III}]\lambda\lambda 4959, 5007$, the second one including the $[\text{N II}]\lambda\lambda 6548, 6583$, $H\alpha$ and $[\text{S II}]\lambda\lambda 6717, 6731$ lines. The spatial sampling was $0''.39$, with a seeing limited resolution of $0''.9-1''.0$ (285-315 pc). The spectral resolution was $3.6 \text{ \AA}$ (i.e. 180 km s^{-1} at $6000 \text{ \AA}$). Details of the observations and data reduction can be found in Ferruit et al. (1999).

3.1. Deconvolution of 3D data cubes

As illustrated in Fig. 1, the major limitation encountered in these observations is the 'moderate' spatial resolution ($0''.9-1''.0$) of our ground-based observations, which is more than ten times lower resolution than what can be achieved by the Hubble Space Telescope (HST). The arcs straddling the nucleus, which are clearly seen in the HST image, are barely visible as such in our reconstructed image. Thus, to improve the spatial resolution of our data, we have deconvolved our continuum subtracted data cubes. We have used a new method developed in a collaboration between the University of Maryland, the CRAL - Observatoire de Lyon and L. B. Lucy (Ferruit et al. 1999, in preparation). This method is based on a classical Richardson-Lucy algorithm, modified to use the HST emission-line images as a guide. This allowed us to gain of a factor $\simeq 3$ in resolution, the final spatial resolution of the $[\text{O III}]$ data cube being $0''.35$.

3.2. Results

Figure 3. HST/FOC $[\text{O III}]\lambda\lambda 4959, 5007$ images of Mrk 573 with contours of the $[\text{O III}]$ velocity field (solid lines, from the TIGER deconvolved data cube) and VLA 3.6 cm radio continuum map (dashed lines, from Falcke et al. 1998) superimposed. The inset shows an enlargement of the nuclear regions (inner $0''.5$ [160 pc]) with only the 3.6 cm radio continuum isophotes superimposed (dashed lines, from Falcke et al. 1998).

Jet-cloud interactions in the emission-line knots ? The upper panels of Figure 2 show the results of a single Gaussian fitting of each spectrum of the deconvolved $[\text{O III}]$ data cube. The gain in spatial resolution permitted by the deconvolution is obvious when comparing the reconstructed image shown in the upper left panel with the image reconstructed before deconvolution of the data cube (Fig. 1, upper right panel). The corresponding $[\text{O III}]$ velocity field is shown

in Fig. 2 (upper right panel). In addition to a smooth velocity pattern with a kinematical major axis along $PA \simeq 60^\circ\text{--}90^\circ$, we detect three strong velocity perturbations spatially coincident with emission-line knots inside the arcs. Each of these three velocity features is located close to one of the three radio knots (Fig. 2, upper right panel) and is also associated with a strong line broadening (Fig. 2, line profiles).

The fact that the emission-line knots are the locus of strong kinematical disturbances and are either spatially coincident or in the immediate vicinity of the radio components suggest that they trace the interaction between the radio jets and clouds in the interstellar medium (ISM) of the galaxy.

The arcs: shocks versus central source photoionization. We have investigated two scenarios for the ionization of the emission-line arcs of Mrk 573. In the first one the arcs are photoionized by the central source, while in the second one they are fast, photoionizing shocks (Dopita & Sutherland 1995).

Figure 4. *Left panel:* Map of the ionization parameter U of the gas, as derived from the $[N\text{ II}]/H\alpha$ ratio, with the isophotes of the HST $[O\text{ III}]$ image (Falcke et al. 1998) superimposed. Black dots indicate regions where the $[S\text{ II}]$ ratio (used to estimate the density of the gas) is consistent with the low density limit to within the errors. *Right panel:* Plot of the number of ionizing photons, Q_{ion} , emitted per second by a compact central source versus radius.

To test the validity of the central source photoionization scenario, we have first mapped the $[N\text{ II}]/H\alpha$ line ratio. Using a relation derived from simple photoionization models computed using Mappings 1c, we have converted this line ratio image into a map of the ionization parameter of the gas (Fig. 4, left panel). From this map, and using the gas density inferred from the observed $[S\text{ II}]$ ratio, we have then computed, as a function of the radius from the nucleus, the production rate of ionizing photons (Q) that a central source should have in order to explain the ionization state of the gas. Instead of the constant or decreasing curve expected for a central source photoionization model, we find that Q increases steeply with the distance from the nucleus (Fig. 4, right panel). This rules out the central source photoionization scenario, unless the photoionizing photon flux of the compact central source has decreased over the last 4000 years.

Exploring the shock scenario, we have then compared the morphological and spectral properties of the arcs, as observed in different emission lines (using both our data cubes and published HST images), with the structure predicted by the photoionizing shock models. We find that they are not photoionizing shocks, i.e. the ionizing photons are not produced locally in the arcs. This, however, does not imply that these ‘external’ ionizing photons come from the central source. Indeed, as the inner knots could trace the interaction between the radio jet and ISM clouds (as suggested by our observations), the ionizing photons could be produced there, by the fast shocks induced by the interaction.

4. Conclusions - Perspectives

The TIGER 3D observations of the central regions of nearby Seyfert galaxies have brought important results on the physical and energetical properties of the ionized gas in these environments. The 3D aspect of these observations has been a key to understand the kinematics and the ionization state of the gas. The major limitation of the TIGER observations, their moderate spatial resolution, has been overcome using a new, dedicated deconvolution method. We are now continuing our program of observations using the new IFU OASIS, which is mounted behind the adaptive optics bonnette of the CFHT and will therefore allow us to reach spatial resolutions of a fraction of an arc second on Seyfert galaxies like NGC 4151.

Based on observations with the NASA/ESA *Hubble Space Telescope* obtained at the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy Inc., under NASA contract NAS 5-26555, and on observations collected at the Canada-France-Hawaii Telescope which is operated by CNRS of France, NRC of Canada and the University of Hawaii. PF acknowledges support by NSA under grants NAGW-3268 and NAG8-1027, by the National Science Foundation under grant AST 9527289, and by a grant from NASA administered by the American Astronomical Society.

References

- Bacon, R., et al. 1995, A&AS, 113, 347
- Dopita, M. A., & Sutherland, R. S. 1995, ApJ, 455, 468
- Falcke, H., Wilson, A. S., & Simpson, C. 1998, ApJ, 502, 199
- Ferruit, P., & Pécontal, E. 1994, A&A, 288, 65
- Ferruit, P., Adam, G., Binette, L., & Pécontal, E. 1997a, NewA, 2, 345
- Ferruit, P., Pécontal, E., Wilson, A. S., & Binette, L. 1997b, A&A, 328, 493
- Ferruit, P., Wilson, A. S., Falcke, H., Simpson, C., Pécontal, E., & Durret, F. 1999, MNRAS, accepted
- Nelson, C. H., & Whittle, M. 1995, ApJS, 99, 67
- Pécontal, E., Adam, G., Bacon, R., Courtès, G., Georgelin, Y., Monnet, G. 1990, A&A, 232, 331
- Pécontal, E., Ferruit, P., Binette, L., & Wilson, A. S. 1997, Ap&SS, 248, 167

Integral Field Spectroscopy with Optical Fibers: Studies of the Circumnuclear Regions of Nearby Active Galactic Nuclei

S. Arribas^{1,2} & E. Mediavilla¹

¹*Instituto de Astrofísica de Canarias, 38200, La Laguna, Tenerife, Spain*

²*Consejo Superior de Investigaciones Científicas, Spain*

Abstract. Optical fiber systems have played an outstanding role in the development of integral field spectroscopy (IFS). This review focuses on two main subjects: i) the generic possibilities offered by optical fibers in IFS and ii) specific studies performed with these types of systems. These studies have been mainly related to the circumnuclear region of nearby AGNs. Three main results will be commented on with some detail: i) the presence of several kinematically/physically distinct gaseous components, ii) the ionized gas and stellar 2D velocity fields, and iii) the detection of off-centered nuclei in some galaxies. This two-fold perspective (generic possibilities—specific results already obtained) should give a full idea of the potential of this technique.

1. Introduction

Integral field spectroscopy (IFS) provides a spectrum *simultaneously* for each spatial sample of a two-dimensional field. This technique has received much attention in recent years due to its advantages with respect to classical *sequential* 2-D spectroscopic techniques (e.g., long-slit scans, Fabry-Perots, etc.) when studying relatively small extended objects. Several approaches to IFS have been developed, based on the use of optical fibers (e.g., Vanderriest & Lemonier, 1988), microlenses (Bacon et al., 1995), micro-mirrors (Weitzel et al., 1996), or mixed solutions (e.g. Allington-Smith et al., 1997).

Optical fiber systems have played an outstanding role in the development of IFS. These instruments are based in the use of optical-fiber bundles with the following basic layout: one end of the bundle has the fibers distributed in a two-dimensional array, and is connected to the focal plane of the telescope; the other end of the bundle has the fibers aligned at the entrance of the spectrograph simulating the slit (*pseudo-slit*). Hence, when the telescope points to an extended object, each fiber transmits the light coming from a region of the object and a set of spectra (one per fiber/region) are simultaneously recorded on the detector. In this way, the spectral and the spatial information is recorded *simultaneously* (under the same atmospheric and instrumental conditions).

This basic arrangement can be improved with the use of microlenses, which have two main advantages: i) light losses due to the packing-fraction of the fibers are avoided, and ii) the effects due to the focal ratio degradation (FRD)

produced by the fibers are minimized. However, for 4-m class telescopes the use of microlenses has been difficult due to the focal-plane scale involved. For the 8-m class telescopes these problems are relaxed.

These systems have been used in many areas such as comets (e.g., Malivoir et al., 1990), quasars (e.g., Chatzichristou et al., 1999), gravitational lenses (e.g., Mediavilla et al., 1998), population analysis of galaxies (Peletier et al., 1999), etc. However, they have been especially useful when studying the circumnuclear regions of AGNs. Therefore, after commenting on the most relevant instruments (Section 2) and the generic advantages of the use of optical fibers in IFS (Section 3), the main results obtained for nearby AGNs will be commented on in some detail (Section 4).

2. The Instruments

The idea of using optical fibers for transforming from a two-dimensional geometry at the focal plane of the telescope into the one-dimensional geometry at the entrance of the spectrograph was first proposed more than 40 years ago by Kapany (1958). However, he was thinking of improving the spectral resolution in coudé spectroscopy rather than in 3D observations (c.f., Vanderriest, 1995). The first serious test of 3D observations with optical fibers was done by Vanderriest (1980) at the CFHT. However, several years had to pass until the first generation of instruments of this kind were built. These instruments were: SILFID/ARGUS (Vanderriest & Lemonier, 1988) at the CFHT, DENSEPACK (Barden & Wade, 1988) for the 4 m (KPNO), HEXAFLEX (Arribas, Mediavilla, & Rasilla, 1991) for the WHT, and MPFS (Afanasyev et al., 1995a, and references therein) for the 6 m, which was the first instrument combining the use of fibers and microlenses. In any case, all these instruments were basically experimental, which had the merit of producing the first scientific results and, therefore, of showing the community the potential of this technique.

Now there are other instruments of this type running in different observatories: 2D-FIS (WHT; García et al., 1994), MOS/ARGUS (CFHT; Vanderriest, 1995), INTEGRAL (WHT; Arribas et al., 1998), and WIYN-DENSEPACK (WIYN-KPNO; Barden & Ingerson, 1998), some of them being common-user facilities. In addition, the first generation of fiber instruments working at near-infrared wavelengths are now available (e.g., SMIRFS, Haynes et al., 1998; CHOSI-IFU Kenworthy et al., 1998).

In any case, a great advance is going to happen in a near future when the new generation of instruments (many planned for 8-m class telescopes) come into full operation (e.g., PMAS-3.6 m Calar Alto, Roth & Laux, 1998; TEIFU-WHT, Haynes et al., 1998; GMOS-Gemini, Allington-Smith et al., 1997; GIRAFE-VLT; NIRMOS & VIRMOS-VLT, Le Fevre et al., 1998; SINFONI-VLT, Thatte et al., 1998; MRS-IFU-HET, Bershadsky et al., 1998; OAIFU, Guerin, 1998; OSIRIS-Keck, Larkin, 2000; AUSTRALIS-VLT, Taylor, 1998; FMOS-SUBARU, etc.).

3. Advantages and Drawbacks

Most of the advantages of the IFS technique are direct consequences of the simultaneity when recording spatial and spectral information. The simultaneity

not only implies a more efficient way of observing but, more importantly, it guarantees a great homogeneity in the data. In addition, with IFS systems there is no need for an accurate centering of the object in the slit or to adapt the slit width (spectral resolution) to the seeing conditions; neither are the data affected by “slit effects” when determining radial velocities (Vanderriest 1995). Moreover, the effects due to differential atmospheric refraction can be well corrected a posteriori (Arribas et al., 1999). On the contrary, the main disadvantage of IFS systems with respect to more classical techniques is their relative small field of view.

Optical-fiber systems have, in addition, several advantages with respect to other IFS approaches such as those based on the use of microlenses only (TIGER approach; Bacon et al., 1995) and “image slicers” (e.g., Weitzel et al., 1996; Content, 1998). These advantages are due to the great flexibility offered by optical fibers for extracting the relevant information from the focal plane, and for arranging it properly on the detector (see below). On the contrary, the main drawback of the fiber (+ microlenses) approach is related with the difficulties when aligning the fibers with the microlenses, especially at cryogenic temperatures. In any case, some promising solutions have been proposed (Tecza & Thatte, 1998). In the following subsections different aspects of the flexibility offered by optical fibers in IFS are commented on in more detail.

3.1. Configuration in the Focal Plane

Optical fibers allow ad hoc configurations to be defined in the focal plane for specific programs. For instance, in cases where sky subtraction is an important issue, a relatively large number of fibers may be devoted to this aim. In addition they may be arranged with a proper geometry (see, for instance, García et al., 1994). A further example was proposed by Bershadsky et al. (1998) for studying galaxy kinematics. There, the fiber bundles have a fine sampling at the center, and a coarse sampling outwards. The basic idea is to have complete spatial information in the inner regions, where the velocity field is expected to be more complex, and coarse information in the outer regions where more regular behavior is expected.

3.2. Entrance of the Spectrograph

At the entrance of the spectrograph, fibers can also be arranged in several ways. If required, they may be distributed along two parallel pseudo-slits, in such a way that no free space is left between adjacent spectra in the detector (see, for instance, Kenworthy et al., 1998). This configuration should be equivalent to the one used in long-slit (or image-slicer) observations. However, adjacent fibers can be separated at the entrance of the spectrograph offering the possibility of separating the spectra in the detector. This possibility is relevant in order to minimize optical cross-talk among different spectra due to the PSF of the spectrograph. Note that the effects of the cross-talk in IFS are not equivalent to an isotropic loss of spatial resolution (such as that due to the seeing), but may introduce rather complex artifacts (García et al., 1998).

Optical cross-talk inside the spectrograph also depends on the difference of intensity between adjacent spectra. That is why the way in which the spectra are ordered on the detector (way in which the fibers are ordered in the slit) is

also relevant for minimizing optical cross-talk. Optical fibers can be ordered in the desired way and, accordingly, they allow cross-talk to be minimized in specific or generic types of observations. For instance, if the object has a more or less isotropic intensity distribution (as is the case for many astronomical objects), the *ring* ordering proposed by García et al., (1994) is preferable to the *image-slicer* configuration (e.g., Guérin, 1998).

The way in which the fibers are ordered at the entrance of the spectrograph is also relevant for equalizing the output intensities (see Section 3.5).

3.3. Arranging the Spectral and Spatial Information

Spectral elements can be interchanged with spatial elements (and vice versa) in a very easy way thanks to the use of such configurations as that presented in Figure 1. On the left is represented the standard use of the 3D space. Here, the fibers fill a *pseudo-slit* at the entrance of the spectrograph, and one dimension of the spectrograph is used for recording the spectral information of each fiber/spatial element. In this configuration the detector pixels are used for recording *long* spectra of a relative small number of spatial elements. On the right part of this figure, the number of spatial elements has been increased by a factor of three. At the entrance of the spectrograph the fibers are arranged in three *pseudo-slits*. In order to avoid overlapping between the spectra, a filter is used to restrict the spectral coverage. Therefore, in this case, the number of spatial pixels is relatively large, but the *length* of each individual spectrum is smaller. The important point is that to change from the configuration defined on the right to the one defined on the left only a mask is required to block the outer parts of the focal plane, and to remove the filter.

3.4. Interchanging Configurations

The multi-bundle arrangement design (Arribas, Mediavilla, & Rasilla, 1993) allows one to interchange bundles very easily with different configurations. The idea is to connect simultaneously at the entrance of the spectrograph several bundles with different configurations, and to connect at the focal plane the desired bundle (blocking the entrance of the others). This type of arrangement allows one, for instance, to modify the spatial sampling/resolution at the entrance without the need for pre-optics.

This design has proven to be very useful in some instruments. For instance, the INTEGRAL system at the WHT (Arribas et al., 1998a) has four optical-fiber bundles with different configurations, which can be interchanged in a few seconds.

An extreme case using the capabilities of the optical fibers in this respect is the “fiber switch-yard” for AUSTRALIS (Taylor, 1998).

3.5. Increasing the Dynamic Range

Recently, we have presented a new technique for performing two-dimensional (integral-field) spectroscopy of objects with a large range in intensity (Arribas et al., 1998b). It is based on the capability of optical fibers to discretize the intensity distribution at the telescope focal plane. This allows not only reformatting, but also the equalization of fiber output intensities matching them to the dynamic range of the detector. This technique, which we have called equalized

Figure 1. This figure represents two possible uses of the 3D space with optical fiber systems. On the left is represented the standard use of the 3D space. Here, the fibers fill a *pseudo-slit* at the entrance of the spectrograph. In this configuration the detector pixels are used for recording *long* spectra of a relative small number of spatial elements. On the right, the number of spatial elements has been increased by a factor of three. At the entrance of the spectrograph the fibers are arranged in three *pseudo-slits*. In order to avoid overlapping between the spectra, a filter is used to restrict the spectral coverage. Therefore, in this case, the number of spatial pixels is relatively large, but the *length* of each individual spectrum is smaller. The important point is that to change from the configuration defined at the right to the one defined at the left only a mask is required to block the outer parts of the focal plane, and to remove the filter. Therefore, these configurations can be interchanged quickly and easily.

Figure 2. Reconstructed images in R (a), I (b), and $H\alpha$ (c) from 3D observations based on equalized integral field spectroscopy of the binary system HD 167605 A+B (see text). The two stars have about $\Delta R=3.8$ mag, and $\Delta I=3.1$. Note the high contrast in the $H\alpha$ map. This figure clearly indicates that while in the primary $H\alpha$ appears in absorption, for the companion this feature is shown in emission.

integral field spectroscopy (E-IFS), permits spectral information to be obtained, as well as accurate relative photometry and astrometry. This technique may be of interest in several fields (AGNs, Brown dwarfs, star-forming regions, etc.), which require 2D spectroscopy of regions with a high range in intensities.

In Fig. 2 we present the results of this technique when observing the environment of the K2 V star HD167605, for which recent *HST*/NICMOS observations in coronagraphic mode have revealed the presence of a faint companion. We have also detected this star, whose spectral type (M4 V) was determined, together with its relative distance ($\Delta\alpha=+0.80''$ and $\Delta\delta=+0.86''$) and photometry ($\Delta R=3.8 \pm 0.1$ mag, and $\Delta I=3.1 \pm 0.1$ mag) with respect to the primary, confirming that these stars form a binary system. For the companion we have also found the presence of $H\alpha$ in emission, which indicates that this binary system is relatively young.

To summarize this section we can say that optical fibers offer great flexibility for extracting the relevant information from the focal plane of the telescope and arranging it properly on the detector.

4. Results on the Circumnuclear Regions of AGNs

Three-dimensional studies on the circumnuclear regions of AGNs have been performed using a number of observational techniques. Regarding sequential approaches, the work by Wilson and collaborators based on long-slit scans deserves mention (see Wilson et al., 1989 and references therein). Much effort has also been made using other 3D techniques (F-P, TTF), which have been commented on in detail in this conference (see, for instance, contributions by

Bland-Hawthorn, Cecil, Veilleux, and Tadhunter in these proceedings, and references therein). Concerning integral-field spectroscopic techniques, it is worth mentioning the work with TIGER (see Ferruit in these proceedings and references therein), and with the instrument 3D (see Maiolino in these proceedings and references therein). Here we are going to concentrate on studies performed with optical-fiber systems.

IFS techniques are especially suitable for studying the circumnuclear region of AGNs. This is due to the fact that their morphology and kinematics are very complex as a consequence of the strong influence of the active nucleus. With different degrees of depth, somewhat less than 30 nearby AGNs have been observed with optical fiber-systems so far. Here we will concentrate on three main types of results: i) the presence of several gaseous components, ii) the stellar and ionized gas velocity fields, and iii) the presence of off-centered nuclei.

4.1. Presence of Several Gaseous Components

Studies of the narrow emission lines prior to 3D observations (mainly based on nuclear data) have shown that these lines are very often asymmetrical with wings, shoulders, secondary peaks, or other evidence of substructure. One of the most remarkable result of 3D studies in general (and IFS-fiber results in particular) is that all the galaxies analysed have narrow emission-line profiles which exhibit substructure in the circumnuclear region also (see, for instance, the cases of NGC 5728, Arribas & Mediavilla, 1993; NGC 3227, Mediavilla and Arribas, 1993; NGC 1068, García Lorenzo et al., 1999).

This finding implies that several gaseous systems inhomogeneously distributed in space appear projected in the observer's line of sight, hindering the obtaining and interpretation of 2D maps of features such as velocity, intensity, line-quotients diagnostic, etc. This obviously represents a major difficulty.

However, since fiber systems allow one simultaneously to obtain 2D spectroscopic data regularly distributed over the circumnuclear region (under the same atmospheric and instrumental conditions), in some cases it is possible to track the properties of individual components. In fact, a multi-component line analysis has been attempted for several galaxies (NGC 5728, Arribas & Mediavilla, 1993; NGC 3227, Mediavilla & Arribas, 1993; NGC 1068, Arribas et al., 1996; NGC 7331, Mediavilla et al., 1997; NGC 3516, Arribas et al., 1997), which has permitted the identification of physically/kinematically distinct gaseous components, some of them clearly related to the nuclear activity.

Thanks to this deprojection, it has been possible to analyze the general properties of the different components. The following behavior is very often found: i) one component has a regular kinematic pattern, which may be well explained by rotation and has a relatively low ionization, ii) another component indicates the presence of an outflow and in most cases has a higher ionization (this component is generally located in the innermost region), and iii) if more components are present, they are found in relatively small regions and are difficult to interpret (e.g., in MRK 509 an extended broad Balmer emission was reported by Mediavilla et al., 1998a). Therefore, we are talking about a complex environment with at least two kinematic components and ionization sources.

A good example is the case of NGC 4303 (Colina & Arribas, 1999). In Fig. 3 we present the *HST* image obtained in the UV continuum (2200 Å), together

Figure 3. The inner region of NGC 4303. An *HST* image obtained in the UV continuum (2200 Å), together with several emission-line maps inferred from IFS using the 2D-FIS system.

with several emission-line maps inferred from IFS using the 2D-FIS system. The correspondence between these different types of images is clear. It is interesting to note the different ionization conditions between the nucleus and the region towards the SW (compare, for instance, the $H\beta$ and $[O\ III]$ maps). Comparing the $H\alpha$ and $H\beta$ maps, a clear idea of the relative extinction over this region can also be inferred. In NGC 4303 there are two main kinematic subsystems (see figure 4 of Colina & Arribas, 1999). One is well represented by rotation, and the other indicates strong radial motions.

4.2. Stellar and Ionized Gas Velocity Fields

The 2D velocity fields for the ionized gas have been obtained for a quite large number of galaxies. In Fig. 4 several examples of ionized gas velocity fields are given. As mentioned before, some of the components show a regular pattern, with the kinematic axes well aligned with the photometric ones, which suggests rotation. This is the case for the velocity field presented for NGC 4303, which corresponds to the low-ionization component. However, on the whole, most of the velocity fields are distorted, indicating the presence of radial motions. In some cases, as NGC 985 (Arribas et al., 1999b) the velocity field is only slightly distorted in spite of the fact that this galaxy is likely to be the result of a two-galaxy collision. In others, such as NGC 1068, radial motions dominate

Figure 4. Ionized gas velocity fields for NGC 1068 (García-Lorenzo et al, 1997, 1999a); NGC 4303-Low ionization component (Colina & Arribas, 1999); NGC 2992 (García-Lorenzo et al, 1999b); NGC 4151 (Mediavilla et al., 1991; NGC 985 (Arribas et al, 1999b); NGC 3227 (Mediavilla & Arribas, 1993, Arribas & Mediavilla, 1994); NGC 3516 (Arribas et al., 1997); NGC 7331 (Mediavilla et al, 1997); NGC 5728 (Arribas & Mediavilla, 1993), NGC 6500 and NGC1052 (del Burgo et al., 1998).

the ionized gas velocity field (see, Arribas et al., 1996; García Lorenzo et al., 1997). Further examples of ionized gas velocity fields are given by Afanasiev et al. (1995b) for Mrk 573, Afanasiev & Shapovalova (1994) for NGC 4151, and Afanasiev et al. (1995c) for NGC 4258.

The stellar velocity field of AGNs is generally more difficult to obtain. However, for about ten galaxies it has been already obtained thanks to fiber systems. These are: NGC 1068 (García-Lorenzo et al., 1997), NGC 2992 (García-Lorenzo et al., 1999), NGC 985 (Arribas et al., 1999), NGC 7331 (Mediavilla et al., 1997), NGC 3516 (Arribas et al., 1999), NGC 6500 and NGC 1052 (del Burgo et al., 1998), M31 (del Burgo et al, 1999). When these velocity fields are compared with their corresponding ionized gas velocity fields a clear decoupling is found in many cases, which suggests that most of the ionized gas is kinematically decoupled from the stellar component.

4.3. Off-centered nuclei

Another remarkable result obtained from these data is the non-coincidence between the optical nucleus and the kinematic center. In the Seyfert 2 galaxy

NGC 5728 (Arribas & Mediavilla, 1993), this disagreement could be explained as the result of obscuration hiding the true active nucleus. This explanation was supported by *HST* images (Wilson et al., 1993), which revealed a clear bi-conical structure, the apex being in very good positional agreement with the kinematic center. These results strongly support the unified models, according to which the central engine (an accretion disk) is surrounded by a torus or disk of blocking material which, complemented or not by a larger structure, collimates the ionized radiation into two oppositely directed cones. For Seyfert 2 galaxies the line of sight is not within the radiation cones, and the true nucleus cannot be seen directly.

These results obtained in NGC 5728 suggested that a method for finding the true nucleus in Seyfert 2 galaxies could be as simple as the determination of the kinematic center. However, in NGC 3227 (Mediavilla & Arribas, 1993) the Seyfert 1 nucleus was found to be in positional disagreement with respect to the kinematic center inferred from the ionized gas. In the case of the Seyfert 2 galaxy NGC 1068 (Arribas et al, 1996; García-Lorenzo et al, 1997), in spite of the high symmetry of the ionized gas velocity field, which means a high accuracy when determining the kinematic center, the location of the obscured nucleus is in disagreement with other determinations (see Thatte et al., 1997). Other cases of reported off-centered nuclei from IFS are NGC 3516 (Arribas et al., 1998), and M31 (Bacon et al, 1994; del Burgo et al., 1999).

All this evidence (the presence of several gaseous components, decoupling between the stellar and ionized gas velocity fields, and off-centered nuclei) strengthens the case for galaxy merging as the origin of the nuclear activity, a hypothesis which still needs additional investigation of individual objects. In this respect, it is worth pointing out the work by Chatzichristou & Vanderriest (1995) and Chatzichristou et al. (1998)

Thanks are due to Begoña García-Lorenzo, and C. del Burgo, who provide comments and some figures. We also acknowledge the help by Terry Mahoney and Ramon Castro in the correction and edition of this manuscript.

References

- Afanasiev, V.L. & Shapovalova, A.I., 1994, in *Mass-Transfer Induced Activity in Galaxies*, (Ed. I. Shlosman), Cambridge University Press, p95.
- Afanasiev, V.L., Vlasiouk, V.V., & Green, R.F., 1995a, in *Tridimensional Optical Spectroscopic Methods in Astrophysics*, (Ed. G. Comte & M. Marcelin), ASP con. ser., Vol 71, p. 266.
- Afanasiev, V.L., Shapovalova, A.I., Burenkov, A.N., Dodonov, S.N., & Vlasiouk, V.V., 1995b, in *Tridimensional Optical Spectroscopic Methods in Astrophysics*, (Ed. G. Comte & M. Marcelin), ASP con. ser., Vol 71, p. 261
- Afanasiev, V.L., Courtes G., Dodonov S.N., Plana H., in *Tridimensional Optical Spectroscopic Methods in Astrophysics*, 1995c (Ed. G. Comte & M. Marcelin), ASP con. ser., Vol 71, p. 278.
- Allington-Smith J.R., Content R.C., Haynes R., Lewis I.J., 1997, *Proc. SPIE*, 2871, 109.
- Arribas, S., Mediavilla E., Rasilla J.L. 1991, *ApJ*, 369, 270.
- Arribas, S. Carter, D., Cavaller, L., del Burgo, C., Edwards, R., Fuentes, J., García, A., Gentles, B., Herreros, J. M., Jones, L., Mediavilla, E., Pi, M., Pollacco, D., Rees, P., & Sosa, N. 1998a *Proc. SPIE*, 3355, 821.

- Arribas, S., Mediavilla E., Fuensalida J.J. 1998b, ApJL, 505, 43.
- Arribas, S. & Mediavilla E., 1993, ApJ, 410, 552.
- Arribas, S., Mediavilla E., Rasilla, J.L., 1993 in *Fiber Optics in Astronomy II*, (Ed. P. Gray), ASP conf. ser., Vol. 37, 322.
- Arribas, S. & Mediavilla E., 1994, ApJ, 437, 149.
- Arribas, S., Mediavilla, E., Garcia-Lorenzo B., del Burgo C., 1997, ApJ, 490, 227.
- Arribas, S., Mediavilla, E., Garcia-Lorenzo B., 1996, ApJ, 463, 509.
- Arribas, S., Mediavilla, E., Garcia-Lorenzo B., del Burgo C., Fuensalida J.J., 1999a A&ASS, 136, 189.
- Arribas, S., Mediavilla, E., del Burgo C., García-Lorenzo B., 1999b, 511, 680.
- Bacon, R., Emsellem, E., Monnet, G., Nieto, J.L., 1994, A&A, 281, 691.
- Bacon, R., Adam, G., Baranne, A., Courtes, G., Dubet, D., Dubois, J.P., Emsellem, E., Ferruit, P., Georgelin, Y., Monnet, G., Pecontal, E., Rousset, A., and Say, F. 1995, A&AS, 113, 347.
- Barden S.C., Wade, R.A., 1988 in *Fiber Optics in Astronomy*, (S. Barben), ASPcon.ser, Vol 3, p. 113.
- Barden, S.C., & Ingerson, T.E., 1998 in *Fiber Optics in Astronomy III* (Eds. S. Arribas, E. Mediavilla, and F. Watson), ASP con.ser., 152, 60.
- Bershady, M.A., Andersen, D., Ramsey, L., & Horner S., 1998 in *Fiber Optics in Astronomy III* (Eds. S. Arribas, E. Mediavilla, and F. Watson), ASP con.ser., 152, 253.
- Bland-Hawthorn, J. 2000, this volume.
- Cecil, G. 2000, this volume.
- Colina, L. & Arribas, S., 1999, ApJ, 514, 637.
- Content, R., 1998, SPIE, 3356, 122.
- Chatzichristou & Vanderriest, 1995, A&A, 298, 343.
- Chatzichristou & Vanderriest, M. Lehner, 1998, A&A, 330, 841.
- Chatzichristou, E., Vanderriest, C., Jaffe, W., 1999, A&A, 343, 407.
- del Burgo C., Arribas, S., Mediavilla, E., and García-Lorenzo, B., 1998, in *Fiber Optics in Astronomy III* (Eds. S. Arribas, E. Mediavilla, and F. Watson), ASP con.ser., 152, 189.
- del Burgo C., Arribas, S., Mediavilla, E., García-Lorenzo, B., 1997, IAU 184, p 85.
- Ferruit, P. 2000, this volume.
- García-Lorenzo, B., Mediavilla, E., Arribas, S., del Burgo, C., ApJL, 483, 99.
- García-Lorenzo, B., Mediavilla, E., Arribas, S., 1999a, ApJS, 518, (in press).
- García-Lorenzo, B., Arribas, S., Mediavilla, E., 1999b, ApJ (in preparation).
- García, A., Rasilla, J.L., Arribas, S., Mediavilla, E., 1994, SPIE, vol. 2198, 75.
- García, A., Rasilla, J.L., Arribas, S., Mediavilla, E., 1998 in *Fiber Optics in Astronomy III* (Eds. S. Arribas, E. Mediavilla, and F. Watson), ASP con.ser., 152, 203.
- Guerin J., 1998 in *Fiber Optics in Astronomy III* (Eds. S. Arribas, E. Mediavilla, and F. Watson), ASP con.ser., 152, 282.
- Haynes R., Allington-Smith, J. & Lee, D., 1998 in *Fiber Optics in Astronomy III* (Eds. S. Arribas, E. Mediavilla, and F. Watson), ASP con.ser., 152, 289.
- Haynes R., Content, R., Allington-Smith, J., & Doel, P., 1998 in *Fiber Optics in Astronomy III* (Eds. S. Arribas, E. Mediavilla, and F. Watson), ASP con.ser., 152, 193.
- Kapany, N., 1958, in *Concepts of Classical Optics*, (Ed. J. Strong), Feeman publish. (San Francisco).
- Kenworthy, M.A., Parry, I.R., and Ennico, K.A., 1998 in *Fiber Optics in Astronomy III* (Eds. S. Arribas, E. Mediavilla, and F. Watson), ASP con.ser., 152, 300.
- Larkin, J. 2000, this volume.
- Le Fevre O., et al., 1998, SPIE, 3355, 8.
- Maiolino, R. 2000, this volume.

- Malivoir, C., Encrenaz, T., Vanderriest, C., Lemonnier J.-P., & Kohl-Moreira, J.-L., 1990, *Icarus*, 87, 412.
- Mediavilla E., Arribas, S., Rasilla, J.L., 1992, *ApJ*, 396, 517.
- Mediavilla E., & Arribas, S. 1995, *MNRAS*, 276, 579.
- Mediavilla E., & Arribas, S. 1993, *Nature*, 365, 420.
- Mediavilla E., Arribas, S., García-Lorenzo B., del Burgo, C., 1997, *ApJ*, 488, 682.
- Mediavilla E., Arribas, S., García-Lorenzo B., del Burgo, C., 1998, *ApJ*, 494, L13.
- Mediavilla E., Arribas, S., del Burgo, C., Oscoz, A., Serra-Ricat, M., Alcalde, D., Falco, E.E., Goicoechea, J.L., García-Lorenzo, B., & Buitrago, J. 1998, *ApJL*, 503, 27.
- Peletier R. F., Vazdekis, A., Arribas, S., del Burgo, C., García-Lorenzo B., Gutierrez, C., Mediavilla, E., Prada, F., 1999, *MNRAS* (in press).
- Roth, M.M., & Laux, U., 1998 in *Fiber Optics in Astronomy III* (Eds. S. Arribas, E. Mediavilla, and F. Watson), ASP con.ser., 152, 168.
- Taylor, K., 1998 in *Fiber Optics in Astronomy III* (Eds. S. Arribas, E. Mediavilla, and F. Watson), ASP con.ser., 152, 261.
- Tadhunter, C. 2000, this volume.
- Thatte et al., 1998, *SPIE*, Vol. 3353, 704.
- Thatte N., Quirrenbach, A., Genzel, R., Maiolino, R., Tecza, M., 1997, *ApJ*, 490, 238.
- Tecza M., & Thatte N., in 1998 in *Fiber Optics in Astronomy III* (Eds. S. Arribas, E. Mediavilla, and F. Watson), ASP con.ser., 152, 271.
- Vanderriest C., 1980, *PASP*, 92, 858.
- Vanderriest C., and Lemonnier J.P., 1988, in *Instrumentation for ground-based Astronomy*, (Ed. L. Robinson) New York, Springer-Verlag, p.304
- Vanderriest C. 1995, in *3D Optical Spectroscopic Methods in Astronomy*, ASP cof. Ser., 71 (Eds. G. Comte, M. Marcellin), San Francisco, ASP, 209.
- Veilleux, S. 2000, this volume.
- Weitzel, L., Krabbe, A., Kroker, H., Thatte, N., Tacconi-Garman, L. E., Cameron, M., and Genzel, R. 1996 *A&AS*, 119, 531.
- Wilson, A.S., Wu, X., Heckman, T.M., Baldwin, J.A., Balick, B., 1989, *ApJ*, 339, 729.
- Wilson, A.S., Braatz J. A., Heckman T.M., Krolik J.H., Miley G.K., 1993, *ApJ*, 419, L61.

Infrared 3-D Observations of Nearby Active Galaxies

R. Maiolino¹, N. Thatte², A. Alonso-Herrero³, D. Lutz², A. Marconi¹

¹*Osservatorio Astrofisico di Arcetri, Firenze, Italy*

²*Max-Planck-Institut für Extraterrestrische Physik, Garching, Germany*

³*Steward Observatory, Tucson, Arizona*

Abstract. We present multi-wavelength imaging observations of three nearby and famous active galaxies obtained with of NICMOS, ISOCAM and the MPE near-IR integral field spectrometer. The data reveal a variety of features and properties that are missed in optical studies and in traditional IR monodimensional spectroscopy.

1. Introduction

Infrared observation of active galaxies (starbursts and AGNs) have greatly improved our understanding of these systems, not only because they are usually heavily obscured at optical wavelengths, but also because the infrared bands offer a wealth of indicators that often do not have an equivalent in the optical range. Over the past few years, new IR instruments and observing facilities have provided two-dimensional spectroscopic information of these systems through integral field spectroscopy and narrow band imaging. New integral field spectra obtained with the MPE 3D near-IR spectrometer, NICMOS-HST narrow and broad band images and ISOCAM-CVF spectra of three nearby and famous active galaxies allowed a detailed investigation of their nuclear region revealing new features and properties, and showing that optical observations and traditional IR spectroscopy only provide a limited view of this class of objects. Here we summarize some of the results based on these data.

2. NGC1068

This famous galaxy is regarded as the archetype of Seyfert 2 galaxies. The [OIII]5007Å map shows an ionization cone to the North-East of the nucleus, and a much fainter counter-cone to the South-West. The morphology and the dynamics of the Narrow Line Region (NLR) appear affected by the radio jets (eg. Axon et al. 1998). The radio jets also feed two large scale radio lobes.

We observed the nuclear region of NGC1068 with 3D, the MPE near-IR integral field spectrometer (Weitzel et al. 1996), assisted by ROGUE, a first order adaptive optics system (Thatte et al. 1995), at the William Herschel Telescope (WHT) and at the Anglo Australian Telescope (AAT).

Various IR lines and powerful hot dust emission were detected (Thatte et al. 1997). Here we focus on the properties of the [SiVI]1.96μm high excitation

Figure 1. NGC1068. Left: integrated $[\text{FeII}]1.64\mu\text{m}$ emission obtained with the MPE-3D integral field spectrometer overlaid on the 6cm VLA map of the NE lobe. Right: $[\text{FeII}]1.64\mu\text{m}$ and $[\text{SiVI}]1.97\mu\text{m}$ line profiles in the two ionization cones.

coronal line and the $[\text{FeII}]1.64\mu\text{m}$ line that instead traces the low ionization region. Both lines are detected in the NE ionization cone and, at a lower level, in the SW cone. Overall, the morphology of their integrated emission is similar to the $[\text{OIII}]\text{-HST}$ map, though at a lower resolution. In Fig. 1 we show the $[\text{FeII}]$ map, which has the characteristic cone-like morphology. Given that both these lines are observed in the ionization cones it is tempting to assume that they are emitted by the same population of ionized clouds. However, the kinematics of the gas as traced by the two lines points to a more complex scenario. In Fig. 1 we show the profile of the $[\text{FeII}]$ and $[\text{SiVI}]$ lines, with respect to the host galaxy systemic velocity. In the NE cone $[\text{SiVI}]$ is blueshifted while $[\text{FeII}]$ is redshifted and the shift between the two lines is several 100 km/s. In the SW cone the situation is exactly reversed. The large velocity difference between the two lines indicates that they come from different populations of ionized clouds.

We think that the radio lobes affect the kinematics and the ionization structure of the large scale NLR, similar to that observed on smaller scale clouds near the radio jet (Axon et al. 1998). Fig. 2 schematically illustrates our model. In the NE radio lobe (right-hand side in Fig. 2) the upper part of the expanding bow shock accelerates the gas in our direction, hence its emission is blueshifted. This gas is mostly out of the galactic gas disk and, therefore, it has low density, hence high ionization parameter that favors the emission of high ionization species such as Si^{+6} . On the opposite side the bow shock enters the dense gas of the galactic disk, that is therefore redshifted. Here the bow shock rapidly loses its energy, a fraction of which destroys dust grains returning Fe into the gas phase, thus increasing the $[\text{FeII}]$ emission. Also, the high density characterizing this region decreases the ionization parameter, while the column of gas and dust in these equatorial regions hardens the ionizing photon flux; both these effects favor the emission of low ionization lines typical of the transition regions, such

as [FeII]. The symmetric model for the SW cone explains the velocities of [FeII] and [SiVI] inverted with respect to the NE cone.

Figure 2. Proposed model for the Narrow Line Region of NGC1068

Although the bow shock returns Fe into the gas phase and produces the observed redshift of the low ionization gas, probably it does not play a major role in exciting the line. Indeed, the [FeII] map has a cone-like morphology, indicating that the nuclear X-ray source is probably responsible for exciting this line.

Finally, the comparison of the radio and [FeII] maps in Fig. 1 further supports our model. The lower end of the radio lobe in Fig. 1 nicely matches the upper end of the [FeII] cone, we think that this dividing line traces the region where the radio lobe enters the galactic gas disk. Therefore, the dense gas south of this line has been Fe-enriched and redshifted, as actually observed. A more detailed discussion of these results is given in Thatte et al. (in prep.).

3. Circinus

At a distance of only 4 Mpc, the Circinus galaxy is the closest *bona fide* Seyfert 2 known. Its proximity makes this galaxy an excellent candidate with which to tackle various issues related to the interaction between active nuclei and their circumnuclear region. Within this context, the mechanism responsible for feeding AGNs on the 10–100 pc scale is a debated issue. Shlosman et al. (1989) and Wada & Habe (1992) proposed that when the mass of the nuclear gas disk is a significant fraction of the dynamical mass (>20%) it becomes gravitationally unstable and forms a gaseous bar that can drive gas into the innermost region to fuel the AGN. Circinus has an extremely gas-rich nuclear region and, therefore, it is an optimal candidate to search for this nuclear gaseous bar.

Broad band and narrow band NICMOS-HST observations of the nuclear region of Circinus were obtained. Fig. 3 shows the H-K color map of the nuclear region (Maiolino et al. 1999). The galaxy major axis is at P.A. $\approx 25^\circ$ and the SE is the near side of the disk. The PSF of the nucleus, dominated by hot

dust emission (Maiolino et al. 1998), has been subtracted in Fig. 3. In the circumnuclear region the H-K map traces the effect of dust reddening. The most interesting feature of this map is the L-shaped dusty feature located to the South-East of the nucleus. We identify the radially extended part of this feature with the gas bar expected to feed the AGN in this system according to models. The deprojected length of the gas bar is about 100 pc. The gaseous nature of this bar is inferred by the lack of a stellar counterpart even in the K band light, where extinction is greatly reduced.

Our interpretation can be checked by looking at the kinematics of the molecular gas. Shocks and torques on the leading side of the bar should remove angular momentum from the gas which should result in a strong velocity gradient and radial inflow motions in this region. We observed the nuclear region of Circinus with 3D+ROGUE at the AAT. Fig. 3 shows the velocity field of the H₂ line at 2.12 μ m. The velocity field of the molecular gas along the leading side of the bar is characterized by a strong velocity gradient and a highly redshifted component (i.e. radially inflowing), just as expected by models. This map not only supports our interpretation of the gas bar, but directly shows radial inflow motions that are probably responsible for the AGN fuelling.

Figure 3. Circinus. Left: H-K NICMOS map; redder colors correspond to darker regions. Right velocity field of the H₂2.12 μ m line obtained with the MPE-3D spectrometer. The cross marks the K-band nucleus.

The large amount of molecular gas driven into the central region also triggers nuclear star formation. Indeed, both Pa α and [FeII] narrow band NICMOS images reveal diffuse circumnuclear emission out of the ionization cone that very likely traces recent star forming activity. A young nuclear stellar population is also inferred from the low mass-to-light ratio as derived from the CO stellar bands at 2.29 μ m observed in the 3D data (Maiolino et al. 1998).

4. NGC4945

Within the context of the starburst-AGN connection, NGC4945 is one of the most spectacular examples of a system where both phenomena coexist. In the

optical and in the IR this edge-on, nearby (3.6 Mpc) galaxy appears as a heavily obscured starburst. However, this galaxy is one of the brightest extragalactic sources at 100 keV, thus revealing the presence of an active nucleus that is heavily absorbed along our line of sight ($N_H \simeq 5 \times 10^{24} \text{cm}^{-2}$).

Figure 4. NGC4945. Top: NICMOS images of Pa α (1.87 μm), H $_2$ (2.12 μm) and Pa α equivalent width; the cross marks the location of the H $_2$ O maser. Bottom: ISOCAM-CVF spectrum of the nuclear region and profiles of the [NeII]12.8 μm line and 14 μm continuum emission along the major axis.

NGC4945 was observed with NICMOS both with broad and narrow band filters (Marconi et al., in prep.). The Pa α image in Fig. 4 reveals a nuclear starburst probably distributed in a ring (about 100 pc in radius) on the galactic plane. Note that this nuclear region is completely obscured at optical wavelengths. The K-band image (not shown) shows no point like source close to the location of the H $_2$ O maser identified by Greenhill et al. (1997). Also, the CO-index map (derived from the NIC2-F237M image) does not provide evidence for a significant dilution of the CO stellar features due to emission of hot dust heated by the AGN. Such emission is instead clearly seen in other similarly obscured AGNs (eg. NGC1068 and Circinus). This indicates that the AGN is obscured along our line of sight even at 2.3 μm . The H $_2$ 2.12 μm emission line is mostly distributed above the galactic plane (Fig. 4) in a cavity created by the starburst superwind and observed also in the J band images (Moorwood et al. 1996). The lower gas density in the cavity also results in a lower Pa α emission with respect to the continuum, hence lower Pa α equivalent width. Interestingly, the cavity traced by the EW(Pa α) has a cone-like morphology (Fig. 4), similar to the morphology of the NLR in several Sy2 galaxies. However, no UV-ionizing

photons from the AGN reach the gas in the cavity. Indeed, even though plenty of gas is present in the outer parts of the cavity, as traced by the H_2 line and dust filaments, this gas is little ionized and with a spectrum typical of LINERs (possibly excited by the superwind shock). Therefore, the AGN radiation must be absorbed even in the direction of the cone-like cavity. Very likely, the AGN is in a very early stage, still enshrouded in a 4π dusty shell a few pc in size. Given the huge radiation pressure in this region, the dusty shell cannot survive very long. When the dusty shell will break out, then the AGN UV radiation will photoionize the gas within and above the cone-like cavity, transforming the latter into an ordinary ionization cone.

Although the active nucleus is completely obscured even in the K band, we might see its dust emission in the mid-IR, both because the obscuration is further reduced and because warm dust typically emitting at these wavelengths (~ 100 K) is located at larger distances. For this reason ISOCAM-CVF observations of NGC4945 were obtained. Fig. 4 shows the CVF spectrum of the nuclear $3''$. This spectrum shows typical starburst features (strong PAH, [NeII] $12.8\mu\text{m}$ and [ArII] $7\mu\text{m}$ emission), but the saturated silicate absorption feature at $10\mu\text{m}$ indicates that the mid-IR emission comes from a heavily obscured region, with $A_V > 50$ mag. There is some hot dust continuum emission at $14\text{--}16\mu\text{m}$. However, as shown in Fig. 4, this continuum emission is resolved along the galaxy major axis and its distribution is similar to that of [NeII], therefore also the continuum emission is mostly due to the starburst activity. The active nucleus is therefore hidden and undetected even in the mid-IR. Higher angular resolution mid-IR observations are required to check if a weak point-like source might be present on the nucleus.

These results were obtained within a wide collaboration that includes also S. Anders, R. Genzel, D. Macchetto, A. Moorwood, E. Oliva, A. Quillen, M. Rieke, G. Rieke, E. Schreier, E. Sturm, L. Tacconi-Garman and D. Tran.

References

- Axon, D.J., Marconi, A., Capetti, A., Macchetto, F.D., Schreier, E., Robinson, A. 1998 ApJ, 496, L75
- Greenhill, L.J., Moran, J.M., Herrnstein, J.R. 1997 ApJ, 481, L23
- Maiolino, R., Krabbe, A., Thatte, N., Genzel, R. 1998 ApJ, 493, 650
- Maiolino, R., Alonso-Herrero, A., Anders, S., Quillen, A., Rieke, M.J., Rieke, G.H., Tacconi-Garman, L.E. 1999 ApJ, submitt.
- Moorwood, A.F.M., van der Werf, P. P., Kotilainen, J. K., Marconi, A., Oliva, E. 1996 A&A, 308, L1
- Shlosman, I., Frank, J., Begelman, M.C. 1989 Nature, 338, 45
- Thatte, N., Kroker, H., Weitzel, L., Tacconi-Garman, L.E., Tecza, M., Krabbe, A., Genzel, R. 1995 Proceedings of the SPIE, 2475, 228
- Thatte, N., Quirrenbach, A., Genzel, R., Maiolino, R., Tecza, M. 1997 ApJ, 490, 650
- Wada, K., Habe, A. 1992 MNRAS, 258, 82
- Weitzel, L., Krabbe, A., Kroker, H., Thatte, N., Tacconi-Garman, L. E., Cameron, M., and Genzel, R. 1996 A&AS, 119, 531

The Kyoto Multimode 3-D Spectrographs

H. Ohtani, T. Ishigaki¹, T. Hayashi², S. Ozaki, T. Hattori, H. Sugai, M. Ishii

Department of Astronomy, Kyoto University, Kyoto 606-8502, Japan

M. Sasaki

Shimonoseki City University, Shimonoseki 751, Japan

1. Instrument

1.1. Kyoto 3D-I

The first generation instrument, Kyoto 3D-I, an area spectrograph in optical region is in commission at the 188cm telescope of the Okayama Astrophysical Observatory (Ohtani et al. 1994, 1998). This instrument is equipped with four modes of area spectroscopic method, i.e., filter imager, long slit for slit scan, imaging Fabry-Perot interferometer, and integral field spectrograph (IFS) (Figure 1). Selection of a particular mode is made by moving the slide table to insert an appropriate focal unit (filter, slit, or aperture plate with enlarger) and switching the ray path with a movable folding mirror. The filter imager and the long slit modes are quite usual ones, and, therefore, these are not described here.

Two Fabry-Perot etalons of Queensgate Instruments are available; one is a tunable narrow band filter and the other is for measurement of velocity field. The etalon parameters are given in Table 1. The observing field of view is 4.6' square. The etalon temperature is stabilized within $\pm 0.5^\circ\text{C}$ to suppress drifts of the transmitting wavelength within one tenth of the width of the Airy profile.

IFS is of the TIGER (Bacon et al. 1995) type, but has a dual-channel enlarging optics before the microlens array; one channel acquiring a target area of an extended object and the other an 'uncontaminated' sky apart from the object by 3.7', simultaneously. A 1".3 square field is subtended by a single lenslet of the microlens array, of which 7 x 11 and 7 x 2 lenslets are allocated to the target and the sky, respectively. Thus, ninety one spectra in total can be obtained within the wavelength range of 4000 to 7000 Å.

1.2. Kyoto 3D-II

The second generation instrument, Kyoto 3D-II, similar to the first but with much higher performances than it is under development. This spectrograph is

¹Present address: Department of Applied Physics, Graduate School of Engineering, Hokkaido University, Sapporo 060-8628, Japan

²Present address: Toyama Science Museum, Toyama 934-8084, Japan

Table 1. Parameters of Fabry-Perot etalons

Etalon	Gap	Wavelength	Finesse	Resolution	Remark
	(μm)	Region ($\text{\AA}$)	at 6328 $\text{\AA}$	R at H α	
No.1	3-8	4000-7000	25	250-600	Tunable Filter
No.2	55 ± 3	4000-7000	40	6800	Velocity Field

intended to be mounted on a 2-meter telescope under construction at Haleakala by Tokyo University and also on the SUBARU.

The optics of the new spectrograph is designed to have excellent image quality and high transmission from 3600 to 9000 $\text{\AA}$; about 40 percent even at 3600 $\text{\AA}$ is expected with a multilayer AR coat. An EEV thinned 2K x 2K CCD with an AR coat is used. Image scales of the Fabry-Perot imaging modes are 0.3" and 0.06" per pixel for the 2 meter telescope and the SUBARU, respectively.

In the IFS mode, a crossed-cylindrical microlens array of quartz manufactured by LIMO is used. The array format is 37 x 37. The pitch corresponds to 0.4" at the 2 meter telescope and to 0.07" at the SUBARU. The latter is optimized for the case when adaptive optics becomes available.

This instrument will be completed in a year. The readers should refer to Sugai et al. (1998) for more details of this spectrograph.

2. Astronomical Results with Kyoto 3D-I

2.1. Observations with the 188cm telescope

Indication of Accelerated Gas Outflow from the Nucleus of NGC4151 A Fabry-Perot interferometric observation at [OIII] 5007 $\text{\AA}$ line was made to obtain velocity field around the nucleus of NGC4151. Line-splitting increases with distance in NE and SW from the nucleus, and suddenly diminishes at 3" in the both directions. This may be due to existence of accelerated gas outflows from the nucleus along the inner surfaces of the bipolar radiation cone embedded in the the galactic disk (Ohtani et al., 1997).

Possible Shock Interaction between the Molecular Ring and the Radio Lobe in NGC1068 A map of [SII]/ H α intensity ratio was obtained from monochromatic images obtained with the tunable filter. Distribution of relative intensity of [SII]/H α is anticorrelated with that of [OIII]/ H α which has the biconical shape along NE-SW axis. In addition, an arc feature is recognized in the [SII]/H α map between the molecular (or starburst) ring and the head of the SW radio lobe. From the shape and [SII] enhancement, this feature seems due to shock interaction between them. (Ishigaki et al., 1998)

Emission Line Imaging of LINERs A survey program of line imaging of LINER galaxies is in progress. In nuclear regions of some objects, there have been detected extended emission line regions whose shape of Balmer line image is different from that of forbidden line images. This may be indication of existence of anisotropy of nonthermal ionizing radiation from the nucleus such as seen in Seyferts. An example, NGC4457, is presented in Figure 2.

Figure 2. Emission line images of the LINER galaxy NGC4457

IR Luminous Interacting System with AGN Extensive 3D spectroscopic observations with imaging Fabry-Perot and slit scan were carried out for the IR luminous interacting systems Mrk266 and NGC7592.

For the former, gases in the largest tidal tail extending to the north have been found to be ionized by a power law continuum which is probably originated from the Seyfert nucleus of the SW galaxy of the system. The nucleus is inferred to be one of the most powerful Seyfert nuclei.

For the latter object, one component has been found to be a disk galaxy with active starbursts. In the other component, a Seyfert 2 galaxy, has been detected a biconical ionizing radiation cone and a starburst ring surrounding the nucleus.

Details of these studies are presented elsewhere in the present conference (Ishigaki et al. 1999, and Hattori et al. 1999).

Velocity Field of the Bipolar Planetary Nebula Hb4 An integral field spectrogram of the bipolar planetary nebula Hb4 was obtained. The velocity field was analyzed (Figure 3). The northern part and the southern part of the main body around the central star are approaching to and receding from us, respectively, while the northern jet and the southern one are receding and approaching, respectively. This suggests that the main body is an expanding disk along whose axis the bipolar jet was ejected.

2.2. Super-Wide-Field Observation

A telephoto lens of focal length 300mm can be attached to the spectrograph instead of the telescope. The spectrograph with the lens is mounted on the tube of a 60 cm telescope. Then, a very wide panoramic 3D-spectral view is obtained. The observing parameters in this case are tabulated in Table 2.

The Rosette Nebula: Shock Interaction with a Supernova Remnant Line images of [SII] and H α were obtained with the tunable filter. [SII]/H α intensity ratio

Figure 3. Bipolar planetary nebula Hb 4. (a) H α image (Schwarz et al., 1992), (b) and (c) H α velocity field

Table 2. Super-Wide-Field

Mode	F.O.V.	Spatial resolution
F-P	7 x 7 degree	1.0 arc min
IFS	18 x 28 arc min	2.5 arc min

is large in three parts of the peripheral regions of the Rosette. In the two among the three, the SE and the NW parts, molecular clouds are detected and, therefore, these parts are partially ionized regions like the Orion Bright Bar. In the other, the NE part, neither CO nor HI radio line is detected. The Monoceros Loop, a ring of faint filamentary nebulosities seen between the Cone Nebula and the Rosette apparently contacts here with the latter. From these, the [SII] enhancement in the NE part is interpreted as the result of shock excitation due to collision with the Loop. This also implies that the Loop is the remnant of a supernova exploded in the Monoceros star forming complex.

Global Ionization Structure of the Orion Nebula 3D spectrum for the whole of the Orion Nebula was obtained in the IFS mode and maps of various line ratios were made. From these, radial variations of the line intensity ratios can be traced up to the periphery of the Nebula, i.e., nearly to the surface of the Strömgren sphere. So far, no observation of such a large scale was performed. Distribution of [NII]/H α ratio is presented in Figure 4 where a general tendency of increase of the ratio near the periphery is seen. This is consistent with results of theoretical model calculations. Farther study will be useful to examine details of the UV spectral energy distributions of the exciting stars.

References

Bacon, R. et al., 1995, A&AS, 113, 347.

Figure 4. Distribution of $[NII]/H\alpha$ ratio for the whole area of the Orion Nebula. The denser the gray scale, the ratio is the smaller

- Ishigaki, T. et al., 1998, in Central Regions of the Galaxy and Galaxies, IAU Symposium No.184, 119.
- Ishigaki, T. et al., 2000, this volume.
- Hattori, T. et al., 2000, this volume.
- Ohtani, H. et al., 1994, in Instrumentation in Astronomy VIII, SPIE Symp., Vol. 2198, 229
- Ohtani, H. et al., 1998, in Optical Instrumentation in Astronomy, SPIE Symp., Vol. 3355, 750.
- Ohtani, H. et al., 1997, in Emission Lines in Active Galaxies: New Methods and Techniques, IAU Col.159, 355.
- Schwarz, H. E., 1992, A&AS, 96, 23.
- Sugai, H. et al., 1998, in Optical Instrumentation in Astronomy, SPIE Symp., Vol. 3355, 665.

Integral Field Spectroscopy with GEMINI and Space Telescopes

~~Jeremy~~ Allington-Smith, Robert Content, ~~Roger~~ Haynes & David Robertson

*Astronomical Instrumentation Group, University of Durham Physics
Dept, South Rd, Durham DH1 3LE, UK*

Abstract. We describe work on instruments for integral field spectroscopy (IFS) for the GEMINI telescopes, together with prototypes built for use with AO and in the near-IR under implementation on 4m telescopes. We also briefly describe progress on a spectrograph for the Next Generation Space Telescopes. Finally, we highlight recent work on systems for multiple IFS.

1. Introduction

Fig. 1 summarises the main techniques of dispersive IFS. The *lenslet-only* approach (e.g. TIGER, Bacon *et al.* 1985; SAURON, Miller *et al.*, these proceedings) has been widely used because of its relative simplicity but compromises the length of spectrum to avoid overlaps between spatially-independent spectra. The *fibre-lenslet* technique (F-L; see references below) uses fibres to reformat the field into a pseudo-slit to eliminate spectrum overlaps. Lenslet arrays coupled to the fibres provides unit fill factor and efficient coupling to the telescope. This design maximises spectrum length and allows more efficient utilisation of the detector surface (Allington-Smith & Content 1998). The slicer approach (e.g. 3D; Weitzel *et al.* 1996 and references below) retains spatial information along each slice and so makes optimum use of the detector surface. Since the optical surfaces can be made of the same material as the mount, it is particularly suitable for use in cryogenic instruments.

2. Prototypes on 4m telescopes

We have built two F-L IFUs for use on 4m telescopes. Both use fused multimode step-index silica fibre arranged in an matrix of microtubes. Arrays of lenslets are used at both the input and output to adapt the beam to that of the telescope and spectrograph. These are custom-made to match exactly the pitch of the fibre mounting.

SMIRFS is an uncooled near-infrared system coupled to the CGS4 cryogenic 1-5 μ m spectrograph on UKIRT. It is limited to 72 spatial elements by the spectrograph optics which are arranged to give a 6x4 arcsec field with 0.6 arcsec sampling. An example of data obtained with SMIRFS on NGC4151 is given in

Figure 1. Summary of different techniques of dispersive IFS.

Fig. 2. The throughput is $\sim 50\%$ in both the J and H bands (Haynes *et al.* 1998, 1999).

TEIFU, the *Thousand Element Integral Field Unit (TEIFU)* contains two groups of fibres in bundles of 504. It is designed primarily to exploit the sharpened images produced by the ELECTRA AO system (Buscher *et al.* 1996) via switchable scales of 0.25 and 0.13 arcsec/element. The two sub-fields may be combined into a single field of 54 square arcsec (0.25 arcsec sampling) or given a variable separation by means of a segmented pickoff mirror. This arrangement gives considerable flexibility in background subtraction since one field may be dedicated to background estimation and the two may be beam-switched while maintaining lock on the reference star. The first bundle is now complete, with commissioning expected in June 1999. It will be adapted for use with the common-user NAOMI AO system being developed from ELECTRA.

3. GEMINI integral field spectrographs

GMOS, the *GEMINI Multiobject Spectrographs* (Murowinski *et al.* 1998, Allington-Smith *et al.* 1997) will be equipped for IFS by sliding the IFU into the input focal plane in place of a slit mask, using the same interchange mechanism. The required field area is 50 square with $\sqrt{2}$ aspect ratio and 0.2 arcsec sampling. The GMOS-IFU will require at least 1500 elements which represents a big increase over TEIFU. The baseline design (Fig. 4) envisaged an object field made up of a single bundle, but experience with TEIFU suggests an alternative design in which three bundles of ≥ 500 fibres are optically combined into a single field. A background field will be provided at a fixed separation of 1 arcmin to facilitate accurate background subtraction, especially when used with AO.

GNIRS, The *GEMINI Near-infrared spectrograph* (Elias *et al.* 1997), is a cryogenic 1-5 μm instrument. It will convert to IFS mode by the insertion of IFUs into the relayed focal plane after the fore-optics. Two units are available

Figure 2. Data on NGC4151 obtained with SMIRFS

Figure 3. Schematic showing the optical principle of TEIFU and its twin-field pickoff.

Figure 4. Mechanical packaging of GMOS IFU (baseline design).

with sampling of 0.04 and 0.15 arcsec over fields of 1.4 and 14 square arcsec respectively. The number of spatial samples is limited by the 1024x1024 detector format. The IFU uses an *Advanced Image Slicer (AIS)* design developed from 3D (Weitzel et al. 1996) by Content (1998). The optical surfaces (Fig. 5) will be diamond-turned from aluminium, the same material as the mount, to minimise thermal problems (Fig. 6).

4. IFS with the Next Generation Space Telescope

For NGST, a combined *integral field/multiobject spectrograph (IFMOS)* has been studied for ESA by a European consortium for possible inclusion in the science module. This provides two modes: LR — 1800 square arcsec at 0.19 arcsec sampling intended for low spectral resolution MOS where every spatial element in the field is sampled (appropriate near the galaxy confusion limit); and HR — 10 square arcsec at 0.05 arcsec sampling for detailed mapping of individual objects at high spectral resolution.

Since the total number of spatial samples is 58,000, only an AIS image slicer design is feasible since it makes optimum use of the detector format and minimises losses due to diffraction. Our design includes anamorphic magnification so that a 2x1 pixel region on the detector projects to a square region on the sky to provide correct spectral sampling without compromising spatial coverage. The instrument requires no moving parts for simultaneous coverage of two octaves within 1-5 μ m (one octave for the HR mode). (See the talks by Olivier LeFevre and Robert Content in these proceedings for further information.)

5. Multiple IFS

The techniques discussed above can be applied to multiple IFS whereby a modest number of IFUs are deployed within a larger field. The multiplex gain must be

traded-off against the sub-field size, sampling and spectrum length because of the limit imposed by the total number of detector pixels available. We are studying two such systems.

FMOS is a MOS/IFS fibre visible/near-IR spectrograph proposed for SUBARU. The baseline IFS system has 15 sub-fields of ~ 5 arcsec with a sampling of 0.3 arcsec. These are deployable within the 30 arcmin prime focus field. The 3000 IFS fibres are routed via a connector block incorporating a visible/near-IR dichroic splitter into a set of near-infrared and visible-light spectrographs. For this application, an F-L system gives great flexibility in deployment, similar to that of MOS fibre positioners such as Autofib-2 and 2dF.

GIRMOS is a concept for a MOS/IFS 1-2.5 μ m instrument for GEMINI. Each IFU would provide a field of 3 arcsec with 0.15 arcsec sampling. The spectrum length, appropriate for $R > 6000$, would require ≥ 4000 pixels in the dispersion direction. Since the spectrograph must be cryogenically cooled, it is not clear if fibres can be used. Consequently, we are studying an AIS system using diamond-turned metal optics. The deployment mechanism is under development (see the talk by Gillian Wright in these proceedings).

We thank our colleagues on these projects including David Lee, Ian Parry, Peter Hastings, Gillian Wright, Olivier Le Fevre, Eric Prieto, Naranjan Thatte, Roland Bacon, Winfried Posselt, David Crampton, Roger Davies, Rick Murowinski, Terry Purkins, Hiroshi Karoji to name only a few. We also acknowledge the support of our colleagues in the SUBARU and GEMINI telescope projects.

References

- Allington-Smith, J., *et al.*, 1997. *Experimental Astronomy*. 7, 293-299.
- Allington-Smith, J. & Content R., 1998. *PASP*, 110, 1216-1234.
- Bacon R. *et al.*, 1995. *A&A Suppl.* 113, 347.
- Buscher D., *et al.* 1996. 'ESO/OSA Topical Meeting on Adaptive Optics, p63'
- Content, R., 1998. *SPIE* 3354, 187-200.
- Haynes, R., *et al.* 1998. *SPIE* 3354, 419-430.
- Haynes, R., *et al.* 1999. submitted to *PASP*.
- Murowinski, R., *et al.*, 1998. *SPIE* 3355, 188-194.
- Weitzel, L., *et al.* 1996 *A&AS*, 119, 531

Integral Field Spectroscopy of NGC 2992

B. García-Lorenzo¹, S. Arribas^{1,2}, and E. Mediavilla¹

*Instituto de Astrofísica de Canarias, 38200-La Laguna, Tenerife, Spain
Consejo Superior de Investigaciones Científicas*

Abstract. Two dimensional spectroscopy with optical fibers has been obtained for the circumnuclear region of NGC 2992. The observed spectral range includes emission and absorption lines that enable the ionized gas and stellar kinematics to be obtained.

1. Introduction

NGC 2992 is a nearby galaxy at a distance of ~ 31 Mpc. This allows its circumnuclear region ($1'' \sim 150$ pc) to be sampled in great detail. NGC 2992 is considered as a hybrid between an intermediate Seyfert and a starburst galaxy, the latter being induced by the interaction with NGC 2993. NGC 2992 shows prominent signatures of interaction with its companion. The radio emission shows a figure-of-eight structure along a position angle of about 150 degrees.

Kinematic studies of the circumnuclear region of this galaxy have focused on the ionized gas. Previous long-slit observations at different position angles report asymmetrical emission-line profiles in the circumnuclear region and complex kinematics (Colina et al. 1987; Márquez et al. 1998).

2. Instrumentation and observations

We used the new fiber optical system INTEGRAL to perform two-dimensional spectroscopy of NGC 2992 (Figure 1). INTEGRAL links the Nasmyth focus of the WHT with the WYFFOS spectrograph. It currently has three fiber bundles which can be interchanged on-line during observations (more details in Arribas et al. 1998).

NGC 2992 was observed on 1998 April 4 and 5 using the standard bundle 2 (SB2) of INTEGRAL. SB2 has fibers $200 \mu\text{m}$ in diameter which are arranged in two groups in the focal plane. One forms an array with a spatial coverage of $16'' \times 12''$ and the other forms a ring of $90''$ in diameter. When we observe an extended object with this system, each fiber of the central array receives light from the object, while the ring fibers should collect background light. However, in the case of NGC 2992, some ring fibers receive light from the tidal tail (figure 2). We observed two overlapped spectral ranges, $4360 - 5830 \text{ \AA}$ and $5675 - 7125 \text{ \AA}$ with non-photometric conditions and a seeing of about $1''$.

Figure 1. INTEGRAL is an optical fiber system for performing 2D spectroscopy of extended objects at the 4.2 m William Herschel Telescope (WHT). It is mounted at the GHRIL Nasmyth focus together with newly built acquisition, guiding, and calibration units. It makes use of the specially designed fiber spectrograph WYFFOS. This system allows up to six bundles to be mounted simultaneously. It currently contains three science oriented fiber bundles, any one of which can be easily and quickly placed in the telescope beam. Their spatial resolution elements (fiber core diameters) are $0''.45$, $0''.9$, and $2''.7$, respectively. Hence, depending on the prevailing seeing conditions the instrument can be easily optimized for the scientific program (Arribas et al. 1998).

Figure 2. The interacting system NGC 2992-2993. The spatial distribution of the SB2 fibers at the focal plane has been overplotted on NGC 2992.

Figure 3. (a) The 189 spectra from the central region of NGC 2992 in the spectra range 4950-5080 Å, including the [OIII]λλ4959, 5007 emission lines; (b) continuum map in the spectral range 5100-5150 Å; (c) velocity field of the ionized gas in the central region of NGC 2992 inferred from cross-correlation in the spectral range including the [OIII]λλ4959, 5007 emission lines; and, (d) velocity field of the stellar component in the central region of NGC 2992 inferred from cross-correlation in the spectral range of MgIb absorption lines.

3. Preliminary results direct from data

Figure 3(a) shows spectra in the [OIII] $\lambda\lambda$ 4959, 5007 emission lines after reduction. From any spectral feature (intensity, velocity, etc) we can recover maps using the Renka & Cline two-dimensional interpolation method (E01SAF and E01SBF routines of the NAG Fortran Library Routine Document).

The continuum map (Figure 3(b)) clearly shows a dust lane crossing the galaxy. The isophotes in the observed region are elongated along a position angle of about ~ 36 degrees, whereas the photometric major axis angle is $\sim 15^\circ$ (at large scales). This twist may be due to the interaction with its companion NGC 2993.

The stellar velocity field (Figure 3(d)) has been obtained by cross-correlation of the MgIb absorption lines in the spectral range 5140 – 5260 Å and could be well explained by a rotating disk around the optical nucleus. Its major kinematic axis seems to be in a direction nearly parallel to the photometric major axis in the circumnuclear region.

The ionized gas velocity field has been obtained cross-correlating the [OIII] emission lines in the spectra range 4940 – 5100 Å (figure 3(c)). The major kinematic axis of the ionized gas seems to be aligned with that corresponding to the stars along P.A. $\sim 35^\circ$. Thus, the velocity field of the ionized gas could also be explained in terms of a disk rotating around the optical nucleus. This suggests that stellar and ionized gas components are coupling in the circumnuclear region of NGC 2992. However, the ionized gas kinematics show some peculiarities with respect to the regular rotation. The minor kinematic axis twists in S-shape, and there is a high-velocity region at the south-west, but the rotational component seems to be the dominant one. These peculiarities could be related to asymmetries in the emission line profiles.

4. Conclusions

Two-dimensional spectroscopy of the central region of the Seyfert galaxy NGC 2992 has been obtained. Preliminary analysis of the data indicates that the major axis of the ionized gas and the stellar velocity fields agree with the inner photometric major axis. Therefore, the ionized gas and the stellar disks seem to rotate in a same plane.

We appreciate Terry Mahoney's help with the manuscript. The WHT is operated on the island of La Palma by ING in the Spanish Observatorio del Roque de Los Muchachos of the Instituto de Astrofísica de Canarias. This work has been partially supported by the Spanish Dirección General de Investigación Científica y Técnica (PB93-0658).

References

- Arribas, S. et al. 1998, SPIE, 3355, 821
- Colina, L., Fricke, K.J., Kollatschny, W., & Perryman, M.A.C. 1987, A&A, 178, 51
- Márquez, I., Boisson, C., Durret, F., & Petitjean, P. 1998, A&A, 333, 459

Tridimensional Spectroscopy of the Interacting System NGC7592

T. G. Hattori⁽¹⁾, M. Yoshida⁽²⁾, H. Ohtani⁽¹⁾, T. Ishigaki⁽³⁾, H. Sugai⁽¹⁾,
T. Hayashi⁽⁴⁾, S. Ozaki⁽¹⁾ and M. Ishii⁽¹⁾

(1) *Department of Astronomy, Kyoto University, Kyoto 606-8502,
Japan*

(2) *National Astronomical Observatory of Japan, Tokyo 181-8588,
Japan*

(3) *Department of Applied Physics, Graduate School of Engineering,
Hokkaido University, Sapporo 060-8628, Japan*

(4) *Toyama Science Museum, Toyama 934-8084, Japan*

1. Introduction

Many statistical studies established that galaxy interactions can enhance the star-forming and AGN activities in galaxies. Although theoretical calculations and numerical modeling have shown that gravitational interactions can drive interstellar gas down to nuclear regions to initiate the activities, the diverse nature of activities in interacting galaxies is not fully understood.

With detailed optical spectroscopy, one can investigate the physical states of ionized gases, which are essential information to reveal the nature of the activity, and the kinematical properties of the interacting system. As interacting systems are very complex in their morphological appearance, 3D spectroscopic observations may offer important keys to understand the effects of interaction on activities.

NGC7592 is an interacting system with three prominent peaks in optical broad-band image (e.g. Mazzarella and Boroson 1993) and also is a luminous infrared galaxy ($L_{IR} \sim 2 \times 10^{11} L_{\odot}$). The peaks show strong emission lines. From the ratios of these emission lines, one of the peaks is classified as Seyfert 2 (north-western nucleus) and others are classified as HII nuclei. In order to investigate the star-forming/AGN activities and their relation with the interaction in this system, an extensive tridimensional spectroscopy was executed. We also utilized an available HST image.

2. Observations

Observations were carried out with the 188cm telescope at the Okayama Astrophysical Observatory.

H α and [OIII] images were obtained by using Kyoto 3D-I in the mode of imaging Fabry-Perot interferometer with a low dispersion etalon as a tunable filter (see Ohtani et al. in this proceedings). The bandpasses of these observations are 21.1Å and 20.4Å, respectively, for the H α and [OIII].

Figure 1. Observed properties in an area around the Seyfert 2 nucleus. North is up and east to the left. The nucleus is at (0,0). The area is indicated in Fig. 2

Long-slit spectral scans were obtained by using a Cassegrain spectrograph in the Spectro-NebulaGraph mode (Kosugi et al. 1995). The observed wavelength regions were 4600-5400Å and 6200-7000Å, with a spectral resolution of $\sim 3\text{\AA}$. The target object was scanned in a direction perpendicular to the slit orientation (P.A. 90°).

3. Results and Discussion

3.1. Structure around the Seyfert Nucleus

Fig. 1 shows the observed properties around the Seyfert 2 nucleus.

1. The highly excited gases are distributed along nearly north-south direction ([OIII]/H α).
2. In the narrow-band images from which we subtracted the unresolved nuclear component, both H α and [OIII] show elongations and they are nearly perpendicular each other.

Figure 2. $H\alpha$ image. The square in this figure represents a region of panels shown in Fig. 1

3. In an HST broad-band visual image, a structure elongated to the same direction as the $H\alpha$ image is seen.
4. The emission line profiles of $H\alpha$ and [NII] change gradually from blue-ward asymmetry in the east of the nucleus to red-ward asymmetry in the west. Spectra along P.A. $\sim -20^\circ$ have similar profiles and radial velocities to one another. This indicates a rotational motion around an axis of P.A. $\sim -20^\circ$

From the analysis of emission line ratios, Rafanelli and Marziani (1992) showed that the optical spectrum from the Seyfert nucleus is a mixture of lights from two components, HII regions and the NLR (and/or the ENLR). According to their analysis, the HII regions make a substantial contribution to the $H\alpha$ flux. Because the gaseous component ionized by the AGN seem to distribute along the N-S direction, we consider that the E-W elongations in the $H\alpha$ and the HST images correspond to the other component, namely the HII regions. If these HII regions surround and rotate around the nucleus, the line profile distribution shown in Fig. 1 can be also explained.

Thus, these results suggest that the Seyfert nucleus is surrounded by a star-forming disk or ring and illuminates gases in the direction perpendicular to the plane of the star-forming region.

3.2. Nature of the Host Galaxies

Fig. 3 shows the $H\alpha$ mean velocity field, and color distribution obtained from continuum images which are free from emission lines. The smoothness of the velocity field in the eastern region suggests that north-eastern and southern HII nuclei belong to a common galaxy. In the color image, the western region which contains Seyfert nucleus shows redder color with respect to the eastern region. The border of the two regions is fairly distinct, and it may be the actual border of a western galaxy and a eastern one.

The velocity field in the eastern region seems to indicate a rotation in a disk galaxy. We compared the data with simple models to determine parameters (1)center of rotation, (2)P.A. of the major axis, and (3)inclination of the disk. The velocity field was fit well with rotation models, and two parameters were determined with sufficient precision: P.A. of the major axis and the position of the center along the major axis. These determined parameters agree with the P.A. of the major axis and the peak position in the K-band image by Bushouse

Figure 3. The $H\alpha$ mean velocity field and color distribution. The superposed contour map is the $H\alpha$ image.

and Werner (1990). The agreement supports the above statement that the velocity field in the eastern region represents a rotation of a disk galaxy, and, if this is correct, the center of the galaxy lies at the K-band peak, which is $\sim 2''$ ($\sim 1\text{kpc}$) from the optical HII "nucleus" ($H\alpha$ peak).

For star-forming regions in this interacting system, we find the followings.

1. In the eastern galaxy, star-forming regions are distributed in a very patchy and asymmetric manner, which may be a result of heavy extinction, and extend over 10kpc scale.
2. The western galaxy have redder color and more compactly distributed star-forming regions than the eastern galaxy.

A similar situation is seen in the interacting system NGC7469/IC5283 studied by Marquez and Moles (1994). They suggested that the difference in the distributions of star-forming regions could be the result of different galaxy type (i.e. the different response to the dynamical perturbation).

In NGC7592, it is difficult to explain the redder color in the western galaxy by dust extinction, because it would require a uniform dust distribution in a very large scale. Therefore, the color difference suggests that the western galaxy is an earlier type galaxy than the eastern one. If this is correct, the same mechanism may cause the difference in the star-forming properties between the eastern and western galaxies as for the NGC7469/IC5283 system. However, these considerations are still a matter of conjecture and further observations are needed to confirm them.

References

- Bushouse, H. A., & Werner, M. W. 1990, *ApJ*, 359, 72
 Kosugi, G., Ohtani, H., Sasaki, T., Koyano, H., Shimizu, Y., Yoshida, M., Sasaki, M., Aoki, K., & Baba, A. 1995, *PASP*, 107, 474
 Marquez, I., & Moles, M. 1994, *AJ*, 108, 90
 Mazzarella, J. M., & Boroson, T. A. 1993, *ApJS*, 85, 27
 Rafanelli, P., & Marziani, P. 1992, *AJ*, 103, 743

First Results with a Wide-Field Near-Infrared Integral Field Unit

T. M. Herbst

*Max-Planck-Institut für Astronomie, Königstuhl 17, 69117 Heidelberg,
Germany*

Abstract. This paper presents first astronomical results using a near infrared integral field unit (IFU) designed for very high throughput ($> 90\%$), many spatial pixels (128×128), and a wide field of view (> 1 arcminute). Each spectrum contains approximately twenty spectral channels. We achieve the high efficiency and wide field of view by placing an array of microlenses in the focal plane wheel of a pre-existing camera/spectrograph. A narrow-band filter prevents overlap of the individual traces. First trials at the telescope produced very promising results.

1. Spatial Field vs Spectral Coverage

In an imaging spectrograph, the relatively small size of visible wavelength and, especially, infrared detectors forces a tradeoff between spatial and spectral coverage: the product of the number of spatial and spectral channels must be less than or equal to the number of detectors on the array. A typical imaging spectrograph may have of order 15×15 spatial channels (covering $5'' \times 5''$) and approximately 1000 spectral elements. This balance between spatial field and spectral coverage may not represent the best compromise from the viewpoint of the astronomer, however. There are many astrophysical problems in which a larger field is paramount. For example, detecting extragalactic HII regions or determining the velocity dispersion of a galaxy cluster requires ~ 20 spectral channels (and not 1000) spanning a single emission line. Our integral field unit represents a significant departure from existing and planned instruments, in that it provides many spatial pixels and a very wide field of view, but with relatively few spectral channels (see figure 1). This larger, one arcminute field of view is well-matched to real astrophysical targets and allows a measurement of "true" sky.

2. Optical Concept

Our integral field spectrograph works in combination with Omega Cass, a large format near infrared camera on the Calar Alto 3.5 meter telescope. We take advantage of a unique aspect of this camera/telescope combination: Omega Cass can operate at $f/10$, and the telescope has an $f/45$ secondary. Omega Cass uses a "classical" collimator-camera optical path. The telescope focal plane falls on a movable field stop wheel located just inside the cryostat window. A

Figure 1. Comparison of the new integral field unit (gray) with existing and planned instruments. The two visible wavelength instruments, Gemini/GMOS and TIGER, have been offset 30" along the "field of view" axis for clarity.

four-element collimator lens forms a pupil image, where filters and grisms are installed. A second, four-element camera lens re-images the telescope focus on the 1024×1024 HgCdTe detector.

Converting this instrument into an integral field spectrograph is relatively straightforward. A 128×128 array of microlenses is formed in an epoxy substrate using a negative master (similar to the production of a replica grating). This array of lenses lies in the focal plane of the telescope and forms multiple images of the telescope pupil in the focal plane of the camera. In normal operation, the telescope and camera focal planes coincide – for the imaging spectrograph, the telescope focus must be moved upward approximately 4 mm. See figure 2 for clarification.

With the $f/45$ secondary on the 3.5 meter telescope and 3.8 mm focal length lenslets, these "micropupils" are essentially small, 1.8 pixel diameter, images of the primary mirror of the telescope diverging at $f/10$ (see figure 2). The image on the detector also resembles the grid of pupil images shown in figure 2 (also see figure 3). Inserting one of the Omega Cass grisms into the optical path disperses the light, resulting in an array of 128×128 spectra on the detector. Rotating the microlens unit and using a narrow band (1%) filter as an order sorter prevents overlap, giving up to 20 channels per spectrum per sky position. See Herbst (1998, SPIE 3354, p. 720) for a more complete description. This overall scheme resembles the TIGER imaging spectrograph, although the latter instrument adopts the smaller area/many spectral channels approach (figure 1).

3. Laboratory Tests

Very little is known about the cold performance of microlens arrays, including, for example, whether the epoxy lenses will survive being cooled to 77 K. To ad-

Figure 2. Micropupil formation in the focal plane of the infrared camera.

dress this situation, we installed a sample microlens array from Adaptive Optics Associates into a small test cryostat. In addition to survivability tests, we measured the focal length shift between ambient and liquid nitrogen temperatures using a visible-wavelength CCD camera. This shift arises due to changes in the refractive indices of the epoxy and substrate with temperature, as well as from changes in the shape of the lenses due to contraction.

The laboratory tests were very promising. We found that:

- The sample microlenses survived two cool-downs (and one vacuum leak which formed ice on the lens array!). The same device has since survived multiple thermal cycles in the infrared camera.
- The focal length shortened by roughly 0.7% between ambient and liquid nitrogen temperature (77K).
- The array produced diffraction limited images both warm and cold.

4. First Observations

We installed the sample microlens array into the focal plane wheel of Omega Cass in autumn 1998 and attempted observations at the 3.5 m telescope on Calar Alto. Despite poor weather and unrelated problems with the detector, we managed to perform some test observations of astronomical targets. Figure 3 shows the array of micropupils produced by a bright standard star (left), and the corresponding spectra after inserting the grism. Figure 4 contains an image of the well-known planetary nebula NGC 7027 taken in the $2.12 \mu\text{m}$ quadrupole line of molecular hydrogen. The spectrogram reveals line emission superimposed on the dust continuum in the central regions, and pure H_2 emission in the periphery.

Note that the H_2 emission line is not centered in the continuum spectrum isolated by the filter. This was due to a problem, since corrected, with the narrow band filter. The IFU is clearly a powerful tool for investigating complex astronomical sources, as well as complex astronomical instruments! Further observations with the IFU are scheduled for summer 1999.

Figure 3. First observations with the near infrared IFU. A bright star produces distinct micropupils (left) and spectra (right).

Figure 4. IFU image of the planetary nebula NGC 7027 in the 2.12 μm line of molecular hydrogen. Note the complex mix of line and continuum emission.

The disturbed structure of the nuclear region of Mkn298

A. Rifatto¹, P. Rafanelli^{1,2,3}, V. Afanasiev⁴, S. Dodonov⁴, K. Birkle⁵, P. Böhm³, T. Boller⁶, M. Radovich², G. Richter³ and J. Vennik⁷

¹*Osservatorio Astronomico di Capodimonte, Napoli, Italy*

²*Dipartimento di Astronomia, Università di Padova, Padova, Italy*

³*Astrophysikalisches Institut Potsdam, Potsdam, Germany*

⁴*Special Astrophysical Observatory, Zelenchukskaya, Russia*

⁵*Max-Planck-Institut für Astronomie, Heidelberg, Germany*

⁶*Max-Planck-Institut für Extraterrestrische Physik, Garching, Germany*

⁷*Tartu Observatory, Tõravere, Estonia*

Abstract. The 6 m telescope of the Special Astrophysical Observatory equipped with the Multi Pupil Fiber Spectrograph (MPFS) has been used to perform 3D-spectroscopy of the nuclear region of the galaxy Mkn298 (IC1182). The array consisted of 16×16 fibers, covering a region of angular size $16'' \times 16''$. The distribution of the emission in the [OIII] $\lambda 5007$ line shows two distinct bright knots, while the distribution of the emission in the nearby continuum ($\lambda 5200$) is characterized only by a single peak, the nucleus of the galaxy, the position of which is misaligned with both [OIII] knots. The emission line ratios in the spectra of the two knots suggest the presence of gas photoionized by both thermal and nonthermal sources, like in starburst regions close to an active nucleus. The kinematics of the gas component indicates that the two detected knots are either rotating around the nucleus or moving in opposite directions relative to it. The interpretation in terms of circumnuclear starbursts partly covering an active nucleus is discussed.

1. Introduction

Mkn298 ($z = 0.034$, scale ~ 660 pc/arcsec, for $H = 75$ km s⁻¹ Mpc⁻¹) is a galaxy, the classification of which as AGN is controversial in literature. Koski (1978) and Veilleux & Osterbrock (1987) classified Mkn298 as a Seyfert-2 galaxy with extremely small [NII] $\lambda 6583$ /H α line ratio. Conversely Viegas-Aldrovandi & Gruenwald (1990) enclosed Mkn298 in a sample of LINERs. Optical variability was claimed by Metik & Pronik (1981) who observed variations of the relative intensity of [OIII] $\lambda\lambda 4959, 5007$ and H β in the periods 1969–70 and 1977. More recently, Rafanelli et al. (1999) found that Mkn298 is a soft X-ray source, showing variability in the time scale of days.

Figure 1. MPFS 16×16 array overlayed to the continuum subtracted $H\alpha$ image of Mkn298. North at the top, East to the left

2. Spectral Observations and Results

Spectra of the nuclear region were taken using the Multi Pupil Fiber Spectrograph, MPFS, at the 6 m telescope of the Special Astrophysical Observatory of the Russian Academy of Sciences at Zelenchukskaya. The array, consisting of 16×16 microlenses, each of angular size $1'' \times 1''$, covered an area of $\sim 16'' \times 16''$ centered on the nucleus and oriented at P.A. = 150° , as shown in Figure 1 where the position of the array has been drawn on a continuum subtracted $H\alpha$ narrow band image. The 256 spectra obtained in this way have a typical resolution of $\sim 4 \text{ \AA}$, and cover the spectral range 4450–7100 $\text{\AA}$.

Slices of the MPFS data cube in correspondence of the [OIII] $\lambda 5007$ line and of its nearby continuum [OIII](cont) have been plotted as [OIII] contour map overlayed to the [OIII](cont) image (Figure 2). The distributions of the $H\alpha$ (Figure 1) and of the [OIII] $\lambda 5007$ (Figure 2) emitting gas are similar, both characterized by two distinct emitting knots, aligned at P.A. = 115° and separated by $3.3''$. We call them SEK (South East Knot) and NWK (North West Knot) on the basis of their relative position. No double structure is present in the [OIII](cont) image which shows a single emitting nucleus lying between SEK and NWK at $0.8''$ from the peak of SEK (Figure 2). Since the $H\alpha$ angular sizes of SEK and NWK are $3.6''$ and $2.1''$ respectively, the center of the nucleus lies within or behind SEK (Figure 2).

Dynamically SEK and NWK are moving in opposite radial directions (Figure 3) like two knots moving each other away or rotating around an axis at P.A. $\sim 40^\circ$, through the continuum emitting nucleus.

All spectra obtained from the fibers covering NWK and SEK show strong absorption features, such as the MgI b blend $\lambda 5175$, the FeI+CaI blend $\lambda 5269$ and the NaI D line $\lambda 5893$, due to the underlying galaxy. An absorption-line template spectrum was therefore subtracted by matching the absorption features, and then the emission line fluxes were measured. The line ratios proposed by Veilleux & Osterbrock (1987) were then derived from the measured line intensi-

Figure 2. Morphology of the galaxy Mkn298 in the continuum near $\lambda = 5200 \text{ \AA}$ (image) and in the [OIII] $\lambda 5007$ line (contours).

Figure 3. Image of Mkn298 in the [OIII] $\lambda 5007$ line with overlaid the isoveLOCITY contours.

ties and plotted in the diagnostic diagrams which report $[\text{OIII}] \lambda 5007 / \text{H}\beta$ versus $[\text{SII}] \lambda \lambda 6716 + 6731 / \text{H}\alpha$, $[\text{NII}] \lambda 6583 / \text{H}\alpha$, $[\text{OI}] \lambda 6300 / \text{H}\alpha$.

The $[\text{NII}] \lambda 6583$ lines are systematically too faint, $\log([\text{NII}]/\text{H}\alpha) < -0.3$, in both knots to justify the presence of a nonthermal ionizing source as it is suggested by the intensity ratios $[\text{SII}]/\text{H}\alpha$ and $[\text{OI}]/\text{H}\alpha$. A similar situation could be reproduced also combining thermal photoionization with shock ionization (Baldwin et al. 1981). The coexistence of bursts of star formation ($L(\text{H}\alpha)_{\text{SEK}} = 2.5 \cdot 10^{41} \text{ erg s}^{-1}$, $L(\text{H}\alpha)_{\text{NWK}} = 1.8 \cdot 10^{41} \text{ erg s}^{-1}$) together with a nonthermal photoionizing source and/or with shock ionizing sources seems then to be quite likely.

ROSAT PSPC data give an absorption corrected luminosity $L(0.1\text{--}2.4 \text{ keV}) = 1.3 \cdot 10^{42} \text{ erg s}^{-1}$, a value typical of Seyfert-1 nuclei, for which it is $\bar{L}_{\text{S1}}(0.1\text{--}2.4 \text{ keV}) = 5.6 \cdot 10^{42} \text{ erg s}^{-1}$ (Rush et al. 1996; Rafanelli et al. 1997).

3. Conclusions

The optical and soft X-ray features of the nuclear region of Mkn298 can be summarized as follows:

- a) presence of two gaseous regions (SEK, NWK), both emitting in the $[\text{OIII}] \lambda 5007$ and $\text{H}\alpha$ lines ($L(\text{H}\alpha) \sim 10^{41} \text{ erg s}^{-1}$, typical of starburst regions) and both photoionized by a mixture of thermal and nonthermal radiation;
- b) presence of a continuum nuclear source aligned neither with SEK nor with NWK;
- c) presence of a soft X-ray source of luminosity $L(0.1\text{--}2.4 \text{ keV}) = 1.3 \cdot 10^{42} \text{ erg s}^{-1}$, typical of Seyfert-1 nuclei.

On the basis of these data we suggest for the nuclear region of Mkn298 a scenario in which two giant starbursts SEK and NWK are hiding an active nucleus, the radiation of which partially ionizes them.

References

- Baldwin, J.A., Phillips, M.M., & Terlevich, R., 1981, *PASP*, 93, 5
- Koski, A.T., 1978, *ApJ*, 223, 56
- Metik, L.P., & Pronik, I.I., 1981, *Astrofizika*, 17, 629
- Rafanelli, P., Piro, L., Radovich, M., Rifatto, A., et al. 1997, *Mem. SAI*, 68, 301
- Rafanelli, P., Afanasiev, V., Birkle, K., Böhm, P., et al. 1999, *Mem. SAI*, (in press)
- Rush, B., Malkan Matthew, A., Fink Henner, H., & Voges, W., 1996, *ApJ*, 471, 190
- Veilleux, S., Osterbrock, D.E., 1987, *ApJS*, 63, 295
- Viegas-Aldrovandi, S.M., & Gruenwald, R.B., 1990, *ApJ*, 360, 474

3-D Spectroscopy of High-z Radio Galaxies with the IFUs TIGER and OASIS

Brigitte Rocca-Volmerange (1), Gilles Adam(2), Pierre Ferruit(3),
Emmanuel Moy (1),

(1) *Institut d'Astrophysique de Paris*, (2) *CRAL-Observatoire de Lyon*
(3) *Physics Dept, Univ. of Maryland*

Abstract. We report the study of the morphology and kinematics of emission nebula of high- z (from 0.5 to 3.8) radio galaxies with the IFUs TIGER and OASIS at CFHT. These 3D observations allow us to identify the nature of the various observed components (e.g. to distinguish between continuum and line emission structures) and to map them individually. For all the objects, even the most distant, we find roughly symmetrical radial velocity distributions with amplitudes ranging from 200 to 600 km s⁻¹ (rest frame). A scenario in which the ionized gas takes part to a large scale rotation is consistent with observations but would require confirmation. Evidence for obscuration by dust (which is believed to play a role in the general history of these galaxies) is found in some regions. In the near future, additional constraints on the geometry and the kinematics of these nebula, and on the various sources of ionization (in particular on the related star formation rate) will come from the improved spectral and spatial resolution of OASIS/CFHT.

1. Introduction

Various mechanisms, such as shocks and photoionization by the active nucleus or by massive stars, can account for the ionisation of the gas of distant radio sources. Identification of the main processes of nebular emission would strongly help to understand the formation and evolution of radio sources, as well as their connection with active galaxy nuclei (AGN) and normal galaxies. However, at high redshift, the evolution and cosmology effects combined with lowered spatial resolution, make their identification difficult. In this context, 3D spectroscopy of radiogalaxies with integral field units (IFUs) mounted at large telescopes and associated with adaptative optics is a unique tool to identify emission lines and continua, line widths and line ratios, velocity profiles. We have therefore undertaken a program of 3D spectroscopy of radio galaxies with the IFUs TIGER and OASIS aiming at understanding the physics associated with the nebular and the stellar components.

2. Kinematics, ionisation levels and physics from emission lines

The measurements of the intensity and width of strong emission-lines (Hydrogen, Oxygen, Carbon, Nitrogen, Neon) allow to build the velocity field and the corresponding map of velocity dispersion of the gas content. The observed gaseous kinematics can then be interpreted, e.g. in terms of rotation or dynamical interactions. Likely associated to the feeding process of the AGN, rotation brings constraints on the angular momentum, accretion mass and environmental properties. The spatial distribution of the line emission with respect to the radio emission and the galaxy itself, gives critical information on its triggering mechanism. Moreover, the measurement of line widths and ratios or the detection of specific lines allow the identification of zones of low and high ionisation (Villar-Martín et al, 1999).

3. Evolved and young stellar components from continua, dust and absorption lines

Another goal is to search traces of star formation in radiogalaxies. Some arguments suggest radiogalaxies could be the cocoon of elliptical galaxies. Multiple galaxy companions of the central radio source could finally spiral down into the potential well to form a giant elliptical galaxy. They can be dated from their continua, dust amount, absorption lines, with the help of the recent model PEGASE (Fioc and Rocca-Volmerange, 1997) in its new version with metal and dust effects. Another recent improvement of PEGASE is its connexion with the photoionization model CLOUDY (Ferland, 1996). Metallicity, dust and Lyman continuum photons are the common parameters of the two codes (Moy and Rocca-Volmerange, in preparation). The information on the sites of stellar formation inferred from our observations and subsequent modeling with PEGASE will be used to investigate the various mechanisms proposed to trigger stellar formation in these environments (jet-cloud interaction, turbulence, ...). PEGASE will also bring constraints on datation, needed to establish some hierarchical effect in the radio galaxy evolution.

4. The integral field spectrographs

TIGER and OASIS are integral field units (IFU) successively mounted at CFHT and based on the dispersion of micropupils focused from a grid of microlenses (Bacon et al, 1995).

The IFU OASIS at the 3.60m CFHT telescope is an improved version of TIGER. Equipped with a grid of 1200 micro-lenses, OASIS has a spectral resolution ranging from 600 to 3300, a field of view from 1.6"x1.2" to 15"x12" and the possibility to use adaptative optics. The spatial sampling varies from 0.04" to 0.41" with a 0.087 to 0.47 nm/pxl spectral sampling. Its overall spectral coverage is $\Delta\lambda = 415\text{-}1037$ nm. The coupling with adaptative optics makes this instrument unique. More information may be get from the WEB site of OASIS : <http://www.cfht.hawaii.edu/Instruments/Spectroscopy/OASIS/>.

The data reduction has been performed using a dedicated software, XOasis, developed at the CRAL-Observatoire de Lyon and adapted from the TIGER

software (Rousset, 1992). The wavelength calibration accuracy is about $1 \text{ \AA}$ in the low resolution configuration used. Astrometric errors on the following objects are limited to $0.06''$ on (α, δ) and to 5° on P.A.

5. Observations

Distant radio galaxies with $0.5 < z < 4$ are ideal targets for IFUs operating in the visible because of their size well suited to the field of view (a few tens of arcsecs), of the strength of emission lines (more than $10^{44} \text{ erg s}^{-1}$ for the $\text{Ly}\alpha$ line), of the intense far-UV emission redshifted to the optical or near-infrared.

- The radiogalaxy **3C435A** ($z = 0.471$), (Fig. 1) was observed with TIGER. It shows completely different morphologies in the [OII] and [OIII] emission lines, as well as in the stellar continuum. Its complex velocity field shows velocity changes of up to 400 km s^{-1} over 100 kpc (for a flat universe). The emission zones of ionized gas and a high number density of old evolved galaxy companions are well identified along the radio axis. A young, extremely blue component is observed in a perpendicular direction. The most striking property comes from the large ages (respectively 8 and 10 Gyrs) attributed to the continua of the North and central components, both corresponding to old ellipticals at $z = 0$ (Rocca-Volmerange et al, 1994).

Figure 1. Grey-scale [OII], [OIII] and continuum reconstructed images with the isophotes of the 1.4 GHz VLA map of McCarthy et al, 1989 overimposed. Spectrum of the central component is also shown in miniature. Flux is in arbitrary units

- The radiogalaxy 4C41.17 ($z=3.8$), (Fig. 2) has also been observed with TIGER (Adam et al, 1998). A striking feature is the curvature of the Ly α isophotes with a deflection angle of about 20° . The strong Ly α peak observed along the radio jet is not located at the core (Carilli et al, 1995). A sub-mm emission (Dunlop et al, 1994) is associated to a dust cloud on the core line of sight. These two features were confirmed on the high resolution HST image (van Breugel et al, 1998). The velocity field shows symmetrical values of 200 to 300 km s $^{-1}$ (redshift corrected), while the velocity dispersion may reach 300 to 500 km s $^{-1}$ (same conditions). The surprising regularity and the classical values of kinematics do not show a catastrophic scenario as an intense merging or coalescence process would do.

Figure 2. The Ly α image of the radiogalaxy 4C41.17 and its environment, reconstructed from individual spectra. North is up and East left. Scale units are in arcsecs and the grey scale is logarithmic

- The radiogalaxy 3C247 ($z= 0.749$), (Figure 3) was the first object observed with OASIS (article in preparation). The image shows an [OII] lobe which is not aligned with the radio jet and a large velocity gradient ($\Delta V = 500$ km s $^{-1}$), which can be interpreted by a large scale gaseous rotation.

Figure 3. The [OII] reconstructed image of 3C247, observed with OASIS superimposed to its HST continuum (Best et al, 1997)

6. Conclusion and future

3D spectroscopy confirms its ability to study ionisation levels, kinematics and shocks over the whole radio galaxy. In relation with dynamics, strong constraints on the formation and evolution processes can be derived. The improved spectral resolution of OASIS allows the study of line profiles with consequences on the ionisation levels. The sensitivity of the new generation of telescopes may give hopes to detect and to date continua of galaxy companions, possibly confirming a hierarchical merging process proposed by semi-empirical models of galaxy formation. The future will be related to far-IR emissions (ISO, SIRTf), to H₂/CO submillimeter observations and metal content in relation with dust amounts.

References

- Adam, G., Rocca-Volmerange, B., Gérard S., Ferruit, P., Bacon, R., 1997, *Astron. Astrophys.*, 326, 501
- Bacon et al, 1995, *Astron. Astrophys.*, 113, 347
- Best, P.N., Longair, M. S., Röttgering, H.J.A., *Mon. Not.R.Astron. Soc.*, 292, 758
- Carilli, C. L., Röttgering, H., van Ojik, R., Miley, G., van Breugel, W.J., 1997, *Astrophys. J. Sup. Ser.*, 109, 1
- Dey, A., van Breugel W., Vacca, W., Antonucci, R., 1997, *Astrophys. J.*, 400, 698
- Dunlop J., Hughes D., Rawlings, S., Eales S., Ward M., 1994, *Nature*, 370, 347
- Fioc M., Rocca-Volmerange B., 1997, *Astron. Astrophys.*, 326, 950
- Rocca-Volmerange B., Adam G., Ferruit P., Bacon R., 1994, *Astron. Astrophys.*, 292, 20
- Rousset, A., 1992, Thèse Université J. Monnet, St Etienne, France
- van Breugel et al, , 1997, in "the most distant galaxies", Röttgering et al, eds, Amsterdam , p. 49
- Villar-Martín M., Tadhunter C., Morganti R., Clarck N., Killeen N., Axon D., 1998, in press

Kinematics of the Ionized Gas in Irregular Galaxies: IC 1613 and NGC 4449

¹M. Valdez-Gutiérrez, ²M. Rosado, ²L. Georgiev, ²C. Lacey, ³J.
Borissova, ⁴R. Kurtev

¹*Instituto Nacional de Astrofísica, Óptica y Electrónica, Tonantzintla,
Puebla, México*

²*Instituto de Astronomía, UNAM, México*

³*Institute of Astronomy, Bulgarian Academy of Sciences, Bulgaria*

⁴*Department of Astronomy, Sofia University, Bulgaria*

Abstract. In this contribution we present preliminary results of a kinematical study based on $H\alpha$ Fabry-Perot interferometry of the irregular galaxies IC 1613 and NGC 4449.

1. Introduction

Irregular galaxies account for a substantial fraction of all galaxies and they are also dominant among the actively star-forming galaxies. Because they lack spiral density waves, they are excellent laboratories in which we can examine the star formation process in an alternative scenario. For this reason, the nearby irregular galaxies are excellent targets for the study of stellar and gas content. In this work, the goals are to obtain the general velocity field in $H\alpha$ for the nearby irregulars IC 1613 and NGC 4449, as well as, to get kinematical information around specific sources like HII regions and supernova remnants (SNRs). Our optical data of NGC 4449 are complemented with radio data at 1448.438 MHz. Very accurate positions have been determined from this radio study for 17 sources: 4 thermal, 6 non-thermal and 7 sources where the spectral index error was too large to determine the identification (Lacey 1998). We aim to analyse and to compare the optical data with the radio ones in the search of SNR kinematical signatures. For IC 1613, we are presenting $H\alpha$ Fabry-Perot interferometry of the northeast corner. Recently Georgiev et al. (1999) have determined sizes and positions for the young associations catalogued by Hodge et al. (1990, hereafter HLG) embedded at the optical center of IC 1613 and found new ones. In this contribution, we have obtained kinematical information for the global radial velocity field of IC 1613 as well as for 14 positions corresponding to OB associations from Georgiev et al. (1999) as a preamble for a more extensive study of these overlapping giant shells.

Figure 1. Different channels of the radial velocity field of the nearby irregular galaxy NGC 4449. The radial velocities in km/s are given at the bottom left corner of each panel. North is up and east to left.

2. The data

The data were obtained with the scanning Fabry-Perot interferometer (SFPI) PUMA (Rosado et al. 1995, 1999) on December 5-6 1996. PUMA is an interferometer ET-50 from Queensgate Instruments with a servostabilization system and a focal reducer coupled to the SFPI. The galaxies IC 1613 and NGC 4449 were observed with PUMA at the $f/7.9$ Ritchey-Chrétien focus of the 2.1m telescope at the Observatorio Astronómico Nacional at San Pedro Mártir, México. The setup was composed of a $f/3.95$ camera, and a 1024×1024 TEK CCD detector. The bandpass of the $H\alpha$ filter was $20 \text{ \AA}$, and the spectral range was 847 km/s ($18.6 \text{ \AA}$). The field covered $10' \times 10'$. The phase calibration was done taking data cubes of an hydrogen lamp before and after the galaxy exposure. This lamp was centered at $6562.78 \text{ \AA}$. The data cubes are composed of 48 steps, with integration time of 60 seconds for each step giving a total of about 1 hour to obtain a full data cube. Reductions are carried out with the dedicated software CIGALE (standing for CINématique des GALaxiEs - Le Coarer et al. 1993). The integration of the 48 channels gives the total monochromatic emission for each pixel (CIGALE enable us to separate monochromatic emission from continuum emission - Laval et al. 1987). To extract the kinematical information from the data cubes, $H\alpha$ line profiles for each region of interest has been constructed. Gaussian curves have been fitted to the nebular observed profiles and the sky lines and the instrumental profile are well fitted by an Airy function.

Figure 2. Different channels of the radial velocity field of the Local Group irregular galaxy IC 1613. The radial velocities in km/s are given at the bottom left corner of each panel. North is up and east to left.

3. Results

For NGC 4449, we have detected emission for heliocentric velocities between +100 and +400 km/s (Fig. 1). Our velocity map for the entire field in $H\alpha$ shows a well defined decreasing gradient in radial velocities along the optical bar from northeast to southwest reproducing very well previous kinematical results from other authors (e.g., Crillon & Monnet 1969). Some of the non-thermal sources reported by Lacey (1998) show interesting kinematical properties as asymmetric and composite profiles. All but one (source 4) of the non-thermal sources coincide with HII regions cataloged by Sabbadin & Bianchini (1979) as illustrated in Fig. 3, left panel. We have obtained radial velocity profiles of these sources by integrating in boxes of different sizes. The velocity dispersions in these regions are supersonic, having values between 13 and 38 km/s and there is no correlation between the velocity dispersion of the sources and their position inside the galaxy. Finally, the sources have a range of heliocentric velocities from +160 to +247 km/s. For IC 1613 (Fig. 2), we have detected emission for heliocentric velocities between -100 and -350 km/s, in agreement with previous kinematical results from other authors (e.g., Meaburn et al. 1988). Our radial velocity map for the entire field in $H\alpha$ shows a distribution of velocities where the kinematical signatures of the super giant shells are clearly seen. In our map, the region that exhibits the most extreme radial velocity has -340 km/s and it is identified as the region 13, field S of HLG. As a first step for a more detailed and complete kinematical study, we have determined kinematical parameters for 14 OB associations (Georgiev et al. 1999) in the northeast field of this galaxy. This field, corresponds to field L in HLG. The 14 selected regions (Fig. 3, right panel) correspond to associations 2, 7, 10, 11, 14, 15, 16, 17, 19, 21, 23, 24, 25 and 28

Figure 3. Left panel: radio sources positions for NGC 4449 (Lacey 1998). The sources 1, 2, 3, 5 and 6 correspond to the HII regions 59, 187, 36, 200 and 209 from Sabbadin & Bianchini (1979). Right panel: OB associations for IC 1613 (Georgiev et al. 1999).

from Georgiev et al. (1999). We have obtained radial velocity profiles of these sources by integrating in boxes of different sizes. Some of these regions have line profiles with a complex structure that are not fitted by a single Gaussian curve and the velocity dispersions are supersonic, varying from 30 to 50 km/s and have a range of heliocentric velocities from -155 to -246 km/s. This kind of profiles along a shell is indicative of gas movements onto or from the star forming regions. The kinematical study combined with the study of the OB associations (Georgiev et al. 1999) will give us important clues about the objects that are driving these overlapping giant shells.

A more detailed and extensive study of these irregular galaxies is in progress.

This project is supported by grants from CONACyT and DGAPA (México).

References

- Crillon, R., Monnet, G. 1969, *A&A*, 1, 449
- Georgiev, L., Borissova, J., Rosado, M., Kurtev, R., Ivanov, G., Koenigsberger, G. 1999, *A&AS*, 134, 21
- Hodge, P. A., Lee, M. G., Gurwell, M. 1990, *PASP*, 102, 1245 (HLG)
- Lacey, C. 1998, private communication
- Le Coarer, E., Rosado, M., Georgelin, Y., Viale, A., Goldes, G. 1993, *A&A*, 280, 365
- Laval, A., Boulesteix, J., Georgelin, Y. P., Georgelin, Y.M., Marcelin, M. 1987, *A&A*, 175, 199
- Meaburn, J., Clayton, C. A., Whitehead, M. J. 1988, *MNRAS*, 235, 479
- Rosado, M. et al. 1995, *Rev. Mexicana Astron. Astrofis. Conf. Ser.*, 3, 268
- . 1999, in preparation
- Sabbadin, F., Bianchini, A. 1979, *PASP*, 91, 280

Molecular Hydrogen Imaging of Star-forming Regions

S.K. Ramsay Howat¹, A. Chrysostomou², M.G. Burton³ and P.W.J.L. Brand⁴

¹ *UK ATC, Royal Observatory, Blackford Hill, Edinburgh EH9 3HJ.*

² *Joint Astronomy Centre, 660, N. A'ohoku Place, Hilo, HI 96720.*

³ *School of Physics, UNSW, Sydney, New South Wales 2052, Australia.*

⁴ *Institute for Astronomy, University of Edinburgh, Royal Observatory, Blackford Hill, Edinburgh EH9 3HJ, U.K.*

Abstract.

Observations of excited emission from molecular hydrogen have been used extensively for studies of star-forming regions for more than a decade. The characteristic response of the molecular gas to different excitation mechanisms, coupled with the in-depth theoretical modelling that has been carried out, make the near-infrared spectrum of H₂ a powerful diagnostic of the physical conditions in these regions. Early observations were able to establish whether molecular gas was excited by shocks or by radiation from young stars, using spectrometers with large beams. With improvements in theoretical understanding of the H₂ spectrum, observers were able to infer different density components within a molecular cloud from the NIR spectrum. Recent observations of the Orion Bar region made at the UK Infrared Telescope have combined Fabry-Perot imaging spectroscopy with observations at high angular resolution to examine the structure of the clouds in more detail. These observations reveal a complex filamentary structure on sub-arcsecond scales and provide direct evidence for high density clumps in the region.

1. Introduction

The observations presented in this paper are the result of an investigation of the physical structure of the Orion Bar. The aims of this project were to map the density structure in the Bar, searching for direct evidence of high density "clumps", and to investigate whether these clumps control the structure of the Bar. Our method of addressing these questions was to ratio two images of the Bar in the $v=1-0$ S(1) and $v=2-1$ S(1) emission lines of molecular hydrogen. This ratio has been shown to be sensitive to density (Burton, Hollenbach and Tielens 1990, Sternberg and Dalgarno 1989, Draine and Bertoldi 1996). In a radiatively excited cloud of density greater than a critical density of $n \sim 5 \times 10^4 \text{ cm}^{-3}$, collisions between molecular and atomic hydrogen preferentially de-excite the 2-1 S(1) line. The 1-0 S(1)/2-1 S(1) line ratio increases above the value predicted

for pure fluorescence ($=1.9$). In this work, this sensitivity of the line ratio to density is used as a tool to uncover the physical structure in the region. This approach has also been used by Ryder et al. (1998) in a study of Parsamyan 18.

The Orion Bar is the classic example of a photo-dissociation region (PDR), where the chemistry and physical structure are controlled by the UV radiation from young stars (see e.g. Tielens et al. 1993). Early PDR models treated such region as a slab of material of single density (Tielens and Hollenbach 1985a). Among the first observations to challenge this simple model were those of M17 SW by Stutzki & Guesten (1990) and of the Orion Bar by Parmar et al. (1991), both of which showed emission from molecular gas occurring at greater depths into the cloud than predicted. A two density component model was proposed; a low density "interclump medium" (ICM) and a number of unresolved "clumps". Emission from the dense clumps can explain the intensity of emission from the regions, whilst allowing photons to penetrate to greater depth. Further evidence for these clumps is found in observations by Tauber et al. (1994), Hogerheijde et al. (1995), van der Werf et al. (1996) and Usuda et al. (1996), amongst others. A broad agreement on the density structure emerges from the imaging observations with 10% of the gas thought to exist in clumps with density $\sim 10^6 \text{ cm}^{-3}$ embedded in a more rarified ($\sim 10^4 \text{ cm}^{-3}$) ICM.

A source of continued debate is whether the clumps control structure of the Bar, or are of secondary importance compared with the interclump medium. Tauber et al. found that the barely resolved clumps in their CO data played a negligible role in determining the structure of the cloud. In contrast, van der Werf et al. (1996) concluded that the high density clumps do control the cloud structure.

Conclusions drawn from recent spectroscopic observations of H_2 emission in the NIR by Luhmann et al (1998) and Marconi et al (1998) are at odds with each other and with the conclusions of the imaging studies. Luhman et al find that the spectrum can be modelled by a single high density component ($\sim 10^6 \text{ cm}^{-3}$). Marconi et al. find no evidence for densities higher than $\sim 10^4 \text{ cm}^{-3}$. Both groups confirm the conclusion of Burton et al. (1991) that there is no evidence for shock excitation in the NIR emission spectrum of H_2 .

To a greater or lesser extent, each of these studies suffers from beam dilution effects. By observing the Bar at high angular resolution, we aimed to resolve the clumps on sub-arcsecond scales and to clarify the importance of a two clump model in characterising the Orion Bar.

2. Observations

The observations were made with the UKIRT telescope on 1998 January 16-19. The infrared camera, IRCAM was used with a pixel scale of 0.286 arcsec and a 72 arcsec field of view. The field was centred on the position selected by Hayashi et al (1985). A 350 km s^{-1} Fabry-Perot etalon from Queensgate instruments was used to isolate the lines. In a region with such a high line-to-continuum ratio, the rejection of continuum from the line pass-band is important to maximise signal/noise on the line flux. The method of observing was to observe the line, off-set to sky, re-tune the FP to nearby continuum, observe the continuum then offset to sky and make a measurement of the sky at the continuum wavelength.

Figure 1. The 1-0 S(1) line of H₂ imaged using a Fabry-Perot.

A quarter of the total observing time is spent observing the source. Over the four nights, the 1-0 S(1) was observed for 45mins and the 2-1 S(1) line for 2hours and 30mins. Frames were “jittered” by 5arcsecs as well as offset to the sky. In the reduction, each frame was divided by a flat-field frame derived from the median of the sky observations. The flat-field frame used was specific to the FP setting. Individual frames were registered using the many field stars present before constructing the final mosaic. The 1-0 S(1) and 2-1 S(1) images are shown in Figures 1 & 2.

3. Discussion and analysis

The images reveal a complex filamentary structure on scales as small as 0.5arc-sec (equivalent to 10^{-3} pc). Comparing the $v=1-0$ S(1) and $v=2-1$ S(1) frames, changes in the relative intensity are seen, providing immediate evidence for changing densities across the region. For example the northernmost filament running parallel to the Bar ((40,115) - (50,110) using the coordinates in Figure 1) is the brightest 1-0 S(1) feature. The brightest 2-1 S(1) feature is centred on (40,90).

A contour plot of the ratio of the two line frames is shown in Figure 3. The line images were placed on the same reference frame and were smoothed to increase the signal/noise in the 2-1 S(1) image before ratioing. Typical values of the ratio, line intensity and implied density for the brightest structures in the region are given in Table 1. The density has been derived from the models of Draine and Bertoldi (1996), assuming an intensity of the UV field for the region of 5×10^4 larger than that of the ISM (Tielens and Hollenbach 1985b).

Figure 2. The 2-1 S(1) line of H_2 imaged using a Fabry-Perot.

The areas where the density is highest are immediately apparent as two bright ridges running from (40,120)-(50,110) and (30,90)-(40,60), using coordinates from Figure 3. The filament bridging these features, (60,110)-(30,90), is fainter in the ratio map, despite being prominent in the 1-0 S(1) intensity map. The filaments are surrounded by gas with a line ratio of ~ 3 , a value typical of low density fluorescence, though this must be interpreted cautiously as the signal/noise on the 2-1 S(1) line is low in many of these regions.

The anti-correlation between changes in intensity and density can be understood as follows. As the density increases, collisional de-excitation occurs for both the $v=1$ and $v=2$ levels, but the effect is stronger for the $v=2$ levels. In addition, the $v=1$ -0 S(1) line may also be collisionally excited. Hence the line ratio tends to increase with increasing density, and may be expected to be correlated with increasing 1-0 S(1) line strength. However, the flux per pixel depends on the volume of gas present in the pixel, as well as on the density of the gas. If the clumps in the region were fully resolved, the changes in intensity and line ratio would be correlated. Where a high density region is of low intensity, the pixels contain unresolved clumps. This highlights the power of the approach taken here, where the density is measured directly rather than inferred from changing line intensity. Regions of high and low density are indistinguishable in a single line map.

For the whole area observed, the line ratio has been calculated and compared with the 1-0 S(1) line strength. Only those pixels with signal-to-noise on the 2-1 S(1) line of greater than 5 were selected. Ratios in the range $1 \leq 1\text{-}0\text{ S}(1)/2\text{-}1\text{ S}(1) \leq 8$ are measured. The low ratios are consistent with pure fluorescent emission, with no thermal contribution to the spectrum. This is emission from

Figure 3. A contour plot of the line ratios in the Orion Bar.

the ICM with density $\leq 10^4 \text{ cm}^{-3}$. Measured ratios of ~ 8 imply the presence of densities greater than 10^6 cm^{-3} , as required to explain the observed CO 1-0/7-6 and CO 7-6/14-13 ratios from Tauber et al. (1994) and Stacey et al. (1993).

(x,y) from Figure 3	$I_{\text{H}_2} \text{ v=1-0 S(1)}$ $\text{Wm}^{-2} \text{ arcsec}^{-2}$	$I_{\text{H}_2} \text{ v=2-1 S(1)}$ $\text{Wm}^{-2} \text{ arcsec}^{-2}$	ratio	density
(40,120)-(50,110)	$\geq 6 \times 10^{-18}$	1×10^{-18}	6 ± 0.6	10^6 cm^{-3}
(60,110)-(30,90)	4.5×10^{-18}	1.25×10^{-18}	3.5 ± 0.4	10^5 cm^{-3}
(30,90)-(40,6)	3×10^{-18}	0.5×10^{-18}	6 ± 1.2	10^6 cm^{-3}

Table 1. Example line strengths and line ratios from prominent filaments in the Bar maps.

Since these data reveal the high density clumps, they can be used to estimate whether the clumps we observe are responsible for the structure of the Bar more effectively than previous data sets. The ICM density is taken as 10^4 cm^{-3} . Using the column density for H₂ of $5 \times 10^{21} \text{ cm}^{-2}$ as derived from the CO observations of Tauber et al. and a distance to the Bar of 460pc, an optical depth of $1A_V$ should correspond to 70arcsecs. Tielens & Hollenbach (1985a) show that the H₂ peak occurs at an optical depth of ~ 2 , or 140arcsecs into the cloud in the case of the Orion Bar. We see the H₂ emission peak at a distance of 15arcsecs from the ionisation front. This implies that a higher density gas component is controlling the structure of the cloud. The dense filaments of H₂ emission cover an area of approximately 125arcsecs squared or a projected area of $\sim 45\%$ of the Bar.

Assuming that the filling factor is the same along the line of sight, the volume filling factor is 20%, and the clumps are a significant fraction of the cloud. Our data therefore support the conclusion of van der Werf et al. that the clumps control the cloud structure.

4. Conclusions and Prospects

Pseudo-3D spectroscopy of the Orion Bar using Fabry-Perot imaging in key molecular transitions has provided a picture of the density structure of the Orion Bar. Filaments with densities in the range 10^5cm^{-3} to 10^6cm^{-3} are observed and the interclump medium ($n \sim 10^4\text{cm}^{-3}$) detected. The structure of the Bar is found to be controlled by these high density clumps.

FP imaging provided the excellent spatial coverage required to survey this nearby region. Follow-up studies in this and other PDRs will benefit from the development of true 3D spectroscopy via a NIR integral field unit such as that being developed for the UKIRT Imaging Spectrometer, UIST (Wells, Hastings and Ramsay Howat 2000). FP imaging requires 75% of the observing time to be spent on calibrations. This experiment, requiring a total integration time of 13 hours, could be completed in approximately one hour.

UKIRT is operated by the Joint Astronomy Centre on behalf of the UK Particle Physics and Astronomy Research Council. The authors would like to thank Thor Wold for his assistance at the telescope.

References

- Burton, M.G., Hollenbach, D.J. & Tielens, A.G.G.M. 1990, *ApJ*, 365, 620.
 Draine, B.T. & Bertoldi, F. 1996, *ApJ*, 468, 269.
 Hayashi, M., Hasegawa, T., Gatley, I., Garden, R., Kaifu, N., 1985, *MNRAS*, 215, 31.
 Luhmann, K.L., Engelbracht, C.W. & Luhman, M.L. 1998, *ApJ*, 499, 799.
 Marconi, A., Testi, L., Natta, A. & Walmsley, C.M. 1998, *A&A*, 330, 696.
 Parmar, P.S., Lacy, J.H. & Achtermann, J.M., 1991, *ApJ*, 372, L25.
 Hogerheijde, M.R., Jansen, D.J. & van Dishoeck, E.F. 1995, *Astron. Astrophys.*, 294, 792.
 Ryder, S.D., Allen, L.E., Burton, M.G., Ashley, M.C.B. & Storey, J.W.V. 1998, *MNRAS*, 294, 338.
 Stacey, G.F., Jaffe, D.T., Geis, N., Genzel, R. Haris, A.I., Poglitsch, A., Stutzki, J. & Townes, C.H., 1993, *Ap.J.*, 404, 219
 Sternberg, A. & Dalgarno, A., 1989, *ApJ*, 338, 197.
 Stutzki, J. & Guesten, R. 1990, *ApJ*, 20, 513.
 Tauber, J.A., Tielens, A.G.G.M., Meixner, M., and Goldsmith, Paul F., 1994, *Ap.J.*, 422, 136.
 Tielens, A.G.G.M. & Hollenbach, D. 1985a, *Ap.J.*, 291, 722.
 Tielens, A.G.G.M. & Hollenbach, D. 1985b, *Ap.J.*, 291, 747.
 Tielens, A.G.G.M., Meixner, M.M., van der Werf, P.P., Bregman, J., Tauber, J.A., Stutzki, J., & Rank, D., 1993, *Science*, 262, 86.
 Usuda, T., Sugai, H., Kawabata, H., Inoue, M., Kataza, H. & Tanaka, M. 1996, *ApJ*, 464, 818.
 Van der Werf, P.P., Stutzki, J., Sternberg, A., & Krabbe, A. 1996, *Astron. Astrophys.*, 313, 633.
 Wells, M., Hastings, P.R. & Ramsay Howat, S.K. 2000, this volume.

Probing the Origin of Mass-loss in Young T Tauri Stars

C. Dougados, C. Lavalley

*Laboratoire d'Astrophysique, Observatoire de Grenoble, B.P.53X,
F-38041, Grenoble Cedex 9*

S. Cabrit

*Observatoire de Paris, DEMIRM, 61 Avenue de l'Observatoire, F-75014
Paris*

Abstract. We present spectro-imaging observations recently obtained with the TIGER and OASIS instruments at the CFHT, aimed at constraining the ejection and collimation mechanisms in active young pre-main sequence stars (age $\simeq 1\text{--}3$ Myrs). Three stars have been observed so far in the lines of [O I] 6300Å, H α , [N II] 6584Å and [S II] 6717,6731Å with spatial resolutions ranging from 0.5'' to 0.25'' and velocity resolutions of 170 and 100 km s $^{-1}$. The morphology of the inner regions of the wind is revealed with unprecedented details. Constraints on excitation conditions and collimation and acceleration processes are derived.

1. Introduction

Outflow phenomena appear to be intimately connected with the star formation process. While the existence of a link between accretion and ejection is now well established (Cabrit et al. 1990), the detailed physical process involved is still poorly constrained. Magneto-hydrodynamical models, where the magnetic field extracts angular momentum from the accretion process itself, provide to date the most consistent physical description (e.g. Ferreira & Pelletier 1995; Shu et al. 1995). This theoretical framework is now widely used in the context of low mass star formation. There is however no direct observational confirmation that this mechanism is indeed at play. MHD models predict that most of the collimation and acceleration occur over distances less than 100 AU. Constraining the physical properties of the wind at such spatial scales is thus essential to our understanding of the mass-loss process in young stars.

The optically revealed Classical T Tauri stars (cTTs), located in the nearby Taurus cloud ($d=140$ pc), offer an unique opportunity to study jet formation on 100 AU scale. Indeed, although outflow manifestation is most spectacular around the youngest embedded sources (age $< a\text{ few }10^5$ yr), where collimated Herbig-Haro (HH) jets extending over a few parsecs are observed, the presence of substantial circumstellar environment at these early stages prevents the study of the innermost regions of the wind. cTTs are typically an order of magnitude older (age $\simeq a\text{ few }10^6$ yr) and have dissipated most of their parent gas envelope. Strong infrared excesses and blueshifted optical forbidden emission lines

demonstrate that these objects are still actively accreting and ejecting matter, albeit at rates 10-100 times smaller than embedded sources driving jets (Hartigan et al. 1995). The ejection/accretion rate ratio $\simeq 0.1$ is similar in both cases, suggesting that the same driving mechanism is at work. It has been early recognized that the forbidden emission lines observed in cTTs trace the dense ($n_e \simeq 10^4\text{-}10^6 \text{ cm}^{-3}$) and hot ($T > 5000 \text{ K}$) central regions of the wind. Their study has proven however challenging because of both their small spatial extent (a few '' at the distance of Taurus) and their very low contrast with respect to the strong nearby continuum emission (typical equivalent width of $1 \text{ \AA}$). Long-slit spectroscopic technics were first instrumental to reveal jet-like forbidden line emission extended over scales of 3-4'' around a dozen cTTs (Solf & Böhm 1993; Hirth et al. 1997). Strong velocity and excitation gradients were observed. The lack of true 2D information prevented however a firm conclusion on collimation and excitation processes.

We report here on the first 2D spectro-imaging observations of young T Tauri star winds, obtained with the TIGER and OASIS instruments at the Canada France Hawaii Telescope. The velocity resolution of our observations range from 170 (TIGER) to 100 (OASIS) km s^{-1} and the spatial sampling is $0.3''$ for the TIGER observations and $0.16, 0.11''$ for the OASIS ones. Three stars, chosen among the candidates with highest accretion rates, have been observed so far in the lines of [O I] 6300Å, [N II] 6584Å, [S II] 6717, 6731Å and $\text{H}\alpha$. The OASIS instrument is coupled to an adaptive optics bonnette, allowing to reach spatial resolutions in the range $0.1\text{-}0.2''$ (15-30 AU at the distance of Taurus). This unique combination of high spatial resolution and spectral resolution is perfectly adapted to the aims of our study. We illustrate below, through specific examples on two of our targets (DG Tau and RW Aur), the new information derived on the morphology, excitation conditions and collimation mechanisms of the wind.

2. Morphology and Kinematics of the Wind

The first information derived from these data is the sub-arcsecond morphology of the wind. We show in Fig. 1 the reconstructed [O I] line emission map obtained for the star DG Tau. HST images of the DG Tau jet were obtained by Kepner et al. (1993) in [O I] and Stapelfeldt et al. (1997) in broad band R. These images suffer respectively from a degraded psf and a low contrast between continuum and line emission, thus preventing a detailed study of the inner jet regions. Our spectro-imaging observations not only allow an accurate subtraction of the continuum, but also provide a simultaneous estimate of the point spread function. We thus can significantly increase the spatial resolution through deconvolution techniques.

Our DG Tau [O I] image reveals for the first time the detailed structure of the wind on subarcsecond scales. The jet-like morphology of the emission is clearly confirmed. The emission is dominated by two knots on the blueshifted side, the outer one having a clear bow-shaped morphology reminiscent of the bow-shock structures observed in younger Herbig-Haro jets. The jet is resolved transversally beyond 70 AU and seems to be precessing: it shows a sinuous beam and its FWHM increases steadily up to 90 AU at $1''$. Strong kinematical

Figure 1. DG Tau reconstructed $[O\ I]\ 6300\text{\AA}$ line emission map obtained with the TIGER instrument in November 1994 at the CFHT. The line map has been deconvolved by the continuum map with the Lucy algorithm. Resulting spatial resolution is $0.35''$. The cross locates the centroid of the continuum emission. Average spectra of selected regions in the jet are also shown on the right. A two gaussian decomposition of the profiles is indicated.

gradients are observed over the central arcsecond (150 AU) at which point the line profiles become narrower and symmetrical. Kinematical and morphological arguments suggest that the outer knot is an internal working surface, analogous to those observed in younger HH flows. Additional details about the analysis of these data can be found in Lavalley et al. (1997). Efficiently constraining the excitation conditions requires however studying more than one line.

3. Excitation conditions

We reobserved the DG Tau jet in January 1998 with the OASIS instrument in the $H\alpha$, $[N\ II]$ and $[S\ II]$ lines. Poor seeing conditions resulted in achieved spatial resolutions of $0.45''$. OASIS allows to observe these three lines in one single spectral configuration. Relative astrometry as well as line ratios can thus be reliably obtained. We show in Fig. 2 the derived emission maps in the low ($< 250\text{ km s}^{-1}$) and high ($> 250\text{ km s}^{-1}$) velocity channels for the blueshifted flow. A new peak, prominent in the $H\alpha$, $[N\ II]$ and $[O\ I]$ high-velocity channels, is revealed at $1''$. One dominant process invoked for the formation of knots in HH jets is internal working surfaces produced by variability in the ejection velocity and/or direction (Raga et al 1990). Indeed, the relative line fluxes observed in the two bright high velocity knots at $1''$ and $3''$ indicate that they are likely excited by shocks of speed $\simeq 100\text{ km s}^{-1}$. An additional argument in favor of an internal working surface nature is the detection of a proper motion of $0.2''$,

Figure 2. DG Tau reconstructed OASIS line intensity maps in two different velocity channels: $> 250 \text{ km s}^{-1}$ (upper panels), $< 250 \text{ km s}^{-1}$ (lower panels). The cross locates the centroid of the continuum emission. Contours decrease by factors of 2 from 83 % of peak intensity. Spatial resolution is $0.45''$.

corresponding to a tangential velocity of 160 km s^{-1} , for the outer bow-shaped knot. Thus time variability seems to play a dominant role in the excitation process beyond 140 AU in this jet. Timescales in the range 5-10 yrs are indicated by the typical spacings of $1\text{-}2''$. The presence of such strong shocks would provide a natural explanation for the enigmatic excitation gradients observed along cTTs jets by Hirth et al. (1997). They also cast doubts on current mass-loss rate estimates, which assume uniform excitation (HEG95).

Different properties are observed in the inner regions ($d < 140 \text{ AU}$) of the jet. A very strong velocity gradient occurs in this central region, in accordance with line profiles predicted by current magneto-hydrodynamics models (Cabrit et al. 1999, Shang et al. 1999). Large velocities, reaching up to 500 km s^{-1} , as well as strong excitation gradients are also observed within that spatial scale. Line emission in [O I] and [S II] is dominated there by a strong low-velocity component, in contrast with the outer regions. The [O I] emission peak appears stationary. The [S II] 6716/6731 line ratio is saturated indicating electron densities in excess of 10^4 cm^{-3} . The low excitation line ratios as well as the stationary [O I] emission peak indicate in these central regions excitation mechanisms markedly different from the ones in the outer knots. The detailed study of this central region is however hampered by the limited spatial resolution of this data set.

4. Probing the innermost regions of the jet

Under good seeing conditions, we obtained in December 1998 with the OASIS instrument spectro-imaging data on the inner regions of the jet associated with the multiple system RW Aur. These observations achieve spatial resolu-

Figure 3. RW Aur reconstructed [S II] line emission map. White contours show the continuum emission from the binary RW Aur-A and B. Spatial resolution is $0.25''$. The lack of detected [S II] emission towards RW Aur-A may be an artefact due to the large photon noise produced by the continuum there. Three spectra taken at the position of RW Aur-A and at two positions along the blueshifted jet are shown on the right.

tions of $0.25''$ and will allow to study with unprecedented detail the innermost regions of the jet (30-150 AU), where we have seen that most of the excitation and kinematical gradients occur. RW Aur is a hierarchical triple system with the primary A separated by 200 AU from the close binary B & C. Only A and B are detected in the optical. They are both actively accreting T Tauri stars. A large fraction of young stars are found in multiple systems with separations comparable to the size of individual accretion disks. The extent to which multiplicity affects the physics of the accretion process is at present not well established. The study of the mass-loss process associated with a young multiple system additionally bears on this issue.

We recently obtained high resolution narrow band images of this jet in the lines of [S II] and [O I] with the adaptive optics system PUE'O at the CFHT (Dougados et al. 1999). These images reveal that the jet originates from the primary component, RW Aur-A, which is also the most active. The jet is detected over $10''$ and appears remarkably straight with a constant FWHM of 50 AU. As in the DG Tau case, the emission is dominated by knots separated typically by a few arcseconds. We show in Fig. 3 the [S II] line emission map as reconstructed from the OASIS data. The jet appears already collimated within 50 AU of the central source. Emission in the jet is continuous up until $1''$ then breaks up into knots. Strong excitation gradients are also observed in this central region. Detailed comparison of the morphology, kinematics and excitation conditions in the inner jet with predictions from MHD models are under way. The striking morphological similarity of this jet with the one detected around the single star CW Tau seems to indicate that the mass-loss originates from regions in the disk

well inside the potential well of the primary and dynamically unperturbed by the companion.

5. Conclusion and perspectives

Spectro-imaging coupled with high angular resolution techniques have proven extremely powerful to study the origin of mass-loss in young stars. Using the TIGER and OASIS systems operated at the CFHT, we obtained the first detailed images of the inner jet regions associated with three active T Tauri stars (CW Tau, DG Tau, RW Aur). The main results to date are the following: all three detected jets present multiple strong emission knots separated by a few arcseconds; density collimation and acceleration processes occur on scales < 50 and 140 au respectively, compatible with predictions from current MHD models; excitation conditions in the outer regions ($d > 140$ au) indicate the presence of strong shocks produced by time variability in the ejection process. These properties are remarkably similar to the ones observed in the younger HH jets, confirming that ejection mechanisms are indeed similar.

The high angular resolution OASIS data set recently obtained will allow a study of the acceleration regions of the flow with unprecedented detail. Deriving reliable conclusions on the mass-loss process around young stars will of course require significantly extending the sample. This will only be possible if access to a 3D spectrograph operating in the optical is guaranteed on an 8m class telescope. No significant infrared lines (such as H_2 or $[Fe II]$) are indeed detected in these jets.

We would like to thank particularly the TIGER/OASIS team in Lyon for all the support they have provided throughout this program.

References

- Cabrit, S., Ferreira, J., Raga, A., 1999, A&A, 343, L61
- Cabrit, S., Edwards, S., Strom, S., Strom, K., 1990, ApJ, 354, 687
- Dougados, C., Cabrit, S., Lavalley, C., Ménard, F., 1999, A&A, submitted.
- Ferreira, J., Pelletier, G., 1995, A&A, 295, 807
- Hartigan, P., Edwards, S., Gandhour L., 1995, ApJ, 452, 736 (HEG95)
- Hirth, G., Mundt, R., Solf, J., 1997, A&AS, 126, 437
- Kepner, J., Hartigan, P., Yang, C., Strom, S., 1993, ApJ, 415, L119
- Lavalley, C., Cabrit, S., Dougados, C., Ferruit, P., Bacon, R., 1997, A&A, 327, 671
- Raga, A., Canto, J., Binette, L., Calvet, N., 1990, ApJ, 364, 601
- Shang, H., Shu, F., Glassgold, A., 1998, ApJ, 493, L91
- Shu, F., Najita, J., Ostriker, E., Shang, H., 1995, ApJ, 455, L155
- Solf, J., Böhm, K., 1993, ApJ, 410, L31
- Stapelfeldt, K., Burrows, C., Krist, J., 1997, in *Herbig-Haro Flows and the Birth of Low Mass Stars*, IAU 182, ed. B. Reipurth & C. Bertout (Kluwer: Dordrecht), p. 355.

QSO's and the 3-D Universe

B.J. Boyle

*Anglo-Australian Observatory, PO Box 296, Epping, NSW 1710,
Australia*

Abstract. Despite almost four decades of extensive study, the true nature of QSOs, in particular their relationship to the galaxy population, remains tantalisingly elusive. I review the most recent evidence which increasingly suggests that there may indeed be a straightforward relationship between radio-quiet QSOs and galaxies. Through studies of QSOs with the new generation of 3D instrumentation, we may therefore hope to cast fresh insight on formation processes occurring in the high redshift Universe.

1. Introduction

QSOs have long been seen as enigmatic objects that, while interesting in their own right, have limited usefulness as tracers of the high redshift Universe due to their uncertain relation to galaxies. Indeed, as galaxy redshifts (Hu et al. 1999) now outstrip those of QSOs (Fan et al. 1999), one might be forgiven for thinking that QSO studies had increasingly little relevance to our understanding of nature of the Universe at moderate-to-high redshifts.

However, QSOs are still the most readily detectable class of astronomical object in the key redshift range $z \sim 1 - 2$ in which the star formation rate in the Universe is inferred to peak (see e.g. Madau et al. 1996). Indeed, the cosmological evolution of QSOs appears to track the evolution of the star-formation rate (Boyle & Terlevich 1998), suggesting that the two phenomena are closely related. Recent theoretical work by Kauffman & Haehnelt (2000) also suggests that the evolution of black holes and QSOs is 'inextricably' linked to the hierarchical build-up of galaxies.

Within the next few years, two major new surveys of QSOs will be released: the 2dF QSO survey (Boyle et al. 1999) and the Sloan Digital Sky Survey (SDSS, Margon et al. 1999). As of October 1999, the 2dF QSO survey comprises over 7000 $B < 20.8$ QSOs, over one quarter toward completing its goal of 25000 QSO redshifts (see http://msowww.anu.edu.au/~rsmith/QSO_Survey/progress.html, for an up-to-the-minute progress report). Figure 1 shows the current distribution of QSOs in the 2dF survey. For comparison, the latest (Dec 1998) compilation of QSOs/AGN by Veron & Veron-Cetty (<http://vizier.u-strasbg.fr/viz-bin/VizieR?-source=VII/207>) lists just over 11500 QSOs in total, with the largest single homogeneous sample of QSOs comprising 'only' ~ 1000 QSOs (Hewett et al. 1995). By the time these conference proceedings are published, the number of QSOs identified by the 2dF survey alone should

Figure 1. The distribution of 2dF QSOs with currently measured redshifts in the 2dF QSO Redshift Survey. The survey comprises two $75^\circ \times 5^\circ$ strips; one in the North Galactic Cap (NGC) and one across the South Galactic Pole (SGP). The top wedge is the NGC area at $\delta = 0^\circ$, while the bottom wedge is the SGP region at $\delta = -30^\circ$. The rectangular strips show the current coverage projected onto the sky. From Croom et al. (1999)

have exceeded the total number in the Veron & Veron-Cetty catalogue. The SDSS will all shortly begin its survey for 10^6 QSOs with $B < 19$ (see Margon 1999), but has already proved its success in finding large numbers of high redshift ($z > 3.5$) (Fan et al. 1999).

A primary goal of these surveys is to use the observed properties of QSOs to derive information about galaxy formation or large-scale structure. The extent to which they will be successful in this aim will depend on how well we understand the connection between QSOs and galaxies. This review therefore focusses on the observational properties of QSOs¹ in relation to galaxies. Throughout this review, all cosmological quantities are calculated assuming an Einstein de-Sitter Universe with $H_0 = 50 \text{ km s}^{-1} \text{ Mpc}^{-3}$.

2. QSO host galaxies

2.1. Low redshifts

Ground-based optical imaging studies of the host galaxies of low redshift QSOs over the past 20 years (see e.g. Simkin, Su & Schwarz 1980, Smith et al. 1986, MacKenty 1990, Zitelli et al. 1993, Kontilainen & Ward 1994) have been limited by the spatial resolution attainable from the ground; $1 \text{ arcsec} \equiv 2.8 \text{ kpc}$ at $z = 0.1$, and no strong consensus has been reached over the general properties of QSOs host galaxies from such studies.

Ground-based infra-red imaging studies of QSOs have yielded a clearer picture (Dunlop et al. 1993, McLeod & Reike 1994, Taylor et al. 1996, McLeod et al. 1999). Infra-red studies not only benefit from the improved seeing in the H and K bands, but also because of the increase dominance of the red host galaxy against the blue nucleus. Such studies reveal that powerful QSOs inhabit luminous ($L > L^*$) and massive ($r_{1/2} > 10 \text{ kpc}$) galaxies. These studies also find that radio-loud QSOs are found exclusively in early-type galaxies. Taylor et al. (1996) also found early-type galaxies acting as hosts for almost half of the radio-quiet QSOs in their sample, challenging the existing orthodoxy that radio-quiet QSOs are predominantly found in spiral galaxies.

With the excellent imaging performance provided by the COSTAR-corrected optics, a number of QSO host galaxy studies have recently been carried out with the HST (Bachall et al. 1997, Boyce et al. 1997, McLure et al. 1999). These studies focus on the brightest ($M_B < -23$) QSOs at low redshift and confirm earlier results that these QSOs are indeed preferentially found in large, luminous galaxies. These studies also find that a significant fraction of QSOs are found in interacting or tidally-disrupted systems (Bahcall et al. 1997, see Figure 2), providing circumstantial evidence for enhanced merger activity fuelling QSOs. However, McLure et al. caution that the incidence of similar levels merger activity in luminous galaxies without an active nucleus is not itself well known.

McLure et al. (1999) also confirm that a significant fraction ($\sim 50\%$) of the radio-quiet QSOs imaged have elliptical hosts, claiming tentative evidence

¹In this review the term QSO is used to denote any broad emission-line object (or sample that includes such objects) brighter than $M_B = -23$. For all other cases the term active galactic nucleus (AGN) is used.

Figure 2. A selection of QSOs (including PSF star) from the HST imaging survey of Bachall et al. (1997). Many are found in interacting or disrupted systems.

Figure 3. A selection of QSOs from the HST imaging survey of Schade et al. (2000). These QSOs are, on average, ten times less luminous than those imaged by Bahcall et al. (1997) and do not appear in interacting or disrupted systems.

Figure 4. Inferred black-hole mass (M) as a function of bulge mass (M_{bulge}) for a variety of nearby galaxies. The best-fit line corresponds to the relation $M = 0.006 M_{\text{bulge}}$. Reproduced from Magorrian et al. (1998).

that the fraction of bulge-dominant systems in radio-quiet QSOs increases with luminosity. In addition, McLure et al. demonstrate that the observed correlation between host galaxy and nuclear luminosity is consistent with the relation derived by Magorrian et al. (1998) in which the nuclear black hole mass comprises 0.6% of the spheroid mass (see Figure 4). In agreement with results from ground-based infra-red studies (see e.g. McLeod et al. 1999), McLure et al. infer that the brightest QSOs can radiate at up to 20% of their Eddington luminosity.

In a joint ground-based/HST imaging survey of 76 X-ray selected QSOs, Schade et al. (2000) find that 55% of radio-quiet QSOs at lower luminosities ($-21 > M_B > -24$) are also found in spheroid-dominant systems. However, Schade et al. (2000) find no evidence for any enhanced merger activity in these QSOs; none of the 76 QSOs imaged were in interacting systems (see Figure 3). Indeed, after accounting for the bias toward bulge-dominant systems, the host galaxies in this survey exhibited identical luminosities and sizes to those of field galaxies. Schade et al. (2000) found a weak correlation between nuclear and host galaxy luminosity, with inferred nuclear luminosities between 1–5% of the Eddington luminosity. In contrast to the McLure et al. (1999) results, Schade et al. (2000) found no correlation was between the nuclear luminosity and the bulge-dominance of the host galaxy.

However, the bulge-dominance amongst host galaxies for radio-quiet QSOs does not extend to the lowest luminosity AGN. An HST study of over 300 $M_B > -21$ active galaxies carried out by Malkan et al. (1998) found that the vast majority of Seyfert 1 nuclei ($> 80\%$) inhabit disk galaxies. Malkan et al. (1998) also found little evidence for any enhanced merger activity in these AGN;

Figure 5. The starburst/QSO UN1025-0040 discovered by Brotherton et al. (1999). Left panel: Keck *K*-band image of UN1025-0400. Right panel: Keck spectrum of the object, including decomposition into instantaneous starburst (ISB) and QSO component.

in agreement with the growing body of observational evidence (Dultzin-Hacyan et al. 1999, de Robertis et al. 1998) that Seyfert 1s do not exhibit any evidence for enhanced interactions with companion galaxies within 100kpc.

Malkan et al. (1999) found that 30% of their Seyfert 1 galaxies contained no point-source nuclear component, ascribing this to (patchy?) obscuration of the central source. Schade et al. (2000) also find that 10% of their brighter QSO sample also no nuclear component at the $M_B > -17$ level.

2.2. High redshift

At high redshifts ($z \sim 2$), imaging studies have been restricted to the most luminous QSOs. Many of these objects do show evidence for extended (up to 100 kpc) luminous galaxies based either on broadband optical/near IR (Lehnert et al. 1999) or narrowband redshifted $\text{Ly}\alpha$ (Heckman et al. 1991) imaging studies. None of these surveys to date comprise a large, representative sample of QSOs from which to draw meaningful statistical inferences regarding the QSO population as a whole. Most surveys have focussed on radio-loud QSOs, although evidence of a large luminous ($M_R = -25$) host galaxies has been found around a few radio-quiet QSOs at high redshift (Aretxaga et al. 1995, Hutchings 1995). Aretxaga et al. (1998) have argued that the properties of these host galaxies are consistent with an extrapolation of the luminosity-size relation for HII regions at low redshift.

One of the most spectacular discoveries of an AGN host galaxy at intermediate redshifts is the powerful post-starburst host galaxy of the $z = 0.634$

radio-loud QSO UN J1025-0040 (Brotherton et al. 1999, see Figure 5). Optical spectroscopy and optical/near-IR imaging of this object are consistent with a $M_B = -21.9$ AGN in a $M_B = -24.6$ 400 Myr $10^{11} M_\odot$ post-starburst galaxy. The infra-red properties of this object are consistent with other ultra-luminous infrared galaxies (ULIRGs, see Sanders & Mirabel 1996). This object may represent a transitional phase in the life of a QSO between dust-enshrouded birth (as a ULIRG) and maturity (as a 'normal' QSO). At low redshifts, tunable filter imaging has also provided some spectacular results. Shopbell et al. (1999) have recently discovered a 250 kpc nebula around an X-ray-selected radio-loud QSO at $z = 0.089$.

2.3. Summary

Both these latter objects are very much exceptional cases. Overall, imaging studies have shown that at low redshifts QSOs appear to lie in a relatively normal galaxies. The correlation between nuclear and host galaxy luminosity results in more luminous QSOs lying in larger and more luminous galaxies. Radio-loud QSOs appear to exclusively inhabit elliptical systems, while radio-quiet QSOs are found in both elliptical and spiral systems. At least half of the $M_B < -21$ radio-quiet QSOs/AGN are found in bulge-dominant systems. Otherwise the properties of QSO host galaxies are consistent with the field. There is little evidence from recent comprehensive imaging studies for enhanced interaction/merger activity occurring in the vast majority of QSOs/AGN. Any such activity appears confined to the brightest QSOs.

Diffraction-limited imaging in near-infrared offered by the latest generation of 8m-class telescopes will allow host galaxy studies to extend to more representative samples of QSOs at $z > 1$. However, an equally promising route is the use of spatially-resolved spectroscopy to probe the central regions of QSO host galaxies. Spectroscopy (e.g. by using the strength of the Ca triplet) can be used to determine the relative contribution from stellar/non-thermal processes (host galaxy/nuclear component) in the central regions. Some impressive results in the nuclei of low-redshift AGN have already been obtained using integral field IR spectroscopy with the MPE 3D instrument (see e.g. Maiolino et al. 2000). With the light-grasp and spatial resolution of 8m-class telescopes, such observations are likely to make a significant contribution towards our understanding of the physical processes occurring in QSOs and their host galaxies.

3. QSO environments

Much of the pioneering work in field of QSO environments was carried out by Yee and collaborators (Yee & Green 1984, 1987, Ellingson, Yee & Green 1991). Based on broadband imaging studies of both radio-loud and radio-quiet QSOs at $z < 0.6$, they found that the environments of radio-loud QSOs undergo a dramatic evolution at $z \sim 0.4$, from a field galaxy environment (a galaxy clustering scale length $r_0 \sim 10$ Mpc) to an environment consistent with Abell richness class 1 clusters ($r_0 \sim 20$ Mpc). In contrast, the environments of radio-quiet QSOs remained consistent with that of a field galaxy environment throughout the redshift range studied.

Figure 6. QSO-galaxy cross-correlation functions $w(\theta)$ for low redshift radio-quiet QSOs. At a given angular scale θ $w(\theta)$ measures the mean fractional excess in the surface density of galaxies around QSOs. Different symbols are used to denote redshift ranges for various QSO samples: $z < 0.3$ QSOs (asterisks), $0.3 < z < 0.5$ QSOs (triangles), $0.5 < z < 0.7$ QSOs (squares). The solid line is the galaxy correlation function from Maddox et al. (1990), $r_0 = 11$ Mpc. The dotted line is a predicted QSO-galaxy cross-correlation function with $r_0 = 4$ Mpc, equivalent to the correlation length of high redshift galaxies in the CFRS (Le Fèvre et al. 1996). Compilation supplied by Smith (private communication), based on results reported in Smith et al. (1996) & Smith et al. (2000).

3.1. Radio-quiet QSOs

Further broadband imaging studies of low-intermediate redshift QSOs have confirmed this distinction between the galaxy environments of radio-loud and radio-quiet QSOs. At low redshifts ($z < 0.3$), Smith et al. (1996) have shown that the correlation length of galaxy clustering around X-ray selected (predominantly radio-quiet) $M_B > -24$ QSOs is identical to that of field galaxies over the same redshift range (see Figure 6). More recently Smith et al. (2000) have extended this work to higher redshifts ($0.3 < z < 0.7$). They confirm that the level of galaxy clustering around QSOs remains weak, evolving from a QSO-galaxy correlation length of $r_0 \sim 10$ kpc at $z \sim 0.3$ to $r_0 < 4$ Mpc at $z \sim 0.7$. This is consistent with the evolution of the galaxy correlation length at $z < 1$ seen in the Canada-France Redshift Survey (CFRS) by Le Fèvre et al. (1996).

At still higher redshifts $z > 1$, most optical (Boyle & Couch 1991, Croom & Shanks 1999) and near-infrared imaging surveys (Teplitz et al. 1999) have found no evidence for any significant excess of galaxies around radio-quiet QSOs (see Figure 7). Croom & Shanks demonstrate that their null results is con-

sistent with a 2σ upper limit on the QSO-galaxy clustering scale-length of $r_0 < 12$ Mpc. However, imaging studies carried out by Hutchings and collaborators (see Hutchings 1998 and references therein), cite evidence for significant galaxy excesses around some radio-quiet QSOs at $z \sim 2$. The nature of the discrepancy between the results of Hutchings et al. and other imaging studies remains unclear.

3.2. Radio-loud QSOs

In marked contrast to radio-quiet QSOs, broadband imaging optical/near-infrared studies of radio-loud QSOs (Hintzen et al. 1991, Aragon-Salamanca et al. 1995, Hall & Green 1999), have found an enhanced galaxy surface density of up to 100% around radio-loud QSOs in the redshift range $0.9 < z < 2$ (see Figure 7). The inferred space densities of galaxies are consistent with Abell richness class 0 clusters.

In broadband imaging surveys the true space density of galaxies around the QSO must be inferred from the observed excess projected on the plane of the sky. This requires detailed knowledge of the galaxy luminosity function and its evolution with redshift. Any measure of the galaxy clustering scale length around QSOs is therefore dependent on our current imprecise knowledge of the galaxy luminosity function at anything other than the lowest redshifts ($z < 0.3$). Although the relative difference between the galaxy environments of radio-quiet and radio-loud QSOs at high redshift is unaffected by this, any measure of the clustering length of galaxies around QSOs at $z > 0.3$ is subject to increasing uncertainty.

A solution to this problem is offered by tunable filter imaging. Baker et al. (1999) have used the Taurus tunable filter to identify [OII]-emitting galaxies centered around a radio-loud QSO at $z=0.9$. Tuned to the wavelength of a prominent emission line (e.g. $\text{Ly}\alpha$, [OIII], [OII]) at the redshift of a given QSO, tunable filter imaging can be used to identify associated high-redshift galaxies. This technique does not require the costly amounts of telescope time (and aperture) which would need to be invested in multi-object spectroscopy. The success of narrowband searches for $\text{Ly}\alpha$ emitters in the fields of $z > 3$ QSOs (see e.g. Hu et al. 1998), should provide encouragement for those who wish to use the tunable filter technique at yet higher redshifts.

3.3. Summary

Radio-loud and radio-quiet QSOs appear to inhabit different environments over a wide range in redshift ($0.5 < z < 2$). Galaxy clustering around radio-quiet QSOs is consistent with a field galaxy environment over this redshift range ($r_0 < 10$ Mpc), while radio-loud QSOs are found in much richer environments ($r_0 \sim 20$ Mpc). The conversion of observed galaxy excess on the sky to physical excess at the redshift of the QSO is limited by our knowledge of the galaxy luminosity function at high redshifts. Tunable filter/ narrow band imaging provides both a means to identify clusters around QSOs at $z > 1$ and to measure directly the strength of galaxy clustering around such QSOs.

Figure 7. Upper panel: The QSO-galaxy cross-correlation function $w_{\text{qg}}(\theta)$ for $1 < z < 1.5$ radio-quiet QSOs based on deep b_J band photographic plate measurements (Croom & Shanks 1999). No excess is detected. Lower panel: The observed surface density of galaxies detected in the K band around radio-loud QSOs with $z > 1.4$ (Hall & Green 1998). In this case, the mean fractional galaxy excess around QSOs reaches up to 50%, $w(\theta) = 0.5$, at scales less than 10 arcsec.

Figure 8. The QSO optical luminosity function derived from the 2dF QSO redshift survey (Boyle et al. 2000). Bins fainter than $M_B = -23$ may be subject to significant incompleteness.

4. The Cosmological Evolution of QSOs

4.1. Optical

The QSO optical luminosity function (OLF) and its evolution with redshift provides a fundamental measure of the demographics of the QSO population. Based on the major ultra-violet excess (UVX) QSO surveys of the 1980s (Braccesi: Marshall et al. 1983, Palomar-Green: Green, Schmidt & Liebert 1986, Durham-AAT: Boyle et al. 1990), a picture emerged in which the low-intermediate redshift ($z < 2.2$) QSO optical luminosity function (OLF) was modelled by a two-power-law function with a 'break' luminosity at L^* whose redshift dependence was best fit by pure luminosity evolution (PLE), i.e. a uniform increase in luminosity toward high redshift (see e.g. Marshall et al. 1985; Boyle et al. 1988; Hartwick & Schade 1990).

Later QSO surveys, particularly those that focussed on bright magnitudes (LBQS: Hewett, Foltz & Chaffee 1993, EQS: Goldschmidt & Miller 1998, HBQS: La Franca & Cristiani 1997) have reported evidence for a more complex form of evolution in which the bright end of the OLF shows a significant flattening of the bright end of the OLF with decreasing redshift. At the very lowest redshifts studied, the OLF showed a single featureless power law (Kohler et al. 1997).

Figure 9. The evolution of L^* as a function of fractional look-back time (τ). The solid line denotes the evolution of the optical L_B^* derived from 6000 QSOs in the 2dF QSO redshift survey with $0.35 < z < 2.3$. The extrapolation of this model to $z = 5$ is shown by the dotted line. The modelled behaviour of L_R^* for radio-loud QSOs is shown by the dashed line (Dunlop & Peacock 1990).

Most recently, results from the 2dF QSO redshift survey (Boyle et al. 2000, see Figure 8) confirm that PLE remains a valid model for the cosmological evolution of QSOs with $-23 > M_B > -26$ over the redshift range $0.35 < z < 2.3$. The form of evolution is best-fit by a polynomial function of the form:

$$L^*(z) \propto L^*(0) 10^{1.35z - 0.26z^2}$$

The redshift dependence of this functional form is shown in Figure 9. Note that it predicts a strong peak in the evolution of QSOs at $z \sim 2.5$, followed by a significant decline.

Much evidence for this decline already exists. Both Schmidt et al. (1994) and Warren et al. (1994) model the numbers of QSOs identified from multi-colour/grism surveys by dramatic decline in the space density of luminous ($M_B < -26$) at $z > 3$. The rate of decline is equivalent to a factor of 3 drop in the space density of QSOs per unit redshift. The rapid rise in the luminosity density of QSOs at $z < 2$, followed by the subsequent decline at $z > 3$ was shown by Boyle & Terlevich (see Figure 10) to mirror the evolution of the star formation rate derived by Madau et al. (1996). Further critical tests of the decline, particularly for the bulk of the QSO population which lies at fainter magnitudes ($M_B > -26$), will be possible with the ever increasing numbers of QSOs now being discovered at these redshifts by surveys such as the SDSS (Fan et al. 1999).

Figure 10. The redshift evolution of the QSO (solid/dashed line) and galaxy (data points) UV (2800Å) luminosity density. The galaxy UV luminosity density has been scaled down by a factor 40. Reproduced from Boyle & Terlevich (1998).

4.2. Radio

Any decline in the observed space density of QSOs at optical wavelengths may, of course, simply be caused by increasing obscuration at high redshifts (see e.g. Fall & Pei 1993). Indeed, if the peak of QSO formation is associated with the peak in the star formation rate, then it might be expected that gas/dust associated with vigorous star-formation in the host galaxies at high redshift might preferentially obscure QSOs at these redshifts. However, evidence for a decline in the space density of QSOs at high redshifts ($z > 2$) is also found at radio wavelengths (Shaver et al. 1996) unaffected by dust obscuration. At $z < 2$ Dunlop & Peacock (1990) had previously fit the evolution of radio-loud QSOs by a similar polynomial evolution law to that subsequently derived for the optical LF (see Figure 9):

$$L^*(z) \propto L^*(0) 10^{1.18z - 0.23z^2}$$

Combined with the results at high redshift, there inference is that the radio-loud QSO population also peaks in the range $2 < z < 3$.

4.3. X-ray

Although X-ray missions such as ROSAT and ASCA have yielded a dramatic increase in our knowledge of the evolution of the QSO X-ray LF over the past 5–10 years, there are still few X-ray-selected QSOs with $z > 3$. As a result, the nature of any decline at these redshifts is still poorly known.

Figure 11. The Comastri et al. (1995) model (solid line) fit to the X-ray background spectrum. The various curves represent the contribution of unabsorbed (dotted line) and absorbed (dot-dashed lines) objects, labeled in $\log(N_H)$ values. Symbols represent measured values of the X-ray background from a variety of sources; ASCA 1–7keV data (Gendreau 1995), HEAO1-A4 10–100keV data (Gruber et al. 1999), HEAO1-A4 70–500keV bow tie (Kinzer et al. 1997), HEAO1-A2 3–300keV data (Gruber 1992), BeppoSax 2–8keV bow tie (Vecchi et al. 1999), ASCA/ROSAT 1–10keV joint fit (Miyaji et al. 1999). See Comastri et al. (1995) for details.

At lower redshifts, many studies have found that pure luminosity evolution is a good fit to the available data (Boyle et al. 1994, Page et al. 1997). Recently, this picture has been challenged (Miyaji et al. 1998), with more complex evolutionary models required to fit the data. At some level, this discrepancy may come down to what is classified as a QSO in these surveys. The Miyaji et al. survey include both broad emission line (Type 1 AGN/QSOs) and narrow emission line (Type 2 AGN) objects in their analysis, whereas previous studies have traditionally only included broad emission-line QSOs. It is certainly true that increasing numbers of high redshift ‘Type 2-like’ objects are being found in soft X-ray surveys at the faintest flux limits (Almaini et al. 1995, McHardy et al. 1998) or at harder X-ray energies (Ohta et al. 1996, Boyle et al. 1998). These objects have moderate column densities of neutral hydrogen ($N_H \sim 10^{21} - 10^{22} \text{ cm}^{-2}$) and are predicted to appear at faint X-ray fluxes by models which sought to explain the X-ray background by QSOs/AGN (Comastri et al. 1995).

Although ‘Type 1’ QSOs are capable of producing 40% of the X-ray background at 1 keV ($\equiv 1.2 \text{ \AA}$), the spectral shape of the X-ray background is much flatter ($f_\nu \propto \nu^{-0.4}$) than that observed for normal QSOs ($f_\nu \propto \nu^{-1}$). Indeed most of the energy in the X-ray background originates at 30 keV (Fabian & Barcons 1992), at much harder energies than those studied by most X-ray imaging missions (0.1 – 10 keV).

Figure 12. Deep number counts in the $850\mu\text{m}$ band. The solid and dashed lines denote the predicted QSO contribution according to two different evolution models. The symbols denote the observed surface densities in the $850\mu\text{m}$ band from a variety of different surveys. At the deepest limits QSOs can account for up to $\sim 20\%$ of the sources at $850\mu\text{m}$. Reproduced from Almaini et al. (1999).

By assuming an intrinsic spread in N_H values, a mix of unobscured ($N_H < 10^{21}\text{cm}^{-2}$) through to the most heavily obscured ($N_H \sim 10^{25}\text{cm}^{-2}$) QSOs can reproduce the observed X-ray background both in intensity and spectral shape (Comastri et al. 1995, see Figure 11). Current Chandra observations should reveal the full extent of this population at the deepest X-ray flux levels (see <http://asc.harvard.edu> for latest Chandra information).

4.4. Sub-millimetre

Reprocessing of radiation by dust implies that the heavily-obscured QSOs required to explain the X-ray background should also show up in the deep sub-millimetre surveys recently carried out with SCUBA (see e.g. Hughes et al. 1998). Almaini et al. (1999) have calculated that, for reasonable spectral energy distributions, the Comastri et al. model used to reproduce the X-ray background implies that up to 20% of the sources in the deep SCUBA fields are likely to be QSOs (see Figure 12). This is close to the observed fractions from the most recent SCUBA source identification programs (Barger et al. 1999). The far infrared luminosities ($L \sim 10^{12}L_\odot$) of these objects suggest that such objects are the high redshift equivalent of the local ULIRGs. Approximately 20–30% of ULIRGs exhibit evidence for an active nucleus (see Sanders & Mirabel 1996), and thus the obscured QSO population at $z \sim 2$ is consistent with our current observation of the observed frequency of active nuclei in star forming galaxies

at the present day. The origin of the far-infrared/submm emission (AGN *v.* starburst) in these objects remains an outstanding issue.

5. Conclusions

A few years ago the received wisdom regarding the environmental and evolutionary properties of radio-loud and radio-quiet QSOs run thus:

- Radio-loud QSOs live in elliptical galaxies
- Radio-quiet QSOs live in spiral galaxies
- QSOs are frequently seen in interactions/mergers which are likely to be an important fuelling mechanism for QSOs
- QSOs live in richer environments than average galaxies
- QSOs exhibit rapid cosmological evolution which peaks around $z \sim 2$.

These 'observations' were sufficient to stigmatise QSOs as being unrepresentative of the galaxy population at high redshift. Recent results have shown that at least three out of five of these statements are probably incorrect and at least one does not give the full story. Based on what we now know, they may be revised to read:

- Radio-loud QSOs do, indeed, appear to live exclusively in elliptical galaxies
- At least 50% of radio-quiet QSOs ($M_B < -21$) live in bulge-dominant (elliptical) galaxies. Once this bias is taken into account the properties of the host galaxies are identical to field galaxies.
- There is no strong evidence (indeed possibly the reverse) that the frequency of mergers/interactions in the vast majority of QSOs/AGN is enhanced compared to an equivalent sample of field galaxies.
- Radio-loud QSOs live in richer environments than field galaxies, but radio-quiet QSOs live in environments consistent with the field over a wide range in redshift $0 < z < 1.5$.
- QSOs do exhibit rapid cosmological evolution which peaks around $z \sim 2$, (in optical, radio and X-ray regimes) that happens to be very similar to the inferred evolution in the luminosity density of galaxies

Of course, most of those properties are only derived from studies of low redshift (and hence predominantly low luminosity) QSOs. The next important step will be to determine the properties (host galaxy, environment) of representative samples of high redshift QSOs, particularly over the range in redshift ($1 < z < 3$) where both the QSO luminosity density and star formation rate in the Universe appear to peak. Theoretical work which links galaxy formation to the build up of supermassive black holes at these redshifts (Kauffman & Haehnelt 2000, Fabian 2000) make specific predictions for the evolution of QSOs and their environments at these epochs. Many of these predictions will

be directly testable with the forthcoming generation of 3D instruments (tunable filters, multi-slit/IFU spectrographs) on large telescopes making full use of diffraction-limited imaging in the infrared ($0.1 \text{ arcsec} \equiv 1.6 \text{ kpc}$ at $z = 2$). It may well be that a full understanding of galaxy formation will only be possible with a parallel understanding of QSO formation and evolution. 3D observations will undoubtedly play a major role in furthering our understanding of QSOs and star formation at this key epoch in the history of the Universe.

6. Acknowledgements

I would like to thank Joss Hawthorn and Wil van Breugel for organising such an excellent conference. They have also both shown commendable patience in waiting for me to complete this paper. Joss Hawthorn also provided many helpful suggestions on a early draft on this paper. I am also indebted to Omar Almaini, John Bahcall, Michael Brotherton, Andrea Comastri, Scott Croom, John Magorrian and David Schade and my collaborators on the 2dF QSO redshift team for allowing me to reproduce some of their results in this review.

References

- Almaini O., Boyle B.J., Griffiths R.E., Shanks T., Stewart G.C., Georgantopoulos I., 1995, *MNRAS*, 277, 31
- Almaini O., Lawrence A., Boyle B.J., 1999, *MNRAS*, 305, L59
- Aragon-Salamanca A., Ellis R.S., O'Brien K.S., 1995, *MNRAS*, 281, 945
- Aretxaga I., Boyle B.J., Terlevich R.J., 1995, *MNRAS*, 275, L27
- Aretxaga I., Le Mignant D., Melnick J., Terlevich R.J., Boyle B.J., 1998, *MNRAS*, 298, L13
- Bahcall J.N., Kirakos S., Saxe D.H., Scheider D.P., 1997, *ApJ*, 479, 642
- Baker J.C., Hunstead R.W., Bremer M., Bland-Hawthorn J., Athreya R., 2000, *MNRAS*, in press
- Barger A., Cowie L.L., Smail I., Ivison R.J., Blain A.W., Kneib J.P. 1999, *AJ*, 117, 2656
- Boyce P.J., et al., 1996, *ApJ*, 473, 760
- Boyle B.J., Fong R., Shanks T., Peterson B.A., 1990, *MNRAS*, 243, 1
- Boyle B.J., Shanks T., Peterson B.A., 1988, *MNRAS*, 235, 935
- Boyle B.J., Couch W.J., 1993, *MNRAS*, 264, 604
- Boyle B.J., Shanks, T., Georgantopoulos, I., Stewart G.C., Griffiths R.E., 1994, *MNRAS*, 271, 639
- Boyle B.J., Terlevich R.J., 1998, *MNRAS*, 293, 49
- Boyle B.J. Almaini O., Georgantopoulos I., Blair A.J., Stewart G.C., Griffiths R.E., Shanks T., Gunn K.F., 1998, *MNRAS*, 297, L53
- Boyle B.J., Croom S.M., Smith R.J., Shanks T., Miller L., Loaring N., 1999, *Phil. Trans. R. Soc. Lond. A.*, 357, 185
- Boyle B.J., Shanks T., Croom, S.M., Smith R.J., Miller L., Loaring N., 2000, *MNRAS*, submitted
- Brotherton M. et al., 1999, *ApJ*, 520, 87
- Comastri A., Setti G., Zamorani G., Hasinger G., 1995, *A&A*, 296, 1
- Croom S.M., Shanks T., 1999, *MNRAS*, 303, 411
- Croom S.M., Shanks T., Smith R.J., Boyle B.J., Miller L., Loaring N., 1999, Dunk Island Workshop WWW proceedings <http://www.mso.anu.edu.au/DunkIsland/Proceedings/Croom/>
- De Robertis M.M., Yee H.K.C., Hayhoe, K., 1998, *ApJ*, 496, 93

- Dunlop J.S., Peacock J.A., MNRAS, 1990, 247, 19
- Dunlop J.S., Taylor G.L., Hughes D.H., Robson E.I., 1993, MNRAS, 264, 455
- Dulzitin-Hacyan D., Krongold Y., Funes-Guridi I, Marziani P., 1999, ApJ, 513, L111
- Ellingson E., Yee H.K.C., Green R.F. 1991, ApJ, 371, 49
- Fan X. et al., 1999, AJ, 118, 1
- Fabian A.C., 2000, MNRAS submitted (astro-ph/9908064)
- Fabian A.C., Barcons X., 1992, ARAA, 30, 429
- Fall S.M., Pei Y.C., 1993, ApJ, 402, 479
- Gendreau K.C., 1995, PhD thesis, Massachusetts Institute of Technology
- Green R.F., Schmidt M., Liebert J., 1986, ApJS, 61, 305
- Gruber 1992, in *The X-ray Background*, eds X.Barcons & A.C.Fabian, (Cambridge: Cambridge University Press), p44
- Gruber D.E., Matteson J.L., Peterson L.E., Jung G.V., 1999, ApJ, 520, 124
- Goldschmidt P., Miller L., 1998, MNRAS, 293, 107
- Hall P.B., Green R.F., 1998, ApJ, 507, 558
- Hartwick F.D.A., Schade D.J., 1990, ARAA, 28, 437
- Heckman T.M., Lehnert M.D., van Breugel W., Miley G.M., 1991, ApJ, 370, 78
- Hewett P.C., Foltz C.B., Chaffee F.H., 1993, ApJ, 406, L43
- Hewett P.C., Foltz C.B., Chaffee F.H., 1995, AJ, 109, 1498
- Hintzen P., Romanshin W., Valdez F., 1991, ApJ, 366, 7
- Hu E., McMahon R.G., Cowie L.L., 1999, ApJ, 522, L9
- Hu E., Cowie, L.L., McMahon R.G., 1998, ApJ, 502, L99
- Hughes D.H. et al., Nat, 394, 241
- Hutchins J.B., 1995, AJ, 110, 994
- Hutchins J.B., 1998, AJ, 116, 20
- Kaufmann G., Haenelt M., 2000, MNRAS submitted (astro-ph/9906493)
- Kinzer R.L., Jung G.V., Gruber D.E., Matteson J.L., Peterson L.E., 1997, ApJ, 475, 361
- Kohler T., Groote D., Reimers D., Wisotzki L., 1997, A&A, 325, 502
- Kotilainen J., Ward M.J., 1994, MNRAS, 266, 953
- La Franca F., Cristiani S., 1997, AJ, 113, 1517
- Le Fèvre O., Hudon D., Lilly S.J., Crampton D., Hammer F., Tresse L., 1996, ApJ, 461, 534
- Lehnert M.D., van Breugel W.J.M, Heckman T.M., Miley G.K., 1999, ApJS, 124, 11
- MacKenty J.W., 1990, ApJS, 72, 231
- McHardy I. et al., 1998, MNRAS 295, 641
- McLure R.J., Dunlop J.S., Kukula M.J., Baum S.A., O'Dea C.P., Hughes D.H., 1999, MNRAS, 308, 377
- McLeod K.K., Reike G.H., 1994, ApJ, 431, 137
- McLeod K.K., Reike G.H., Storrie-Lombardi L.J., 1999, ApJ, 511, 67
- Madau P., Ferguson H.C., Dickinson M.E., Giavalisco M., Steidel C.C., Fruchter A., 1996, MNRAS, 283, 1388
- Maddox S., Efstathiou G., Sutherland W., Loveday J., 1990, MNRAS, 242, L43
- Magorrian J. et al., 1998, AJ, 115, 2285
- Maiolino R., Thatte N., Alonso-Herrero, Lutz D., Marconi A., 2000 in *Imaging the Universe in Three Dimensions*, eds W. van Breugel & J. Bland-Hawthorn, ASP Conference Series, in press
- Malkan M.A., Gorijian V., Tam R.T., 1998, ApJS, 117, 25
- Margon B., 1999, Phil. Trans. R. Soc. Lond. A 357, 93
- Marshall H.L., Tananbaum H., Avni Y., Zamorani G., 1983, ApJ, 269, 35
- Marshall H.L., 1985, ApJ, 299, 109

- Miyaji T., Ishisaki Y., Ogasaka Y., Ueda Y., Freyberg M.J., Hasinger G., Tanaka Y., 1998, A&A, 334, L13
- Miyaji T., Hasinger G., Schmidt M., 1998, astro-ph/9809398
- Ohta K. et al., 1996, ApJ, 458, L57
- Page M.J., Mason, K., McHardy I.M., Jones L.R., Carrera F.J., 1997, MNRAS, 291, 324
- Sanders D.B., Mirabel I.F., 1996, ARAA, 34, 749
- Schade D.J., Boyle B.J., Letawsky M. 2000, MNRAS, submitted
- Shaver P.A., Wall J.V., Kellerman K.I., Jackson C.A., Hawkins M.R.S., 1996, Nat, 384, 439
- Shopebell P., Veilleux S., Bland-Hawthorn J., 1999, ApJL, 524, 83
- Simkin S.M., Su H.J., Schwarz M.P., 1980, ApJ, 237, 404
- Smith E.P., Heckman T.M., Bothun G.D., Romanshin W., Balick B. 1986, ApJ, 306, 64
- Smith R.J., Boyle B.J., Maddox S.J., 1996, MNRAS, 277, 270
- Smith R.J., Boyle B.J., Maddox S.J., 2000, MNRAS, in press
- Schmidt M, Schneider D.P., Gunn J.E., 1994, AJ, 110, 68
- Taylor G.L., Dunlop J.S., Hughes D.H., Robson E.I., 1996, MNRAS, 283, 930
- Teplitz H.I., McLean I.S., Malkan M.A., 1999, ApJ, 520, 469
- Vecchi A., Molendi S., Guainazzi M., Fiore F., Parmar A.N., 1999, A&A, 349, L73
- Warren S.J., Hewett P.C., Osmer P.S., 1994, ApJ, 421, 412
- Yee H.K.C., Green R.F., 1984, ApJ, 280, 79
- Yee H.K.C., Green R.F., 1987, ApJ, 319, 28
- Zitelli V., Granato G.L., Mandolesi N., Wade R., Danese L., 1993, ApJS, 84, 185

Recent Spectroscopy of the 3C 48 Host Galaxy and a Simple Image Slicer Design

Alan Stockton and Gabriela Canalizo

*Institute for Astronomy, University of Hawaii, 2680 Woodlawn Drive,
Honolulu, Hawaii 96822*

Abstract. We discuss results of recent spectroscopy of the host galaxy of 3C 48. Although this spectroscopy was carried out with a single long slit (on the Keck Low-Resolution Imaging Spectrometer), it is useful to consider ways to retain the advantages of the long-slit configuration while still achieving some multiplex advantage. We describe a simple three-slice image slicer that we have used for QSO host-galaxy spectroscopy and similar applications on University of Hawaii spectrographs.

1. Spectroscopy of the 3C 48 Host Galaxy

Efforts at spectroscopy of QSO host galaxies until recently have been largely confined to simple attempts to detect the presence of stars and to studies of the extended emission-line regions that surround many QSOs. With the introduction of 8–10-m-class telescopes, the emphasis is shifting to studies of stellar populations and attempts to reconstruct the star-formation history in these objects. We have recently completed a detailed investigation of the host galaxy of 3C 48 (Canalizo & Stockton 1999) using the Low-Resolution Imaging Spectrograph (Oke *et al.* 1995) on the Keck II telescope. The results include:

- Radial velocities of the stellar component in 32 independent regions of the host galaxy. These show, *e.g.*, that the NW tidal tail is approaching by as much as 300 km s^{-1} relative to the systemic velocity.
- Models of the stellar populations in the same 32 regions, involving an underlying old stellar component and a recent starburst or post-starburst component, allowing the determination of an age for the younger component (see example in Fig. 1).
- The NW tidal tail has a probable dynamical age older than most of the recent star-forming activity. The delay of much of the star formation until considerably later than the first close passage may mean that at least one of the merging galaxies had a massive bulge (Mihos & Herquist 1996).
- Vigorous star formation is still going on in regions just NE and SE of the quasar. 3C 48 is probably at or near the peak of its starburst activity.
- The extended emission-line gas is distributed mostly in a few discrete regions. Almost all of this emission falls in one or the other of two distinct

Figure 1. Spectrum of a region about $1''$ square in the host galaxy of 3C 48. Shown are the observed data in the rest frame (heavy line), an underlying old stellar population model (lower light line), a young instantaneous starburst model (upper light line), and the χ^2 fit of the sum of the two models to the data (dotted line). The old model is a 10 Gyr old stellar population with an exponentially decreasing star formation rate with an e-folding time of 5 Gyr. The original data have been smoothed with a Gaussian filter with $\sigma = 2 \text{ \AA}$. Emission lines seen in these spectra come from the extended-emission region of the quasar and not from the star-forming regions.

velocity regimes: either (1) within $\sim 200 \text{ km s}^{-1}$ of the systemic velocity, or (2) blueshifted by $\sim 500 \text{ km s}^{-1}$ or more. Densities in the brighter regions are $\sim 150 e^- \text{ cm}^{-3}$, indicating compression by shocks.

- Extremely luminous high-velocity emission close to the quasar and the convoluted shape of the VLBI radio jet (Wilkinson *et al.* 1991) indicate a strong interaction between the radio jet and the ambient gas. Such an interaction is also probably responsible for the bubble-like structure seen $1''$ NE of the quasar in the HST images.

These observations were carried out with a traditional long slit, using six slit positions to cover most of the galaxy. Clearly it would be useful to multiplex spectroscopic studies of this kind with some sort of integral-field unit. For at least some aspects of QSO host galaxy spectroscopy (particularly in obtaining very accurate sky subtraction and in retaining information at all spatial frequencies in the spatial dimension, needed, *e.g.*, for accurate deconvolution), it is very desirable to avoid introducing structured elements, such as fiber optics or lenslet arrays, before the final sampling of the focal plane at the detector, so a device that essentially multiplies the long-slit format is close to ideal. The rest of this paper describes a very simple device that produces in-focus spectra on the detector of three contiguous long slits side-by-side on the sky.

2. A Simple Three-Slit Image Slicer

2.1. Introduction

For a number of years, we have used various realizations of an image-slicer concept, developed by the first author, on UH spectrographs used at the UH 2.2 m telescope and at the Canada-France-Hawaii telescope (CFHT). Although this slicer was briefly mentioned in a review by Picat (1988), it has never been described in any detail. Its principal advantages are that:

- The slicer can simply replace a slit assembly in many existing spectrographs, providing only that there is sufficient clear space below the slit plane.
- Some versions of the slicer can be designed to accomodate relatively fast focal ratios, while keeping the slices in focus (we have used one at the CFHT f/8 Cassegrain focus).
- Since the slicer uses either refraction or total internal reflection for all beam-steering functions, the only light losses are four (coated) air-glass interfaces and the very slight internal absorption of the medium. Efficiencies over wide bandpasses are $\sim 96\%$.
- The slicer can be made with the three slits having different widths. This feature can be useful, *e.g.*, for spectrophotometry, where one might wish to make the outer slits wide to cover essentially all of the light from an object while keeping the central slit narrow to maintain high spectral resolution on the brightest part of the object.

The only significant disadvantage of the slicer design is that there is no very practical way to extend the design beyond three slits.

2.2. Description

For purposes of discussion, it is useful to isolate the three functions of the slicer: (1) reformatting of three contiguous slits to three separated slits, (2) displacement of the outer slits parallel to their lengths, and (3) pathlength compensation to bring all three slits to the same focus.

We have used two approaches to reformatting contiguous slits to separated slits. The first of these, illustrated in Fig. 2, uses tilted plane-parallel plates to shift the outer slits away from the central slit. The advantage of this approach is that it can be used in fairly fast beams. We have used a slicer with this sort of front end at the f/8 Cassegrain focus of the CFHT. The disadvantage is that the corners of the prisms defining the central slit have to be sharp and without chips. These are difficult to manufacture and are vulnerable to damage. In practice, we use a cover plate over the slicer to protect these fragile edges.

Our second approach, shown in Fig. 3, is more robust but is restricted to somewhat slower focal ratios. This technique uses total internal reflection at the interfaces of a high-index glass and a low-index glass in a compound prism. We have had several of these prisms made, with different amounts cut off the top to give different central slit widths. Depending on the masking arrangement, the side slits can be the same width as the central slit, or some different width.

Figure 2. Reformating a rectangular field into three separated slits by the use of tilted plane-parallel plates. In this view, the slit lengths are perpendicular to the plane of the figure, and the heavy lines at the top of the prisms indicate an opaque mask defining the width of rectangular field.

Both of these approaches effectively present the subsequent optics with the appearance of three well separated slits. At this point it is relatively easy to displace the outer slits in opposite directions parallel to their long dimensions and by an amount a little more than their lengths. Again, we use total internal reflection, this time at glass—air interfaces of rhomboid prisms, as shown in Fig. 4. In practice, we combine these rhomboid prisms with the plane-parallel refracting plates used to separate the slits in our first approach described above (Fig. 2).

Note that in both cases, the central slit is exactly like a conventional spectrograph slit, although in our second type of slicer the light for this slit passes through a horizontal plane-parallel plate. However, the paths for the two side slits are longer than that for the central slit because of components perpendicular to the optical axis that have been introduced in separating the slits and in displacing them parallel to their lengths, so the images of these slits will come to a focus in front of the detector. To bring the images of the three slits to a common focus on the detector, we must compensate for these increased path lengths. This can be done by introducing additional optical material into these paths; if this material has a thickness s and an index n , then the focal position is extended by $s(n-1)/n$. Note that much of the previous optical path for these side slits is in glass, so part of the compensation needed is already present.

Aberrations introduced by these slicers (spherical aberration for a plane parallel plate, and, in addition, astigmatism if the plate is tilted) are negligible for the focal ratios at which we have used them. Of more concern may be

Figure 3. Reformatting a rectangular field into three separated slits by total internal reflection in a compound prism. As in the previous figure, the slit lengths are perpendicular to the plane of the figure, and the heavy lines at the top of the prisms indicate an opaque mask defining the width of rectangular field.

the fact that the focus compensation is strictly perfect only for a single wavelength. However, for typical Cassegrain focal ratios, the focus can remain quite acceptable over bandpasses of $\sim 2000 \text{ \AA}$.

Figure 5 shows the output from a slicer used with the University of Hawaii Faint-Object Spectrograph on the CFHT. For certain types of projects, particularly those for which good sky or background subtraction is essential, these slicer designs can increase observational efficiency by nearly a factor of 3 over that of a single long slit, without introducing significant new problems. Slicers based on these designs can be fairly easily retrofitted to existing spectrographs, as long as there is sufficient room below the slit plane.

The development of these images slicers has been supported by the National Science Foundation through various research and instrumentation grants.

References

- Canalizo, G., & Stockton, A. 1999, *ApJ*, submitted
- Mihos, J. C., & Hernquist L. 1996, *ApJ*, 464, 641
- Oke, J. B., et al. 1995, *PASP*, 107, 375
- Picat, J.-P. 1988, in *Instrumentation for Ground-Based Astronomy*, ed. L. B. Robinson (New York: Springer-Verlag), p. 209
- Wilkinson, P. N., Tzioumis, A. K., Benson, J. M., Walker, R. C., Simon, R. S., & Kahn, F. D. 1991, *Nature*, 352, 313

Figure 4. Shifting the slits parallel to their lengths. In this figure, the three separated slits must be visualized as being horizontal, with their lengths parallel to the plane of the figure and stacked one behind another. The path labeled F is that for the slit closest the viewer, M is that for the middle slit (passing through a gap between the two rhomboid prisms), and B is that for the slit farthest from the viewer. The two rhomboid prisms are labeled a , and the two rectangular prisms used for focus compensation are labelled b .

Figure 5. Simultaneous spectra of the radio galaxy Cygnus A obtained with an image slicer in the University of Hawaii Faint-Object Spectrograph mounted on the CFHT. The wavelength ranges covered by the three slits vary slightly because of the optical separation of the slits, even though they cover contiguous regions on the sky.

Rugate Filters: Quasars Beyond $z \sim 7$?

S. Cianci

*Chatterton Astronomy Department, School of Physics, University of
Sydney, NSW 2006, Australia*

J. Bland-Hawthorn

*Anglo-Australian Observatory, PO Box 296, Epping, NSW 2121,
Australia*

J. O'Byrne

*Chatterton Astronomy Department, School of Physics, University of
Sydney, NSW 2006, Australia*

Abstract. Interference filters are commonly used in astronomical instrumentation to isolate discrete spectral bands. Advances in recent years with thin film technology for interference coatings allow for more complex filter designs for specific science applications and better control of 'out of band' interference. A multi-layer thin film filter specifically for passbands at $H\alpha$ and $H\beta$ will be manufactured soon. In tandem with the Taurus Tunable Filter at the AAO, and using charge shuffling (see Bland-Hawthorn & Jones 1998*a, b*), this filter will allow for differential imaging of galactic nebulae and nearby spiral galaxies. Rugate technology (thin film coatings with continuous, periodic variations in refractive index) can provide several discrete high transmission and reflection bands simultaneously, ideal for custom-designed beam-splitter filters. Such filters can be used for searches of high-redshift quasars and other classes of objects. These filters and other specific multi-band/multi-notch filters will be made at Sydney University in conjunction with Francis Lord Optics.

1. Introduction

1.1. Thin film coatings as interference filters

Thin film coatings are not new to astronomical instrumentation, e.g. they are often used as anti-reflection (AR) coatings on optical components, or as interference filters. Most conventional thin film coatings for commercial applications use multi-layer stack structures (Baumeister and Pincus, 1970), consisting of alternating layers of materials with high and low refractive index relative to the glass substrate on which they are deposited. The partial transmission and partial reflection of light at each interface within the multi-layer coating create a complex, wavelength-dependent pattern of high transmittance and reflectance. It is

possible to design an interference filter to provide transmittance and reflectance at desired spectral bandwidths, by selecting particular thin film materials for deposition, and the thicknesses of the layers deposited (see Cianci *et al*, 2000, for more information).

Higher quality, more optically stable coatings can be manufactured by using such techniques as ion-assisted deposition, which has been available for several years (George, 1992, Chapter 3). These techniques, however, are continually improving, as the deposition instrumentation is refined and improved. More accurate monitoring techniques, such as in situ ellipsometry (McFall *et al*, 1996), have also assisted in providing better quality thin film coatings.

1.2. Rugate coatings

Rugate coatings have received considerable attention over the past 15 years, as developments in deposition techniques have allowed such coatings to be manufactured (Johnson & Crane, 1993; Lim *et al*, 1996). A rugate coating is a dielectric coating where the refractive index varies sinusoidally with depth in the coating. The term *rugate* comes from the Latin for 'wrinkled' or 'creased'. The refractive index profile of a rugate filter, $n(z)$, as a function of the thickness of the film, z , can be described by

$$n(z) = n_a + n_v \sin\left(\frac{2\pi z}{\Lambda}\right) \quad (1)$$

where n_a and n_v are the average index and the amplitude of index variation, and Λ is the thickness of a single sinusoidal period (see Figure 1).

Figure 1. Comparison between a conventional multi-layer coating and a rugate coating: (a) the multi-layer film has a step-wise change in refractive index as a function of optical thickness through the filter; (b) in contrast, the refractive index of a rugate coating varies sinusoidally with the optical thickness through the film.

It is difficult to manufacture such coatings since conventional deposition techniques do not easily produce the graded refractive index profile required for the rugate coating. The development of techniques such as plasma enhanced chemical vapour deposition (PECVD) has allowed a closer investigation into the possibilities of rugate coatings for many different applications (George, 1992; Chapter 6). CVD techniques use a vapour/vapour mixture of the thin film materials, which are maintained at a suitable temperature within a chamber containing the substrate. The vapour/vapour mixture reacts to form a solid

film on the substrate surface, and control of the composition of the thin film is obtained by regulating the contents of the vapour mixture. PECVD allows for thin films to be deposited at much lower temperatures than with conventional CVD. The plasma promotes the chemical reactions, as it ionizes and dissociates the gas molecules even at average electron energies (1 - 20eV). At these lower temperatures, degradation of the thin films due to excessive heating is prevented, and a wider variety of materials can be formed.

2. Beam-splitter camera for contrast imaging

The development of rugate coatings provides a new approach to identifying classes of astronomical sources with distinctive spectra, e.g. objects with a series of broad peaks and troughs. In this case, an interference filter could be designed for use as a beam splitter, to transmit the spectral peaks and reflect the troughs. This filter would be part of a ‘beam-splitter’ camera, a possible design of which is shown in Figure 2. Objects matching the beam-splitter profile will have a maximum contrast, $\mathcal{C}$, between the reflected and transmitted images, i.e.

$$\mathcal{C} = (O_T - S_T) [(O_R - S_R)((O_R + O_T) - (S_R + S_T))]^{-1} \quad (2)$$

where O and S are the object and sky fluxes, and the suffices R and T refer to reflection and transmission. The beam-splitter profile is not significantly degraded in fast beams (Bland-Hawthorn *et al*, 1999), and is therefore ideally suited to wide field object searches. To search for different classes of objects, only the beam splitter itself needs to be replaced.

One possible application for rugate beam splitters is for very high redshift quasar searches (e.g. $z > 7$). Haiman and Loeb (1998) suggest that for quasars located just beyond reionization, the Lyman- α ($\text{Ly}\alpha$) forest in the quasar spectra will contain distinct, disjoint Gunn-Peterson (GP) troughs. These are due to absorption by the neutral intergalactic medium (IGM) at wavelengths between the $\text{Ly}\alpha$ wavelength at reionization and the $\text{Ly}\alpha$ wavelength at the source. This trough broadens as sources of progressively higher redshift are observed, as seen in Figure 2 (right). A similar effect could be seen also for higher Lyman series, with some transmitted flux between these troughs.

To investigate this, rugate beam splitters could be designed to transmit the Lyman series flux regions and reflect the GP troughs of such quasar spectra. While the redshift of reionization, z_{reion} is unknown, a series of beam splitters could be used, each designed for a particular quasar redshift and z_{reion} , e.g. a beam splitter designed for $z \sim 8$ and $z_{\text{reion}} = 7$ would give a maximum contrast between transmission and reflection *only* for quasars and reionization corresponding to these conditions (with spectra of the form similar to the 4th shown in Figure 2 (right)). It is important to obtain strong theoretical information corresponding to the form of quasar spectra at these redshifts beyond z_{reion} , which it is hoped cosmologists will provide in the future.

Figure 2. (Left) The rugate beam splitter camera divides the input beam so that the combined flux from all spectral peaks are transmitted to the detector. The spectral troughs are reflected twice before being re-imaged beside the image of the peaks. (Right) Simulated spectra of five quasar sources of different redshift located just beyond z_{reion} (here taken to be 7), when all contaminating effects, such as peculiar velocities, the ionization of the IGM in the vicinity of the source, or absorption by the Ly α forest, are ignored. [Courtesy of Haiman & Loeb, 1998]

3. Other applications for multiple passband interference filters

3.1. H α -H β double-passband filter

As discussed in more detail in Cianci *et al* (2000), a multi-layer blocking filter has been designed with high transmission passbands at H α (656.3nm) and H β (486.1nm), as shown in Figure 3. Using this filter in tandem with a tunable filter (e.g. the Taurus Tunable Filter at the Anglo-Australian Telescope (AAT)), it will be possible to image these two widely separated emission lines at high spectral resolution. Conventionally, this is done using separate blocking filters for each spectral band of interest, but with the multi-band H α -H β filter, only one filter is necessary. This filter will be used for rapid differential imaging of galactic nebulae, by making use of charge shuffling techniques. This will allow for more accurate measurements of the local dust extinction in spiral galaxies, as charge shuffling, using the one optical path, averages out variations in the response of the atmosphere, telescope and detector.

Figure 3. Transmission profile of a 31-layer $\text{TiO}_2/\text{SiO}_2$ coating on a BK7 glass substrate, with GG435 glass to block near-UV transmission and a quarter-wave stack for near-IR rejection.

3.2. Multi-notch anti-reflection coatings.

More complex thin film designs also allow for ‘multi-notch AR coatings’, i.e. coatings that provide very low reflectance at more than one discrete spectral band, as shown in Figure 4. Deposited on the ‘back’ side of an optical component, facing the detector, this type of coating assists in reducing ghost images due to reflections from the detector surface.

Figure 4. Reflectance profile of a 5-layer AR coating design for $\text{H}\alpha$ and $\text{H}\beta$. The coating is designed using TiO_2 , SiO_2 and MgF_2 .

There are many applications for such AR coatings in a wide variety of astronomical instrumentation, as they allow specialist instruments to operate at much higher efficiency in multiple discrete bands. For instance, a single AR coating for both $[\text{OIII}]$ and $\text{H}\alpha$ could be applied to the Planetary Nebula Spectrograph at the AAT (see <http://www.aao.gov.au/local/www/pns/pns.html>), to allow the highest possible throughput in these two bands.

3.3. Multiplex advantage of a long-slit spectrograph.

The multiplex advantage of the RGO spectrograph at the AAT and the Mount Stromlo Observatory 2.3m Double Beam Spectrograph can be significantly in-

creased for the purposes of observing extended, line-emitting sources (e.g. high velocity clouds, spiral galaxies, etc.). Figure 5 shows how a multi-slit (a) can be used in conjunction with a double-band filter (b) to isolate two spectral regions, i.e. $H\alpha$ –[NII] and the [SII] doublet, at four slit positions. In the same manner, multi-band filters can be used in conjunction with the AAT's Low Dispersion Survey Spectrograph (LDSS⁺⁺) in order to identify line emitters at targeted redshifts. The suppression between the bands allows objects to overlap spectrally, thereby affording a further increase in the multiplex advantage.

Figure 5. An illustration of how the multiplex advantage of a long-slit spectrograph can be increased with (a) a multi-slit and (b) a double-band filter. The slits are spaced so that the four double bands are offset in the dispersion direction, with respect to each other, at the detector.

4. Summary

The possibility of manufacturing high quality, optically stable coatings allows for more complex interference filters to be considered. Designs incorporating multiple passbands will be ideal for applications such as high-resolution differential imaging at widely separated emission lines, efficiency improvement of instruments at more than one discrete spectral band, and increasing the multiplex advantage of long-slit spectrographs.

Rugate technology will be important in, for example, a ‘beam-splitter camera’, to image transmitted and reflected discrete spectral bandwidths corresponding to the typical spectra of different classes of objects. This instrument is well suited to object searches, such as high redshift quasars beyond the reionization redshift, brown dwarfs, Swan bands, M dwarfs, etc. Using rugate technology, it is may also be possible to enhance the power of existing specialist cameras (e.g. MAPPIT; Bedding *et al*, 1994) and the new generation of multi-band cameras (e.g. UT dichroic camera; Doi *et al*, 1998).

We wish to thank Dr Zoltán Haiman and Prof Abraham Loeb for the use of their quasar spectra in Figure 2.

References

- Baumeister, P. & Pincus, G. 1970, *Scientific American*, 223, 59
- Bedding, T.R., Robertson, J.G. & Marson, R.G. 1994, *Astronomy & Astrophysics*, 290, 340
- Bland-Hawthorn, J., Dunn, R.B., Gillingham, P.R. & Cianci, S., 1999, in progress

- Bland-Hawthorn, J. & Jones, D.H. 1998*a*, PASA, 15, 44
Bland-Hawthorn, J. & Jones, D.H. 1998*b*, SPIE, 3355, 855
Cianci, S. *et al.* 2000, Imaging the Universe in Three Dimensions, ASP Conf. Series, van Breugel, W. & Bland-Hawthorn, J., Editors
Doi, M. *et al* 1998, SPIE, 3355, 646
George, J. 1992, Preparation of Thin Films, Marcel Dekker, USA
Haiman, Z. & Loeb, A. 1998, submitted to ApJ, July (astro-ph/9807070)
Johnson, W.E. & Crane, R.L. 1993, SPIE, 2046, 88
Lim, S., Shih & S., Wager, J.F. 1996, Thin Solid Films, 277, 144
McFall, W.D., McKenzie, D.R. & Netterfield, R.P. 1996, Surface Coatings and Technology, 81,

Have Most High-Redshift Quasars Been Overlooked?

C. Wolf, K. Meisenheimer, H.-J. Röser, S. V. W. Beckwith, R. Fockenbrock, H. Hippelein, B. von Kuhlmann, S. Phleps, E. Thommes

*Max-Planck-Institut für Astronomie, Königstuhl 17, D-69117
Heidelberg, Germany*

Abstract. We report on quasars in the 16 h-field of the Calar Alto Deep Imaging Survey (CADIS), and include preliminary data from two more fields. In this $10' \times 10'$ field we found six quasars at redshift $z > 2.2$ which are brighter than $R = 22^m$, which is the highest reported surface density of high-redshift quasars within the given limits. Surface densities reported by previous surveys are below one quasar per CADIS field on average, implying that common search methods have overlooked the majority of high-redshift quasars at this magnitude level.

We believe this result to be a consequence of the multicolor database and the search technique used in CADIS, which employs a supervised classification scheme based on a large library of template spectra for stars, galaxies and quasars. This classification also estimates photometric redshifts for galaxies and quasars, producing a true 3-D result for most objects, i.e. their spatial coordinates and redshift (x, y, z) .

Our observations seem to further indicate that the density of high-redshift quasars is also higher than predicted by the luminosity function of Warren et al. (1994) by a factor of three. Our quasars reside in the low-luminosity tail of this luminosity function, i.e. in the part which is not well constrained by the original observations.

1. Introduction

3-D surveys were quite often used to search for quasars. Within this category, there are multicolor surveys and grism surveys recording $F(x, y, \lambda)$ or long-term variability studies recording $F(x, y, t)$. They usually produce a large list of quasar candidates contaminated by other objects to a varying extent. Especially, at $2.2 < z < 3$ quasars display almost stellar colors in common broadband surveys. For imaging surveys follow-up spectroscopy was generally required to proof the true nature of the candidate and determine its redshift.

An ideal 3-D survey for rare objects would simultaneously image a wide field of view with good spectral resolution and coverage. Automated spectral classification schemes could derive a 4-D information vector (x, y, z, class) for each object from such a 3-D database. Presently, the most feasible observational compromise is to use a widefield imaging device and consecutively obtain several mediumband exposures yielding relatively crude spectral information. This approach is taken for the work presented here.

Figure 1. The full CADIS filterset contains three non-standard broadband filters, B , R and K' (not shown), and 13 mediumband filters providing a lot of useful data for multicolor classification. Filters already observed on the 16 h-field are highlighted by thick lines.

2. The Calar Alto Deep Imaging Survey

The Calar Alto Deep Imaging Survey (CADIS) is a deep multicolor survey using three broadband filters (B , R and K') and thirteen mediumband filters in the wavelength range from 390 nm to 930 nm when fully completed (see Fig. 1). For various purposes not described here it is complemented with an emission line survey carried out with a Fabry-Pérot interferometer (Meisenheimer et al. 1998). Images are obtained in ten fields measuring $10' \times 10'$ each. The resulting multicolor database contains faint galactic stars as well as quasars and galaxies over a broad range of redshift at magnitudes of $R < 23^m$ (see Wolf et al. 1999a and references therein). Also, the accompanied K' -band survey establishes one of the largest deep K' -band galaxy surveys and allows to search for unusually red objects.

Observations are performed with the focal reducers CAFOS (Calar Alto Faint Object Spectrograph) at the 2.2-m-telescope and MOSCA (Multi-Object Spectrograph for Calar Alto) and OMEGAprime at the 3.5-m-telescope at Calar Alto Observatory, Spain. For photometric calibration, we established a system of faint spectrophotometric standard stars in the CADIS fields providing us with independence from photometric conditions for regular imaging. The photometric analysis of the individual objects always probes the same effective beam on every survey exposure to insure an accurate measurement of the object colors even under conditions of changing seeing.

3. Classification and Redshift Estimation

Our classification scheme essentially compares the observed colors of each object with a color library of known objects. This library is assembled from observed spectra by synthetic photometry performed on our CADIS filterset. As an input we used the stellar library from Pickles (1998), the galaxy template spectra

from Kinney et al. (1996) and the QSO template by Francis et al. (1991). From this, we generated regular grids of QSO templates ranging in redshift within $0 < z < 4$ and having various continuum slopes and emission line equivalent widths. Also, a grid of galaxy templates has been generated for $0 < z < 2$, and contains various spectral types from old populations to starbursts. We have not yet attempted to find quasars at $z > 4$, since they are believed to be extremely rare and we were not yet able to model their colors well enough.

Objects are classified by locating them in color space and comparing the probability for each class to generate the given measurement. Given the photometric error ellipsoid in the n -dimensional color space, each library object can be assigned a probability to cause an observation of the measured colors. For a whole class, this probability is assumed to be the average value of the individual class members (Parzen's Kernel estimator). From these probabilities we can derive the likelihood of each object to belong to the various classes. Emphasis should be put on the fact, that quasars are selected by a positive criterion (*quasar-like objects*) and not by an exclusive rule (*unusual objects*). The three classes receive equal treatment and their only difference are the volumes they cover in color space.

This statistical approach contains two implicit concepts of unclassifiability. For each object we normalized the relative likelihoods such, that the three classes would always add up to 100%. But if the photometric error ellipsoid contains an overlap of several classes, the likelihood is distributed among them and the true class is *undecidable*. This occurs when the available colors are not discriminating between the classes or when the measurement error is too large. On the other hand, unusual objects which are far away from any library in color space, we call *strange*, in the sense that their measurement is statistically inconsistent with being caused by any library at some level.

Since the galaxy and quasar libraries resemble regular grids in redshift and spectral type, these parameters can also be estimated from the observation. For this purpose, we treat the library as a statistical ensemble generating the measured colors and determine expectation values as well as variances for assessing the quality of the estimate (Wolf et al. 1999b).

So far, we have not completed the entire filterset on any field. Based on the subset exposed on the 16 h-field, we find about 95% of the 500 brightest objects in a field to be classifiable. These objects typically have at least $\approx 20\sigma$ detections in most filters. Down to $R < 23^m8$, only 20 out of roughly 2000 objects are *strange*, i.e. inconsistent with all libraries at more than a 3σ level. A spectroscopic cross-check of more than 200 objects revealed, that 98% of the classified objects at $R < 22^m$ are classified correctly. Especially, there is hardly any contamination among the 22 quasar candidates, out of which 20 were proven to be quasars by spectroscopy.

Our redshift estimation yields mostly good results for galaxies as well as quasars. Quasar redshifts are most reliably estimated at true redshifts of $z > 2$. We have performed Monte-Carlo simulations using the model quasars from the library to predict completeness and accuracy of the redshift estimates. The simulations agree very well with the observations for the photometric redshifts. According to the simulations, we should not miss more than 10% of the library-type quasars with $R < 22^m$.

Figure 2. This R band finding chart shows the quasars found in the CADIS 16 h-field (size $10' \times 10'$, North is up, east to the left). Redshifts are indicated in the image and are mostly above $z = 2.2$.

4. The Quasar Sample

Up to now, 22 QSOs have been spectroscopically confirmed in the CADIS fields 3 h, 9 h and 16 h with two QSOs being found by chance among candidates for starburst galaxies. Only in the 16 h-field all candidates with $R < 22^m$ have been checked. In this field, four QSOs have been found at $z < 2.2$ and six at higher redshift (see Wolf et al. 1999a for a detailed object list). We derive a surface density of 140 ± 70 QSO/ $\square^\circ$ at low redshift ($z < 2.2$) and 215 ± 90 QSO/ $\square^\circ$ at high redshift (for $R < 22^m$).

At low redshift our number agrees well with predictions from contemporary luminosity functions (Boyle et al. 1988) and earlier observations (discussed in Wolf et al. 1999a), which amount to about $100 \dots 150$ QSO/ $\square^\circ$. At high redshift the luminosity function of Warren et al. (1994) predicts ≈ 70 QSO/ $\square^\circ$ with $R < 22$ and $z > 2.2$, about one third of our result. Earlier observations of high-redshift quasars have typically found a surface density for the given limits of only $15 \dots 30$ QSO/ $\square^\circ$, further underlining their incompleteness.

Figure 3. This BRI color diagram shows all unresolved sources with $R < 23^m$ in the 9 h- and 16 h-fields. Objects classified as galaxies by the multicolor algorithm are displayed as grey dots, those classified as stars appear as black dots, and the first nine high-redshift quasars ($z > 2.2$) identified in CADIS are indicated by stars (the three redmost ones are at $z > 3$). Most of them could not be found in broadband surveys due to their stellar colors. The dashed line corresponds to a common selection criterion for $z > 3$ -quasars. It would find only one quasar among a number of candidates which are mostly compact galaxies.

This result is not questioning the validity of earlier work. The low completeness of broadband searches for high-redshift quasars has generally been known and was always difficult to assess. The luminosity function of Warren et al. (1994) was derived from a list of quasars with $R < 20^m$, which hardly constrains the shape of the low-luminosity tail at all. Therefore, disagreements in this domain should not surprise.

This new result for low-luminosity quasars is relevant for several aspects: for the determination of the intergalactic UV background; for studies of the large-scale structure and the inverse effect using the many close-by lines of sight to high-redshift quasars; and for further constraining quasar evolution. The new results challenge the present view of a steep drop-off in quasar abundance beyond $z = 2.5$ (Shaver et al. 1996).

5. Discussion

An elaborate classification technique was used to search for quasars in a multicolor survey that includes many mediumband filters. The classification turned out to be very reliable, also among the quasar candidates. Quasars were found with unprecedented completeness at high redshift, but only one field has been fully checked. Either this field is coincidentally very rich in quasars, or our result suggests that the abundance of low-luminosity quasars at high redshift has been underestimated previously. Given our small number of objects, we can not safely decide between the two options. Work is in progress to collect a larger quasar database in CADIS.

If the 16 h-field is truly representative of the sky, it implies that common broadband search techniques have missed a large fraction of the $z > 2.2$ -quasars. CADIS can still find these objects mainly due to three reasons: its many mediumband filters; its accurate color photometry; and its statistical classification technique (see Fig. 3).

The result motivates a more dedicated survey for QSOs, which is proposed to be carried out with the Widefield Imager at the ESO/MPI 2.2-m-telescope on La Silla. This instrument contains a set of mediumband filters that are optimally suited for quasar search among other purposes. The area of all ten CADIS fields can be covered within a single image. Two fields were chosen to coincide with patches of the sky where important other observations are expected to be done. One is the AXAF South Deep Field and the other one lies in Patch B of the ESO Imaging Survey. The survey proposed here should surpass CADIS in terms of size and homogeneity of the resulting quasar catalogue. Filters are chosen such, that confirmatory follow-up spectroscopy should not be generally necessary and photometric redshifts can be obtained reliably.

This kind of 3-D (x, y, z) survey would greatly benefit from improved 3-D (x, y, λ) imaging capabilities. If, e.g., superconducting tunnel junction (STJ) detectors were presently available and provided the same wide field of view as CCDs, a survey with any number of colors could be carried out in much shorter time. At least the algorithms to process such data have been developed and the approach has proven to be successful.

The authors thank H.-M. Adorf for helpful discussion about classification methods and their fine-tuning. We also thank the Calar Alto staff for their help and support during many observing runs at the observatory. We thank D. Calzetti for kindly making available the galaxy templates in digital form.

References

- Boyle, B. J., Shanks, T., Peterson, B. A., 1988, MNRAS 235, 935
- Francis, P. J., Hewett, P. C., Foltz, C. B., Chaffee, F. H., Weymann, R. J., Morris, S. L., 1991, ApJ, 373, 465
- Kinney, A. L., Calzetti, D., Bohlin, R. C., McQuade, K., Storchi-Bergmann, T., Schmitt, H. R., 1996, ApJ, 467, 38
- Meisenheimer, K. et al., 1998, in *The Young Universe*, ASP Conf. Ser., Vol. 146, p. 134
- Pickles, A. J., 1998, PASP 110, 863
- Shaver, P. A., Wall, J. V., Kellermann, K. I., Jackson, C. A., Hawkins, M. R. S., 1996, Nature, 384, 439
- Warren, S. J., Hewett, P. C., Osmer, P. S., 1994, ApJ, 421, 412
- Wolf, C., Meisenheimer, K., Röser, H.-J., et al., 1999a, A&A, 343, 399
- Wolf, C., et al., 1999b, in preparation

A Very Extended Ionized Nebula around the Quasar MR2251–178

P. L. Shopbell, S. Veilleux

*Department of Astronomy, Univ. of Maryland, College Park, MD 20742
pls@astro.umd.edu, veilleux@astro.umd.edu*

J. Bland-Hawthorn

*Anglo-Australian Observatory, P.O. Box 296, Epping, NSW 2121
jbh@aaoepp2.aao.gov.au*

Abstract. We report the results of deep emission-line ($H\alpha$) imaging of the ionized gas surrounding the low-redshift ($z = 0.0638$) quasar MR 2251–178 using the Taurus Tunable Filter on the AAT telescope. Our observations detect ionized gas out to a radius of 120 kpc, the deepest detection yet in this system. The total mass of ionized gas is $6 \times 10^{10} M_{\odot}$ (upper limit). The large and symmetric extent of the gaseous envelope favors a model in which the filamentary and diffuse emission arises from a large cloud complex, photo-excited by the bright quasar. A crude kinematic analysis does not support an origin for the gas in an interaction with the nearby galaxy G1. We conclude that significant quantities of ionized gas in the galactic environment may go undetected with current observational techniques.

1. Introduction

Recent advances in instrumentation have lead to a number of detections of diffuse, extended ionized gas outside of the standard galactic environment; for example, the extended disks of spiral galaxies (Bland-Hawthorn *et al.* 1997) and clouds in the Magellanic Stream (Weiner & Williams 1996). One particularly successful area of observation has been that of extended stellar and gaseous components surrounding radio-quiet and radio-loud active galaxies and quasars (e.g., Stockton & MacKenty 1987; Tadhunter *et al.* 1994; Hutchings *et al.* 1999). While these components vary greatly in both extent and morphology, the observations suggest that this material may be present in substantial quantities around such systems.

Traditional deep narrowband observations are expensive, in terms of both telescope time and monetary cost for interference filters for modern large-beam telescopes. In order to circumvent this, Fabry-Perot spectrometers are now being employed in a “stare mode” to detect diffuse emission-line gas over small fields and narrow wavelength regions. In this paper, we present deep $H\alpha$ observations of the extended ionized gas surrounding a quasar, MR 2251–178, using a new, etalon-based instrument which has been optimized for the detection of faint,

extended emission-line gas, the Taurus Tunable Filter (TTF)¹. The flexibility of this instrument provides essentially monochromatic imaging at an arbitrary wavelength and bandpass. The high efficiency and tunable nature of the TTF allow the acquisition of multiple images at a range of wavelengths, providing kinematic information as well.

2. Observations

First discovered as a strong and variable X-ray source (Cooke *et al.* 1978), MR 2251-178 was subsequently identified as a quasar at a redshift of 0.0638 (Canizares *et al.* 1978). Optical imagery and spectroscopy have detected two ionized gas components around the quasar: an elongated, highly-ionized circum-nuclear component of diameter ~ 27 kpc, and an extended envelope of faint H α - and [OIII]-emitting filaments out to a radius of ~ 110 kpc (e.g., Bergeron *et al.* 1983; Macchetto *et al.* 1990). Spectroscopic observations have determined that the extended ionized gas component, while undoubtedly associated with the quasar, is not following its rotation pattern (Nørgaard-Nielsen *et al.* 1986; Mulder & Valentijn 1992).

Figure 1 presents our H α imagery of MR 2251-178. In the discussion that follows, we assume $H_0 = 50 \text{ km s}^{-1} \text{ Mpc}^{-1}$ and a corresponding image scale of $1.9 \text{ kpc arcsec}^{-1}$.

2.1. Morphology

Our observations detect H α emission throughout the central arcminute of the field. We confirm two primary ionized gas components in MR 2251-178: The first consists of strong H α emission directly surrounding the quasar out to a radius of ~ 20 kpc, slightly elongated along the axis of the quasar's radio jet (PA $\sim 102^\circ$; Macchetto *et al.* 1990). This component has been identified as an "extended emission-line region" (EELR) of the quasar host galaxy (Mulder & Valentijn 1992).

The second H α component consists of a much larger network of filamentary emission. Diffuse emission is visible surrounding the galaxy on all sides, out to a radius of at least 50 kpc; knots and filaments of ionized gas are visible to radii up to 120 kpc towards the northeast, southeast, and southwest. The morphology of the filaments is very similar to that of spiral arms, with an "arm" clearly visible in each of these three directions. This remarkable symmetry has not been seen in previous, i.e., less deep, imaging observations.

2.2. Kinematics

Comparison of imagery at two separate redshifts (panels *a* and *b* of Figure 1) reveals a broad velocity gradient across the quasar nebula (PA $\sim 40^\circ$). The eastern filaments are generally less redshifted than the western filaments and represent the only gas possibly near the systemic velocity of the quasar. The western filaments in the "red" image appear to be redshifted from the systemic

¹Consult Bland-Hawthorn & Jones (1998, PASA, 15, 44) and the TTF web page for further details on the instrument: <http://www.aao.gov.au/local/www/jbh/ttf/>.

Figure 1. Deep imagery of the field surrounding the quasar MR 2251-178. Panels *a* and *b* are 1200 second exposures of H α at redshifts of 0.0640 and 0.0645, respectively. Panel *c* is an I-band continuum image of the same field, illustrating a number of galaxies in the cluster. Panel *d* is a summed H α image. All panels are 3' in size, with north up and east to the left.

velocity by a value of $\sim 200 - 300 \text{ km s}^{-1}$. The clear difference between the filament morphologies in the blue and red images - particularly in the northeast and the west - implies that essentially all of the detected emission is redshifted relative to the quasar. These conclusions agree well with the published spectroscopic observations of Nørgaard-Nielsen *et al.* (1986). We also confirm that the velocity structure of the extended ionized gas component does not appear to follow the rotation curve of the quasar (Bergeron *et al.* 1983).

2.3. The Mass of the HII Envelope

We derive a total flux from the extended filaments of $1.4 \times 10^{-13} \text{ erg cm}^{-2} \text{ s}^{-1}$. This implies a luminosity of $2.5 \times 10^{42} \text{ erg s}^{-1}$ at the redshift distance of the quasar (384 Mpc). Assuming complete ionization and employing a geometry in which the filaments are approximately as deep as they are wide (i.e., 7 pixels $\sim 5 \text{ kpc}$), we calculate an rms filament density of $\langle n_e \rangle \sim 0.04 \cdot f^{-1/2} \text{ cm}^{-3}$. For a filling factor $f = 1$, this implies an upper limit on the ionized gas mass of $M \sim 6 \times 10^{10} M_{\odot}$, distributed in a volume of $6.1 \times 10^4 \text{ kpc}^3$. These estimates are consistent with previously published values (Nørgaard-Nielsen *et al.* 1986; Macchetto *et al.* 1990).

3. The Origin and Ionization of the HII Envelope

Our observations reveal a tremendous spatial extent for the emission in this system. We also observe a relatively high degree of symmetry in the envelope, including at least three “arms” of ionized gas. This symmetry, together with the organized large-scale kinematics, does not support an interaction model for creation of the ionized gas. The fact that galaxy G1 ($v_r = +1246 \text{ km s}^{-1}$ relative to MR 2251–178) is positioned just north of the quasar, amongst relatively blueshifted filaments, strengthens this argument. Nevertheless, the coarse kinematic structure of the extended ionized gas indicates that it represents a distinct component from the inner counter-rotating EELR, suggesting that the envelope most probably did not originate within the quasar and/or host galaxy. Our observations appear to favor a model in which the extended ionized envelope resides within a large complex of HI gas centered about the quasar.

The source of the ionization of the extended component is difficult to constrain with our limited spectral information. Again, the relative symmetry of the envelope in our observations virtually eliminates ionization mechanisms associated with the quasar jet (PA $\sim 102^\circ$; Macchetto *et al.* 1990). One could conceive of a small population of in-situ hot stars which would provide sufficient ultraviolet flux to ionize the gas yet be faint enough to avoid detection in broadband observations (Macchetto *et al.* 1990). However, a significant amount of *diffuse* ionized gas is present in the nebula, suggesting an exterior ionization source. More importantly, the deep R-band imagery of Hutchings *et al.* (1999) exhibits a faint component extended along the N–S direction, markedly different from the NE–SW axis along which we observe strong emission in H α . They do not detect an excess of continuum emission at the locations of any of the brighter H α knots in the envelope. We therefore concur with previous authors that the primary source of gas excitation is probably the quasar itself.

In summary, our H α observations of MR 2251–178 demonstrate that significant quantities of ionized gas may be present around luminous quasars at flux levels virtually undetectable by standard narrowband imaging techniques. The TTF promises to greatly impact such observational programs by increasing the depth to which diffuse line emission can be detected by a factor of at least 5–10.

References

- Bergeron, J., Boksenberg, A., Dennefeld, M., & Tarenghi, M. 1983, MNRAS, 202, 125
- Bland-Hawthorn, J., Freeman, K.C., & Quinn, P. J. 1997, ApJ, 490, 143
- Canizares, C.R., McClintock, J.E., & Ricker, G.R. 1978, ApJ, 226, L1
- Cooke, B.A., *et al.* 1978, MNRAS, 182, 489
- Hutchings, J.B., Crampton, D., Morris, S.L., Durand, D., & Steinbring, E. 1999, preprint
- Macchetto, F., Colina, L., Golombek, D., Perryman, M.A.C., & Alighieri, S.D.S. 1990, ApJ, 356, 389
- Mulder, P.S. & Valentijn, E.A. 1992, A&A, 259, 435
- Nørgaard-Nielsen, H.U., Hansen, L., Jørgesen, H.E., & Christensen, P.R. 1986, A&A, 169, 49
- Stockton, A. & MacKenty, J.W. 1987, ApJ, 316, 584
- Tadhunter, C., Shaw, M., Clark, N., & Morganti, R. 1994, A&A, 288, L21
- Weiner, B.J. & Williams, T.B. 1996, AJ, 111, 1156

Tunable Filter Imaging of a Quasar Field at $z = 0.9$

Joanne C. Baker

*Astronomy Department, 601 Campbell Hall, University of California,
Berkeley, CA 94720, USA*

Abstract.

Using a unique new instrument — the TAURUS Tunable Filter (TTF) — we have detected an excess of [O II]-emitting galaxies in the field of the radio-loud quasar MRC B0450–221 at $z = 0.9$. A group of 9 galaxies ($I(AB) > 21$) with high estimated star-formation rates ($10 M_{\odot} \text{ yr}^{-1}$ and higher) is found several hundred kpc from the quasar. Four of these have been confirmed spectroscopically to lie at $z = 0.9$, and this in itself corresponds to ten times the expected field density at $z = 0.9$. Probing to fainter levels, more than twenty candidates with rates down to a few $M_{\odot} \text{ yr}^{-1}$ are detected across the field. Together with an excess of red galaxies found within $1'$ of the quasar in broad-band imaging, these data are consistent with the quasar lying in a cluster which is accreting gas-rich galaxies from a neighbouring group at $z = 0.9$.

1. Introduction

We (Baker, Hunstead, Bremer, Athreya) are carrying out a program of deep, narrow-band imaging of the fields of radio-loud quasars using the TAURUS Tunable Filter (TTF: Bland-Hawthorn & Jones 1998*a, b*) on the Anglo-Australian Telescope (AAT). Used on a Fabry-Perot imager, TAURUS, the high-specification TTF etalon can provide high-throughput narrow bandpasses in the range 6–60 Å FWHM anywhere in the wavelength range 6000–10000 Å (see Bland-Hawthorn, this volume). We have used TTF with a bandpass of 7–10 Å to scan the fields of $z \sim 1$ quasars in the [O II] $\lambda 3727$ emission line to investigate the evolution of quasar environments and the star formation properties of galaxies near quasars.

Another primary aim of this project is to detect clusters at $z = 1$ and beyond with a range of richnesses for further study. Measurement of the space density of clusters at $z = 1$ is itself a strong test of cosmological theories. To this end, we are combining TTF imaging with deep broad-band optical and infrared imaging which will find the red, passively evolving ellipticals which populate cluster cores. The TTF observations are critical, as star-forming galaxies are known to be more prolific at higher redshifts both in the field (Lilly et al. 1996; Cowie et al. 1997) and in clusters (Butcher & Oemler 1984; Rakos & Schombert 1995). Also, blue galaxies tend to dominate the outer regions of clusters where sub-structure due to the aggregation of clusters from smaller groups is expected (e.g. Jenkins et al. 1998).

Results are summarised here for the first field studied, MRC B0450-221, a quasar at $z = 0.898$. A detailed analysis is presented in a paper submitted to MNRAS (Baker et al. 1999).

2. Observations and Preliminary Results

The radio-loud quasar MRC 0450-221, at $z = 0.898$ (Kapahi et al. 1998), was observed with TTF at f/8 on the AAT in Feb 1997. We scanned the field in seven adjacent bandpasses of $7 \text{ \AA}$ FWHM (1000s each) either side of $[\text{O II}]$ at the quasar redshift. After standard data reduction in IRAF, aperture photometry was carried out on the TTF frames with FOCAS. The matched catalogues were then compared to find objects with features peaking above 3σ in one or two adjacent frames. Out of 100 objects fainter than $I(AB) = 21$ identified in the co-added TTF data, 31 showed $> 3\sigma$ features (9 with $> 5\sigma$ features) in individual frames. Figure 1 shows zoomed images in the seven frames of some of the emission-line galaxy candidates detected with TTF. Wavelength calibration was carried out using scanned observations of a CuAr arc lamp. Note that for a given frame the wavelength varies quadratically across the field (see Bland-Hawthorn, this volume).

Assuming they lie at $z = 0.9$, the $[\text{O II}]$ equivalent widths of the ELG candidates all lie in the range $W_\lambda = 1-140 \text{ \AA}$ in the restframe, corresponding to equivalent star formation rates of order $1-40 M_\odot \text{ yr}^{-1}$ (Gallagher, Bushouse & Hunter 1989). This is within the expected range for field galaxies at $z \sim 1$ (Cowie et al. 1997; Hammer et al. 1997).

Nine candidates had implied star-formation rates higher than $10 M_\odot \text{ yr}^{-1}$. Following spectroscopy, which was obtained for 5 of these plus another 5 less-significant targets, four have been confirmed to lie at $z = 0.9$. Of the remainder, 3 were emission-line galaxies at lower redshift and 3 are still of uncertain redshift. Even taking only the 4 spectroscopically confirmed candidates at $z = 0.9$, this corresponds to an average overdensity of about ten times that expected in the field (0.4 galaxies with $\text{SFR} > 10 M_\odot \text{ yr}^{-1}$ are expected in a volume of 600 Mpc^3 according to the numbers of Cowie et al. 1997). The overdensity will be higher if more galaxies are confirmed to lie at $z = 0.9$ and if the clustering of the star-forming group within the whole volume sampled is taken into account.

A 5σ excess of red galaxies ($V - I > 2.5$) within $1'$ of the quasar is also found in the broad band images. These red galaxies have colours consistent with ellipticals at $z = 0.9$ that have evolved passively for 3 Gyr or more (using the models of Fioc & Rocca-Volmerange 1997), and which tend to characterise the cores of clusters (Smail et al. 1998). Therefore, our data are consistent with the quasar MRC B0450-221 lying in the core of a cluster at $z = 0.9$. We also detect a group of star-forming galaxies on the outskirts of the cluster, which may be in the process of merging with it.

Figure 1. Examples of ELG candidates in the field of the $z = 0.898$ quasar MRC B0450–221. Sub-images ($12''$ square) have been drawn from the sequence of TTF frames (F1 to F7) spanning $49 \text{ \AA}$ in adjoining $7 \text{ \AA}$ FWHM passbands centred on [O II] at the quasar redshift (i.e. $7074 \text{ \AA}$, F4). All images have been smoothed and are shown on the same linear (wrapping) greyscale. The top panel shows [O II] emission extended to the northwest of the bright quasar (upper left), and slightly redshifted from the published quasar redshift. The other panels are centred on galaxies in the field. Each $7\text{-\AA}$ frame spans 560 km s^{-1} at this redshift. N is up, E to the left.

3. Conclusions

We have carried out a detailed study of the field of the $z = 0.9$ quasar MRC B0450–221 using narrow- and broad-band imaging and spectroscopy. The main conclusions are:

(i) After imaging with a tunable filter, 31 emission-line candidates fainter than $I(AB) = 21$ are found in the field with a significance of greater than 3σ .

(ii) Nine galaxies are detected with star-formation rates greater than $10 M_{\odot} \text{ yr}^{-1}$, corresponding to an average overdensity of more than ten times that expected in the field. These objects are spatially clustered, and the starforming group is offset by several hundred kpc from the quasar.

(iii) Spectroscopy confirms that 4 of the star-forming galaxies with the highest equivalent widths indeed lie at $z = 0.9$.

(iv) The independent discovery of a 5σ overdensity in the reddest (passively evolving) galaxies within $1'$ of and centred on the quasar, suggests the quasar resides in a cluster.

(v) Therefore, the quasar MRC B0450–221 apparently resides in a cluster which is accreting star-forming galaxies from an adjacent group at $z = 0.9$.

We thank Heath Jones for discussions, and staff at the AAT and La Silla. JCB acknowledges support by NASA through Hubble Fellowship grant #HF-01103.01-98A from STSI, operated by AURA, Inc., under NASA contract NAS5-26555.

References

- Bland-Hawthorn, J. & Jones, D.H. 1998*a*, PASA, 15, 44
- Bland-Hawthorn, J. & Jones, D.H. 1998*b*, SPIE, 3355, 855
- Butcher H., Oemler A.Jr, 1984, ApJ, 285, 426
- Cowie L.L., Hu E., Songaila A., Egami E., 1997, ApJ, 481, L9
- Fioc M., Rocca-Volmerange B., 1997, A&A, 326, 950
- Gallagher J.S., Bushouse H., Hunter D.A., 1989, AJ, 97, 700
- Hammer F. et al., 1997, ApJ, 491, 477
- Jenkins A. et al., 1998, ApJ, 499, 20
- Kapahi V.K. et al., 1998, ApJS, ApJS, 118, 327
- Lilly S.J., Le Fevre O., Hammer F., Crampton D., 1996, ApJL, 460, L1
- Rakos K.D., Schombert J.M., 1995, ApJ, 439, 47
- Smail I., Edge A.C., Ellis R.S., Blandford R.D., 1998, MNRAS, 293, 144

The NGST Science Program

H. S. Stockman

*Space Telescope Science Institute, 3700 San Martin Drive, Baltimore,
MD 21218*

J. C. Mather

NASA/Goddard Space Flight Center, Code 685, Greenbelt, MD 20771

Abstract. The Next Generation Space Telescope is the successor to the Hubble Space Telescope and a key element in the NASA Origins program. NGST is designed to address the origins and evolution of galaxies, the early chemical evolution of the universe due to stars, and the processes of star and planet formation. The NGST Ad Hoc Science Working Group has defined a number of specific objectives as part of the Design Reference Mission (DRM). These demonstrate the need for large detector arrays and multiplexing spectrographs (multi-object and 3-D instruments). We review the contents of the DRM and the implications for ongoing studies of NGST instruments by U.S. and international teams.

1. Introduction

The Next Generation Space Telescope (NGST) program finds its technological origins in a series of studies of passively cooled space telescopes. The Edison proposal to ESA (Thronson 1993) and the High-Z and MIRORS proposals to NASA (Stockman 1994, Wade et al. 1995) established the technology of reducing the temperature of a large aperture telescope and instrument complement through radiative cooling. Indeed it is likely that all future infrared (IR) space missions will take advantage of radiative cooling, including *SIRTF*. The scientific goals for NGST, however, were enunciated in the report of the HST & Beyond Committee, chaired by Alan Dressler (Dressler 1996). In the executive summary, the committee emphasized that a large (> 4 m dia.), passively cooled telescope would be the logical successor to *HST* and would enable the study of the origins and evolution of galaxies. Working in the near infrared (NIR) and limited only by the scattering and emission of the zodiacal dust cloud, the soon-to-be-called NGST would be able to detect the earliest star bursts at very high redshift ($z > 10$) as well as the assembly of galaxies in the epoch $z \sim 1-5$. The compelling images ($0.4-0.9 \mu\text{m}$) of the Hubble Deep Field North (HDFN, Williams et al. 1996) have been complemented with NIR band NICMOS images and the HDFs (Williams et al. 1998). These images and the spectroscopic followups with ground-based telescopes (Steidel et al. 1996) have only strengthened the case for a large, space telescope optimized in the NIR. To provide additional guidance in the formulation phase of NGST (previously referred to as Phase A/B),

the Ad Hoc Science Working Group (ASWG) has compiled an extensive set of scientific studies, the Design Reference Mission (DRM), which represent the key science enabled by NGST (<http://www.ngst.stsci.edu/DRM.html>).

NASA, ESA, and CSA are collaborating on the NGST program. NASA has funded two studies (Pre-Phase A) by the TRW Corp and Ball Aerospace Corp. which have qualitatively established the architecture for NGST which is to be developed and launched in 2007-2008. ESA is also contributing to this early definition. All three agencies are supporting science instrument concept studies (Table 1). In summary, all viable architectures involve a large, deployable primary mirror (~ 8 m diameter) and an even larger sunshade to permit the entire telescope structure and instrument complement to cool passively to approximately 30° K. NGST will be launched into a halo orbit about the second Lagrange point, L2. We point the interested reader to recent reviews (Seery 1996) and upcoming monographs describing the NASA Yardstick architecture available on the NGST website, <http://www.ngst.nasa.gov/>. In this review, we discuss those aspects of NGST and the NGST science program that are enabled by and affect the development of 3-D scientific instruments.

Table 1. NASA, ESA, and CSA Instrument Studies

Instrument Concept Study	Principle Investigator	Institution
NGST Payload Study	Jill Bechtold	U. of Arizona, ARC
Imaging Fourier Trans. Spectrometer (IFTS)	James Graham	U. California
MOS w. MEMS Micro-Mir.	John MacKenty	STScI
MEMS Micro-shutter	Harvey Moseley	GSFC
MIR Camera/Spectrometer	Gene Serabyn	JPL
High Contrast Coronagraph	John Trauger	JPL
NIR Spectrometer Study	Olivier Le Fevre	LAS, Marseille
Optical Camera Study	Martin Ward	U. Leicester
PayloadSuite/OTA Study	LeFevre/Alcatel	LAS
NIR MOS/IFS	David Crampton	HIA/DAO
Visible Camera	Paul Hickson	U. British Columbia
IFTS	Simon Morris	HIA/DAO

2. NGST Unique Capabilities

The “core Dressler” program to study the high redshift universe foresaw the importance of capabilities unique to NGST: zodiacal-light-limited backgrounds (< 0.1 MJy ster $^{-1}$), superb, HST-like angular resolution in the NIR (0.06 arcsec) and relatively wide field of view (10–20 arcmin 2). Taken as a whole, NGST can reach a background-limited, high angular resolution image with a given signal to noise (S/N) 10^2 – 10^4 more quickly than an IR-optimized ground-based telescope with laser guide star adaptive optics. This speed advantage, which extends throughout the visible and NIR (0.6–2.2 μ m) is even greater in regions of strong, atmospheric absorption and beyond 2.5 μ m, where the thermal emission of ground-based telescopes and the atmosphere begins to rise precipitously. However, at moderate to high, spectral resolution, the effects of detector noise reduce this advantage. To better describe the regime in which NGST will be the pre-eminent observatory, we have defined the uniqueness space of NGST as a function of wavelength and spectral resolution (Stockman 1998). It is the

Figure 1. The NGST uniqueness space for wide field, medium resolution spectroscopy ($R=1000$). The maximum integration period is 10^5 s; the minimum signal to noise is 10:1. (STScI)

flux regime in which NGST is greater than 100 times faster than a ground-based 8-m telescope such as one of the Gemini or VLT telescopes. We consider the wide-field and single-target regimes separately and assume that 10^5 s is the longest typical accumulation period for data at a given spectral resolution and wavelength setting and a 10:1 S/N. Fig. 1 shows the uniqueness space for medium-resolution spectroscopy for the wide-field capabilities that would be enabled by 3-D technology. Note that we assume that high-angular resolution is required. If 0.4–0.5 arcsec natural seeing is adequate, the advantage of NGST is reduced significantly at spectral resolutions greater than $R > 100$ and in the visible and NIR ($\lambda < 2.5 \mu\text{m}$).

3. Key Science Goals

The DRM programs have five major themes: the structure and geometry of the Universe; the origins and evolution of galaxies; the Milky Way and its neighbors; the physics of star formation; and the origins and evolution of planetary systems. These themes address the higher goal of understanding the origin of the Milky Way, our Solar System, and the astronomical origins of life. In this section, we discuss three of these themes in more detail.

3.1. The Origins and Evolution of Galaxies

Since the HST & Beyond report, *HST* and large telescopes such as the Keck and the VLT have made substantial progress in detecting early galactic structures and what they may imply. The highest confirmed galactic redshift has increased from $z \sim 2$ for high redshift radio galaxies to $z = 5.6$ (Weymann et al. 1999). The UV luminosity density of high redshift galaxies can be interpreted in terms of star formation rates and, when it is adjusted for dust obscuration and selection effects, is remarkably flat in the epoch $z \sim 1-4$. (Steidel 1998). Indeed, the amount of star formation may be sufficient to maintain the ionization of the Universe at epochs prior to the peak contribution by AGN (Madau 1999).

The goals for HST and ground-based telescopes will be to detect galaxies at ever high redshifts and better characterize the number densities and morphologies of galaxies at redshifts, $z \sim 1$. The former goal will involve specialized surveys to find unusually red objects, $H - I > 4-5$. Such objects will be important to understanding when the assembly of spheroid structures is complete. The latter goal will involve moderate resolution NIR spectroscopy ($R \sim 5000$) to work between the OH lines and to measure the dynamics/mass of Milky Way galaxies when they were half their current age. These galaxies will be $H_{AB} \sim 21-23$, with a half-light radius of ~ 0.5 arcsec. NIR spectrometers using excellent natural seeing (0.4 arcsec FWHM) such as the VLT ISSAC (Moorwood et al. 1999) will be perfectly suited to such studies. Because of the OH forest, such studies will highlight galaxies and clusters with redshifts where the chief diagnostic lines may be cleanly separated from sky emission.

The NGST goals for this theme are:

- Detect the formation of the earliest stars, the earliest stellar structures. This involves extremely deep imaging and spectroscopy (10^6 s) to reach magnitudes of $H_{AB} \sim 34$ ($R \sim 5$) and $H_{AB} \sim 31$ ($R \sim 100$) over one or two $4 \text{ arcmin} \times 4 \text{ arcmin}$ fields and at red and NIR wavelengths. At the faintest magnitudes, the sky becomes increasingly confused as faint targets begin to overlap. Ultimately, breaking the confusion-limit will require additional spectroscopic information such as that provided by an integral field spectrograph (IFS) or imaging Fourier transform spectrometer (IFTS).

- Survey the numbers and morphologies of progenitors of galaxies such as the Milky Way between $1 < z < 5$. Obtain star formation rates, metallicities, and masses of the major components ($\sim 1-10\%$ the mass of the Milky Way). This will require the survey of sixteen $4 \text{ arcmin} \times 4 \text{ arcmin}$ fields in the visible-MIR ($0.6-10.0\mu$) and spectroscopic followups in low and moderate resolution ($R \sim 300-3000$, $H_{AB} \sim 28, 24$ respectively). 3-D techniques are important for this goal, since we expect 10-100 targets in this redshift and brightness range per square arcminute.

- Study the evolution of galaxies in the dense environments of groups and clusters ($1 < z < 5$) and surrounding AGN. Similar in observing strategy to the previous goal, these observations will compare these galaxies and progenitors to those in the field for tens of similar fields. Because these targets are selected because of strong clustering, their study is best done with 3-D techniques.

- Determine the number and morphologies of dust-enshrouded, star formation regions in Milky Way galaxies and their progenitors to $z \sim 2$. These regions emit much of their luminosity in the MIR from warm dust in the photo-

Figure 2. The sensitivity of a NGST Deep Field image. $R = 3$, 10^6 s. Spectra of young starbursts and older, established galactic progenitors are also shown. The cosmology adopted is pessimistic: $\Omega_m = 0.2$, $\Omega_\Lambda = 0.0$. (Lilly, U. Toronto)

dissociation regions (PDR) surrounding young star clusters. This goal calls for the study of SIRTf-selected sources as well as those found in the NGST survey fields. Although the number of detectable MIR sources per square arcminute may be large, $n > 100$, bright targets for spectroscopic followup will be relatively rare. However, if these sources are the results of mergers, spatially resolved spectroscopy with an IFS would be of great value.

In Fig. 2, we display the potential sensitivity of NGST for a NGST deep field exposure (10^6 s in a broad band). Also shown are predicted spectra of young starbursts (10^6 solar masses) and older, established progenitors, $0.01 L_*$, 10^8 solar masses and 10^8 yr. ($\Omega_m = 0.0$, $\Omega_\Lambda = 0.0$). With these sensitivities, NGST should detect many hundreds of very high redshift starbursts and Milky Way progenitors in every $4 \text{ min} \times 4 \text{ min}$ field. In the HDFN and HDFS NICMOS observations, we have already detected the population of galaxies whose spectra can be studied with NGST. These sources are faint, $H_{AB} \sim 29$, and are small with a half-light-radius of 0.1 arcsec. To provide discrimination between low redshift dwarf galaxies and high redshift progenitors, the wavelength range of NGST has been extended to $0.6 \mu\text{m}$. Galaxies with $z > 5$ will have their spectra heavily absorbed by the Lyman-alpha forest and will “drop-out” of the NGST R-band images. Galaxies at intermediate redshifts will be readily identified by their spectral energy distributions (SED) (Stiavelli et al. 1999).

For studies of high redshift galaxies, we must consider their apparent size and surface brightness when designing well-matched cameras and 3-D spectrographs. If we use the results of Gardner and Satyapal (1999) based on HST WFPC2 and NICMOS images of the HDF and similar work shown by Ferguson (private communication), we can model galaxies fainter than $H_{AB} > 26$ as having exponential profiles with half-light radii ~ 0.1 arcsec. If we assume that the camera and spectrographs are zodi-background limited, the effective area of these faint targets is equivalent to a top-hat with a 0.35 arcsec diameter. Alternatively, we can choose square apertures with 0.2 arcsec sides and analyze a little more than 50% of the the galaxy light. Using the same criteria as that used for point sources, we can obtain low resolution ($R \sim 100$), continuum spectroscopy to a target brightness of $H_{AB} \sim 28.2$ (25.3 mag arcsec $^{-2}$) and moderate resolution spectroscopy ($R \sim 3000$) to a target brightness of $H_{AB} \sim 24.6$ in the continuum (22.7 mags arcsec $^{-2}$). It is important to note that these surface brightnesses are comparable to that of the zodiacal light background, particularly for high resolution spectroscopy. Only relatively rare and bright galaxies, $H_{AB} \sim 23$ (10 arcmin $^{-2}$), can be spectroscopically resolved into individual 0.2 arcsec elements at the higher spectral resolutions. Emission-line spectroscopy, on the other hand, can resolve the dynamics and star-forming sites of fainter galaxies. This has significant implications for the design of 3-D spectrographs using MEMs or integral field image slicing technologies. Stiavelli (1998a,1998b) has studied this question for both spiral and elliptical galaxies.

3.2. Cosmology and the Chemical Evolution of the Universe

In the next decade, the Microwave Anisotropy Probe (MAP) and Planck will measure the cosmic microwave background on finer scales and to greater sensitivity than *COBE*. By comparing the results of these surveys to the predictions of big bang models, astronomers will constrain the cosmological parameters (H_0 , Ω_Λ , and Ω_m) and identify the most likely scenario for the evolution of structure after the recombination epoch. NGST will be ideal for confirming or constraining these results using weak-lensing surveys and by detecting Type Ia SNe at high redshift ($1 < z < 5$). In addition, NGST will detect Type II SNe to comparable redshifts and thereby provide an independent measure of massive star formation rates and the production of heavy elements at high redshift.

The goals for this theme are:

- Measure the mass of galaxies ($1 < z < 3$): Deep visible and NIR images of (8) 8 arcmin $\times$ 8 arcmin fields will provide statistical mass to light ratios for elliptical, spiral, and irregular galaxies as a function of cosmic time.
- Measure the mass of clusters and proto-clusters of galaxies: A similar observing strategy for 30–50 clusters identified by ground-based surveys, Chandra, and Planck.
- Observe the distribution of dark matter on scales of 1–10 Mpc at $z \sim 1$ –2. Shallow visible and NIR images of a square degree will provide sufficient statistics to determine the distribution of dark matter on scales smaller than those probed by the CMB satellites.
- Observe 50 Type Ia SNe and a similar number of Type II SNe to a redshift of $z \sim 5$. Low resolution spectra ($R \sim 100$ –300) are sufficient to confirm the SNe classification and provide an accurate redshift.

Figure 3. The effects of different cosmological parameters on the flux observed at maximum light for Type Ia supernova at cosmological redshifts. The two independent supernova programs find a trend that is consistent with a flat, Λ -dominated universe ($\Omega_m = 0.2$ and $\Omega_\Lambda = 0.8$; Riess et al. 1999, Perlmutter et al. 1999).

The mapping of dark matter requires wide band imaging. If these fields are obtained in a coordinated and periodic manner during the six months of continuous visibility, these data will also result in the discovery of hundreds of supernovae, mostly Type II at $z < 2$. However, the power of the Type Ia program is at redshifts $2 < z < 5$. Fig. 3 shows the expected magnitude differential expected from different cosmologies using $\Omega_\Lambda = 0.0, 0.2, 0.4, 0.8$ and a constant $\Omega_m = 0.2$. By observing the luminosity at maximum light at the higher redshifts, we can distinguish between straightforward cosmological models and those with third parameters such as unknown supernovae evolutionary effects or uniform cosmological extinction of unknown origin. Only NGST can observe SNe in two bands corresponding to B and V rest wavelengths once they are at redshifts $z > 1.5-2.0$. Two band observations distinguish between different decay and maximum light behaviors. If the NGST NIR spectrometer has a large field of view ($\text{FOV} > 1$ arcmin), it may be possible to obtain confirming spectra and redshifts of SNe as part of the larger high redshift galaxy survey. Graham (2000) has shown how a IFTS can obtain $R = 100$ spectra of a serendipitous SNe with $z \sim 3$ as part of a deep galaxy survey program.

3.3. Star and Planet Forming Regions

The NGST can utilize special capabilities in the NIR and MIR to study the physics of star and planet formation. The goal of these studies is to understand the process of star formation and the origin of the initial mass function. The DRM contains programs that measure the inflow and outflow rates of protostars, Class 0–1, by observing MIR emission line strengths and shape and by actually resolving the inner regions of the accretion disk. The Large Millimeter Array (LMA) will probe the cooler gas, farther from the central star. At $10\ \mu\text{m}$ wavelengths, the VLT Interferometer and the Terrestrial Planet Finder (TPF) will resolve the most central regions and perhaps the origins of the powerful jets.

As discovered by *IRAS*, many young stars display evidence of a remnant protostellar disk. In the MIR, NGST will be capable of detecting such disks about all stars in the Solar neighborhood. These ancient remnants of protostellar nebulae may well provide evidence (e.g. resonance gaps) of planet formation. With a graded Lyot stop in the focal plane, NGST would be capable of detecting the MIR emission from a Jupiter mass planet around stars out to 8 parsecs (a list of 50 candidates to be winnowed by the Space Interferometer Mission (SIM)).

4. 3-D NIR Spectrographs for NGST

The NASA, ESA, and CSA teams are studying various types of 3-D spectrographs for NGST. The Yardstick-type multi-object spectrograph (MOS) would feature a micro-mirror or micro-aperture focal plane array to feed a conventional 2-D spectrograph with adjustable spectral resolutions. The ESA study features two true 3-D image-slicer, grating/grism IFS: one with a large field of view and low spectral resolution and another with a small field of view and moderate spectral resolution, $R = 3000$. An alternative approach, an IFTS, is being studied by NASA and CSA teams. The IFTS is competitive or superior to a conventional camera for all spectral resolutions and could be used with available NIR detectors with little additional development. At spectral resolutions above $R > 100$, the grating or grism dispersive spectrometers are more efficient than a Fourier transform spectrograph but measure a smaller number of targets per field.

In the next year, the international agencies will assess the performance of these and other alternatives in regard to the science mission of NGST. Ultimately, they must reach agreement on the nature of contributions from each international partner based upon scientific priorities, technological readiness (e.g. risk) and cost. The choice of spectroscopic capability(s) will be a difficult but important one. We highlight the generic capabilities of each design in Table 2. Here we have assumed a common $4\text{ K} \times 4\text{ K}$ imaging array for each concept and 6 pixel resolution elements. Not shown are the technical challenges that each technology faces nor their operational complexities.

Based on the submission of the final reports in 1999 September and October and presentations at the NGST Science and Technology Exposition on 13-16 September, the ASWG will make its recommendation of one or more instrument complements to the NGST Project Scientist (John Mather) in late 1999. Although it is too soon to judge the outcome of that process, we can draw some conclusions from Table 2 and the descriptions of the key NGST sci-

Table 2. Near-Infrared Spectrometer/Camera Concepts for NGST. The maximum number of targets per field assumes full field use of a $4\text{ K} \times 4\text{ K}$ array at 6 pixels per resolution element and 10 % of the field for the MOS to avoid confusion.

3-D Technology	Field/Resolution	Spectral Resolution	Max. Targets/Field
MEM-enabled MOS	3 arcmin/ 0.2 arcsec	100-3000	1×10^3
IFS (Low Res)	1 arcmin/0.5 arcsec	300 fixed	1×10^4
IFS (Hi Res)	3 arcmin/0.06 arcsec	3000 fixed	1×10^3
Imaging FTS	3 arcmin/0.1 arcsec	3-300	4×10^6

ence programs. The MEM-enabled MOS spectrometers appear well-matched to a program of low resolution spectroscopy of high-redshift protogalaxies. These are moderately rare (100 arcmin^{-2}) and faint. Such a study benefits from the MOS wide field of view. However, the low resolution IFS may be ideal for the study of faint galaxies in clusters or surrounding an AGN. Its advantage lies in its ability to simultaneously obtain spectra from all sources in the field with no selection bias. For measuring dynamics, star formation, and metal abundances in single galaxies or mergers, the high spectral resolution IFS would be suitable for emission-line spectroscopy and would provide diffraction-limited angular resolution. Finally, the ability to resample the spectral information from an IFTS provides greater sensitivity and “discovery space” in extremely long imaging programs. Thus, the choice from among these concepts is not an easy one. Indeed, one possible NGST instrument complement would be comprised entirely of different 3-D instruments: an IFTS NIR camera, a NIR MOS spectrograph, and a MIR IFS.

We thank the ASWG for the definition of the NGST DRM, upon which many hard decisions will rest. We also thank the organizers of the 3-D conference for making clear that 3-D technology is finding increased and successful roles on ground-based facilities. With the advent of large format, high-performance infrared arrays, we anticipate at least one 3-D instrument to be an important part of the Next Generation Space Telescope. NGST studies are supported by NASA, CSA, and ESA and is a key component of the NASA HQ program, Searching for our Astrophysical Origins.

References

- Dressler, A. 1996, *Exploration and the Search for Origins: A Vision for Ultraviolet-Optical-Infrared Space Astronomy*, AURA,
 Gardner, J. & Satyapal, S. 1999, *ApJ*, in preparation
 Graham, J. 2000, this volume
 Madau, P. 1999, *VLT Opening Symposium*, (Springer-Verlag) in press
 Moorwood, A. et al. 1999, *VLT Opening Symposium*, (Springer-Verlag) in press
 Perlmutter, S. et al. 1999, *ApJ*, in press, astro-ph/9812133
 Riess, A. G. 1999, *ApJ*, in press, astro-ph/9805201
 Seery, B. 1996, <http://www.ngst.nasa.gov/cgi-bin/pubdownload?id=21>
 Steidel, C. C. et al. 1996, *ApJ*, 462, L17

- Steidel, C. C. et al. 1999, ApJ, in press, astro-ph/9811399
Stiavelli, M. 1998a, Science with NGST, ASP Conf. Ser. 133, 279
Stiavelli, M. 1998b, 34th Liege Inter. Astrophys. Colloq., ESA, 71
Stiavelli, M. et al. 1999, A&A, 343, L25
Stockman, H. S. 1994, High-Z, STScI
Stockman, H. S. 1998, <http://www.ngst.nasa.gov/cgi-bin/pubdownload?Id=323>
Thronson, H. A. 1993, EDISON, Didcot: Rutherford Appleton Labs
Wade, L. A. 1995, MIRORS, JPL
Weymann, R. et al.(1999), ApJ, in press, astro-ph/9807208
Williams, R., et al. 1996, AJ, 112, 1335
Williams, R., et al. 1998, BAAS, 193, 7501

An Imaging Fourier Transform Spectrometer for the Next Generation Space Telescope

James. R. Graham

Department of Astronomy, University of California, Berkeley, CA 94720

Abstract.

Due to its simultaneous deep imaging and integral field spectroscopic capability, an Imaging Fourier Transform Spectrograph (IFTS) is ideally suited to the Next Generation Space Telescope (NGST) mission, and offers opportunities for tremendous scientific return in many fields of astrophysical inquiry. We describe the operation and quantify the advantages of an IFTS for space applications. The conceptual design of the Integral Field Infrared Spectrograph (IFIRS) is a wide field ($5'.3 \times 5'.3$) four-port imaging Michelson interferometer.

1. Introduction

When the Next Generation Space Telescope (NGST) begins observing towards the end of the first decade of the next millennium it will usher in a new era of infrared (IR) astronomy. The combination of a deployable 8-meter aperture and an L2 orbit, which will allow the telescope to cool to 30-50K, will enable zodiacal-light limited performance for $\lambda < 10 \mu\text{m}$.

The rationale for placing NGST far from Earth and shielded from the Sun is reduction of the background by factors of up to 10^6 compared to terrestrial environments. This will enable studies of the origins of the structure in the universe, galaxies and quasars, stars, and planetary disks — core mission objectives of the NASA Origins Program. Many of these programs require IR observations: the light from distant galaxies seen at early epochs is redshifted by the expansion of the universe; regions of star formation in our galaxy are obscured by dust in the visible, but are penetrated by infrared radiation; and cool objects such as forming stars and proto-planetary disks emit in the IR.

2. What is IFIRS?

The goal of IFIRS is to obtain ultra-deep, wide field, diffraction limited imagery from near- to mid-IR wavelengths, with flexible spectral resolution (Graham et al. 1998). IFIRS is a Michelson interferometer configured as an imaging Fourier transform spectrometer (IFTS) (See Fig. 1) (Bennett 1993). An interferogram is recorded for every pixel in the field of view, and hence a spectrum can be obtained for every object. Since IFIRS is an FTS, it is both a high-throughput camera and a multi-object or integral field spectrometer. As a camera IFIRS has unusual flexibility of spectral resolution, and a unique pan-chromatic imaging

Figure 1. **Left:** An IFTS is a Michelson interferometer where the telescope focal plane is imaged onto a detector array. An interferogram is recorded for every pixel, and hence a spectrum can be obtained for every object. **Right:** IFIRS has a four-port design that wastes none of the light. Instead of flat mirrors, IFIRS uses cube-corners, which displace the input and output beams.

mode. IFIRS is a four-port interferometer (Fig. 1). One of the disadvantages of a classical flat-mirror Michelson is that the beam-splitter reflects 50% of the light back to the source; a four-port design collects all the light.

2.1. Design of IFIRS

On the object side a collimator illuminates the interferometer with parallel light. The interfering beams are collected by a camera, creating a one-to-one mapping between points in the object and image planes. By placing a detector focal plane array (FPA) at the focus of the camera, each pixel is matched to a single point on the sky. At any given optical path difference (OPD) the image of the sky is modulated spatially by the interferometer's fringe pattern, which encodes the spectral information. By recording images of the sky at different OPDs, the spectrum of each pixel can be reconstructed. The OPD is scanned in discrete steps since FPAs are integrating detectors. The time series from each pixel forms an independent interferogram for every point on the sky within the field of view. These interferograms are Fourier transformed individually yielding a spectral data cube composed of the same spatial elements as the image. The sampling theorem establishes the number and amplitude of OPD steps necessary to recover the spectrum at a given spectral resolution.

The major components of IFIRS, shown in Fig. 2, are the collimator, interferometer, cameras, and FPAs and associated electronics. Two important ancillary subsystems, not shown are the metrology system and the calibration unit.

The collimator is a three mirror anastigmat (TMA) that illuminates a four-port Michelson. There are two input and two output ports. One input port is fed with the sky signal, and the other input can be illuminated by the calibration unit. The interferometer consists of a 50:50 reflecting/transmitting beam splitter, and two cube-corner retroreflectors. The appropriate beam splitter is

Figure 2. The optical layout of IFIRS. For simplicity only the near-IR FPA have been shown at the output ports. The final fold mirror is either a dichroic or flip-mirror which directs the mid-IR light to the mid-IR detectors. For clarity reimager # 2 has been rotated by 180 degrees. The actual design is symmetric about the beam-splitter.

selected using a filter-wheel mechanism. One cube-corner is located on a translation stage, which permits precision control of the OPD. The OPD is monitored using an interferometric metrology system. Each output port of the interferometer feeds identical TMA cameras. A final reflection (an articulated fold mirror or a static dichroic) directs short wavelength radiation to the near-IR FPA and long wavelength radiation to the mid-IR FPA. In summary, the properties of IFIRS are:

The absolute wavelength calibration is provided by the interferometer metrology system, which is referenced to a diode laser. The accuracy should be better than 7.5 GHz (0.25 wavenumbers). This provides the wavelength calibration for both pure and dispersed FTS modes.

Table 1. Observational Capabilities

	NIR Channel	MIR Channel
Bandpass/Detector	0.6-5.6 μm /InSb	5-15 μm /HgCdTe
Maximum Spectral Resolution	1 cm^{-1}	1 cm^{-1}
FOV/Array Format	5.'28/8k $\times$ 8k	2.'64/2k $\times$ 2k
Pixel size/Nyquist λ	0."0386/3 μm	0."0772/6 μm
Wave front error/Strehl	150 nm rms/0.8	150 nm rms/0.9
Throughput	> 0.7	> 0.6
Sensitivity ^a for $R^b = 1/5/100$	0.2/1/35 nJy	13/65/1300 nJy

^a SNR = 10 for a 10^5 s integration. All spectral channels are obtained simultaneously.

^b R is the number of simultaneous spectral channels in the band-pass.

3. Unique Attributes of an IFTS

The signal-to-noise (SNR) performance of 3-d imaging spectrometers equipped with 2-d FPAs is the same for all architectures, in the ideal case of photon shot noise limited operation (Bennett 1995, 2000). This is correct so long as the spectrometers are equipped with the same size FPAs, and the same spatial and spectral degrees of freedom of the astronomical scene are observed.

An IFTS measures the spectrum of every pixel on the field of view and hence provides the only efficient means of conducting *unbiased* spectroscopic surveys of the high- z universe, i.e., without object preselection (e.g., using broad band colors) and without the restrictions imposed by spectrometer slit geometry and placement. Because there is no slit, an IFTS can record spectra for adjacent individual objects, down to the confusion limit (e.g., resolved stellar populations), and for moving targets (e.g., Kuiper Belt objects).

An IFTS also allows spectroscopy over a wide bandpass, typically three octaves, and affords flexibility in choice of resolution ($R = 1\text{--}10^4$). IFIRS uses all reflecting optics, and high-efficiency beam-splitters. The throughput is high and flat across the band-pass. For the near-IR channel the efficiency is approximately 75%, including optics, beam splitter, and detector QE (see Fig 3).

The IFTS has the brightest imagery of any imaging spectrometer design because the full band-pass is transmitted to the FPA. The summed signal from the two output ports is unmodulated and corresponds to the total broad-band photon flux entering the instrument. This pan-chromatic image is peculiar to a four port interferometer. Thus deep, broad-band imaging is acquired simultaneously with higher spectral resolution data over a broad wavelength range. A pan-chromatic image can be formed with a filter-wheel camera by summing the sequence of filter images, but the IFTS pan-chromatic image has a speed advantage factor that is equal to the number of filters used. The IFTS advantages due to broad-band operation are high SNR pan-chromatic imaging for science and also for telescope guiding, tolerance of cosmic rays hits, detector noise, and internal instrument background, and high SNR determination of flat-fields and detector non-linearity.

A four port Michelson interferometer is intrinsically a superb instrument from a calibration point of view. The measured interferogram results from the difference between spectra of sources at the two input ports. Calibration sources

Figure 3. The throughput of IFIRS for the near- and mid-IR channels.

placed at the second input port act as transfer standards for full radiometric calibrations performed on the ground prior to flight. The mid-IR channel is calibrated by varying the temperature of a cold blackbody at the second input that fills the field of view. The near-IR channel is calibrated using a dilute, hot blackbody, that does not produce excessive heating of the instrument. In both cases, a full calibration of each FPA pixel's offset, spectral responsivity, and non-linearity can be conducted by varying integration times or source intensities. This capability is invaluable should the system response change due to exposure to the space environment.

IFIRS also has a hybrid, or dispersed FTS mode. In a regular FTS, spectral information is encoded in the z -direction of the data cube, and there is no mixing of spectral and spatial information. The advantage is that a spectrum is recorded for every pixel on the sky. The penalty is that the photon shot-noise from all spectral channels is present at each frequency. This shot-noise can be reduced by masking the telescope focal plane around objects of interest with a programmable focal plane mask and inserting a prism into the collimated space. The dispersed FTS mode is used to obtain the highest possible sensitivity at high spectral resolution ($R=600$ -10,000) (Bennett, 2000). The slit width does not determine the spectral resolution in the dispersed FTS mode, since the spectral resolution is derived from the interferograms. The dispersed FTS data-cube contains spectra which are tilted with respect to the z -axis. The tilt angle is the arc tangent of the ratio of spectral resolutions of the dispersive element and interferometer. The dispersed FTS has better SNR performance than the pure FTS, and the source density of object slits is higher than for the pure multi-object spectrometer.

4. Performance of IFIRS on NGST

IFIRS records a spectrum for every pixel in the field of view. This means that data produced by IFIRS is three-dimensional, i.e., it forms a data cube consisting of two spatial dimensions (x, y) and one wavelength dimension (z). A horizontal (x, y) slice through a data cube corresponds to a monochromatic image. A line extracted from the data cube in the z -direction corresponds to a spectrum.

To estimate the performance of an IFTS in the NGST environment we have simulated data as follow. An astronomical scene is represented as a noise-free distribution of objects. This input data cube is convolved with the telescope point spread function (PSF). This is a 3-d convolution, since the PSF depends on wavelength. The spectra are then multiplied by the wavelength dependent throughput. From this spectral data cube the interferogram is calculated by a Fourier transform. At each OPD step noise is added. The noise sources treated are photon shot noise from the zodiacal light, photon shot noise due to thermal emission from the telescope, photon shot noise from the target, shot noise due to detector dark current, and detector read-noise. The noisy interferogram cube is then Fourier transformed back into a spectral data cube.

We have used Im & Stockman's (1998) simulation of a blank field at very low flux levels to demonstrate the spatial multiplex advantage of IFIRS. In this simulation we show only 0.02% of the field of view of IFIRS. If we had simulated

IM & STOCKMAN NGST DEEP FIELD SIMULATION

Figure 4. IFIRS data cube for the Im & Stockman (1998) NGST deep field.

the entire IFIRS data cube it would contain spectra for tens of thousands of galaxies.

The spectral energy distributions omit nebular emission lines, so redshifts must be deduced from stellar absorption features and spectral breaks. Simulations of star forming galaxies show that a cluster forming stars at $2.5 M_{\odot} \text{ yr}^{-1}$ for 10^7 yr can be detected in $\text{Ly}\alpha$ emission to $z = 12$. In this example objects with $K = 31.3$ AB mags are detected with $\text{SNR} = 5$ per spectral resolution element, which is sufficient SNR to identify the Lyman break in objects as distant as $z = 12$.

I am indebted to the IFIRS team for their contributions to this report: M. Abrams, C. Bennett, J. Carr, K. Cook, A. Dey, R. Hertel, N. Macoy, S. Morris, J. Najita, A. Villemaire, E. Wishnow, & R. Wurtz. This work was support by NASA.

References

- Bennett, C. L., Carter, M. R., Fields, D. J. & Hernandez, J. A. 1993, Proc. SPIE, 1937, 191
- Bennett, C. L., Carter, M. R. & Fields, D. J. 1995, Proc. SPIE, 2552, 274
- Bennett, C. L. 2000, this volume.
- Graham, J. R., et al. 1998, PASP, 110, 1205
- Im, M. & Stockman, H. S. 1998, ASP Conf. Ser., 133: Science With The NGST, 263

IFMOS: Integral Field Multi Object Spectrograph for the NGST

O. Le Fèvre¹, E. Prieto¹, J. Allington-Smith², R. Bacon³, R. Content²,
S. Cristiani⁴, R. Davies², B. Delabre⁴, R. Ellis⁵, G. Monnet⁴, E.
Pécontal³, W. Posselt⁶, N. Thatte⁷, P.T. de Zeeuw⁸, P. van der Werf⁸

(1) *Laboratoire d'Astronomie Spatiale, Marseille, France*, (2) *Department of Physics, University of Durham, UK*, (3) *Observatoire de Lyon, France*, (4) *European Southern Observatory, Garching, Germany*, (5) *Institute of Astronomy, Cambridge, UK*, (6) *Dornier Satelliten Systeme GmbH, Ottobrun, Germany*, (7) *Max Planck Institute für Extraterrestrische Physik, Garching, Germany*, (8) *Leiden Sterrewacht, Leiden, The Netherlands*

Abstract. Conducting resolved spectroscopy on large samples of very faint objects is a strong requirement for the Next Generation Space Telescope (NGST). We are presenting the IFMOS concept, a 1 to 5 microns integral field spectrograph developed under contract with ESA. Image slicers are feeding compact spectrograph modules with very few moving parts. The wide field survey mode has low spatial – low spectral resolution over a field $\sim 46 \times 40$ arcsec² (0.19 arcsec per resolved element). The high spatial – high spectral resolution mode covers a field 3.8×2.6 arcsec² sampled at 0.05 arcsec, a pointed object mode at the diffraction limit of the telescope. A detailed opto-mechanical design has been produced, meeting the environmental requirements for the NGST payload.

1. Introduction

This study is being conducted by our european consortium of scientific institutes and industry, under contract with ESA.

The goal of this pre-phase A study is to produce a design of an integral field / multi object spectrograph in the 1-5 microns domain for the NGST. The study Phase 1 started in September 1998 to define the science goals vs. the Design Reference Mission established by the NGST project, set up the instrument performance requirements, and conduct a design trade-off to identify the most suitable design. Phase 2, ending in June 1999, is concentrating in the detailed opto-mechanical, thermal, and electrical design. We present here the current status of the instrument design.

2. Science Goals

The 1-5 microns spectrograph will be a workhorse instrument for the NGST. The main programs to be carried out have been identified in the NASA Design Refer-

ence Mission. As such, it is required that the spectrograph offer a broad range of capabilities. We have defined two important areas: low spectral resolution over a wide field to conduct survey type observations, and moderate to high spectral resolution to conduct detailed 2D investigations of selected targets. These two modes cover all the requirements set by the DRM for 1-5 microns spectroscopy. This is in line with recent developments in instrumentation choices for ground based 8m telescopes involving both wide field multi-object capabilities and 3D spectroscopic abilities (e.g. Le Fèvre, 1998; Davies et al., 1996; Thatte et al., 2000).

3. Design Selection

We have conducted a trade-off analysis between several possible concepts including multi-slit spectroscopy, integral field (3D) spectroscopy, and Fourier transform spectroscopy. The comparison between these concepts is presented in Table 1. We have concluded that an integral field concept with wide field and narrow field capabilities was the most efficient design for the NGST, presented limited technological development requirements and was the simplest and most robust to operate.

Table 1. Comparison between possible concepts for multi-object spectroscopy with NGST

Item	Multi-slit	Integral Field	FTS
Field	3'x3'	1'x1'	3'x3'
Spectral Resolution	100-3000	100-3000	5-5000
Spatial/slit sampling	0.1"	0.05 to 0.2"	0.03"
Target selection	need images to pre-select	all targets	all targets
Measure 2D fields	not possible	Yes	Yes
Geometry or time constraints	slits and spectra overlap	None	time scan
Slit losses	slit width smaller than object	None	None
Moving parts	micro-mirrors or moving slits, gratings, filters	None	continuous motion
Target acquisition	slits to 0.01"	pointing to 0.1"	pointing to 0.1"

4. Concept baseline

4.1. Overall concept

The concept is based on a modular approach, with simultaneous low resolution and high resolution observations. Compact dedicated modules include the

Figure 1. (left) IFMOS field of view, with simultaneous wide field and narrow field, (right) Overall optical layout, the fore-optics, slicers, one of 3 low resolution spectrograph modules and the high resolution spectrograph are indicated

Figure 2. Opto-mechanical layout of one complete LR spectrograph and slicer module. The HR spectrograph and slicer feed is shown on the side.

slicer optics, the 3 low resolution spectrograph modules and the high resolution spectrographs. The field of view ranges from $3.8 \times 2.6 \text{ arcsec}^2$ in the high spatial/spectral resolution mode, to $40 \times 46 \text{ arcsec}^2$ in the low spatial/spectral resolution mode, with spectral resolutions from $R \sim 150$ to $R \sim 3000$. In the low spectral resolution mode, the simultaneous wavelength coverage spans 2 octaves from 1.25 to 5 microns. In the high spectral resolution mode, one of 2 octaves can be selected, either 1.25-2.5 or 2.5-5 microns.

The overall concept is presented in Figures 1 and 2. The overall dimensions are 1360mm in length, 1300mm in width and 1050mm in height. Moving parts are kept to a minimum: the low resolution mode has only one motion to allow for refocusing of the blue camera, while the high resolution spectrograph requires only the grating exchange mechanism. This would allow robust and reliable operations.

The capabilities of the spectrograph are summarized in Table 2.

Table 2. IFMOS capabilities

Item	Low resolution mode	High resolution mode
Field of view	$40 \times 46 \text{ arcsec}^2$	$3.8 \times 2.6 \text{ arcsec}^2$
Spatial sampling	0.19 arcsec/pix	0.05 arcsec/pix
Spectral resolution	~ 150	~ 3000
Simultaneous spectral sampling	1.25 to 5 μm	1.25 to 2.5 μm or 2.5 to 5 μm
Throughput	~ 0.6	~ 0.6

4.2. Image slicer

The image slicer concept includes fore-optics to match the image scale to the focal plane and several slicers to re-format the input 2D field onto the 1D spectrograph slits.

Each low resolution spectrograph module has 2 slicer mirrors with 72 slits (54x0.88 mm) each leading to 9 slit units of 16 pupil mirrors (Fig.3). The high resolution spectrograph has one slicer unit with 52 slit (also 54x0.88 mm each) and pupil mirrors. Each slit mirror / pupil mirror couple has to be designed independantly.

See Content (2000) for more details.

4.3. Low resolution spectrograph modules

The low resolution mode of the spectrograph offers a field of view of $40 \times 46 \text{ arcsec}^2$, with a sampling of 0.19 arcsec per sky pixel. The image slicer produces 432 pseudo slits. Three spectrograph modules are taking each 9x16 slits as input. Each spectrograph module is made of a three mirror anastigmat collimator (TMA), a red prism with a dichroic coating dispersing the beam into a TMA red camera, with a reflected dichroic beam going through a blue prism and a TMA blue camera. The field of view could be increased to $66 \times 46 \text{ arcsec}^2$ with 2 more LR modules.

The spectrograph design allows to take the input F/9.6xF/6 beam with the camera producing a F/4xF/2.5 beam going to one 2048^2 18 microns pixel detector. Prisms allow for a spectral resolution ~ 150 without having to deal with multiple orders.

4.4. High resolution spectrograph

The high resolution spectrograph has a field of view of $3.8 \times 2.6 \text{ arcsec}^2$, with a sampling of 0.05 arcsecond per sky pixel. The image slicer feed the spectrograph with 52 pseudo slits. The spectrograph camera produces a F/4xF/9 beam going to a 4096^2 mosaic of 4 detectors each 2048^2 pixels with 18 microns pitch.

5. Performances

The IFMOS offers high throughput performances comparable to other types of spectrograph. Performances on point sources are listed in Table 3.

Figure 3. Low resolution image slicer concept

The overall efficiency is the actual time required to complete a given program. It has to take into account all parameters, including throughput, diffraction and slit losses, multiplex gain, geometric limitations, instrument set-up time, flat fielding accuracy etc., and depends on the type (projected density, size) of targets observed.

Integral field spectrographs are at a great advantage because of:

- No slit losses: spectra can be optimally extracted from the 3D data cube selecting all pixels over the full object extent. This effectively produces the equivalent of a throughput or exposure time gain when compared to classical multi-slit (for which slits cannot be broadened because of degraded spectral resolution).
- No geometric limitations: the spectra of two very closeby objects can be recorded simultaneously, only limited by the image sampling at 0.19 arcsec per resolved element on the sky. There is no restriction on objects that can be observed imposed by the position and extent of spectra of other selected objects.
- Limited diffraction losses and hence low instrument background.
- No need to obtain images to pre-select targets for spectroscopy. 3D techniques have great potential for serendipitous discoveries. A source with little continuum but strong emission lines may escape apriori broad band imaging, while it would be detected from 3D spectroscopy. Moreover, a 3D spectrograph is an imager, since spectra at any wavelength and at a resolution below the nominal resolution can be easily computed from the 3D data-set.
- Limited instrument pointing time, 10 times less stringent than on-target slit positioning.
- Very accurate flat fielding: a wide field 3D spectrograph such as the IFMOS-LR mode can be used in shift-and-add mode similarly to what is being done in imaging.

As an example, we consider a deep survey of 4000 galaxies at the limiting magnitude $AB = 28.5$, requiring a galaxy sampling as close as possible to 100%. About 13000 seconds are required to get $S/N=5$ for point sources. The density of sources is $\sim 10^6$ galaxies/deg², or ~ 150 galaxies in the IFMOS-LR field of view. IFMOS would thus require about 26 pointings to get redshifts of all galaxies. A multi-slit spectrograph with a field 3×3 arcmin² will see ~ 2500

galaxies to the limiting depth, but geometric constraints limit the multiplex gain. With typical galaxy sizes $r_{1/2} \sim 0.1 - 0.15$ arcsec (Thompson et al., 1999), the slit length has to be ~ 3 arcsec in order to maximize light gathering efficiency and get a local sky and instrument background measurement, which implies a maximum of ~ 500 galaxies observed simultaneously. However, simulations show that to observe up to 80% of these sources, and allow for e.g. close pairs measurements, it is required to observe the same field about 3 to 5 times. About 30 pointings are thus required to gather the galaxy sample, *comparable to the time required by the IFMOS-LR*. This is without taking into account the slit losses in the multi-slit configuration, the required time for preliminary imaging, and the longer telescope set-up time to make sure that slits are on-target.

Another pilot program in the DRM is to obtain resolved spectroscopy of a sample of galaxies to measure their velocity fields, and e.g. virial masses. The IFMOS-HR mode allows to do this with great efficiency for one galaxy at a time, while a multi-slit instrument cannot reasonably tackle this complex observing requirement (see e.g. Bacon et al. 2000), and an FTS needs a long time scan to reach the needed resolution over a significant wavelength range.

Table 3. IFMOS performances (point source, S/N=5, 1h)

Mode	λ	AB mag	object	sky (%)	dark (%)	readout (%)	N_{pix}
LR1	2.2	27.4	20	62	12	4	1
LR2	3.6	27.8	22	58	14	4	1
HR1	2.2	24.6	19	0	51	17	2
HR2	3.6	24.2	15	0	62	21	9

6. Summary

The IFMOS concept developed by our consortium is a powerful approach to wide field spectroscopy surveys and spatially resolved 3D spectroscopy in the 1-5 microns domain with the NGST. Comparison with other designs proposed for the NGST require to take into account all the parameters that make up the overall observing efficiency. We believe that the IFMOS concept is over-performing other existing concepts, and yet retains a simple operating scheme with few moving parts and easy set-up.

References

- Bacon, R., et al., 2000, this volume
- Content, R., et al., 2000, this volume
- Davies, R., et al., 1996, proc. SPIE 2871, 1099
- Le Fèvre, O., et al., 1998, proc. SPIE 3355, 8
- Thatte, N., et al., 2000, this volume
- Thompson, et al., 1999, AJ, 117, 17

The Tunable Filter Program for NGST

S. Satyapal, M. A. Greenhouse, R. Barclay, D. Amato, B. Arritt, R. Barry, C. Holt, and J. Kuhn

NASA Goddard Space Flight Center, Greenbelt, MD 20771

T. Higelman and N. Fonneland

Northrop Grumman Corporation, Bethpage, NY

L. Lesnya

Lockheed Martin Corporation

Abstract. We display work in progress on a near-infrared tunable band-pass filter for the wide field camera of the Next Generation Space Telescope (NGST) Integrated Science Instrument Module (ISIM). This Demonstration Unit for Low Order Cryogenic Etalon (DULCE), is designed to demonstrate a high efficiency scanning Fabry-Perot etalon operating in interference orders 1-4 at 30 K with a high stability DSP based servo control system. DULCE is currently the only available tunable filter for low order cryogenic operation in the near-infrared. In this application, scanning etalons will illuminate the focal plane arrays with a single order of interference to enable wide field low resolution ($50 < R < 200$) hyper-spectral imaging over a wide range of redshifts. We discuss the specific science applications in which tunable filters are the instrument of choice.

1. Introduction

The Next Generation Space Telescope will gather unprecedented spectroscopic and photometric data on galaxies out to the highest redshifts. The anticipated number counts in the NGST field of view is large, expected to be close to 10,000 at AB=30 mag (Gardner & Satyapal 1999), requiring an efficient method for 3D imaging. There are several possible instrument concepts for carrying out imaging spectroscopy over a large field of view. These include imaging Fourier Transform Spectrometers (IFTS; Graham et al. 2000), dispersive-based multi-object spectrometers (Mackenty et al. 2000), image slicers (LeFevre et al. 2000), and tunable filters or Fabry-Perots. A critical comparison of these approaches is given by Bennett et al. in this volume. For the specific application of carrying out imaging spectroscopy for a narrow range of redshifts over the entire NGST field of view, or for scientific applications requiring the observation of only several spectral lines in extended sources, Fabry-Perots are the most effective instrument choice. In addition, tunable narrow band filters can easily be combined with other spectrographs in a single instrument package to extend sensitivity, field

of view, or resolution capabilities. For example, when combined with a higher spectral resolution IFTS, narrow-band tunable filters can help improve overall instrument sensitivity when full band wavelength coverage is not needed. Fabry-Perots are extremely compact, are based on relatively mature technology, and can easily be accommodated at a large pupil location or, in some cases, at a field position. They are therefore a powerful and versatile potential instrument component in future observatories such as NGST.

2. CHALLENGES IN DEVELOPMENT OF LOW RESOLUTION INFRARED FABRY-PEROTS

A Fabry-Perot (FP) interferometer or etalon consists of two flat, parallel, semi-transparent plates coated with films of high reflectivity and low absorption. The pass band of the etalon is determined by the separation between the plates which is generally varied using piezoelectric translators (PZT). Fabry-Perots have been used extensively for the last several decades over a wide range of wavelengths. However, these etalons have been traditionally operated at room temperature and at moderate to high resolving powers ($R > 500$). Options for low resolution variable interference filters have typically been restricted to circular variable filters or linear variable filters. These options require an optical design that includes a small (1-2 mm diameter) pupil. With the advent of large aperture wide field telescopes, incorporating a small pupil in the optics design while retaining good image quality and spectral purity over the desired field of view (FOV) is extremely difficult. In such situations, low resolution FPs are an ideal option for narrow-band imaging. They ease size requirements on filter wheels and offer flexibility in choice of spectral resolution.

At infrared wavelengths, etalons must be operated at cryogenic temperatures in order to suppress the background radiation they emit. At these temperatures, PZTs alone do not provide sufficient translation to tune the etalon over the desired orders of interference without becoming large and cumbersome. In addition, low resolution infrared etalons require cavity spacings on the order of a micron. A PZT driven low resolution etalon assembled at room temperature and then operated cryogenically requires a mechanical design that achieves a high degree of thermal expansion compensation to achieve the desired sub micron gap within the limited cryogenic scan range of the actuator. As a result of these technical challenges, no commercial cryogenic low order etalon existed prior to the development of DULCE.

3. NEAR-INFRARED TUNABLE FILTERS FOR NGST

We are currently constructing near-infrared tunable bandpass filters for the baseline wide field camera of the Next Generation Space Telescope (NGST) Integrated Science Instrument Module (ISIM). This Demonstration Unit for Low Order Cryogenic Etalon (DULCE), is designed to demonstrate a high efficiency scanning Fabry-Perot etalon operating in interference orders 1-4 at 30 K with a high stability DSP-based servo control system. The compact design and current aperture goal allows the etalon to be placed at the pupil location in the near-infrared camera of the GSFC Yardstick ISIM. The near-infrared camera in

this ISIM (Greenhouse et al. 1998) has a $4' \times 4'$ field of view that is divided by a pyramid mirror into four identical camera modules that each cover a $2' \times 2'$ FOV. In Figure 1, we display one such module showing the location of tunable and fixed filters. The module contains an Offner relay that produces a 40 mm pupil at the location of the etalon. The filter wheel contains broadband filters for photometric imaging or order sorting the $R=50-200$ tunable filters when they are rotated into the beam. The range of angles passing through the 40 mm aperture etalon is small enough that the shift in wavelength across the array is less than a resolution element.

Figure 1. Each camera channel contains a focussing Offner relay, a filter wheel, and a retractable tunable filter.

4. MECHANICAL DESIGN AND OPTICS CHARACTERISTICS

In Figure 2, we display the first iteration of DULCE currently assembled. Here the 'fixed' ZnSe plate is mounted via three phenolic tabs, which are bonded to the plate and clamped to the aluminum structure with a belvil washer. Relative radial motion between the plate and aluminum is allowed due to a light preload from the washer. The 'movable' plate rests on three diamond turned pads which are machined into the mount. The back of the plate is clamped to the mount via a circular spring which is manufactured to apply pressure only to three points

directly opposite the diamond-turned pads, to prevent warping of the optics surface. Local stresses are minimized by using a ball-bearing between the spring and the aluminum pad that contacts the plate. This mechanism ensures that clamping stresses are always distributed over the whole pad area. Finite element analysis shows that this design allows for secure mounting with negligible optical surface deformation. As can be seen from the figure, the movable plate is chosen to have a larger diameter than the fixed plate, allowing the optical surfaces to achieve the desired sub micron separation without interference from the mounts.

Figure 2. Assembly of DULCE I

The current design allows for mechanical adjustment of the cavity spacing and parallelism. The fixed plate is mounted in an aluminum ring which is threaded on the outer edge to provide coarse adjustment of the plate separation with a 10-20 μm resolution. Parallelism adjustment is accomplished through differential screw assemblies located beneath each actuator (not shown in Figure 2). DULCE I utilizes state of the art mechanically amplified PZT actuators, shown in Figure 3. These actuators, manufactured by Garman Industries, are small (less than one inch length) PZT stacks in an aluminum frame which provides approximately a factor of 6 in amplification as well as preloading of the stack. Therefore, although there is a 75 percent reduction in PZT stroke at cryogenic temperatures, the mechanically amplified stroke at 40 K is 30 μm , approximately 15 orders of interference, with a maximum drive voltage on the PZT of 150 volts. The cryogenic strain and contraction upon cooling of each

actuator was measured using eddy current sensors. The contraction from room temperature to 77 K was measured to be approximately 500 μm .

Figure 3. Magnified expansion low voltage PZT actuator selected for DULCE I yields 30 μm of strain at 30 K.

Gold deposition pads are located on the ZnSe plate surfaces for capacitive sensing of the cavity spacing and plate parallelism. The Mrk I prototype etalon has multilayer dielectric coatings over a 25 mm diameter aperture. These coatings are designed to minimize phase dispersion over the 3-5 μm wavelength range. For low order etalons, the thickness of the coating is comparable to the separation of the plates. Since there is a phase shift upon reflection that is dependent on wavelength, the tuning characteristics of the etalon will not generally be constant with wavelength. In addition, the free spectral range without phase control will be too small, resulting in a greater number of required order sorting filters. NGC has developed low phase dispersion coatings for the 3-5 μm wavelength band which approach the properties of ideal metal reflectors.

5. ELECTRONICS

The electro-mechanical block diagram for DULCE is shown in Figure 4. The FPI tuning frequency will drift slowly over time unless active stabilization of the mirror separation is employed. Cavity stability is required to be within one-tenth of a resolution element, or 3 nm. The capacitance between the pads deposited on the edges of both mirrors changes as the cavity spacing varies. This capacitance can be read out, and used by the controller to generate an error signal that can be fed to the PZT actuators to keep constant plate separation and parallelism. The controller uses a Digital Signal Processor, DSP, for flexibility. Control schemes can easily be modified by changing software. 16-bit analog-to-digital and digital-to-analog converters are used for the sensor and actuator

driver interfaces, respectively. This approach provides for a sensing resolution of about 0.5 nm over the 30 μm stroke. Capacitive sensors were chosen due to their low power dissipation.

Figure 4. Electro-mechanical block diagram for DULCE I

6. CURRENT STATUS AND FUTURE DIRECTIONS

DULCE I is currently assembled and is undergoing laboratory tests. The entire module is only 3.5 inches in outer diameter for a 1 inch clear aperture etalon, and design for an etalon with a larger clear aperture (40 mm) and smaller overall packaging (80 mm) is underway. This Mrk II design will meet NGST flight dynamics requirements and will utilize Si rather than ZnSe plates. We are exploring other options for cryogenic actuators which will provide even greater strokes at 40 K, including superconducting magnetostrictive actuators which provide a factor of 7 increase in cryogenic strain relative to our current actuators.

References

- Bennett et al. 2000, this volume
- Gardner and Satyapal, PASP, in prep
- Graham et al. 2000, this volume
- LeFevre et al. 2000, this volume
- Mackenty et al. 2000, this volume

A Multi-Object Spectrometer using Micro Mirror Arrays

J.W. MacKenty and M. Stiavelli

*Space Telescope Science Institute, 3700 San Martin Drive, Baltimore
Maryland 21218*

Abstract. We present a concept for achieving multi-object spectroscopy using a micro-electronic mechanical system (MEMS) array of small mirrors to synthesize a focal plane slit array. This type of spectrometer is particularly suitable both for space astronomy applications where traditional punch plate designs are impractical and for ground based applications where cryogenic operation is required. We discuss results from our ongoing NASA supported study of a Multi Object Spectrometer design for the Next Generation Space Telescope (NGST) including our technology development program for Micro Mirror Arrays (MMAs).

1. Introduction

Multiplex advantage in spectrometers is often desired in astronomical applications. This has been achieved by a variety of means including integral field designs which obtain spectra of every point within a field (e.g. image slicers, fiber bundles, fourier transform spectrometers, and fabry-perot etalons) and multi-object spectrometer (MOS) designs which obtain spectra of observer specified locations. Successful MOS designs have included robotically positioned optical fibers, which then direct light into a classical long slit spectrometer, and punch plates which define entrance apertures for an imaging spectrometer. This paper discusses a new variant of the latter design.

A multi-object spectrometer implemented with a punch plate at the telescope focal plane provides apertures (or slits) at observer selected locations. The positions of sources are measured prior to the acquisition of the spectra and an appropriate plate produced. The primary advantages of this systems are: (1) large field of view, (2) mechanical simplicity, especially compared to robotic fiber systems, and (3) generally high throughput. The main disadvantages are: (1) the need to prepare punch plates in advance, (2) the potential for overlap of the spectra on the detector without careful planning, and (3) incomplete use of the pixels on the detector array.

2. A MOS Spectrometer for NGST

2.1. Astronomical Motivations

The NASA Next Generation Space Telescope (NGST), expected to be flown in 2008 with participation by the European Space Agency (ESA) and the Canadian Space

Agency (CSA), is intended to explore the earliest periods of galaxy formation. The NGST is planned to have a 8 meter diameter collecting area with all optics cooled below 60 °K. To explore the epoch between redshifts of 3 and 10 (or beyond), NGST is optimized for observations between 0.6 and 5 μm with potential extension to 10 to 20 μm . Observations, most notably with the Hubble Space Telescope, have shown that most sources at very high redshifts are both small and relatively faint. NGST requires high sensitivity, excellent background rejection, and angular resolution comparable to what HST provides in the optical (approximately 0.1 arcseconds).

Since deep imaging surveys with HST in both the visible and near-infrared show that the source density below a realistic magnitude limit for NGST spectroscopy ($K_{AB} \leq 27$ for $R=1000$; see Table 1) is less than about 100 per square arc minute, that a small fraction of these sources are at $z \geq 3$, and that long exposures will be required, a suitable spectrometer requires both high spatial multiplexing capability and a large field of view. To obtain redshifts, kinematics, and compositions, spectral resolutions ranging from $R=100$ to 3000, or perhaps 5000, are necessary (see Stiavelli 1999 and Kennicutt 1999). Further, since NGST is a cost limited mission, it is unlikely to have a large number of scientific instruments. Therefore, its spectrometer must be designed to both serve the highly demanding observations of high redshift galaxies (which may consume a significant fraction of the total mission duration) and serve the broad needs of a diverse observing program.

K_{AB} mag	$R=100$	$R=300$	$R=1000$	$R=3000$
25	0.11	0.48	3.24	23.84
26	0.42	2.16	17.57	141.6
27	1.94	11.4	103.5	
28	10.4	66.6		
29	61.1			
30	375			

Table 1. Expected performance of a slit spectrometer for NGST. Exposure times are given in hours.

2.2. Conceptual Design

In the spring of 1996, during the development of the NGST yardstick instrument complement (Stockman 1997), we proposed the concept of a multi-slit spectrometer quite similar to traditional designs but with its slit mask implemented in reflection through a digital micro mirror array. This design was further developed in our MIROS proposal, an optical-IR MOS for HST (see also Figure 7.8 in Stockman 1997). NGST-MOS is an evolution of this concept. The reflection slits are synthesized on an array of $100 \times 100 \mu\text{m}$ mirrors located in the telescope focal plane. This array would contain 2000×2000 such mirrors, each of which is individually addressed via a multiplexer. These arrays can be constructed using conventional micro-electronic-fabrication techniques and use electrostatic forces to position each mirror into an ON or OFF state. In its ON state, a mirror is tilted approximately 10° to direct light falling upon into the spectrometer. In the OFF state the mirror is tilted the opposite direction to discard the light (see Fig. 1).

Figure 1. The left panel illustrates how a micro mirror array can be used to define slits. The right panel shows, as an example, a simulated $R=300$ observation with the MIROS spectrometer on HST of a portion of the Hubble Deep Field.

2.3. Key questions

There are three key questions our ongoing study is exploring. First, does this type of spectrometer best serve the scientific goals of NGST? Second, will the performance of this design be satisfactory? In particular, can the necessary contrast on the micro mirror array be achieved so that the spectrometer will be detector noise limited for most cases? And, third, can a Micro Mirror Array (MMA) suitable for NGST be fabricated?

3. The NGST-MOS study

3.1. Baseline Design Capabilities

The present design for NGST includes a 2048×2048 Micro Mirror Array with each mirror subtending $0''.1 \times 0''.1$ in the focal plane thus offering a $3'.4 \times 3'.4$ field of view. This spectrometer would be capable of observing approximately 1000 (or more) simultaneous targets depending upon the source density, the wavelength coverage, and the spectral resolution desired. Use of the MMA permits very fast slit reconfiguration times (less than 1 minute) and allows slits to be specifically tailored to target morphologies. The spectrometer would cover a wavelength range of 0.6 to $5.3 \mu\text{m}$ with spectral resolutions of 100, 1000, 3000 (or 5000), plus provide an imaging and target acquisition capability. It will require a 4096×4096 pixel detector array, a grating wheel and a filter wheel.

3.2. Operations

The MMA design permits a broad variety of versatile operational configurations. Unlike many instruments, these are implemented entirely in software rather than requiring complex mechanical devices. These modes may include: (1) simple single slit with its length and width being any integer multiple of the micro mirror size ($0''.1$), (2) multiple identical slits with a filter or Fabre-Perot to define the wavelength range, (3) custom slits tailored to targets (see Fig. 2), (4) slit-

Figure 2. The left panel illustrates the capability of a micro mirror based spectrometer of tailoring slits to target shapes and orientations. In the right panel an image of the Hubble Deep Field South NICMOS field is used to illustrate how in the course of a deep spectroscopic study of faint galaxies, two-dimensional spectroscopic data can be obtained on the brighter ones by stepping the slit.

less spectroscopy and imaging (where all mirrors set to their ON state), (5) blind pointing with relatively large slits for background reduction at low dispersion without target acquisition time overhead, (6) autonomous survey with onboard target selection and acquisition, and (7) a coronagraph mode with multiple bright objects occulted by MMA slitmask (both for imaging and spectroscopy).

Rapidly reconfigurable slit masks offer opportunities which may not have been previously considered. For example, it enables dithering of spectra on the detector array is made possible as are innovative calibration methodologies. Two dimensional spectroscopy for brighter and extended sources may be obtained simultaneously with long integrations on the very faint targets by scanning a long slit over the extended source (see Fig. 2).

3.3. Simulated Performance

The most important MMA requirement is that its scattered light contribution be kept low enough that the achieved signal to noise ratio is limited by either sky background or detector noise and dark current. We are conducting a detailed series of simulations (see Fig. 3) of the optical and science performance of NGST-MOS to arrive at a clear requirement for MMA scattered light. Preliminary indications are promising that this requirement can already be achieved with existing commercial devices.

4. MMA Technology Development

4.1. MEMS: The enabling technology

The MMA array is based on Micro-electronic Mechanical Systems (MEMS) technology. This technology uses standard micro-electronics fabrication techniques and has a significant industrial track record. For example, MEMS technology is the basis for the airbag sensors in well over half of all automobiles. A key optical MEMS

Figure 3. The left panel shows a region from the Hubble Deep Field South NICMOS field of $50''0$ on the side. The right panel shows a simulated image of the scattered light contribution. These simulations allow us to accurately assess the effect of micro mirror scattered light.

product which demonstrates the viability of the MMA approach is the Texas Instruments Corporation Digital Micro-mirror Device (DMD) which underlies their Digital Light Processing products. The DMD is available in sizes up to 1024×768 mirror arrays and is used in video projection displays. It utilizes $16 \times 16 \mu\text{m}$ aluminum mirrors and has been in quantity production since 1996. Studies we have performed with these devices demonstrate excellent optical performance and impressive reliability and ruggedness.

However, to support the highly demanding requirements of the projector marketplace (e.g. very high speed operation to generate gray levels, defect free devices, and low cost for individual parts), these devices are almost certainly not suitable for NGST. Further, NGST imposes additional requirements: most importantly a physically larger mirror to match the NGST plate scale (approximately $100 \times 100 \mu\text{m}$ mirrors), operation at 30°K , and a failure and defect tendency toward their OFF state.

4.2. MMA Development Roadmap

To provide a 2048×2048 mirror array in the NGST focal plane, the NGST-MOS design uses a 2×2 mosaic of 1024×1024 MMAs capable of operation at 30°K . Flight qualified arrays would need to be ready by late 2004. We have started to explore the technology development requirements and will provide a "roadmap" as a product of the NGST-MOS study effort.

To demonstrate the potential of this development process and to anticipate the areas requiring the greatest efforts, we have, together with the Goddard Space Flight Center's Detector Systems Branch, constructed 3×3 prototype arrays over the past year (see Fig. 4). These devices were built with three main goals in mind: (1) to demonstrate that we can construct working devices, (2) to explore construction of torsion bar supported mirrors with $100 \mu\text{m}$ scale, and (3) to demonstrate 30°K operation. As of this writing items 1 and 2 are completed

Figure 4. Scanning Electron-Microscope image of a 3×3 prototype array designed and produced by our group at the Goddard Space Flight Center's Micro-Electronics laboratory. Each element has $100 \mu\text{m}$ scale. Visible are the supports, yoke, and torsion bar. In this prototype the mirror has not been deposited.

and item 3 should be completed in the course of June 1999. For the coming year work on a 32×32 element cryogenic multiplexer is being started. In addition, significant testing has been done to demonstrate the validity of our physical models of device structure and to identify technology challenge areas and limit the design phase space.

The NGST-MOS study effort is funded by NASA Contract NAS5-98167 to STScI. The NGST-MOS study team consists of: J. W. MacKenty (PI), S. Casertano, M. Clampin, H.C. Ferguson, M.D. Rafal, E.J. Schreier, M. Stiavelli (STScI); C.A. Allen, J.P. Gardner, M.A. Greenhouse, J.E. Mentzell, S.H. Moseley, Jr., D.B. Mott, S. Satyapal, P.K. Shu, H. Teplitz (GSFC); V. Buat, D. Burgarella, K. Dohlen, O. LeFevre (LAS- Marseille); C.M. Carollo (JHU), S.D. Collins (UC-Davis), J. Fischer (NRL), R. Genzel (MPI-Garching), R.F. Green (NOAO), I.S. McLean (UCLA), R.L. Smith (UC-Davis), C.C. Steidel (Caltech), M.C. Wu (UCLA), and the Ball Aerospace Corp.

References

- Kennicutt, R., 1999, in "The Next Generation Space Telescope", ESA Publications (Noordwijk, the Netherlands), p.81
- Stiavelli, M., 1999, in "The Next Generation Space Telescope", ESA Publications (Noordwijk, the Netherlands), p.71
- Stockman, H. S., 1997, "The Next Generation Space Telescope", AURA.

Figure 1. LISA's 70 - 120 μ m spectrometer channel.

2. Instrument Concept

The wavelength range covering 2.5 octaves is divided into 3 separate intervals, each of which is fed into a separate grating spectrometer unit. The spectral resolution in each channel varies between 16 and 28 thus covering the whole spectral range with about 36 independent spectral resolution elements. Fully sampled spectra can be obtained by tilting the grating slightly by a fixed amount, such that the spectra are shifted by half a pixel across the array in the dispersion direction. The current design is optimized for 32 $\times$ 32 pixel Ge:Ga detectors with 350 μ m pixel pitch and 16 $\times$ 16 stressed Ge:Ga with 4mm pitch, but can easily be extended to accommodate larger e.g. 64 $\times$ 64 pixel detectors. The only moving part within LISA will be the grating tilt mechanisms.

An imaging spectrometer was given preference over a set of narrow band filters for several reasons:

- Flat-top high transmission narrow-band filters with $\lambda/\Delta\lambda \sim 6-20$ covering the spectral range 40 μ m - 220 μ m are not available. The 3 broad-band ($\lambda/\Delta\lambda \sim 2$) order sorting filter for LISA will have higher transmission and a better blocking than existing narrow-band filter.
- The data sets obtained with an imaging spectrometer are more consistent and always complete. Since this instrument will obtain a data cube during each individual integration, varying observing conditions will effect all data consistently in the same fashion. Such changes during the flight include variation of seeing, air mass, water vapor & atmospheric transmission, residuals in telescope tracking, pointing, image rotation & structural bending, as well as detector performance & temperature variations. Under unfavorable conditions, these modulations will introduce systematic noise into nonmultiplexed data, thus lowering the effective signal/noise ratio significantly.
- The integration time per target is limited to typically 3 hours per target and the risk of unexpected technical events on an airborne telescope is

LISA A low Resolution MIR/FIR Imaging Spectrograph for SOFIA

A. Krabbe and J. Wolf

*German Aerospace Center, Institute of Space Sensor Technology and
Planetary Exploration, Rutherfordstraße 2, 12489 Berlin - Adlershof,
Germany*

Abstract. LISA is a proposed low resolution ($R \sim 15-30$) imaging spectrometer for SOFIA, the American-German Stratospheric Observatory for Infrared Astronomy. Covering the wavelength range from $40\mu\text{m}$ to $220\mu\text{m}$ with three channels, LISA provides diffraction limited spectroscopy for each pixel within a rectangular field of view (FOV). The FOV sizes for the channels are $18 \times 11''$ for 40 to $70\mu\text{m}$, $32 \times 19''$ for 70- $120\mu\text{m}$, and $22 \times 22''$ for 120- $220\mu\text{m}$, each channel upgradable to about twice the linear size.

LISA will be able to address a variety of astrophysical topics, in particular on faint targets. Spatially resolved temperature distribution of dust emission can be studied in the environments of young and evolved stars, star forming regions, as well as in nuclear regions of nearby galaxies. Spectral energy distributions (SEDs) of distant galaxies will be determined in search for cold gas and dust components. LISA's spectral resolution will be high enough to search for broad spectral features (e.g. PAHs, ices). The optical design, the mechanical layout, the estimated sensitivity, and the scientific potential of such an instrument are briefly discussed.

1. Introduction

SOFIA, the Stratospheric Observatory For Infrared Astronomy, will be the biggest and probably the only platform with regular access to the entire MIR and FIR wavelength range between 25 and $250\mu\text{m}$ in both celestial hemispheres for the next two decades. From 2002 on it will open a new era in MIR/FIR astronomy, providing the highest available spatial resolution in combination with excellent sensitivity. SOFIA is a 2.7m telescope onboard a Boeing 747SP flying at an altitude of between 31000 up to 45000 feet during typically 7 hours observing flights. 10 facility, PI and special purpose instruments have already been selected and approved for first light.

Here we present a concept study of LISA, a highly efficient Low resolution Integral field Spectrometer for SPICA, which is considered to become part of the proposed Spectral-Photometric Far-Infrared Camera SPICA (see Wolf et al., this issue) for SOFIA. LISA will be a self-contained subunit of SPICA, complementing it with diffraction limited integral field spectroscopic capabilities at low spectral resolution of about $R \sim 20$.

capability to image and spectroscopically characterize the atmospheres of Earth-like planets at distances up to 50 lightyears.

While forecasting the future is an uncertain business at best, the discovery of biological activity on nearby planets would change our perception of the universe in the next century, just as irreversibly as the Michelson-Morley experiment altered the foundations of our understanding of the universe in this century.

The authors would like to acknowledge the collaboration of M. R. Carter and C. L. Bennett.

Figure 3. Schematic spacecraft orientations for a three-ring synthetic image.

The necessary enabling technology is the pointing of the apertures at the same scene. The HST fine guidance system, which uses several white light, shearing interferometers, achieves a pointing accuracy of 10 nano-radians - which is insufficient for these applications. Instead, a correlation tracking or vector integral projection (VIP) approach is used. VIP co-registration takes the scenes generated by each aperture and correlates them, such that each aperture's pointing system would be adjusted until they stared continuously at the same point, to within the required accuracy. The accuracy of the approach depends on the aperture sizes and the contrast to noise ratio of the scene. For the STARS system, with 10 meter apertures, the VIP approach can correlate two scenes to within one nano-radian. This is sufficient to enable coherent superposition.

4. The Geostationary-Geostationary Michelson (GGM)

Unfortunately, for imaging planets, the STARS system is still insufficient. Instead of linking the telescopes, if they are free-flyers in geostationary orbit acting together as a Michelson stellar interferometer then baselines up to 71,500 km can be achieved which yields a resolution element of 0.017 pico-radians which is more than sufficient to image an Earth-like planet around another star. For the chemical characterization of its atmosphere - a stellar occultation approach would permit spectral measurements of the atmosphere to be performed. The measurement is simultaneously enabled by the high signal to noise ratio of the transmission spectrum and complicated by the scattering of sun light about the planet causing significant dynamic range issues.

5. Conclusions

On the brink of a new century, 3D imaging has taken on new meaning and new roles in astronomy. Imaging Fourier transform spectrometry is one approach that is well suited to low spectral resolution, large field imaging. Combined with Michelson stellar interferometry to produce sparse thinned apertures, a system such as STARS offers significant performance enhancements over single monolithic apertures. Taken to a plausible limit, formation flying two large apertures in synchronous geostationary orbits will permit a system with sufficient

Figure 2. Schematic diagram of the STARS telescope and relay optics.

the lowest cost means of synthesizing an aperture - by varying the separation distance adequate sampling of the Fourier components of the image can be obtained. The key enabling technology is the ability to register the images with sub-pixel accuracy to insure proper interference of the scene energy.

The STARS telescope operates by rotating or translating to each observation position and then co-registering the images to subpixel levels. All successive scenes must be aligned and registered. The coregistration will have to include both x, y and rotation. Light from the sub-apertures is combined via a series of periscopes and combined in a Michelson spectral interferometer. The inclusion of the periscopes in the rotary joints of the armature permits transmission of the light between the apertures and the interferometer (Figure 2).

The spacecraft configuration consists of two lightweight telescopes separated by a bar linkage mechanism that permits a variable telescope separation. A three-axis stabilized spacecraft is located at the center pivot of the mechanism and provides navigation, stabilization, power, communications, command, and control. A Michelson spectral interferometer is located within the central spacecraft at a fixed optical path length from the telescopes for all telescope positions. Rotation is provided by harmonic drives and momentum wheels are used for gross pointing and to react to the rotation of the telescopes. Figure 3 illustrates the spacecraft in plan and scene view and a sequence of telescope positions is illustrated in Figure 4 along with the sampling locations.

Figure 1. Conceptual diagram of the Spaceborne Thinned Aperture Rotating System (STARS) spacecraft configuration.

3. Engineering Practicalities: The Spaceborne Thinned Aperture Rotating System (STARS)

Current space engineering practice establishes certain limits beyond which financial feasibility becomes vague at best - conversely, it is a useful exercise to ask 'how far can one push today's technology'. Such an exercise in space astronomy led directly to the STARS concept - if NGST is feasible, how far could the technology be pushed. As the Keck interferometer will demonstrate and the Space Interferometer Mission (SIM) and the Terrestrial Planet Finder (TPF) will exploit in space, sparse aperture arrays enable higher spatial resolution imaging at lower costs than filled apertures. Of course, there are a myriad of complications that must be overcome to use a sparse aperture, but the necessary technologies exist.

The STARS spacecraft (Figure 1) consists of two space-telescopes coupled in a manner to enable the apertures to be sampled as pairs as the telescope is rotated. To provide adequate spatial sampling, the system could be used at three different diameters. The telescopes would be coupled by articulated booms to a central spacecraft that contains a Michelson spectral interferometer that is used both to combine the beams from the apertures and perform 3D spectral imaging. In this 'artist's visualization, we imagined a pair of Hubble like (shrouded) telescopes, but it conveys the concept:

By utilizing multiple apertures and multiple samples of the same scene from several positions the STARS system will be able to match the performance of a much larger single aperture system. A rotating system of two apertures is

one can look at the universe. In addition, Michelson developed another interferometer - the stellar interferometer which provides significantly improved spatial resolution with sparse (typically pairs of telescopes) apertures. The combination of the two Michelson interferometers - using the spectral interferometer to combine the wavefronts from a stellar interferometer offers a path forward into a new class of 3D imaging systems.

There is a clear trend towards larger apertures and sparse optical telescopes such as the KECK observatory. In addition to a myriad of cosmological and astronomical studies, there is one class of measurements that is outside the capability of our present and planned observatories: the ability to spectroscopically detect the presence of life on a planet around another star. The problem may be cast in the following fashion: life as we know it is based on the photosynthesis-respiration process - the conversion of oxygen into carbon dioxide in a cyclic fashion - but the detection of oxygen (a homonuclear molecule with no infrared spectra) is difficult - so instead one can seek spectral signatures of ozone - the photolysis product of oxygen in our atmosphere. The spectroscopic detection of 'intelligent life' is more difficult, but as we know from our own atmosphere, the spectral signatures of the manmade freons will suffice - since only intelligent life would pollute their atmosphere with something that destroys it.

2. Achieving High Angular Resolution Imaging

A planetary system around a star at 50 light years distance has the following angular resolution characteristics - the radius of an Earth-like orbit is 250 nano-radians, the star 2.3 nano-radians in diameter, and an Earth size planet is 21 pico-radians in diameter. A brief survey of telescopes indicates how difficult a challenge this observation will be. A Hubble-class telescope (2.5 m aperture) has an angular resolution of 0.5 micro-radians, an NGST-class telescope (10 m, for this example) has an angular resolution of 0.12 micro-radians.

From this basis, it is possible to extrapolate beyond NGST to two more advanced concepts.

The Spaceborne Thinned Aperture Rotating System (STARS) - a 100 meter thinned array consisting of a pair of 10 meter apertures on a 100m articulated boom - which would yield an angular resolution of 12 nano-radians (Ironically, only after preparing this paper, was it noticed that this is essentially the KECK interferometer IN SPACE - perhaps the above error (10m rather than 8m for NGST was a small, but necessary, Freudian slip). But 12 nano-radians is still insufficient to image an Earth-like planet. While this may seem a little far out, combined properly, we have the necessary technology, and to satisfy Martin Harwitt concern about budgets - a system such as STARS delivers an order of magnitude increase in performance for a cost increase of 2-3 times the cost of NGST.

To achieve the necessary spatial resolution to image a planet, larger baselines are needed, and from an orbital mechanics point of view, a Geostationary to Geostationary Michelson (GGM) interferometer -two free flying apertures in opposing GEO orbits acting together as a Michelson Stellar Interferometer, with a baseline of 71,500 km yields an angular resolution of 0.017 pico-radians - sufficient to image that 'pale blue dot' around another star.

Beyond the NGST (Next Generation Space Telescope)

M. C. Abrams, D. D. Gregory, and R. Hookman

*ITT Industries, Aerospace/Communications Division, P. O. Box 3700,
Fort Wayne, IN, 46801-3700*

Abstract. The advent of space missions such as Next Generation Space Telescope (NGST), Space Interferometry Mission (SIM), and Terrestrial Planet Finder (TPF) provide the building blocks for new approaches to 3D imaging of the universe. Imaging spectrometers are vying for the opportunity to expand our perspective on IR astronomy on NGST, and spatial interferometry is the key enabling technology for SIM and TPF. Combining these technologies offers another leap in capability and performance. The technical feasibility and scientific potential of combining spatial and spectral interferometry will be addressed in the context of enabling new frontiers in astronomy.

1. Introduction

In the light of the exciting applications and potential of 3D imaging in astronomy, it may be worth taking a flight of fantasy - grounded in the capability of today's technology. Martin Harwit has suggested that success in astronomy will come from designing programs that are within our national financial means and that often the major advances in this century have been a result of exploiting obsolete military hardware - radio astronomy after World War II and gamma ray astronomy following the observations of the VELA program are but two examples. Within that context, I would like to look beyond NGST - to the next generation of space-based astronomical telescopes.

There is a peculiar nature to the last years of a century - the 'fin du siecle' - the arrogance and ennui that occurs at the 'end of the century'. Just over 100 years ago A. A. Michelson (with Morley) attempted to validate 'modern' electromagnetic theory by observing the motion of the Earth through the ether - the medium in which electromagnetic wave propagated. The experiment yielded a null result - there was no doppler shift and hence no ether. In the first years of this century, Einstein explained that the speed of light is constant, and our understanding of the physical universe changed irreversibly. At the brink of another century, another component of Michelson's pioneering work (the stellar interferometer) will enable a new revolution in our understanding of the nature of the universe.

Related papers (Bennett, Maillard, and Graham) have provided a foundation for a new application of Michelson's spectral interferometer - the stigmatic imaging properties of the Fourier transform spectrometer permit simultaneous wide field imaging and spectroscopy - which transforms the manner in which

imaging implies an acquisition of *zpd*, this facility is used also in this mode, in such a way that the first frame is recorded exactly at *zpd*, which also simplifies the subsequent data processing. As in broad-band imaging mode, IFS of the two incoming adjacent fields can be made in two successive operations. Note also that with an IFTS, deep broad-band imaging of a field is always available with spectro-imaging by summing all the frames of the interferometric cube or, in the transformed spectral cube, by selecting the spectral range afterward.

3.3. Servo-system

The two modes described above require a control of the *opd*. The main variation of *opd* is obtained by a motion of one of the cat's eyes, mounted on a parallelogram and displaced by an inchworm actuator. For that, a comparison beam from a reference source is generally needed. As the *opd* is short, and taking into account the stable environment in space, capacitance micrometers as used for stabilized Fabry-Pérot should be seriously considered for controlling the distance in each arm between the beamsplitter and the retro-reflectors.

4. Conclusion

This instrument fully uses the property that an FTS is a direct imager. It offers the unique capability of Integral Field Spectroscopy over a very large spectral range (0.6 to 5 μm), on a $3' \times 3'$ field, and deep broad-band imaging over the same field size, all in a single instrument. The chosen design is oriented toward the detailed spectroscopic study of clusters of galaxies, of star forming regions and of extended single objects, as in a given observing time, spectra with the optimum S/N ratio are obtained for all the points of the field. Thus, this instrument can perform a rich variety of scientific NGST goals.

References

- Adam, G., Rocca-Volmorange, B., Gerard, S., Ferruit, P., Bacon, R. 1997, A&A, 326, 501
- Allington-Smith, J., Content, R., Haynes, R., Lewis, I. 1998, SPIE, 3355, 196
- Arribas, S. et al., 1998, Fiber Optics in Astronomy III, ASP Conf. Series, S. Arribas, E. Mediavilla and F. Watson (ed), 152, 149
- Fowler, A.I., Gatley, I., McIntyre, P., Vrba, F.J., Hoffman, A.W., 1996, Infrared Detectors for Remote Sensing: Physics, Materials, and Devices, R.E. Longshore, J.W. Baars (Eds), Proc. SPIE, 2816, 150
- Graham, J.R., Abrams, M., Bennett, C., Carr, J., Cook, K., Dey, A., Najita, J., Wishnow, E. 1998, PASP, 110, 1205
- Haynes, R., Content, R., Lee, D., Allington-Smith, J. 1998, SPIE, 3354, 788
- Krabbe, A., Thatte, N., Kroker, H., Tacconi-Garman, L. 1997, SPIE, 2871, 1179
- Lefèvre, O., Vettolani, G., Saisse, M., Maccagni, D., Mancini, D., Picat, J., Mellier, Y. 1998, SPIE, 3355, 8
- Maillard, J.P. 1995, Tridimensional Optical Spectroscopic Methods in Astrophysics, IAU Co. 149, G. Comte, M. Marcelin (ed), ASP Conf. Series, 71, 316
- Tecza, M., Thatte, N., Krabbe, A., Tacconi-Garman, L. 1998, SPIE, 3353, 394

Figure 1. Optical layout of an Imaging FTS for NGST. The interferometer itself is contained in a volume of $\sim 950 \times 950 \times 200$ mm. From the use of cat's eyes (M4-M5-M6 assembly) the input and the output beams are in two planes, parallel to the figure, with axis shifted of ~ 100 mm. M1 and M9 are single piece mirrors (360×180 mm), on top of each other, for the two input and the two output beams. The mirrors to block either EI or EII and to direct the flux from a cold sphere into the instrument are not drawn for clarity. They should be on top of M10 and its symmetrical counterpart.

EII. The stability and the accuracy of zpd defines the photometric accuracy. For imaging through the whole InSb range with a photometric precision better than 10^{-2} , zpd must be defined and kept within ± 10 nm.

3.2. Spectro-imaging mode

Due to the scientific goals of NGST, a low resolution of the order of 100 is considered. For an FTS the limit of resolution $d\sigma$ is constant. A $d\sigma$ of 20 cm^{-1} corresponds to $R = 100$ at $5 \mu\text{m}$ and 500 at $1 \mu\text{m}$, i.e. the recording of 400 spectral elements over the 1 to $5 \mu\text{m}$ range. The required maximum path difference (opd) is equal to 0.30 mm , i.e. a maximum travel of 0.15 mm . However, without changing at all the overall size of the instrument, an $opd \simeq 1 \text{ mm}$, thus, giving a provision of a three times higher spectral resolution for special applications, should be wise to adopt. The opd must be scanned step-by-step. A free spectral range ($f\text{sr}$) for the case in which one would like to cover in once the 1 to $5 \mu\text{m}$ range supposes $f\text{sr} \geq 10^4 \text{ cm}^{-1}$. A step size equal to 300 nm is adapted for this bandpass. The 0.3 mm opd will be scanned with 1000 steps, which requires to store ~ 72 Gigapixels for a scan. In order to limit the data cube size for a higher resolution, the bandpass should be limited. As the broad-band

2. Project of a large field IFTS for NGST

The rationale of an IFTS for NGST has been discussed by Graham *et al.* (1998). In this paper are presented the possible capabilities of a spaceborne IFTS with a near-infrared ($1 - 5 \mu\text{m}$) and a mid-infrared channel ($5 - 15 \mu\text{m}$). The principle of an IFTS can work in any wavelength domain. However, the project outlined below, derived from the experience of an astronomical IFTS called BEAR, on the CFH Telescope (Maillard 1995), is intended for the $1 - 5 \mu\text{m}$ region only. This full domain is well covered by InSb arrays with a high quantum efficiency (Aladdin technology, Fowler *et al.* 1996). In this range the thermal background seen by NGST (zodiacal emission) remains very low, which minimizes the reduction of sensitivity of an FTS covering in one shot a very large spectral range. The project is based on the potential availability of $2\text{K} \times 2\text{K}$ arrays, with a pixel size of $18.5 \mu\text{m}$, mounted in mosaic of 3×3 arrays. The sampling scale is chosen equal to $0.03''/\text{pxl}$ (diffraction-limited sampling at $2.4 \mu\text{m}$), which corresponds to a field of view of $3.0' \times 3.0'$.

Fig. 1 shows a possible optical layout for a large field IFTS, at low spectral resolution, based on a dual-input, dual-output design, which presents the major advantage of recovering the two complementary output beams from one aperture. This configuration implies the mounting of two detector arrays, one for each output image. This concept is intended for an object-oriented instrument. The single-input single-output design, which is the other option, potentially simpler, is more adapted to a survey-oriented instrument as a field double in size is possible for the same total number of pixels.

Due to the already large field size, the collimator and the camera mirrors consist of a combination of two aspherical mirrors. The retro-reflectors are cat's eyes, systems easier to adjust than cube corners. The pupil image (size 80 mm) is projected on the beamsplitter made of two parallel plates (one on each beam) in CaF_2 or BaF_2 , with a lateral offset to correct for the translation of the reflected beam with respect to the transmitted beam. The optical path is made of 10 mirrors plus the beamsplitter. With gold-coated mirrors ($R=0.99$) and a beamsplitter with T and R between 40 and 60 % over the 1 to $5 \mu\text{m}$ range, an overall optical efficiency of 87% can be obtained. The tolerances on all the optical elements must be defined in order to obtain a peak-to-valley variation of $\lambda/10$ on the difference of the two interfering wavefronts over the field at $0.6 \mu\text{m}$. With this constraint, the fringe contrast is reduced by only 2% at $1 \mu\text{m}$ and 0.2% at $5 \mu\text{m}$. The final throughput can be $\simeq 75\%$ for the whole spectral range.

3. Modes of operation

As a direct imager, the same instrument can be used as a camera and an IFS.

3.1. Broad-band imaging mode

With EII blocked and replaced by a cold sphere (Fig. 1) and the interferometer strictly at *zpd*, EI is sent toward one output and the other one is dark. Then, one array only is read. In this latter case, the imaging of two close fields with the same pointing is possible. The two entrances are imaged identically in the focal plane of the telescope. They are alternatively closed to image EI and then

Concept of an Imaging FTS and Broad-band Imager for NGST

J.P. Maillard, J.C. Renault

Institut d'Astrophysique de Paris - CNRS, Paris, France

Abstract. The field size in the various systems developed for Integral Field Spectroscopy (IFS) is limited by the number of spectra which can be placed on the detector array. None of them are direct imagers as is an Imaging FTS which then, can fully exploit the spatial resolution of the NGST over a $3' \times 3'$ field. Therefore, the same instrument is usable simply for broad-band imaging. In this latter mode, the FTS must be kept at zero path difference (*zpd*) and be able to switch to a step-scan mode for spectro-imaging. Such a project is outlined with a dual-input, dual-output design, for the 1 to 5 μm domain, as a concept for a NGST instrument.

1. Advantages of an IFTS

Lenslet array (prototype TIGER, Adam *et al.* 1997), fiber reformatter (e.g. INTEGRAL, Arribas *et al.* 1998), or a combination (e.g. GMOS under construction for Gemini, Allington-Smith *et al.* 1998, VIRMOS for VLT, LeFèvre *et al.* 1998) and image slicer (Krabbe *et al.* 1997) are the three major techniques to make IFS with a slit spectrometer. For a brief comparison :

a) Lenslet and fiber are generally considered for visible instrument. However, fiber systems up to 2.5 μm are under development (SMIRFS on UKIRT, Haynes *et al.* 1998 and SINFONI for VLT, Tecza *et al.* 1998). Only image slicers, which are pure mirror systems, are applicable beyond 2.5 μm .

b) The ultimate field size is determined by the number of spectra which can be placed onto the detector array. With $2\text{K} \times 2\text{K}$ arrays this limit can reach up to 2K points, equivalent to a square field of $\sim 45 \times 45$ samples. The sampling scale is constrained by the minimum size of the segmenting elements. To take advantage of the diffraction-limited resolution in the space context, these systems can work on fields of only a few arcseconds.

c) All these systems have in common the need in their data processing for field reconstruction, as they are not **direct imagers**. The precision is necessarily limited by the mechanical characteristics of the device to segment the entrance field.

An Imaging FTS (Maillard 1995) is the only solution which :

- can exploit the spatial resolution of an NGST over a large field,
- does not increase in complexity with the number of field points,
- does not need a field reconstruction in the data processing.

Figure 1. a) SEM image of a single shutter blade, b) Sequence of the microshutter opening (sideview). Microshutter blades are represented by the shaded rectangles. Shaded circles are torsion bars. Two blades are engaged to have one pixel open, c) Opening sequence of the single shutter model. Shutters first are brought to close contact, and a voltage is applied to engage the shutters. The lower shutter is attached to the bulk of a membrane, the upper shutter is welded to a needle-manipulator. The upper shutter is moved along the arc formed by the radius-vector connecting their torsion bars. Frame three shows the shutter blades at 90° to the membrane (full open position). After the voltage is reset, shutters return to initial position (frame four).

3. Double-shutter actuation design

Crucial part of the design is the actuation scheme. After examining various proposed designs, we selected an actuation scheme which we call "double-shutter" (Fig. 1b).

An embedded electronic circuit on the surface of the microshutter membrane will perform addressing. Double-shutter actuation allows to translate macro-motion of a large structure, the membrane itself, into micromotion of the shutters (Fig. 1b,c). To implement this, two identical microshutter arrays are rotated 180 degrees with respect to each other and brought in close physical contact. Addressing and selecting is performed electronically by applying a voltage between a shutter blade and its counterpart on the actuation membrane. Once blades are selected and engaged the whole actuation membrane is moved with actuators outside of the active area and all selected shutters open. De-selection will cause an open shutter to return back into the closed position (Fig. 1c, last frame).

4. Summary

The device being developed in this project will be a large format mask for astronomical multi-object spectroscopy. Major achievement from the technological point of view is the solution of the large deflection angles problem in a close-packed array of elements. Combination of the double shutter actuation scheme with electronic addressing allows random shutter access. Since our device works in transmission, it achieves minimal scattered light and maximum possible contrast of the spectra. Our further plans include developing small and then larger shutter arrays with testing these devices on the telescopes.

This project is supported by grant from NGST GSFC 98-01.

References

- Garcia, E. J., Greenhouse, M. A., Mott, B., Kuhn, J. L., 1999, AAS Meeting 194 , 91.07
- MacKenty, J.W., 1999, AAS Meeting 194, 91.05
- Moseley, S.H., et al., 1998, AAS Meeting 193, 35.18

iciency (80% or better for the ratio of shutter area to total area). The device is being developed as a low temperature mask for the infrared spectrometer on the NGST, but it can also be used at optical wavelengths on the ground-based telescopes.

2. Microshutter design and tests

Shutter blade and shutter suspension are machined as a single piece out of a 1 μm thick membrane. The suspension is a torsion beam, which allows for blade deflection up to 180 degrees (Fig. 1a). The membrane can be single crystal silicon, silicon-nitride or any other material with sufficient strength. Experiments with both silicon and silicon nitride showed good mechanical performance of the shutter. The membrane sits on a frame, a structured wafer, which allows transmission through the shutters. To select and engage the shutters, the voltage is applied using an addressing microcircuit, which is implemented on the frame and the shutters themselves.

The most important issues in the microshutters development are

- the design of the single shutter,
- the actuation scheme: double-shutter with shutters selection and engagement by electronic microcircuit.

2.1. Single shutter

Asymmetric suspension of the microshutters on a torsion bar attached to the shutter only in the middle, allows to have longer torsion bar and therefore reduce the stresses. Sample tests and mechanical analysis proved that this design allows full deflection of the microblade out of the light path (Fig. 1). The blade can be rotated all the way up to 180°. The crucial difference of our device compared to any existing light beam manipulating devices: large angles of deflection on a compact packaging.

Shutters were made with FEI (focused ion beam machine) at the Lab for Ion Beam Research and Application (LIBRA) of the University of Maryland (UMd). Single crystal Silicon wafers were prepared and thinned down to about 1 μm by Christine Allen of the GSFC Detector Systems Branch. Usage of the ion beam milling allowed very quick turnaround and provided very fast progress from the idea inception to the implementation, test, refinement and development of the design. Within few months we were able to progress from scratch to have a fully operational single element microshutter (Fig. 1c).

Mechanical tests of the shutters strength and rigidity have also been carried out at the LIBRA, UMd. Resonant frequency measurements allowed us to measure the stresses in the shutters. Finite element analysis results agreed well with the measurements of the resonant frequencies and demonstrated uniform distribution of the internal stresses. Although all tests have been carried out at room temperature, there are no factors that would prevent this shutters to operate at cryogenic temperatures.

Programmable 2-Dimensional Microshutter Arrays

S.H.Moseley (GSFC), R.K.Fettig, A.S.Kutyrev (Raytheon ITSS),
C.W.Bowers, R.A.Kimble (GSFC), J.Orloff (UMd), B.E.Woodgate
(GSFC)

LASP, GSFC NASA, Greenbelt MD20771

Abstract.

We are developing a two-dimensional array of microshutters which can be used as a high efficiency, high contrast field selection device for a multi-object spectrometer. The device is a close-packed array of shutters, with a typical size of 100 microns square and area filling factor of up to 80%. Each shutter, made of single crystal silicon with an appropriate optical coating, pivots on a torsion flexure along one edge. Each of the shutters is individually selectable. An original double-shutter mechanism is employed for actuation. Since the device works in transmission, there is no loss of contrast due to diffraction from the edges of unactuated pixels. When working in reflection, the device can also be used as a micromirror array.

1. Background

There are different approaches to the problem (Garcia et al., 1999, MacKenty 1999, Moseley et al. 1998) aiming at building large format arrays (typical 1000 by 1000 elements) of very close packed individually controlled active elements, mirrors or shutters. The task is extremely challenging and devices developed so far remain complex and expensive to manufacture. The New Generation Space Telescope (NGST) gave a strong boost to these studies. Primary wavelength region of NGST operation is the near-infrared and devices have to be cooled to cryogenic temperatures, which makes the technological problems yet even more complex.

With the advent of the micro electro mechanical systems (MEMS) era, the development of micromechanical array devices, that allow to redirect light by individually controllable tiny mirrors, has been progressing at a fast pace. Practically all research and development programs in this area are concentrated on micromirrors. So far there has been no large array of transmissive microshutters developed, e.g. arrays of shutters, that are being developed and built by CSEM Co., Switzerland, have a low area filling factor and have small number of shutters (<http://www.csem.ch/microsystems/>). These devices employ a resonant excitation technique of the shutters and designed for operation at room temperature.

In our project we are pursuing the goal of creating large fully addressable microshutter arrays of small elements (100 μ m typical dimension) with high ef-

Figure 2. The working principle of the image slicer. A rectangular field in the sky is split into six individual stripes, which are realigned to form two parallel entrance slits at the input of a grating spectrograph.

slightly higher compared to a ground based observatory. Under such conditions most efficient and multiplexing observing techniques are strongly advised. In addition, spectral multiplexing avoids overheads created by the need to frequently changing the setting of the instrument. Under less than optimal observing conditions, consistent spectral information is at least as important as enhanced spatial resolution.

- Most targets under consideration only cover a fraction of the detector area, in particular if even larger arrays become available. Imaging spectroscopy uses the whole array since it distributes spectral as well as spatial information across it and thus maximizes the observing efficiency at any time.
- An imaging spectrometer does not require an excellent pointing of the telescope to put a target on a narrow slit. Its field of view (FOV) is much wider and the pointing effort is more comparable to an instrument using narrow band filters.
- The imaging capabilities of such a spectrometer can be used to acquire the target or to check its position during the integration. A two dimensional image can always be extracted by collapsing the spectral information and rearranging the pixels according to their proper location on the sky.

3. Optical design

The basic design of LISA is that of a long slit spectrometer (Figure 1) with two parallel input slits. The spectrometer is designed such that both entrance slits are dispersed into nonoverlapping spectra on the array. A system of optical flat mirrors, the image slicer, rearranges the two slits into one rectangular FOV on the sky. Figure 2 explains the slicing technique on the slicer for the 70 - 120 μ m wavelength range. Figure 3 shows a close up view of the image slicer

Figure 3. Close-up view of the image slicer. 12 plane mirrors rearrange the image information. Compare with Figure 2.

itself. For a 32×32 Ge:Ga detector it will be composed of two sets of 6 gold coated plane mirrors and a mechanical support structure all made of diamond turned aluminum. The first set of mirrors, located in an intermediate focus of the telescope, slices a rectangular image into 2×3 strips and deflects them into different directions while the second set of mirrors picks up light from the individual strips and aligns them into two contiguous long strips such that the pupils of the 6 single strips are coincident. The distribution of the spectra across the detector is illustrated in Figure 4.

Since the optical train is telecentric at the location of the entrance slit (see Wolf et al. 1998), the recombined 6 single strips are optically indistinguishable from two single long slits. The fact that the resulting slits look stair-like shifts the spectra on the detector by only a small amount with respect to each other. Each single mirror of the first set is about 12 mm long and 1.2 mm high. The area centers of the second set of mirrors all lie on a rotating parabola such that the focus of the parabola coincides with the telescope focus and the position of the entrance pupil of the unsliced image is retained at $-\infty$. The FOV of LISA was chosen 10×6 pixels. Such a rectangular format is more flexible since most of the targets do not show circular symmetry. A quadratic FOV covering 8×8 pixels is another possible configuration which can be realized with 2×4 slitlets if each slitlet is 8 pixels long.

The current basic parameters of the 3 channels are given in Table 1 for two different detector sizes. The spectrometer itself will be a modified Cerny-Turner type with the grating in off-axis Littrow configuration. The tilted exit slit can be accommodated for by rotating the slicer with respect to the optical axis by

Table 1. Current basic parameters of LISA

Range	40-70 μm		70-120 μm		120-220 μm	
Pixel scale [$''$]	1.8		3.2		5.6	
$\Delta\lambda[\mu\text{m}]$	2.3		4.0		6.5	
Detector	Ge:Ga		Ge:Ga		sGe:Ga	
Detector matrix [pixel ²]	32 $\times$ 32	64 $\times$ 64	32 $\times$ 32	64 $\times$ 64	16 $\times$ 16	32 $\times$ 32
Pitch [μm]	350		350		4.0	
Scale at entrance slit [$''/\text{mm}$]	2.8		2.8		2.8	
Field of view [pixel]	10 $\times$ 6	21 $\times$ 12	10 $\times$ 6	21 $\times$ 12	4 $\times$ 4	10 $\times$ 6
$''$	18 $\times$ 11	38 $\times$ 20	32 $\times$ 19	38 $\times$ 20	22 $\times$ 22	56 $\times$ 34
Slitlet height [mm]	0.7		1.2		2.0	
Slitlet length [mm]	7	14	12	24	8	20
$f_{\text{collimator}}$ [mm]	360		360		360	
f_{camera} [mm]	193		108		180

Figure 4. Distribution of the spectra across the focal plane of LISA on a 32 $\times$ 32 pixel detector. The two entrance slits are dispersed on the same detector but do not overlap. The stair-like geometry of each entrance slit is reflected in the spectral offsets between individual slitlets.

Figure 5. SOFIA's point source sensitivity for broad band imaging and low-resolution spectroscopy. Curves for the SIRTf mission are given for comparison. The triple-dot-dashed curve denotes the increase in sensitivity of SOFIA for surveys due to spatial multiplexing for 64×64 pixel Ge:Ga and a 32×32 stressed Ge:Ga array. The "Cirrus limit" curve is a rough estimate (see e.g. Franceschini et al. 1991 and Herbstmeier et al. 1998 for details). The far infrared sensitivity for a spectral resolution of 30 is denoted by the dot-dashed curve.

the off-axis angle (not shown in Figure 1). Collimator and camera mirror are off-axis parabolas. Diffraction effects at the slicer have to be taken into account since in the worst case the height of a slitlet is only about 10 wavelengths across. Behind the slit the beam will flare to $f/11$ instead of $f/30$ perpendicular to the slit. This leads to an elliptical pupil on the grating with an axis ratio of about 2.5:1. The background on the detector, however, will not be enhanced since the cold Lyot-stop further up in the beam has already limited the background flux.

4. Scientific potential

The current design of LISA is based on the experience with 3D, a near-infrared imaging spectrometer (Krabbe et al. 1996, Weitzel et al. 1996). The expected optical transmission of the slicer will be about 95%, the area covering factor of the detector will be 80% yielding an optical efficiency of 75%. If the target fits within LISA's FOV, the multiplex gain over a set of filters is about 13. Between 12 and 14 spectral positions are recorded per integration under background limited conditions. An instrument using n narrow band filters across the spectral range $40 - 220 \mu\text{m}$ will be about $n/3$ times less efficient provided that the other losses roughly balance each other. With larger arrays the multiplex advantage of an imaging spectrometer will be even higher by about the ratio of the number of pixels.

A model calculation on SOFIA's point source sensitivity using conservative assumptions is shown in Figure 5. For a spectral resolution of 16 instead of 4 the sensitivity curves are for 2.5σ instead of 5σ and 1 hour integration. The broad band SEDs of IRAS10214+4724 and 3C-234 redshifted to $z=1$ are shown to illustrate SOFIA's sensitivity. All IRAS sources lie well above these curves. Between $40\mu\text{m}$ and $100\mu\text{m}$, LISA's sensitivity ranges between 40mJy and 70mJy per beam.

LISA will be able to determine the spectral energy distribution (SED) as well as broad spectral features in galactic and extragalactic targets. On extended objects, the spectral images can be decomposed into components of different temperature and spatial morphology. Applications are widespread. In star forming regions, the interaction between young stars and their environments, their outflows and dynamical evolution (e.g. T Tauri stars, HH objects, AeBe stars) can be investigated in detail. Dust disks and the evolution of protoplanetary conditions can be studied on Vega like main sequence stars (e.g. Vega, βPic). Not much is known about the expanding envelopes of more massive evolved stars during their AGB phase and evolution into protoplanetary nebulae. Interaction between ejecta from violent supernovae explosions and interstellar matter leads to progressing shockfronts, which heat up and process dust and gas of the surrounding molecular clouds. In our solar system low resolution spectroscopy will be an important tool in analyzing the composition of solids and ices on planetary surfaces and in their atmospheres as well as on asteroids and in comets.

The dust content and temperature composition of external active and luminous galaxies like Seyferts, starbursts, or interacting systems will be investigated by obtaining the (spatially resolved) SEDs of the different components and thus constraining the physical processes involved in the energy production. A special population are cool and/or very red objects (e.g. VERO's). Emphasis will be on identifying and studying the colder components in these targets. LISA will thus complement the camera with an efficient tool for analyzing the broad spectral characteristics of targets, which will be discovered and/or studied with SPICA.

References

- Krabbe A. 1996, in: Proc. of SPIE conference on: Imaging Spectrometry II, M.R. Descour & J.M. Mooney, vol. 2819, Denver, 243
- Franceschini A. et al. 1998, A&AS, 89, 285
- Herbstmeier U. et al. 1998, A&A, 332, 739
- Weitzel L., Krabbe A., Kroker H., Thatte N., Tacconi-Garman L.E., Cameron M. & Genzel R. 1996, A&AS, 119, 531
- Wolf J., Schubert J., Anheyer H., Driescher H., Hanna K., Kirches S., Paul E., Rabanus D., Rösner K. & Krabbe A. 1998, in Proc. of SPIE on: Infrared Astronomical Instrumentation, A.M. Fowler, vol. 3354, 952

SPICA: IR-Camera for SOFIA

Jürgen Wolf, Alfred Krabbe and David Rabanus

*German Aerospace Center (DLR), Institute of Space Sensor Technology
and Planetary Exploration, Rudower Chaussee 5, 12489 Berlin,
Germany*

Abstract. The spectral-photometric infrared camera SPICA has been proposed as a German science instrument onboard the Stratospheric Observatory for Infrared Astronomy. Using the 2.5 m SOFIA telescope SPICA will image at wavelengths from 20 to 220 μm with unprecedented spatial resolution in the far-infrared ($\lambda/D = 1.7$ arcsec at 20 μm to 16 arcsec at 200 μm). The instrument will also contain 3 imaging spectrometer channels (LISA, Low Resolution Integral Field Spectrometers for SPICA) with a spectral resolution $\lambda/D \sim 20$.

1. Introduction

The Stratospheric Observatory for Infrared Astronomy SOFIA is under development in the U.S. and Germany to start science operations in the year 2002. The German built 2.5 m Cassegrain/Nasmyth telescope will be flown onboard a Boeing 747SP at altitudes up to 45,000 ft. Above most of the atmosphere's water vapor, the entire IR and FIR spectral range will be available for observations, except for some narrow absorption bands. A number of science instruments are being developed in both countries to provide the observatory with a comprehensive suite of observing tools. The spectral-photometric infrared camera will contain 3 spectral channels for continuum imaging and 3 low-resolution imaging spectrometers. SPICA has been proposed to complement the two German instruments of high spectral resolution that are currently under development: the line spectrometer FIFI-LS (Poglitsch 2000) and the heterodyne receiver GREAT.

2. Instrument Concept

The major advantage of SOFIA as an IR/FIR observing platform will be its superior spatial resolution. Compared to its predecessors IRAS and ISO (0.6 meter telescopes) and to the Kuiper Airborne Observatory (0.9 meter telescope), SOFIA will offer 3-4 times sharper images in an 8 arcmin diameter field of view. SPICA plans to apply three large format detector arrays with pixel sizes of $1/2(\lambda/D)$ at 20 μm , 50 μm and 130 μm , respectively, to obtain fully sampled diffraction limited images:

A Si:As or Si:Sb blocked impurity band (BIB) detector array with 256×256 pixels will cover wavelengths out to 28 μm (As) or 40 μm (Sb). The projected

pixel size on the sky will be $0.8 \times 0.8 \text{ arcsec}^2$, resulting in a $3.4 \times 3.4 \text{ arcmin}^2$ field of view. For wavelength coverage the Sb-doping is preferred, however, the final choice of material will depend on individual detector performance, namely quantum efficiency, and their commercial availability.

A Ge:Ga photoconductor array with 32×32 pixels will work in the 40 to $120 \mu\text{m}$ range. The projected pixel size on the sky will be $2.0 \times 2.0 \text{ arcsec}^2$, resulting in a $1.1 \times 1.1 \text{ arcmin}^2$ field of view. Actually, to fill SOFIA's focal plane, a much larger array of 128×128 pixels covering $4.3 \times 4.3 \text{ arcmin}^2$ of sky is desirable. However, as arrays for that wavelength range are not commercially available, their custom development is a major technological effort. Therefore, a 32×32 array is being baselined with the larger formats up to 128^2 pixels as a future goal.

In the longest wavelength range of $120 \mu\text{m}$ to $220 \mu\text{m}$, an array of Ge:Ga photoconductors under uniaxial stress will be applied. For the most efficient use of SOFIA's observing time, a 64×64 pixel array with a $5.8 \times 5.8 \text{ arcmin}^2$ field of view would be needed. Again, due to the necessary custom development and the rather complicated technology of stressed photoconductor arrays, a modest size array of 16×16 pixels with a $1.4 \times 1.4 \text{ arcmin}^2$ field of view is being baselined with an option for future upgrades to larger formats as they become available.

The mid and long wavelength arrays of SPICA require a customized cryogenic readout electronics which operates at temperatures as low as 1.8 K with very stringent performance specifications on noise, stability, charge handling capacity and power dissipation.

The optics of SPICA will be purely reflective and operated at LHe temperature to reduce thermal emission. The imaging concept includes two steps, (i) from the telescope focus to an intermediate focus including a Lyot stop for background and straylight reduction and a filter wheel, and (ii) from the intermediate focus to the detector arrays with individual back end mirrors to set the plate scales. At the intermediate focus, the light path can be routed into the three integral field spectrometers LISA (Krabbe & Wolf 2000).

3. Project Status

The overall concept of the instrument as been developed at the DLR institute in Berlin. Architectural designs have been developed for the Ge:Ga and stressed Ge:Ga arrays and first prototypes of array modules were built. Detailed design work has been done on the cryogenic readout electronics in collaboration with other efforts in that field. For operating the arrays a state-of-the-art drive electronics and data acquisition system has been obtained and tested with a bare multiplexer of the future short wavelength array. The instrument optics has been designed and ray traced to a high level of confidence and a preliminary design of the cryostat for SPICA exists. This work has been summarized by Wolf et. al. (1998).

Partners to the DLR institute for the development and scientific use of SPICA are Astronomisches Institut der Ruhr-Universität Bochum, MPI für Astronomie, Heidelberg, MPI für Radioastronomie, Bonn, and University of

Florida, Gainesville. Unfortunately, full scale funding for the project is currently not available. However, we are continuing our efforts to start the project.

References

- Poglitsch, A. 2000, this volume
Krabbe, A. & Wolf, J. 2000, this volume
Wolf, J. et al. 1998, Proc. SPIE, Vol. 3354, 952, Infrared Astronomical Instrumentation, ed.
A. Fowler

3-D Spectroscopy with SIRTf

L. Armus

*SIRTf Science Center, California Institute of Technology, 310-6,
Pasadena, California 91125*

Abstract. The Space Infrared Telescope Facility (SIRTf) is scheduled for a December 2001 launch. The three science instruments on SIRTf are the Infrared Array Camera (IRAC), the Infrared Spectrograph (IRS), and the Multiband Imaging Photometer for SIRTf (MIPS). The IRS produces both low ($R \sim 60 - 120$) and high ($R \sim 600$) resolution spectra over the $5 - 40\mu\text{m}$ and $10 - 40\mu\text{m}$ ranges, respectively, in either a Staring or Spectral Mapping mode. Since the IRS slits range in length from $12''$ to nearly five arcminutes, SIRTf has the capability of producing 3D spectra over relatively large areas of the sky.

1. SIRTf Basics

SIRTf represents the first step in NASA's Origins program. Having an 85cm f/12 Beryllium mirror cooled to 5.5K, together with a new generation of detector arrays, SIRTf will provide background limited performance from $3 - 180\mu\text{m}$ and be diffraction limited at $6.5\mu\text{m}$. After a December 2001 launch on a Delta rocket, SIRTf will move into a unique Earth-trailing, Solar orbit for its expected five year lifetime. This orbit provides for excellent sky access and visibility, with 35% of the sky visible at any one time for more than 40 days at a time. Its lightweight, cryogenic optics, high performance passive cooling, and state of the art infrared detectors will allow SIRTf to be used to study debris disks around thousands of stars, normal galaxies beyond $z = 3$, and Ultraluminous Infrared Galaxies (ULIRGs) beyond $z = 5$.

2. The SIRTf Instruments

There are three instruments on SIRTf. From shortest, to longest wavelength, these are the Infrared Array Camera (IRAC), the Infrared Spectrograph (IRS), and the Multiband Imaging Photometer for SIRTf (MIPS). The Infrared Array Camera (Giovanni Fazio, PI) provides wide field imaging in four bands centered at 3.6, 4.5, 5.8, and $8\mu\text{m}$. IRAC uses its four 256×256 SBRC arrays to sample two, 5×5 arcminute fields in the SIRTf focal plane simultaneously and, through the use of two dichroic beamsplitters, each field is viewed in two filters. The IRAC includes a cold shutter for determination of the dark current and total sky brightness, internal flood and transmission calibrators, and it can be operated in a sub-array mode intended to improve dynamic range.

The Infrared Spectrograph (Jim Houck, PI) delivers the primary spectrographic capability of SIRTf. The IRS uses four modules to cover the $5 - 40\mu\text{m}$ range at two resolutions. The two low resolution modules operate at $R = 60 - 120$ over the full $5 - 40\mu\text{m}$ range. The two high resolution modules operate at $R \sim 600$ over the $10 - 40\mu\text{m}$ range. There are four IRS slits in the SIRTf focal plane, which feed light to the four modules. The low resolution slits are $3.6'' \times 109.2''$ (for the $5 - 15\mu\text{m}$ spectrum) and $9.7'' \times 302.6''$ (for the $14 - 40\mu\text{m}$ spectrum). The low resolution slits are each sub-divided into two, in order to generate the 1st and 2nd order spectra for each low resolution module. An entire low resolution spectrum of an object is obtained by moving to each of the four sub-slits, in turn. Due to the true longslit nature of the low resolution slits, spatially-resolved spectra will be routinely produced by the IRS. The high resolution slits are $5.3'' \times 11.8''$ (for the $10 - 20\mu\text{m}$ spectrum) and $11.1'' \times 22.4''$ (for the $19 - 40\mu\text{m}$ spectrum). To facilitate spectroscopy of faint sources, the IRS is also equipped with a pickup feature that can image an approximately one arcminute field around a target position in either a blue ($13 - 18\mu\text{m}$) or red ($18 - 24\mu\text{m}$) filter, find the target of interest, then offset the target to the proper slit. This pickup feature can be used on the spectroscopic target itself, or a secondary (brighter) source whose offset from the primary spectroscopic target is known precisely.

The Multiband Imaging Photometer for SIRTf (George Rieke, PI) provides imaging at $24\mu\text{m}$, $70\mu\text{m}$, and $160\mu\text{m}$. The $24\mu\text{m}$ and $70\mu\text{m}$ channels sample two 5×5 arcminute fields in the SIRTf focal plane, through the use of a 128×128 pixel Si:As array, and a 32×32 pixel Ge:Ga array, respectively. The $160\mu\text{m}$ channel samples a 0.5×5 arcminute field, through the use of a 2×20 pixel, stressed Ge:Ga array. An internal, cryogenic scan mirror in the MIPS allows mapping in all three bands simultaneously, for efficient surveying at long wavelengths, and it also allows modulation of the signal on the Ge arrays. In addition to broad-band filter imaging, and SED mode provides $R \sim 15$ spectrophotometry from $52 - 99\mu\text{m}$ with the MIPS.

3. Spectral Mapping with the IRS

There are two modes of operating the IRS. Staring mode is the basic "point and shoot" mode, where the wavelength range, integration time and source position(s) are specified by the observer. Spectral Mapping mode functions in an identical way as the Staring mode, yet the telescope is commanded to scan in the cross-slit direction at a specified rate (from $0.01 - 1.00'' \text{ sec}^{-1}$) in order to build a data cube over the specified area of the sky. The telescope scan rate is fixed by the selected integration time and width of the chosen slit, such that two spectral samples are returned for each slit crossing. In other words, the array is read out twice in the time it takes for the slit to scan across its own width. The length of the scan on the sky is determined by the duration of a single read, the width of the slit, and the total integration time. A one hour observation (a single scan leg) with the short wavelength part of the low resolution slit during which the array is read out at its fastest rate (8.0 seconds between reads), covers an area of approximately $109'' \times 13.5$ arcmin on the sky. Similarly, a one hour observation with the long wavelength part of the low resolution slit at the same

readout rate, covers an area of approximately $303'' \times 36.4$ arcmin on the sky. Longer readout times directly translate into smaller areas on the sky for a given, fixed total observation time because the telescope moves at a slower rate. Since the longest readout time for the IRS is 512 seconds, deep integrations are built up through multiple exposures (or passes) over a given part of the sky.

4. Spectral Sensitivities and Science Prospects

The spectrophotometric limits for the IRS are as follows: for the low resolution modules, the continuum sensitivity (5σ , 500 seconds) at $15\mu\text{m}$ is $\sim 1\text{mJy}$, while for the high resolution modules, the (unresolved) emission line sensitivity (5σ , 500 seconds) at $15\mu\text{m}$ is $\sim 10^{-22}\text{Wcm}^{-2}$. The IRS will therefore be approximately 1 – 2 orders of magnitude more sensitive than the SWS on ISO. The IRS will advance the study of the physical conditions of atomic and molecular gas in Galactic and extragalactic objects via emission and absorption line diagnostics. In dusty environments (e.g. the Ultraluminous Infrared Galaxies first identified by Soifer et al. 1984, Houck et al. 1985 and Sanders et al. 1988) the wavelength coverage and sensitivity of the IRS will allow the measurement of a large number of emission lines of varying strength and ionization potential (e.g. [ArII], [SIII], [NeIII], [NeV], [OIV], etc.) which are critical for understanding the underlying energy sources in these heavily obscured systems (see Spinoglio & Malkan 1992). As was demonstrated by Genzel et al. (1998), the ratio of high-to-low ionization mid-IR features can be used to discriminate between an AGN or hot stars as the dominant energy source in these luminous galaxies.

While the high resolution IRS spectra will be ideally suited to studies of narrow atomic emission features, the low resolution spectra will be very sensitive to overall continuum shape, and broad emission and absorption features, such as the PAH lines seen in the near and mid-IR spectra of active and starburst galaxies. These lines, which can be excited by single photons, provide a sensitive probe of the ionizing environments in Galactic and extragalactic systems, because the small dust grains which produce them are easily destroyed in harsh environments (such as those immediately surrounding an AGN).

The Spectral Mapping mode of the IRS will provide real 3D spectroscopic measurements of extended sources, such as star forming complexes and molecular clouds, supernova remnants, and nearby starburst and interacting galaxies. Full spectral maps of active galaxies will shed light upon the starburst-AGN connection in the mid-IR, as well as provide galactic “templates” for studies at high redshift where ground-based slits typically encompass large fractions of the target. Since it is known that active galaxies at low redshift (e.g. NGC 1068, NGC 7469, Mrk 231, etc.) can have spatially distinct AGN and starburst components (Wynn-Williams, Becklin & Scoville 1985; Planesas, Scoville & Myers 1991; Mazzarella et al. 1994; Hamilton & Keel 1987), it is natural to suspect that some galaxies at high redshift will also have multiple sources of ionizing photons, and that their spectra might reveal the signatures of AGN, hot stars, and even large-scale shocks. These signatures will probably be most easily detected by mapping the change in the narrow atomic line flux ratios as a function of position, or nuclear distance. However, maps of nearby galaxies and star-forming regions in the PAH features would also be invaluable for refining

models which describe the chemistry of the ISM and the excitation of atoms and molecules at the interfaces of dense, molecular clouds. Thus, galactic spectra built up with the Spectral Mapping mode of the IRS will be important for disentangling the various components in the integrated spectra of more distant, unresolved objects.

Due to the intricacies of synchronizing telescope motion and array readout, as well as properly calibrating the Spectral Mapping data, we are currently planning on testing and refining the Spectral Mapping mode of the IRS during the first $\sim 2-6$ months of the mission (Dec. 2001 – June 2002). It is expected, therefore, that the full capabilities of this mode will be available for routine scheduling approximately nine months after launch.

References

- Genzel, R., Lutz, D., Sturm, E., Egami, E., Kunze, D., Moorwood, A.F.M., Rigopolou, D., Spoon, H.W.W., Sternberg, A., Tacconi-Garman, L.E., Tacconi, L., & Thatte, N. 1998, *ApJ*, 498, 579.
- Hamilton, D., & Keel, W.C. 1987, *ApJ*, 321, 21.
- Houck, J.R., Schneider, D.P., Danielson, G.E., Beichman, C.A., Lonsdale, C.J., Neugebauer, G. & Soifer, B.T. 1985, *ApJ*, 290, L5.
- Mazzarella, J.M., Voigt, G.M., Soifer, B.T., Matthews, K., Graham, J.R., Armus, L., & Shupe, D. 1994, *AJ*, 107, 1274.
- Planesas, P., Scoville, N., & Myers, S.T. 1991, *ApJ*, 369, 364.
- Sanders, D.B., Soifer, B.T., Elias, J.H., Madore, B.F., Matthews, K., Neugebauer, G., & Scoville, N.Z. 1988, *ApJ*, 325, 74.
- Soifer, B.T., Helou, G., Lonsdale, C.J., Neugebauer, G., Hacking, P., Houck, J.R., Low, F.J., Rice, W., and Rowan-Robinson, M. 1984, *ApJ*, 283, L1.
- Spinoglio, L., & Malkan, M.A. 1992, *ApJ*, 399, 504.
- Wynn-Williams, C.G., Becklin, E.E., & Scoville, N.Z. 1985, *ApJ*, 297, 607.

CCD Innovations

~~John~~ L. Tonry

~~Gerard~~ A. Luppino

Institute for Astronomy, University of Hawaii, Honolulu, HI 96822

1. Introduction

Unlike talks during the day, I'd like to try to stimulate discussion, so please break in at any point. I'm not really an instrument builder, even less a person who specializes in detector technology. As far as I'm concerned ZEMAX is a darn good editor, and 1/4-20 is 5. However, I've found that it's possible for scientists to make some significant improvements in even well developed technology such as CCD's and I'd like to share my experiences with you.

CCDs have come a long way in 20 years. Figure 1 shows a TI-800 chip which was the state of the art from 1980-1985. By modern standards it was a pretty awful chip. As Joe Miller mentioned the other day, it was loaded with traps and pockets, it had poor red QE, it was quite noisy, it was only about 1 cm², and it had a lot of dark current unless it was kept quite cold. Nevertheless, it was by far the best device available in that time period and was worth many \$100k, if you could get one (Caltech and JPL had a virtual monopoly, since it was developed using NASA money for WFPC).

Today, CCDs cost about \$1-2k per cm², the silicon is *much* better so that the incidence of pockets can be as low as a few per cm², they have much improved red response and fringing, their amplifiers easily reach 2 e- noise performance at 50kpix/sec, and are well below 10 e- even to 1Mpix/sec readout speeds.

The typical size of CCDs these days is 2k×4k which corresponds to a physical size of about 3×6cm if the pixel size is 15μm. One reason for this is driven by silicon technology. Astronomical CCDs are processed on wafers less than or equal to six inches in diameter. Figure 1 illustrates one example of a 150mm wafer whose area is maximally utilized by filling it with six 3×6cm 2k×4k CCDs (with room for a few smaller devices in the corners). There is no reason that CCDs cannot be made with pixels as small as 10 μm (with smaller full well), but this 3×6cm size is something we will probably be stuck with for a while.

The yields of these devices can be as high as 50% (although 25% might a number you'd bet on) which means that it has not been common to try to make larger, monolithic CCDs (e.g. 6×6cm) instead of putting a number of the 3×6cm devices into a mosaic. The 2k×4k CCID20 illustrated in Figure 1 is three-side buttable, with about a 70 μm gap between butted edges. We have a design for the fourth edge which could require as little as a 2.5 mm space for serial register and bonding pads. EEV is also producing such a "four-side" buttable chip.

Gerry Luppino, Chris Stubbs, and I have a lot run of CCDs currently being processed at Lincoln Labs which consists of four CCID20s (three side buttable

Figure 1. A TI-800 CCD is overlaid on a six inch wafer produced by Lincoln Labs which carries six $2k \times 4k$ CCID20 CCDs (as well as some small OTCCDs in the corners).

$2k \times 4k$), two CCID28 OTCCDs (which I will describe to you in detail), and four "MIDOT"s, three side buttable, which are $1 \times 1.3k$ Orthogonal Transfer devices about which I will not say more, except to note that we may have a lot of these devices and they are very well suited to applications such as Al Diercks describes in his poster on multicolor CCDs.

Figure 2 shows the CCID28 OTCCD in some detail, and I'll describe its features to you. First of all, you will note that it is divided into four sections with a mirror symmetry about the middle. The two sections near the serial register are $0.5 \times 2k$ and the two central sections are $1.5 \times 2k$. The entire CCD is made of orthogonal transfer pixels with no discontinuity in the metrology, but each of these four sections has independent parallel clocks. If you want to tie all four sets of clocks together you can treat this CCD as a monolithic $2k \times 4k$.

I will be using two of these chips as a "hitch-hiker" camera on the bent Cassegrain port on the UH 2.2-m telescope, and I will use the 512 sections as guide star areas. While the science image accumulates in the central $3k \times 4k$, a patch of CCD containing a guide star somewhere in the 512 sections gets clocked out rapidly, and the guide star motion is used both to guide the telescope and do the image motion compensation with orthogonal transfer. This chip has an innovative serial dump gate which permits the serial register to be erased instantly. This is helpful for the "shutterless video" mode of reading out a guide star patch rapidly.

Along with Barry Burke and Dick Savoye at Lincoln Labs, I came up with the idea of an Orthogonal-Transfer CCD (PASP 109, 1154 [1997]). The basic idea is simple: A three-phase CCD normally clocks the charge down to the serial register, and it can equally well clock the charge up by a different phasing of the clock signals. The side-to-side location of pixels is defined by a permanent channel stop implant. It occurred to us that a different layout of the gates which

Figure 2. A schematic of the orthogonal transfer CCID28 2k×4k CCD.

had a 90 degree symmetry would permit us to have normal 3-phase operation if the fourth gate were held low, but would also enable a side-to-side 3-phase operation if the horizontal gates were held low. Thus, this device allows us to move an accumulating image within the silicon to arbitrary location without incurring any readout noise penalty. Figure 3 illustrates the layout of these OT gates. The only additional cost of creating such a device instead of a conventional 3-phase device is that you need to put down a fourth layer of poly, for which an additional mask must be made, and there might be a decreased yield because of inter-gate shorts. Both of these costs have proven to be negligible. It is possible to compensate for image motion this way at a very high rate, at least 10kHz per pixel if you have suitable information of where to move things since it just involves parallel clocking. It is possible to shift the charge by fractional pixels if desired.

The basic loop we use for observing is simple. A guide star is selected, the shutter is closed while the array is erased, and then the exposure begins. After a brief integration time which depends on the brightness of the guide star (on the 2.2-m telescope we can collect enough photons from a star at $m=20$ in one second, $m=17.5$ in 0.1 sec, $m=15$ in 0.01 sec, etc), a small patch around the guide star is read out. This centroid of this image is calculated, compared with the previous iteration, and any shift is compensated for by changing the location of the guide star patch and shifting the science image in the science portion of the array. A suitably filtered version of these shifts is fed back to the telescope control system to provide guiding.

Figure 3. The four-layers of gates which permit orthogonal transfer within a CCD

A point which is not often appreciated is that tip-tilt correction doesn't need to be terribly fast since the power spectrum of image motion is a steeply falling function of frequency. If you need to do true adaptive optics, it is necessary to remove image motion to a very high degree because it is such a large component of wavefront distortion. However, if removal of image motion is the only correction being applied, you can usually remove most of it by working at only a few Hz (depending on the size of your aperture, and any telescope shake). Also, the isokinetic angle over which image motion is coherent is quite large, generally at least a few arcminutes. Thus we expect that we should be able to find a suitable guide star for image motion compensation at 10Hz or faster over virtually all the sky, and should a bright guide star not be available, it is always possible to run a bit more slowly.

Luppino and I have been working with Lincoln Labs to improve the QE of CCDs, illustrated in Figure 4. Once a photon gets into the silicon, the QE in the blue is determined by how effectively electrons released near the surface are swept down to the charge collection region since the optical depth for blue light is only a few hundred angstroms. In the red, the QE gently rolls off beyond about 700nm because the silicon starts getting transparent and the probability of a photon interacting during its passage through the silicon is diminished. At 800, 900, and 1000 nm the optical depth in cold silicon is 30, 50, and 250 μm . The actual process by which a photon creates an electron-hole pair involves a phonon process, so silicon also becomes more transparent as it gets colder.^{3*}

By using high resistivity silicon we are able to make the electric fields created by the gates extend deeper into the silicon, which allows us to stop the thinning process at a 40 μm thickness and still have a fully depleted CCD. This additional thickness can double the QE in the far red over a more conventional 20 μm thickness. We also can put an aluminum reflector over the gates during CCD production before the wafer is turned over, epoxied to its carrying wafer, and thinned. This also enhances the red QE by giving a double pass for some photons. It doesn't double the QE since there is a fair bit of reflection anyway off of the silicon-insulator-polysilicon-SiO₂ structure of the gates, but as Figure 4 illustrates, it does give a significant boost.

Figure 4. Theoretical predictions of the quantum efficiency of CCDs with various treatments

The situation in the blue is somewhat different. The thinning process, whether mechanical or chemical, damages the crystalline structure of the silicon and creates a lot of traps at the surface. These must be removed or they will generate a lot of dark current and pockets, and this is done at Lincoln Labs by annealing the thinned wafer with brief laser pulses. In addition, it is necessary to provide a repulsive force on electrons created at the surface to drive them down to where the field from the gates can hold them, and this is done in a variety of ways including permanent implants, thin, conductive coatings, or UV flooding and cooling. At Lincoln Labs, this surface treatment is a boron implant which creates a negative charge at the surface. Unfortunately, because of the melting from the laser annealing, this implant extends well below where UV and far blue photons are absorbed, so most of these electrons diffuse to the surface and are lost, leading to the falling QE curve seen in Figure 4. Working with LL, we are exploring new techniques to minimize the dead layer and significantly increase the UV QE, with the limit set primarily by the AR coating choice.

CCD AR coatings are an important area for improved CCD performance. Up until recently, little work has been done to develop new AR coating materials. Substantial gains may be possible with multilayer coatings, but this work has so far been limited by the lack of information on the optical constants and deposition properties of various potential coating materials. We find that the published optical constants are far from reliable, and depend on various deposition parameters. In effect, these numbers must be measured for a given deposition system and technique to be sure AR coating models reliably predict the coating performance. Work is proceeding in this area with the superb facili-

Figure 5. Theoretical predictions of reflectivity of CCDs with various AR coatings, overlaid on a plot of atmospheric emission lines.

ties available at LL. The potential benefits are enormous as outlined below. The “20 μ m EPI” curve in Figure 4 shows the QE as a function of wavelength for a 2-layer AR coating (“blue”) which Figure 5 reveals to have very respectably low reflectivity compared to a typical HfO₂-MgF₂ two layer coating. Nevertheless the amplitude of QE variations in the red which causes fringing in spectrographs and red imaging is quite large. With knowledge of the night sky airglow spectrum we have designed a different coating which has null reflectivity at 850 nm, and the “40 μ m” curve in Figure 4 and the “yellow” curve in Figure 5 show how much this can improve the fringing performance. Unfortunately we use TiO₂ for this low fringing AR coating which becomes increasingly opaque to UV light and is the major reason for the observed falloff of QE in the UV for the design shown. It is straightforward to use a simple, single-layer UV transparent AR coating material (e.g. HfO₂) and maintain very high QE in the blue, but at the cost of increased fringing in the red. We are currently actively pursuing designs of a three (or more) layer coatings which will emphasize both red and blue performance.

The Lincoln Lab and EEV CCDs are now achieving very good noise figures, $< 2e^-$ for speeds of 100kpix/sec, and $< 5e^-$ up to 1Mpix/sec. They do this by a drastic reduction of the capacitance of the sense node (gate of the output mosfet), hence the responsivity can be as large as 10–20 μ V/ e^- . This puts extreme demands on the output FET, since we require it to be a source follower with low noise and linearity to better than 1% when it may be putting out as much as a volt or two. It is often the case that the drain voltage must be carefully adjusted not for noise performance so much as to find a place where the linearity is good. This is clearly unsatisfactory, so we are experimenting

on the LL CCID28 with amplifiers at opposite corners which have a feedback capacitor. Then by providing an off-chip gain stage (with some sort of op amp) and feedback we can dramatically improve the linearity, stability, and dynamic range of this amplifier.

I cannot give you a complete description of all the CCD efforts in the world. Commercial vendors such as SITe and EEV continue to produce better and better CCDs for lower and lower cost per unit area, and efforts are ongoing at LBL to make $100\mu\text{m}$ thick CCDs which are fully depleted by application of a surface bias voltage, people at Xerox are working on *extremely* large CCDs made of amorphous silicon, Mike Lesser is achieving good things with thinning technology, and other vendors produce imagers with interesting characteristics. A particularly promising technology is that used by Rockwell for HgCdTe IR array multiplexers. With a proper material for converting photons to electrons, these multiplexers permit non-destructive readout of arbitrary pixels at high speed and low noise. These devices may be what eventually supplant CCDs. What I've tried to emphasize here is that there are real advances yet to be made in optical detector technology, and it's possible for a person such as myself who has a scientific need and a few good ideas to make a difference. I strongly encourage people to try to avoid the vendor-client mentality and to work with the CCD manufacturers to help them improve the quality of CCDs. Even such a simple thing as using a night sky spectrum to design an AR coating is something which will only come from scientists.

A final topic I'd like to tell you about is something which is obviously very pertinent to this conference on 3D imaging. I've been thinking recently about how to make a multi-color CCD. The scientific application I have in mind is low resolution ($R = 20\text{--}40$) coverage of the entire optical spectrum (and near IR with a different detector). This can be extremely useful for photometric redshifts, discovery of classes of objects such as QSOs, etc. Essentially I'm trying to create an instrument which is capable of doing a Sloan-scale survey but with $R = 40$ instead of $R = 5$.

The idea illustrated in Figure 6, is essentially to construct a microlens array which sits immediately over an array of detector pixels, but make each microlens a focusing grism (or grens). Thus each microlens defines an image pixel, and focuses all the light hitting the microlens onto different detector pixels thereby isolating different colors onto different pixels. This is very much like the lenslet micropupil array followed by a dispersive element (such as TIGER), but very much reduced in complexity and scale (and spectral resolution) so that you can imagine making a master microgrism template, cheaply duplicating it in plastic, and gluing it onto each CCD which you want to so equip. I've only begun calculating models of how this might perform, but so far the results are quite encouraging. There is a lot of parameter space to make this thing work so that it should be possible to actually implement it using an actual technology such as binary optics.

Figure 6 illustrates one implementation, where I assume $10\mu\text{m}$ CCD pixels and $60\mu\text{m}$ microgrisms (which is 0.1 arcsecond on Gemini). A single " $2\text{k}\times 4\text{k}$ " CCD then has about 500×1000 image pixels and subtends an area of about 1×2 arcminutes on Gemini. Obviously you can mosaic as many devices and cover as much area as you have money to spend. The performance of an $R=12$ design

Figure 6. An array of "microgrisms" placed over arrays of CCD pixels separate *all* the light in an image pixel into different colors and focuses it onto different CCD pixels.

Figure 7. The bandpasses from an implementation of the microgrism array. The spectral resolution is limited by the quality of focus of these microlenses, and evidently this implementation cleanly separates 50 nm bandpasses.

is quite good. Figure 7 shows how well you can separate the entire optical in to 50nm bandpasses, assuming that you use a dichroic to separate the optical spectrum into chunks which are less than a factor of 2 in wavelength to avoid second order confusion. It would be easy to push this implementation to $R = 20$, and I think that $R = 40$ is about where the "focus" of this microgrism would become comparable to the bandpass, hence failing to provide any real color discrimination at higher resolution.

I'm looking forward to investigating the various technologies which might be used to build such a device. Obviously binary optics techniques might be suitable, but I'm particularly impressed by the possibilities which might arise from using a volume phase holographic (VPH) grating, because you could imagine that a VPH gelatin could be applied directly to a CCD, the entire focussing-dispersing structure could be holographically laid down in the gelatin, developed, and then covered with protective glass. I've only begun to think about the possibility of cross-dispersing, but it is conceivable that VPH technology could also open up some real opportunities in that direction as well.

The Astronomical Potential of Optical STJs

M.A.C. Perryman and A. Peacock

Astrophysics Division, ESA-ESTEC, Noordwijk, The Netherlands

Abstract. The current state of development of superconducting tunnel junctions (STJs) as high-efficiency, energy-discriminating photon counting detectors is reviewed. Arrays of tantalum junctions have now reached a state of maturity such that serious astronomical applications can be considered over the wavelength range from 300 nm to 2 μ m. Results of our first astronomical observations, of the Crab pulsar, are presented, using a 6×6 array of tantalum STJs at the 4.2-m William Herschel Telescope on La Palma. Each array element provides photon counting capability recording up to $\sim 10^3$ photons $\text{pix}^{-1} \text{s}^{-1}$ with an arrival time accuracy of about 5 μ s, and providing a wavelength resolution of about 100 nm. This is the first time that STJs have been used in astronomy, and it confirms the feasibility of fast, high efficiency, low noise photon counting, with photon energy resolution. The agreement between the theoretical tunnel-limited resolution and laboratory measurements, for both niobium and tantalum devices, provides a strong indication that lower critical temperature superconductors will provide considerably improved energy resolution in the near future. Such devices based on hafnium, with a critical temperature of ~ 130 mK, are already under development.

1. Principles of the STJ

The possibility of intrinsic determination of individual photon energies in the optical range was first reported by Perryman et al. (1993), who proposed the application of STJ technology to optical photon counting. Incident photons break Cooper pairs responsible for the superconducting state. Since the energy gap between the ground state and excited state is only a few meV (rather than ~ 1 eV in the case of semiconductors), each individual photon creates a large number of free electrons, in proportion to the photon energy. The first experiments demonstrating single optical photon counting with energy resolution were reported by Peacock et al. (1996) using STJs, and further developments have been described by Peacock et al. (1997). Similar results using superconducting transition-edge sensors (TES) as micro-calorimeters have recently been reported (Cabrera et al. 1998), including first observations of the Crab pulsar by Romani et al. (1998).

The absorption of a photon in a superconductor is followed by a series of fast processes which involve the breaking of Cooper pairs by energetic phonons created by the hot electrons produced as the system relaxes after the initial photo-absorption. The essential result of this cascade is that the photon energy

is converted into a population of free charge carriers, or quasi-particles, in excess of any thermal population. For typical transition metals the time scale of this conversion process ranges from ns (niobium) to μ s (hafnium). At sufficiently low temperatures (typically $0.1 T_c$) the number density of thermal carriers is very small, while the average number of excess carriers N_0 created as a result of the photo-absorption process can be written as $N_0(\lambda) \sim 7 \times 10^5 / (\lambda(\text{nm}) \Delta(T/T_c))$, where the temperature-dependent energy gap $\Delta(T/T_c)$ is in meV. Thus in tantalum with $T \ll T_c$ ($T_c = 4.5$ K, $\Delta \sim 0.66$ meV) the initial mean number of free charge carriers created is $N_0(\lambda) \sim 10^6 / \lambda(\text{nm})$.

The variance on $N_0(\lambda)$ depends on the variance in the partition of the photons energy between productive phonons (i.e. phonons with an energy $\Omega > 2\Delta$ which can break Cooper pairs) and phonons which are essentially lost from the system ($\Omega < 2\Delta$). The latter population evolves with time as the average energy of the increasing quasiparticle population relaxes towards the bandgap through quasiparticle phonon emission. $\langle N_0 \rangle$ depends on the superconductor bandgap Δ and the 'Fano factor' F , where $F \sim 0.2$ for elemental superconductors such as niobium and tin (Kurakado 1982, Rando et al. 1992). Expressing this variance in terms of the wavelength resolution gives $d\lambda_F(\text{nm}) \sim 2.8 \times 10^{-3} \lambda^{3/2} (F \Delta(T/T_c))^{1/2}$, which therefore represents the fundamental Fano-limited wavelength resolution of any superconductor. Thus a superconductor such as tantalum with $T \ll T_c$ irradiated with photons of wavelengths covering the ultraviolet to the near infrared (100 and 1000 nm) yields $d\lambda_F \sim 1.1$ and 34 nm respectively (Peacock et al. 1997, Peacock et al. 1998).

The quasi-particles produced through photo-absorption can be detected by applying a dc potential across two such films separated by a thin insulating barrier, forming a superconducting tunnel junction (STJ). This potential bias favors the transfer of quasi-particles from one film to the other through quantum mechanical tunneling across the barrier. The detector signal is therefore represented by the current developed by this tunnel process. After initial tunneling, a quasiparticle can tunnel back across the barrier, therefore contributing many times to the overall signal (Grey 1978). On average, each quasiparticle will contribute $\langle n \rangle$ times to the signal through an average of $\langle n \rangle$ tunnels before it is lost from the system, for example through traps (Poelaert et al. 1998). Hence the mean number of effective charge carriers $N = nN_0$. The multiple tunnel process leading to n , the average number of tunnels per quasiparticle, is also subject to statistical fluctuation (Goldie et al. 1994). The fluctuations due to the Fano and tunnel processes can be added in quadrature, and the overall limiting resolution for a perfectly symmetrical STJ can be written as $d\lambda_T(\text{nm}) \sim 2.8 \times 10^{-3} \lambda^{3/2} \Delta(T/T_c)^{1/2} (F + 1 + 1/n)^{1/2}$ for $n \geq 2$. Figure 1(a) illustrates this tunnel junction-limited resolution for a number of elemental superconductors. Typically $n \sim 10 - 100$, depending on the details of the STJ.

In optical and UV spectroscopy high resolution normally implies a resolving power $\lambda/\Delta\lambda > 10^4$. From Figure 1(a) it is clear that none of the classical superconductors forming the basis of current STJs under development (those based on Nb, Ta, Al, Mo or Hf) will achieve such a resolving power. In fact a superconducting critical temperature $T_c \ll 100 \mu$ K would be required to achieve a resolving power of 10^4 , suggesting the development of STJs based on superconductors such as rhodium. However, the temporal characteristics associated with

Figure 1. (a) The tunnel-limited resolution of a number of elemental superconductors as a function of wavelength from the X-ray to the near infrared. The experimental data derived from niobium and tantalum STJs is also shown. (b) The charge spectrum from irradiation of a tantalum STJ by photons of wavelength 1183 nm. The various orders from the grating monochromator are easily discernible, and indicate the excellent linearity of the device.

the production of the free excess charge carriers is also a function of the critical temperature, while phonons with $\Omega > 2\Delta$ have wavelengths significantly larger than the thickness of the film. Thus such low temperature superconductors may well be significantly slower in their overall time response.

This moderate spectral energy resolution nevertheless comes with very high quantum efficiency across the UV, optical and into the near IR, excellent (if not state of the art) timing capabilities, and low noise characteristics (read-out noise is non-existent, and dark noise can theoretically be decreased to very low levels). These characteristics make the detector valuable for the study of rapid temporal variability of isolated objects such as pulsars, white dwarfs, and interacting binaries. Other applications can similarly be foreseen: for faint object redshift determination, atmospheric correction including adaptive optics, interferometric path delay determination, as an order separator in high-resolution spectroscopy, etc.

2. Current Performance of STJs

The key factors described in this section involve the basic performance of tantalum based STJs building on the earlier work in niobium. The validation of the basic equations for N_0 and $d\lambda(T)$ with both tantalum and niobium devices give some confidence in the ultimate development of lower temperature STJs such as those based on hafnium (Kraft et al. 1999).

At optical and UV wavelengths, where the photon energy is relatively small, spatial effects on the resolution are unimportant. Rather, the signal is low, such that the signal-to-noise ratio is the dominant factor governing the measured

resolution. At these wavelengths the photons enter the detector through a substrate, which can either be sapphire or magnesium fluoride, depending on the short wavelength cut-off required (Peacock et al. 1998). The theoretical efficiency of a tantalum device deposited on a sapphire substrate with this mode of illumination is very high. All photons are absorbed in the high-quality epitaxial tantalum base film. Efficiencies of $\sim 70\%$ from 200–600 nm are expected, and are limited at the short wavelength by the cut-off of the sapphire substrate. Such efficiencies have been experimentally confirmed.

To illustrate the broad-band response of this type of photon counting spectroscopic detector, Figure 1(b) shows the charge spectrum from a single tantalum based device when illuminated with optical light via a grating monochromator. This grating response covers four orders from 296–1183 nm, i.e., from the UV to the near IR. Not only are the various orders well resolved but the charge output as a function of wavelength can be precisely determined leading to a measured wavelength linearity which is very high. These types of measurements allow the determination of the wavelength resolution across a broad wave band, and are shown for both tantalum and niobium based devices in Figure 1(a).

The STJ based on tantalum or niobium has now been developed to a stage where practical small format arrays (6×6 pixels) have been produced, which provide similar performance to optimised single devices. Such arrays (Figure 2) are now being developed into instruments for ground-based optical astronomy (Rando et al. 1999). The performance of these arrays at UV and soft X-ray wavelengths are such that practical instruments can now be considered for space-based applications. The key specific points which have been experimentally demonstrated can be summarized as: (a) tunnel-limited resolutions have been achieved at optical and UV wavelengths; (b) high efficiency has been shown at UV and optical wavelengths; (c) resolutions are within a factor of two of the tunnel resolution in the optical; (d) high-speed photon counting has been demonstrated, at count rates up to about 10 kHz.

Areas where our effort is ongoing for optical applications are the production of larger format arrays, the development of large format application-specific readout electronics, and the development of lower T_c devices, with targets of $\Delta\lambda \sim 1$ nm in the not-too-distant future.

3. Observations of the Crab Pulsar

PSR B0531+21 in the Crab Nebula was first observed as an optical pulsar by Cocke et al. (1969), and provides an excellent target for verification of the system's astronomical performance. Along with PSR B0833–45 in Vela (Wallace et al. 1977), PSR B0540–69 in the LMC (Middleditch & Pennypacker 1985) and more recently PSR B0656+14 (Shearer et al. 1997) and possibly Geminga (Shearer et al. 1998) it remains one of the few pulsars observed to emit pulsed optical radiation.

While the ~ 33 ms period pulsar has been extensively studied at all wavelengths including the optical (e.g. Percival et al. 1993, Eikenberry et al. 1996, Nasuti et al. 1996, Eikenberry & Fazio 1997, Gull et al. 1998, Martin et al. 1998) it continues to be important in providing new insights into the nature of the pulsar emission mechanism: the pulse profile shape, the separation of the

Figure 2. A high-resolution image of the 6×6 tantalum STJ array used for the first astronomical observations. Each junction is $25 \times 25 \mu\text{m}^2$.

primary and secondary emission peaks by 0.4 in phase and recent results on the energy dependence of the pulse shape over the infrared to ultraviolet range (Eikenberry et al. 1996, Eikenberry & Fazio 1997) provide a challenge to theoretical models in which γ -rays created through curvature radiation interact with the pulsar magnetosphere to produce the X-ray, ultraviolet, optical and infrared pulsations through a variety of energy-loss mechanisms. A photon counting detector with energy resolution in the optical offers an important possibility to examine further the energy dependence as a function of pulse phase.

Our prototype 6×6 tantalum STJ array covering an area of $\sim 4 \times 4 \text{ arcsec}^2$ was operated at the Nasmyth focus of the William Herschel Telescope on La Palma in February 1999. Photon arrival time information was recorded with an accuracy of about $\pm 5 \mu\text{s}$ with respect to GPS timing signals; while the latter is specified to remain within $1 \mu\text{s}$ of UTC, typical standard deviations are much less (Kusters 1996). Observations of the Crab Pulsar were made on 4–6 Feb, although modest seeing ($> 2 \text{ arcsec}$), especially poor on the first two nights, and a significant number of unstable junctions, meant that total intensities could not be determined reliably. Our present analysis is restricted to the signal extracted from just 6 pixels (corresponding to an indeterminate but small (~ 0.1) fraction of the overall PSF), and a consideration of the resulting pulse profile and its energy dependence (Perryman et al. 1999).

The wavelength response is intrinsically very broad (from $< 300 \text{ nm}$ to $> 1000 \text{ nm}$) but was restricted in the present system to about $300\text{--}700 \text{ nm}$, as a result of the atmosphere ($\sim 300 \text{ nm}$) and the optical elements required for the suppression of infrared photons ($\sim 700 \text{ nm}$). The detector quantum efficiency is around 70% across this wavelength range, limited by the device/substrate geometry rather than by the intrinsic detector response. Count rate limits are about $10^3 \text{ photons s}^{-1}$, determined by the output stage electronics, although

Figure 3. (a) pulse profile for the Crab pulsar from the 6 Feb 1999 data over the range 310–610 nm (128 phase bins). Photon statistical error bars are included. During the 10 min interval of best seeing, the total counts summed over the 6 active pixels corresponds to about 1000 photons s^{-1} , roughly half of which originate from the unpulsed component (including sky) and half from the pulsed component. (b) normalized 32-phase bin pulse profiles for the 6 Feb data divided into the $E_1 = 310\text{--}410$ nm (upper curve) and $E_2 = 500\text{--}610$ nm (middle curve) energy channels, with $\lambda_{\text{eff}} \sim 378$ and 553 nm respectively. The upper curve has been shifted vertically by 1.0. The resulting colour ratio versus phase, constructed as $E_2 - E_1 / E_2 + E_1$ ($E_1 = 310\text{--}410$ nm and $E_2 = 500\text{--}610$ nm), is shown in the lower panel; the ordinate is the integrated photon content of each of the summed energy channels (arbitrary scale).

the device relaxation time is much faster, being below $\sim 10 \mu\text{s}$ for the present device. The current wavelength resolution, ~ 100 nm at 500 nm, is driven by system electronics and residual thermal background (IR) radiation, although the intrinsic response of the Ta junctions is some factor of 5 better than the present performance.

Data from the STJ are archived in FITS format with the photon records defined by their arrival time, x, y pixel coordinate, and energy channel in the range 0–255. Channels 50–100 cover $\Delta\lambda \sim 610 - 310$ nm, with $\lambda(\text{nm}) \sim 1238.5 / (m \times N_{\text{ch}} + c)$, where N_{ch} is the channel number, $m \sim 0.04$, and $c \sim 0.03$. Energy calibration was performed using an internal calibration source before and after the target observations, and verified using narrow-band filter observations of a standard star. Photon arrival times were translated to the solar system barycentre using the JPL DE200 ephemeris, taking into account gravitational propagation delay. Our reference timing ephemeris for the Crab pulsar used the 15 Feb 1999 (MJD = 51224) values of $\nu = 29.856\,514\,436\,4$ Hz and $\dot{\nu} = -374\,886.90 \times 10^{-15} \text{ s}^{-2}$ taken from the radio ephemeris of Lyne et al. (1999). Consistent periods were obtained, with a precision of typically 5×10^{-8} s, from a period search of the timing data.

Figure 3(a) shows the light curve from the 50 min of data obtained on 6 Feb, acquired with a time resolution of 5 μ s, and folded into 128 phase bins ($\sim$ 250 μ s per bin), with an arbitrary origin of zero phase, and without background subtraction (impractical due to the combination of poor seeing, small field, and fraction of unstable junctions; the overall system response is undetermined for similar reasons, combined with uncalibrated losses at the derotator entrance aperture). Examination of the light curve, including the peaks, at finer temporal resolution down to about 30 μ s per phase bin, reveals no significant sub-structure persisting over the observation interval.

Figure 3(b) shows the data divided into two separated energy channels, corresponding to $E_1 = 310\text{--}410$ nm and $E_2 = 500\text{--}610$ nm (energy channels 77–99 and 50–61 respectively). The profiles are normalized to the same relative intensities in the peaks, and are displaced vertically. For this choice of energy bins, the ratio of photons is roughly $E_2:E_1 = 4:1$. The resulting colour ratio, constructed as $E_2 - E_1 / E_2 + E_1$ and folded at the pulsar period, is shown in the lower part of the figure.

Wavelength-dependent variations across the peaks has been noted by Eikenberry et al. (1996), based on observations spanning the UV to the infrared K band, and by Sandoval et al. 1998. Not surprisingly given the form of these reported variations, no significant variations with pulsar phase are evident in our data. From the nebular emission-line spectrum reported by Davidson (1979) the [O II] doublet (λ 3726, 3729) lies within our extracted blue passband, while the red passband selected comprises none of the strong emission lines in the red part of the spectrum ([N II] 6584, H α 6563, [N II] 6584, [Si II] 6717, 6731), and only marginally collects photons from the [O III] doublet (4959, 5007). Our *a posteriori* choice of energy channels for this ratio provides sufficient wavelength separation to avoid ‘contamination’ of each bin given the low energy resolution of the device. Further discrimination of the nebular contributions is limited by our presently modest energy resolution, which is also insufficient to confirm the reality of the broad absorption feature near 5920 Å, so far noted only by Nasuti et al. (1996).

These observational results can be compared with predictions from models in which the observed flux in a given phase interval is a combination of emission from effectively disjointed physical regions, in which the sections contributing to the emission in a given phase interval depend on the viewing geometry (e.g. Cheng et al. 1986a, Cheng et al. 1986b). This mixing of physical regions acts to average the total emission, and predicts that observed properties such as the energy ratio remain constant, or change modestly but rapidly close to the emission peak (Romani 1996). Higher S/N STJ data will be required to probe the small, rapid spectral index variations across the pulse peaks reported by Sandoval et al. (1998).

As the first application in astronomy of a superconducting tunnel junction detector capable of providing intrinsic wavelength resolution in the optical, the results are a modest indication of the technology’s capabilities for the future. Significant improvements in performances, and in particular in the wavelength resolution, are expected from the use of lower critical temperature superconductors in the future.

Acknowledgments

We acknowledge the contributions of the ESTEC/ESA team involved in the STJ detector development, in particular D. Goldie, R. Hart, S. Kraft, D. Perez, A. Poelaert, N. Rando, and P. Verhoeve; and to those involved in the first astronomical observations, in particular N. Rando, F. Favata, S. Andersson, J. Verveer. We acknowledge the assignment of telescope time at the William Herschel Telescope, and the excellent support given to the instrument's commissioning, in particular by P. Moore and C.R. Benn.

References

- Cabrera B., Clarke R. M., Colling P. et al. 1998, *Appl. Phys. Lett.* 73(6), 735
Cheng K. S., Ho C., Ruderman M. 1986a, *ApJ* 300, 500
Cheng K. S., Ho C., Ruderman M. 1986b, *ApJ* 300, 522
Cocke W. J., Disney M. J., Taylor D. J. 1969, *Nat* 221, 525
Davidson K. 1979, *ApJ* 228, 179
Eikenberry S. S., Fazio G. G. 1997, *ApJ* 476, 281
Eikenberry S. S., Fazio G. G., Ransom S. M. et al. 1996, *ApJ* 467, L85
Goldie D. et al. 1994, *Appl. Phys. Lett.* 64, 3169
Grey K. 1978, *Appl. Phys. Lett.* 32, 392
Gull G. R., Lindler D. J., Crenshaw D. M. et al. 1998, *ApJ* 495, L51
Kraft S. et al. 1999, *SPIE* in press
Kurakado M. 1982, *Nucl. Instr. Meth.* 196, 275
Kusters J. A. 1996, *Hewlett-Packard Journal* December 1996, 60
Lyne A. G., Pritchard R. S., Roberts M. E. 1999, <http://www.jb.man.ac.uk/pulsar/crab.html>
Martin C., Halpern J. P., Schiminovich D. 1998, *ApJ* 494, L211
Middleditch J., Pennypacker C. 1985, *Nat* 313, 659
Nasuti F. P., Mignani R., Caraveo P. A., Bignami G. F. 1996, *A&A* 314, 849
Peacock A. et al. 1998, *A&AS* 127, 497
Peacock A., Verhoeve P., Rando N. et al. 1996, *Nat* 381, 135
Peacock A., Verhoeve P., Rando N., Perryman M. A. C., Taylor B. G., Jakobsen P. 1997, *A&AS* 123, 581
Percival J. W., Biggs J. D., Dolan J. F. et al. 1993, *ApJ* 407, 276
Perryman M. A. C., Favata F., Peacock A., Rando N., Taylor B. G. 1999, *A&A* 346, L30
Perryman M. A. C., Foden C. L., Peacock A. 1993, *Nuc. Inst. Meth. A* 325, 319
Poelaert A., Kozorezov A., Peacock A., Wigmore K. 1998, *Phys. Rev. Lett.*
Rando N. et al. 1992, *Nucl. Instr. Meth.* A313, 173
Rando N. et al. 1999, *SPIE* in press
Romani R. W. 1996, *ApJ* 470, 469
Romani R. W. et al. 1998, *BAAS* 30, 1266
Sandoval J. L. et al. 1998, *BAAS* 30, 1420
Shearer A., Golden A., Harfst S. et al. 1998, *A&A* 335, L21
Shearer A., Redfern R. M., Gorman G. et al. 1997, *ApJ* 487, L181
Wallace P. T. et al. 1977, *Nat* 266, 692

5000 by 5000 Spatial by 15000 Spectral Resolution Elements: First Astronomical Observations with a Novel 3-D Detector

C. U. Keller

National Solar Observatory/NOAO¹, Tucson, AZ 85719

Abstract. I present the first three-dimensional observations of the Sun obtained with a novel, highly wavelength-sensitive imaging detector. The all-optical device has a spectral resolution of 0.001 nm over a band of 15 nm centered at 634 nm and a spatial pixel size as small as a few microns squared over an area of several square centimeters. It is based on a dye-doped polymer film immersed in liquid helium that is exposed like film and can store true 3-D images with high spatial and spectral resolution. The optically stored information is read out with a tunable laser and a fast CCD camera. The characteristics of this novel detector have been studied with extensive laboratory tests, while observations of the Sun provided the first astronomical data. These initial experiments pave the way for astronomical detectors that improve the efficiency of spectroscopy of extended sources by orders of magnitudes.

1. Introduction

The ideal astronomical detector would collect all information carried by all photons that hit the detector. Such a detector would not require filters, spectrographs, or polarimeters. In particular, the detector should

1. be wavelength sensitive with a spectral resolution of several 100,000
2. be sensitive to linear and circular polarization
3. cover the whole focal plane
4. have small spatial elements on the order of a few μm
5. cover the spectral range transmitted by the Earth's atmosphere
6. have a quantum efficiency of 1.0
7. allow measurements at high temporal cadence

While such an ideal detector has not yet been invented, we have demonstrated a technology that satisfies requirements 1. and 4., and partially satisfies 2., 3., and 7 (Keller et al., 1994, 1995). Those initial experiments showed the high

¹Operated by the Association of Universities for Research in Astronomy, Inc. (AURA) under cooperative agreement with the National Science Foundation

Figure 1. The change in the structure of chlorin when illuminated by light.

wavelength sensitivity of the detector but did not provide true 3-D data. Here we present the first truly three-dimensional observations of the Sun recorded with the same technology. All results reported here were obtained at the Laboratory for Physical Chemistry of ETH Zurich, Switzerland.

2. Spectral Hole-Burning Device (SHBD)

Color film is a true 3-D detector, although with a very low spectral resolution of about 3 due to the 3 different dyes. Since the number of dyes cannot be increased by much more than a factor of 100, more dyes would only provide a very limited spectral resolution. Another approach consists in using a single type of dye molecule with a very narrow absorption line and embed it in an amorphous plastic film. Such a film provides many (electrically) different environments. Suitable molecules can exhibit large shifts of their absorption lines depending on their environment. If the doped film is cooled to very low temperatures, the environment, and therefore the spectral position of the absorption line of a particular molecule, is stable in time. Of course, the dye molecule must change its structure once it has absorbed a photon. After such a photo-reaction, the molecule is no longer sensitive to photons within the original spectral range.

In the past, dye-doped plastic films have been used to store digital data. By illuminating with discrete, monochromatic light, the spectral absorption characteristics at a single spatial position can be permanently changed. This is called persistent spectral hole-burning (see Moerner, 1988, Sild and Haller, 1988, and on the web at

<http://decryptor.uoregon.edu/~mosswww/memory/pshb.html>

for more information). The spectral 'holes' are detected by measuring the absorption spectrum at each spatial location with a tunable laser. While optical data storage applications use spectral hole-burning in a digital mode where each spatial and spectral element defines a single binary digit, we use it to store and read out continuous, analogue spectra at each spatial location. In analogy with charge-coupled devices (CCD), we call this technology the Spectral Hole-Burning Device (SHBD).

2.1. Chlorin in polyvinylbutyral (PVB)

The dye-doped polymer film used for our experiments consists of polyvinylbutyral (PVB) doped with chlorin (2,3-dihydroporphyrin). At 1.5 K the photon-induced transition of chlorin in PVB (see Fig. 1) has a spectral width of only 0.0002 nm. The absorption band of chlorin is centered at 633.5 nm and has a useful width of about 15 nm. The back-reaction can be induced by illuminating the film with white light or heating to about 10 K. The chlorin-in-PVB system has been well studied in the past and has been successfully used for holography and optical computing (e.g. De Caro et al., 1991).

2.2. Properties of the experimental SHBD

The experimental SHBD has the following properties. The numbers reflect the corresponding items in the introduction.

1. spectral resolution of several 100,000
2. sensitive to linear polarization because of the dipole moment of the transition
3. spatial size of about 2 cm by 2 cm limited by the size of the cryostat
4. spatial resolution element on the order of a few μm (as revealed by holography)
5. covers about 15 nm around 634 nm
6. quantum efficiency $< 0.1\%$
7. if the wavelength range is limited, an externally applied electrical field can shift the absorption lines via the Stark effect, therefore allowing rapid collection of images (c.f. De Caro et al. 1991).

3. Observations

3.1. Test with iodine spectrum

Initial tests to assess the wavelength sensitivity of the device have been based on iodine spectra obtained by illuminating the SHBD with the light from a broad-band source that has passed an iodine cell. After the illumination, the SHBD was read out by measuring the absorption with a tunable dye laser. A reference spectrum was obtained by measuring the iodine absorption directly with the tunable laser. Figure 2 shows the reference spectrum along with the SHBD spectrum. The comparison between the two spectra shows the very high spectral resolution of the SHBD. The same setup was used to record a fiber-fed solar spectrum (see Keller et al. 1994 for details).

3.2. Solar rotation

The Sun is the astronomical object that is best suited to be studied with a SHBD because the high intensity requires only a small telescope, and observations are feasible with the low quantum efficiency material. Furthermore, spectral and spatial structures have been well studied with other instruments, which allows for easy comparison with existing data.

Figure 2. Comparison of iodine reference spectrum with an iodine spectrum obtained with the experimental SHBD (from Keller et al. 1995). Notice that the wavelength interval covers less than $1\text{\AA}$.

A simple 15-cm solar telescope was built to record true 3-D images of the Sun. The light from the telescope passes a broad interference filter whose transmission overlaps the absorption band of chlorin. This prevents the photochemical back-reaction of the photo product. The dye-doped polymer is located in superfluid liquid helium at a temperature of 1.5 K. Based on experience with the iodine and fiber-fed solar spectra, an exposure time of 240 s was chosen. For read-out, a tunable dye-laser illuminated a diffuser in front of the SHBD, which reduces the coherence of the laser light and therefore reduces interference fringes in the various windows of the cryostat and the CCD camera. For each wavelength position of the laser, the film's transmission was recorded with a fast CCD camera running at up to 25 frames per second. The scene detected by the CCD camera includes the solar image as well as a reference area that does not contain dye molecules. The latter area serves as a reference channel to detect fluctuations of the illuminating laser. A more detailed description of the optical setup can be found in Rosselet et al. (1995). The CCD camera was connected to a high-speed multiprocessor computer, which provides real-time preprocessing and data storage (see Graff et al. 1995). During exposure, the CCD camera is used to guide the telescope and measure the illuminating solar flux using the light that is transmitted by the film.

Figure 3. Solar rotation as seen with the SHBD.

The data shown here has a 3-D volume of 284 by 393 spatial pixels and 46,000 spectral positions. The spatial sampling is only limited by the small CCD. Figure 3 shows the spatial velocity distribution due to solar rotation as measured in a single spectral line. The velocity measurement could be much improved by using many lines simultaneously and by properly removing fringe patterns due to the various optical windows.

4. Outlook

The initial observations obtained with the experimental SHBD clearly show the potential of this new technique for astronomy. The big advantage compared to other wavelength-sensitive detectors outlined in these proceedings is the inherently high spectral resolution. Since the SHBD is also sensitive to linear

polarization adding two more dimensions (degree of polarization and angle), a SHBD is really a five-dimensional detector.

The next logical step is the development of a prototype SHBD for use with astronomical telescopes. To this end, we need to find better dye molecules that offer

- signal-to-noise ratio of at least 100
- quantum efficiency larger than 1%
- spectral band pass of at least 10 nm
- be in the spectral range of 350 to 1100 nm

Apart from using persistent spectral hole-burning as a 3-D detector, it could also be used as a versatile, narrow-band filter that can transmit or block several wavelengths simultaneously. An interesting application of this would be an OH suppression filter (Oliva 1995, personal communication).

I am grateful to Armel Rosselet and Werner Graff for performing all the experiments, which were funded by the Swiss National Science Foundation, and to U. Wild, R. Gschwind, A. Renn, and J.O. Stenflo for their continuous support.

References

- De Caro, C., Renn A., Wild U.P., 1991, *Appl. Optics* 30, 2890
- Graff W., Rosselet A., Wild U.P., Gschwind R., Keller C.U., 1995, Handling of huge multispectral image data volumes from a Spectral Hole Burning Device (SHBD), *SPIE Proceedings* 2480, 445-452
- Keller C.U., Gschwind R., Renn A., Rosselet A., Wild U.P., 1995, The Spectral Hole-Burning Device: a 3-dimensional photon detector, *A&AS* 109, 383
- Keller C.U., Graff W., Rosselet A., Gschwind R., Wild U.P., 1994, First Light for an Astronomical Three-Dimensional Photon Detector, *A&A* 289, L41
- W.E. Moerner (Ed.), *Persistent Spectral Hole-Burning: Science and Applications*, *Topics in Current Physics* 44, Springer-Verlag, Berlin (1988)
- Rosselet A., Graff W., Wild U.P., Keller C.U., Gschwind R., 1995, Persistent spectral hole burning used for spectrally high-resolved imaging of the sun in solar astronomy, *SPIE Proceedings* 2480, 205-212
- O. Sild, K. Haller (Ed.), *Zero-Phonon Lines and Spectral Hole-Burning in Spectroscopy and Photochemistry*, Springer-Verlag, Berlin (1988)

Novel Interferometer Spectrometer for Sensitive Stellar Radial Velocimetry

David J. Erskine¹ and Jian Ge²

Lawrence Livermore Nat. Lab., Livermore CA 94550

Abstract.

We describe a new kind of stellar radial velocimeter based on the series combination of a wide angle Michelson interferometer and a disperser, and which we call a fringing spectrometer. The simplest instrument response of the interferometer produces smaller instrumental noise, and the low resolution requirements of the disperser allows high efficiency and creates an etendue capability which is two orders of magnitude larger than current radial velocimeters. The instrument is compact, inexpensive and portable. Benchtop tests of an open-air prototype shows the short term instrumental noise to be less than 0.76 m/s. A preliminary zero point drift of 4 m/s is already competitive with traditional instruments, in spite of the lack of obvious environmental controls and a known interferometer cavity length drift. We are currently installing cavity stabilization and other improvements that will lead to testing on starlight.

1. Introduction

Velocity noise in current stellar Doppler velocimeters used to detect extrasolar planets typically ranges 3 to 13 m/s, with the best noise from the Lick Observatory and Keck HIRES grating spectrometers (Butler 1996). This noise level is insufficient to reliably detect a saturn-like planet, which would have a 3 m/s amplitude velocity signature, and insufficient to resolve stellar photosphere dynamics having 1–2 m/s amplitude analogous to solar 5-minute ringing. A component of this observed noise is due to photon statistics, another due to the source such as unresolved motion of photosphere, and the remainder due to instrumental noise.

Our goal is to develop a new kind of stellar radial velocimeter which has lower instrumental noise and which could be made very light efficient. Both goals can be achieved by combining a wide angle Michelson interferometer in series with a disperser. The simple and sinusoidal instrument response of the interferometer produces smaller instrumental noise, and the low resolution requirements of the disperser allows it to be optimized for high efficiency instead of narrow point spread function (PSF). Ancillary benefits include low cost, com-

¹erskine1@llnl.gov

²ge1@llnl.gov

pactness and portability. These alone make the device attractive, as it allows smaller institutions to join in the search for extrasolar planets with lower capital investment.

A further advantage is that the instrument etendue capability is two orders of magnitude larger than for current radial velocimeters (9×10^{-4} sr cm² vs. 5×10^{-6} sr cm²). In addition to allowing multiple targets, this indirectly improves efficiency by allowing the use of tolerant wide diameter optical fibers, so to reduce clipping the wings of a blurry star image. This makes it more feasible to use poorer optical quality but inexpensive large collecting area "light buckets" for Doppler spectroscopy. The surplus etendue can also be used to illuminate several interferometers in parallel. When these have different path length differences, the additional fringe shift information can determine the spectral lineshape asymmetry shift simultaneous to the Doppler shift. This provides additional information about the photosphere dynamics, which is increasingly relevant when seeking velocity signatures in the ~ 3 m/s range.

Current stellar radial velocimeters are based on the diffraction grating. Their instrumental noise is due to difficulty in accurately characterizing the PSF, so that its convolving effects can be removed from the spectral data. Since a grating is a N-beam interference device, where N is in the hundreds or thousands, its PSF has at least N degrees of freedom. That is, the PSF is not a simple mathematical function. Furthermore, the PSF varies with average wavelength across the spectrum so that a family of PSF model functions are needed, and the PSF varies in complicated ways under the insults of air convection, thermal gradients and coating degradation.

2. Fringing spectrometer instrument

Our approach is to use a 2-beam interference device (an interferometer) to resolve the Doppler shifts, so that the relevant instrument response has only 3 degrees of freedom: fringe phase, amplitude and offset. Since the interferometer instrument response is sinusoidal, it is a standard mathematical function. The vastly simpler instrument response allows smaller velocity noise. Under the same insults of air convection, thermal drifts, coating degradation etc., the instrument response remains the same mathematical form. This constancy gives the data analysis algorithm robustness, as well as processing speed. Our preliminary codes (written in high level macro language, not yet optimized for speed) process at about 1 minute CPU time per velocity datum on a Macintosh. In comparison, Butler and Marcy et al. report data processing speed of 2 hours per datum on a Sparc10 workstation (Butler 1996).

However, an interferometer used alone on broadband absorption spectra such as starlight will produce poor signal to noise, because fringes from neighboring spectral lines overlap and confuse their phases. This has discouraged previous use of broadband interferometry for this application. Our solution is to separate the major spectral lines by combining the interferometer in series with a low resolution disperser, to produce fringing spectra. We call our device a fringing spectrometer or spectrally dispersed interferometer.

Figure 1 shows a photograph of our first prototype fringing spectrometer whose purpose is to test the ability to measure broadband fringe shifts with

Figure 1. Our first prototype fringing spectrometer is the series combination of an iodine vapor cell, interferometer, and conventional 0.6 meter grating spectrometer.

sunlight or convenient laboratory sources. The source light conducted through an optical fiber can pass through an iodine vapor cell inserted into the path. This provides an absolute reference absorption spectrum. The light passes through a wide-angle Michelson interferometer (Hilliard 1966) having approximately 11.5 mm path length difference. The interferometer imprints fringes to the beam which are imaged to the slit of the disperser, the fringes being perpendicular to the dispersion direction.

In contrast with Fourier transform spectrometers, our instrument has a nominally fixed path length difference and detects the Doppler shift through the change in fringe phase, averaged over all spectral channels. (For an 11.5 mm delay and 540 nm average wavelength, a 1 m/s Doppler shift corresponds to 1/14000th of a fringe.) However, one of the interferometer mirrors is attached to a PZT actuator so that the path length can be dithered in small ($\approx 90^\circ$) phase steps. Making repeated measurements while stepping the overall phase helps distinguish true fringes from similar appearing intensity variations due to other causes. However, it is possible to make measurement in a single exposure if the CCD images are sufficiently corrected for flat field.

Presently we use a disperser that disperses in one dimension. However we anticipate that the technique may eventually be used with dispersers such as echelle grating spectrometers that disperse in two dimensions and have the advantage of wide bandwidth.

The disperser initially used for our prototype was a conventional 0.6 meter spectrometer (Jobin-Yvon HR640) shown in Fig. 1. This had the advantages of adjustable slit width and center wavelength, but was not especially fast ($f/8$), compact or efficient. Presently, we are substituting for our disperser a high efficiency (80% for unpolarized) volume phase holographic grating, having 90 nm bandwidth, fast $f/1.8$ optics and very compact size (0.008 m^3). With this dis-

Figure 2. a) Snippet of fringing spectrum when slits are narrow so that underlying interferometer comb can be resolved. b) is model, showing moiré effect between absorption lines and interferometer comb, which creates vertically sinusoidal fringes c).

perser the overall instrument could be made the size of a TV-set, and portable. The prototype instrument was assembled rapidly from off the shelf components for a few thousand dollars (outside of the CCD detector). This is in contrast with current radial velocity spectrometers such as the Keck HIRES spectrometer, which cost several million dollars and are the size of a kitchen. Their high cost is due to the need for structural rigidity between the massive optical components spaced at large distances. In contrast, the key component of the fringing spectrometer is the interferometer, which has optical components spaced only a few mm from each other.

3. Fringing spectroscopy technique

The fringing spectrometer technique is based on the moiré effect. This occurs between the fringe comb generated by the interferometer and the absorption or emission lines of the source. A small segment of a fringing spectrum is shown in Fig. 2a, taken with slits narrow enough to resolve the interferometer comb. As modeled in Fig. 2b, the overlay of the comb with the solar absorption lines creates moiré fringes having sinusoidal character in the direction along the slit (Fig. 2c). Essentially, the moiré effect heterodynes the high spatial frequencies of the spectrum (which contain the Doppler shift and lineshape asymmetry information) to low spatial frequencies which are resolved by the disperser. The "carrier" frequency, which is the underlying interferometer comb, needs to be filtered away to reveal the moiré fringe. Low pass filtering can be accomplished numerically, or optically by using wider spectrometer slits. Fig. 3 shows a fringing spectrum taken with wider slits showing the moiré fringes without resolving the underlying interferometer comb. The use of wider slits has the advantage of increasing the flux for blurry sources, although it can diminish the fractional

Figure 3. A snippet of fringing spectrum measured when slits are wide so that underlying interferometer comb is not resolved, showing the moiré fringes.

Figure 4. A "whirl" is a meandering vector versus wavelength that describes the moiré fringe phase and amplitude, and is the fundamental data format for our technique during analysis. Under a Doppler shift the starlight whirl component rotates relative to the iodine whirl component.

fringe amplitude. Because these are competing effects, the optimum slit width will depend on the etendue presented by the source and the mean feature to feature distance in the spectrum.

The fundamental format of the spectra used for analysis is a vector versus wavelength channel, which we call a "whirl" (Fig. 4), where each vector describes the moiré fringe phase and amplitude. Let us denote the reference spectrum "iodine", and the source "starlight" regardless of their true nature. Basis whirls, which are whirls of the starlight alone, and the iodine alone illuminated by featureless white light, are measured separately. Then for each velocity datum the starlight is measured through the iodine vapor cell so that the net whirl contains both stellar and iodine components. The change in rotation of the starlight whirl component relative to the iodine whirl yields the change in Doppler velocity after scaling (14 km/s per cycle).

4. Preliminary results

In order to measure the noise floor of our instrument we desire a velocity source that is clean on a 1 m/s scale. Sunlight has the problem of fluctuating with 1–2 m/s amplitude due to its 5-minute photosphere oscillations. Instead, we use the bromine absorption spectrum (backlit by incandescent light) over a 12 nm bandwidth. This is sufficiently distinct from the iodine absorption spectrum, used as reference. Figure 5 shows that during the first 16 minutes the apparent velocity of the stationary bromine source is measured with a repeatability of 0.76 m/s. In the first 16 minute data set approximately 5×10^9 photoelectrons were used for each velocity datum, distributed over 2500×63 pixels in the wavelength and slit directions, respectively. The fluxes were 4.6 and 3.5×10^9 photoelectrons in the next data sets, taken 3 hours, and another 11 days later, but still using basis whirls of the first set. From Monte Carlo simulations we conclude that the observed velocity scatter for the first data set is dominated by photon noise and that therefore the instrumental component could be several times smaller than 0.76 m/s. How much smaller is difficult to assess accurately until we can repeat the measurement with much higher flux. This level of noise compares very favorably with the instrumental noise of the Lick Observatory and Keck HIRES spectrometers, which we infer from Butler's report (1996) to be $\simeq 2$ m/s. (They report seeing 3 m/s total error with a calculated 2 m/s photon noise contribution, and we assume the errors subtract in quadrature.) However, a strict comparison is unfair at this preliminary stage since the instruments are observing different sources under different conditions.

The second and third data sets of Fig. 5 taken after 3 hours and again at 11 days shows a zero point drift of not more than 4 m/s. This is already competitive compared to the 3–13 m/s typical errors of fully matured traditional spectrometers, in spite of the lack of environmental controls in our prototype, which consisted of bare optics on an optical table. We made no effort to prevent air convection, stabilize the cavity length, nor regulate temperature and pressure of the bromine vapor, which also effects the result through the variation of its absorption depth. In similar conditions we have observed the unstabilized interferometer cavity to wander as much as 1 fringe over 2 minutes. Thus it is remarkable that the prototype's zero point drift is already competitive with traditional spectrometers where the environment is more controlled. Improvements to the apparatus are now being made which should diminish this drift, including active stabilization of the cavity length and increasing the disperser bandwidth sevenfold to 90 nm.

5. Conclusions

Our interest up until now has been simply to verify the fundamental mathematics of the fringing spectroscopic technique and postpone for the moment the other important issues of long term stability and light efficiency. The low short term velocity noise we demonstrated in our bromine repeatability tests confirms that the fundamental mathematics is sound. Now we are turning our attention to the latter two issues and preparing the instrument for tests on starlight. We are optimistic that these tests will demonstrate superior advantages of using

Figure 5. Noise floor test using bromine vapor back lit by white light as a stationary source that is quiet on a 1 m/s scale. Three 16 minute data sets were taken, with flux per datum estimated at 5.2 (circles), 4.6 (triangles) and 3.5×10^9 (rectangles) photoelectrons. For all sets the vertical axis zero refers to the average velocity of the first set, and the bromine and iodine basis whirls were from the first set. Short term repeatability was 0.76 m/s, which we estimate is dominated by photon noise. After 3 hours, a zero point drift of 4 m/s is evident. Future stabilization of interferometer cavity should improve this.

spectrally multiplexed interferometry to measure broadband Doppler shifts to the 1 m/s level and beyond. At the minimum, the instrument's lower cost and portability facilitates the participation of smaller institutions in the extrasolar planet Doppler survey, and its very large etendue allows use at sites having less than ideal atmospheric seeing conditions, or at telescopes designed to provide low cost but blurry stellar images.

Thanks to Mike Rushford for assistance, and to Prof. Gibor Basri for lending us a bromine cell. This work was supported by Laboratory Directed Research and Development funds. Work performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract W-7405-Eng-48.

References

- Butler, R. P., Marcy, G.W., Williams, E., McCarthy, C. Dosanjh, P., & Vogt, S. 1996, *PASP*, 108, 500
 Hilliard, R. L., & Shepherd, G. G. 1966, *J. Opt. Soc. Am.*, 56, 362

Image Slicing with Infrared Fibers

J. E. Larkin¹, A. Quirrenbach², & J. R. Graham³

¹*Department of Physics and Astronomy, University of California, Los Angeles, 8371 Math Sciences, Los Angeles, CA. 90095-1562*

²*Center for Astrophysics and Space Science, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0111*

³*Department of Astronomy, University of California, Berkeley, Berkeley, CA 94720*

Abstract.

We are proposing to build a new integral field instrument (OSIRIS) for use with the Keck Adaptive Optics System. It will utilize a large (1024 element) fiber optic bundle to slice the field and feed a standard infrared spectrograph with a spectral resolution of 5000. To improve the fill factor of the fiber bundle, we plan on coupling it to a matched lenslet array. The spectrograph will have three plate scales of 0''.05, 0''.10 and 0''.20 per pixel, and full broad band spectra will fit on a single 2048×2048 infrared array. Other innovations include using some of the fibers for pseudo-slits and sky measurements and non-uniform spacing for the "linear" feed to the spectrograph.

1. Design Goals

Among the most important design goals for the OSIRIS instrument was to have a sampling close to the diffraction limit of the Keck Telescope (0''.02 to 0''.07 in the near-IR). The only ways of achieving this is to either have very small physical sampling (0''.05 = 36.4 μ m at F/15) or a large focal ratio (0''.05 = 1 mm at F/412). The first option requires new technologies and one must worry about diffraction effects deteriorating image quality. The second option make the spectrograph very large and potentially costly. As you'll see, we've opted for a compromise of moderate focal ratio and sampling size.

Many of the potential science targets are on the order of one arcsecond across and one of our design goals was to encompass all or most of a target in a single exposure. With samplings on the order of 0''.05, this requires at least a 20×20 grid of spatial samplings or at least 400 individual spectra. To avoid spectral overlap, for most geometries this requires at least a 1024 square array.

The instrument also needs to have a spectral resolution of at least 4000. This requirement is set both by the spectral sampling needed on many of the science targets (~ 75 km/s) and also on our desire to suppress OH emission lines. Between 1 and 2.2 μ m the dominant source of background is the forest of atmospheric OH lines (e.g. Herbst, 1994). The continuum emission actually

reaches a minimum near 1.6 microns where scattered light and thermal emission cross. Because the OH lines are narrow and discrete, we can separate them by going to high spectral resolution. At a resolution of 4000 approximately 90% of our pixels will be free of bright emission lines and will thus have greatly reduced backgrounds.

A conflicting requirement is broad band spectral coverage (z, J, H and K). This is an efficiency requirement so that each setting has $\Delta\lambda/\lambda \sim 0.2$. Coupled with the high spectral resolution, this requires approximately 2000 spectral pixels.

Finally, the spectrograph needs to work throughout the near infrared from 1 to 2.5 microns are therefore requires full cryogenic performance (77 K). This requirements places many difficulties with the design including the need for cryogenic optical couplings, cryogenic motors and mechanisms, temperature compensation in the optical alignment and restrictions on the optical components that are available.

2. Image Slicing Technique

With the requirement of diffraction limited sampling mentioned above, we found that we had three viable image slicing techniques. The first would be to use micromirrors (MEMs) to carve up the field. This is a promising avenue, but we have not seen any concrete designs for coupling a micromirror to a spectrograph. One of the largest problems with these devices is that each mirror tips about an individual axis so the optical path lengths from different elements are different. This makes it difficult to collimate the light. Micromirrors also require complex electronics and software to control. Other concerns such as fill factor, repeatability, cryogenic operation and technical maturity also made us decide against micromirrors.

A second slicing solution would be the use of a lenslet array in a Tiger style instrument (Bacon et al. 1995). That is use the pupil images formed by a lenslet array placed at the focus to feed a spectrograph. The lenslets are rotated so that their spectra don't overlap. This is a good option if the desire is for large spatial coverage with modest spectral coverage. The difficulty arises in that spectra from the bottom of the array must be prevented from overlapping with spectra from the top or middle of the array. One solution would be to make a fairly long and narrow lenslet array (10×50 for example) but that would not meet our goal of 1" spatial coverage in both axes.

The third slicing option we considered is a fiber optic bundle. These have now matured into a somewhat standard commercial product in which fibers can be arranged in complex geometric patterns at the input, and then reorganized into a linear (or pseudo-linear) pattern for feeding a spectrograph. They offer the advantages of a fine sampling scale (100 μm), no spectral overlap between fibers, and the ability to dedicate some fibers to simultaneous measurements of the sky. The primary disadvantage is modest fill factors ($\sim 40\%$). To compensate for this we have combined options two and three and designed around a fiber bundle and matched lenslet array placed at the AO focus. The lenslet is used to achieve a much higher fill factor, while the fibers gives increased sampling flexibility and a more optimized arrangement into the spectrograph.

Figure 1. Input fiber pattern showing central 28×28 bundle, 132×1 horizontal and vertical "slits" and the four sky patches at the extreme corners.

Fiberoptic manufacturers have begun creating elaborate fiber bundles for high speed switching applications, and we have identified one company, Fiberguide Industries, that can produce almost arbitrary fiber patterns. The basic process is to define a hole pattern in a computer and use it to direct a laser to cut a precision set of holes in a thin metal plate. The hole spacing can be as small as 150 microns and positional accuracies are better than 4 microns. The fibers are then individually polished and inserted into the holes. The output end is done similarly with essentially any mapping of input to output hole patterns. Fiberguide Industries has an infrared transmissive fiber (Anyhroguide G) that operates from 0.4 to 2.4 microns and has a temperature range from -190°C to 350°C ; ideal for our application. We anticipate a fully cooled optical assembly to just above LN2 temperatures.

Based on certain optical limitations it is difficult to use more than one fiber per two spatial pixels, even with underfilling of each pixel. We are therefore proposing a 1024 fiber bundle with a fairly elaborate input and output arrangement. The primary input component will be a 28×28 fiber rectangular bundle (784 fibers total). Extending from this bundle in the four cardinal directions are 52 fibers yielding pseudo-slits of total length $52 + 28 + 52 = 132$ fibers. These "slits" can be aligned along an extended object's major and minor axis to extend the spatial coverage of the instrument. The final 32 fibers are placed in four bundles of 4×2 fibers each at the extreme corners to provide simultaneous sky coverage for many compact objects. The arrangement is shown in figure 1.

At the output the fibers are staggered to allow a horizontal spacing of 100 microns while maintaining the minimum 150 micron fiber-fiber spacing. Since an arbitrary fiber spacing is possible with bundle technology, we plan on an output pattern that is not exactly evenly spaced and that has a gentle curve from one end to the other. This allows us to remove much of the spectral curvature and distortion of the optics that is inherent in slit spectrographs. So it is hoped

that the spectra on the detector will have a very uniform spacing and spectral features will lie roughly along rows of the device.

2.1. Lenslets

A fundamental problem with fiber bundles is the low fill factor due to the round shape of the fibers and the thick claddings and jackets. We are proposing to remove this problem by placing a matched lenslet array in front of the fiber bundle. The fibers are then at the pupil images of the lenslets making the instrument a pupil spectrograph. Adaptive Optics Associates can manufacture custom lenslets with the desired 150 micron spacing with better accuracy than the fiber bundle. Rectangular lenslets have > 98% fill factor and should have high coupling efficiency. For the infrared, AOA recommends their sapphire substrates with the lenslets formed from infrared transmissive epoxy. AR coatings will be applied to both faces.

A potential problem is the alignment of the lenslets with the fiber bundle and maintaining this alignment when cold. We hope to alleviate this problem by rigidly mounting the lenslets to the face of the fiber bundle. The lenslets can be fabricated such that their focal plane is at their back surface. By fastening this to the fibers, alignment should remain good during cooldown. We will investigate several fastening strategies including ir-epoxy, and mechanical clamping.

For most of the plate scales we recommend, the alignment requirement is relaxed a bit because the spot size generated by each lenslet is significantly smaller than the fiber core itself. So we underfill each fiber and small misalignments don't create light losses. The spot size for each lenslet is essentially the pupil diameter formed by the lenslet since each lenslet is in the focal plane. The pupil size is given by:

$$pupil\ size(microns) = \frac{lenslet\ focal\ length(microns)}{input\ focal\ ratio}$$

As shown below, the input focal ratio for the primary fiber spacing of 0''.05 per fiber is F/61. To achieve a pupil size under the 50 micron core size, we need a lenslet focal length less than 3.05 mm. For the coarser plate scales of 0''.1 and 0''.2 per fiber, the pupil would be too big, however, resulting in a loss of light. At the same time too small of a focal length will result in a very fast focal ratio at the detector. We are therefore recommending a focal length of 1.5 mm, which matches the 0''.1 pupil with the fiber core. It does mean that there will be a 75% loss of light for the coarsest scale of 0''.2 per fiber. As will be shown, a focal length of 1.5 mm for the lenslets requires a focal ratio at the detector of F/3.6 which is difficult but doable.

3. Spectrograph Design

3.1. Front-End Reimaging Optics

The angular scale that corresponds to each fiber is determined by the F/# of the beam as it reaches the lenslet array. A simple relationship can be obtained:

$$\frac{\Delta\theta('')}{\Delta x(\mu m)} = \frac{0.206}{D(m)F/\#}$$

where $D(m)$ is the diameter of the telescope in meters, $F/\#$ is the focal ratio at the lenslets, and $\Delta\theta/\Delta x$ is the angular scale in arcseconds per micron. So for a 150 micron lenslet separation, a 10 meter telescope and a focal ratio of 15, the angular scale per lenslet is 0.206 arcseconds. This is a good sampling for a non-AO mode, but with a diffraction limit at 1 micron of under 0.025 arcseconds, it is necessary to change the plate scale before the lenslet array. It is also difficult to form a good cold pupil or baffle system after the fiber bundle, so the primary baffling is performed in the front-end optical elements.

TABLE 3.1.
Plate Scales

Magnification	Focal Ratio at Lenslets	Angular Scale per Fiber	Central Field	Length of "slits"
1:1	F/15	0''.206	5''.76	27''.2
1:2	F/31	0''.100	2''.8	13''.2
1:4	F/62	0''.050	1''.4	6''.6
1:10	F/155	0''.020	0''.56	2''.64

Table 3.1. gives the four possible plate scales and their corresponding focal ratios. The fields of view of the central 28×28 fiber patch and the lengths of the "slits" formed by long 1×132 fiber stripes is also given in arcseconds. For full sampling of the diffraction limit, we would need a scale per fiber of $\sim 0''.01$ but then the field of view becomes very small, and we believe that for that scale, the other AO instrument NIRC2 is more appropriate in most scientific applications. But a field of $0''.56$ is actually not too small to consider for certain applications like faint field galaxies where the diameters can be under $0''.5$. But the scale that we recommend as the primary choice is a 1:4 magnification that yields $0''.05$ fibers and a $1''.4$ central field of view. This is ideal for most of the scientific programs and although it slightly undersamples the diffraction limit at all of the selected wavelengths, it does maximize the energy per fiber for point sources and allows for very high resolution on extended sources. The "slit" lengths of $6''.6$ are also quite reasonable for investigating the major and minor axes of extended sources such as AGN and the nuclei of Ultraluminous Infrared Galaxies.

There will, however, be occasions where a larger field of view will be desirable so coarser samplings should be considered. Since the front end will consist of a pair of infrared doublets that are of very low cost compared to the overall instrument, we are recommending that a simple turret be used to rotate in different reimaging optics to select different plate scales. We recommend three selectable scales for the instrument of 0.2, 0.1 and 0.05 arcseconds per fiber in a three position turret. Good repeatability of this mechanism is important but not critical since any motion of the lens assemblies would simply change the field of the fiber bundle by a small amount. So no spectral shifts or loss of light would occur due to position errors. The turret also provides the future option that additional scales can be created, and three selected for each run. This would give a powerful non-AO capability to the instrument as well, although there can be significant light loss at the fiber coupling with the coarsest plate scales.

Figure 2. The blaze function for a grating with a blaze wavelength of 6.35 microns. Note how well each order fits within the atmospheric transmission.

3.2. Grating

For maximum reproducibility, we have selected to use a single non-rotating grating. But we still desire high efficiency across each infrared spectral band, and no order overlap within each broad band. This is achievable by blazing the grating to a blaze wavelength of 6.35 microns. The third order then peaks at 2.12 microns, fourth at 1.59 microns, fifth at 1.27 microns and sixth at 1.06 microns. These are ideally situated in the K, H, J and z bands, respectively. A blaze function for such a grating is shown in figure 2.

3.3. Filters

Since an entire broad band fits onto the detector at once, narrow band filters are not needed. So the only filters required are broad band K (2.0-2.4 microns), H (1.5-1.85 microns), J (1.15-1.35 microns) and z (1.00-1.12 microns). We plan on placing these in a filter wheel just in front of the grating, essentially at the pupil plane. This gives uniform wavelength blocking across the field. Ray tracings for standard one inch filters have demonstrated that flatness requirements are quite modest since filters begin as highly polished substrates. Most surface deviations are then due to warping which are matched between the front and back surface.

3.4. Detector

Essentially every advantage of the integral field unit is tied to the use of a 2048×2048 infrared device. Good OH-suppression requires a spectral resolution of 4000 or higher. It also requires a highly reproducible spectral format to assist with software extraction of uncontaminated pixels. Together these requirements mean full band coverage without a moving the grating. For a resolution of 5000, this requires roughly 2000 pixels in each spectral channel. If a 1024 array were used, then either the resolution would need to be halved with considerable loss

of scientific capability and essentially all OH-suppression, or the grating would need to be scanned across the array which would decrease observing efficiency and degrade spectral extractions.

The spatial coverage is also dependent on the number of fibers in the bundle. We've found it difficult to use more than one fiber for every two columns of the array without significant spectral overlap. If a 1024 array were used, then this would correspond to only 512 fibers arranged in a central 20×20 fiber bundle. With $0''.05$ fiber sampling, the field would be one arcsecond on a side. This is marginal for many of the science goals. A 2048 array would allow 1024 fibers arranged in the 28×28 fiber bundle with 240 fibers for the extended linear "slits" and sky patches described above. This is a significant improvement in field size and relaxes many of the pointing accuracies required for targets such as the Kuiper belt objects.

Rockwell has recently begun fabricating multiplexers for a new Hawaii 2 infrared array that has a 2048×2048 format. Quantum efficiencies are expected to be greater than 80% for the 1-2.5 micron version, with a read noise of ~ 2.5 electrons. The detector pitch will be 18 microns yielding a total array physical size of 37 millimeters. The multiplexer can support either 4 or 32 readout channels with pixel clocking rates of roughly a megahertz. Since all of science goals make use of the OH-suppression, it is anticipated that long exposure times will be the rule, so it is not necessary to support a fast readout mode. We therefore believe that a one hertz maximum frame rate is more than sufficient allowing us to use only 4 readout channels. This greatly reduces the cost of the electronics.

4. Science Drivers

The OSIRIS instrument combines four powerful capabilities: high spatial resolution with the Keck adaptive optics system, medium resolution ($R \sim 5000$) infrared spectroscopy, integral field capability, and OH-suppression. Taken together these capabilities allow us to dig much deeper into many pressing astrophysical problems. Below we have identified several scientific programs which should greatly benefit from this instrument.

4.1. Extragalactic Science

Faint field galaxies like those observed in deep imaging surveys are very compact, often with radii well below $1''$. Integral field observations of these objects can determine internal dynamics, star formation rates, redshifts, and Hubble Type simultaneously. These observations are crucial in determining the merger rates and evolutionary processes in the first generations of galaxies. Although the surface brightness per fiber will be quite low, the backgrounds are also down by a factor of 400 compared to a square arcsecond so many of these objects can be observed in reasonable timescales (hours).

Another related group of objects are the host galaxies of quasars. Only with high spatial resolution can the faint disks of galaxies be seen in the glare of quasars. Recent observations suggest that mergers and galaxy interactions may play a key role in the feeding of material in the central supermassive blackhole. An integral field instrument can study not only the emission features of the

Figure 3. Faint field galaxies become very small at high redshift and can have core diameters well below $1''$. As this pair of images from Roger Thompson (Univ. of Arizona) shows, the infrared and optical morphologies are sometimes quite different.

quasar itself, but also measure the morphology, and spectrum of the galaxy and determine star formation rates, masses, metallicities, etc...

Active Galaxies are also an ideal target for the integral field spectrograph. At the resolution of the proposed instrument ($0''.05$) the narrow line regions of some nearby Seyfert Galaxies should be resolvable. This would allow us to directly test the standard model in which AGN have a supermassive blackhole, an accretion disk and two populations of clouds (narrow line and broad line) which orbit at different distances from the center. OSIRIS should be able to spatially and spectrally separate out the narrow line clouds from the more centralized broad line region. OSIRIS should also be able to study nuclear star formation in AGN and spectrally image the interaction between the central engine and surrounding regions.

4.2. Planetary Science

Although the planets are vastly closer than the extragalactic targets, many solar system targets do have interesting structures on subarcsecond scales. Obvious targets include the moons of the Jovian planets and asteroids. The largest moons, such as Titan are roughly $1''$ across, and there is intense interest in their surface chemistry. Titan has a dense methane cloud layer which can best be penetrated in the infrared, and appears to have a surface covered in hydrocarbons. With OSIRIS, one could measure surface chemistry at over 500 independent locations within a few hours. The moons IO and Ganymede are also of great interest.

Kuiper belt objects present a different opportunity for OSIRIS. Although we don't anticipate resolving these tiny objects, there is still significant gain to be had by going to high spatial resolution. Chief of these is a great reduction in

Figure 4. As these recent Hubble Space Telescope images show, the host galaxies of quasars are often highly disturbed and very compact. Only with integral field spectroscopy on subarcsecond scales can these galaxies be studied in detail.

the background. The integral field nature of OSIRIS also reduces the need for accurate tracking and positioning since there are effectively no slit losses.

4.3. Galactic Objects

Within our Galaxy, there are countless varieties of objects, many of which have structural and spectral properties ideally suited for OSIRIS. Perhaps chief among these are faint stellar objects and brown dwarfs. For free floating objects, the reduced background of the high spatial resolution would allow for more sensitive measurements. For faint objects near brighter companions, the pseudo-coronagraphic nature of an integral field instrument (and the capability to put in a real coronagraphic set of stops) gives great advantage in determining the spectral properties. OSIRIS not only offers the ability to study such objects but also to find them in the first place. By imaging the central 1" to 2" around stars, one could use a spectral filter to search for objects that are spectrally different from the primary. Such a filter could be formed from a weighted summed set of spectral channels where the brown dwarf or low mass companion would be expected to be significantly different from the primary, such as in the methane or water bands.

The Galactic center is a unique environment where massive stars are born in compact clusters, and where orbital motion is directly observable around the central blackhole. It is also highly obscured with approximately 30 magnitudes

of visual extinction. Infrared spectra encompassing the central cluster ($<1''$ in diameter) could determine the 3-dimensional kinematics as well as the spectral types of this unique grouping of stars.

5. Instrument Summary

The OSIRIS instrument is an integral field infrared spectrograph designed for the Keck Adaptive Optics System. It will utilize the latest infrared detectors and state-of-the-art fiber bundle technologies to achieve high spatial resolution, integral field imaging, very faint object sensitivity, moderate resolution spectroscopy ($R \sim 5000$) and a limited long slit capability. All of this is achieved with an innovative yet straightforward and easy to use instrument.

Because of the great amount of information in each frame, only very limited observing modes are required. The large array means that a full broad band spectrum is contained in each exposure. By proper grating selection this is further simplified in that a single fixed grating can be used for all bands (z, J, H and K). The band is easily selected by changing the order sorting broad band filter. Thus, within the spectrograph portion of the instrument, the only moving part is a simple filter wheel with five positions corresponding to z band, J band, H band, K band and a closed position. In each position, a complete spectral image is obtained with 2048 spectral channels and 1024 spatial channels.

Within each exposure is an $R \sim 5000$ spectrum of a 28×28 pixel patch of the sky, plus vertical and horizontal slit coverage of 1×132 pixels. At this resolution, OH emission lines that dominate the background in the J and H windows are well separated. By coadding only those spectral channels without OH-emission lines (about 90% are clean) a very deep infrared image can be extracted for the 28×28 pixel field ($1.4'' \times 1.4''$ in the primary adaptive optics mode). But in addition to this image, full spectral coverage is of course present.

The OSIRIS instrument is not only powerful scientifically but also quite modest in cost and complexity. It draws upon much of the experience of previous Keck instruments especially NIRSPEC (McLean et al, 1998). Due to a simpler detector readout, it actually has simpler electronics requirements than the NIRSPEC instrument allowing us to essentially clone most of the vital electronics. Because of the expected low backgrounds and long exposures, the detector can be clocked out relatively slowly, about 1 hertz per frame, meaning that we only need to use a four channel readout scheme. We also anticipate only two or perhaps three cryogenic mechanisms of fairly simple design.

The OSIRIS design study is supported by a grant from the California Association for Research in Astronomy which owns and operates the Keck Observatory.

References

- Bacon, R., et al., 1995 A&AS, 113, 347
- Herbst, T. M., 1994 PASP, 106, 1298
- McLean, I. S., et al., 1998, SPIE Proc., 3354, 566

Advanced Image Slicers from the Laboratory to NGST

~~Robert~~ Content

*Astronomical Instrumentation Group, University of Durham Physics
Dept, South Rd, Durham DH1 3LE, UK*

Abstract. The new design of image slicer developed at Durham University for two-dimensional area spectroscopy has now been adopted for many spectrographs, existing or in project. The unit acts as a coupler between the telescope and a spectrograph to reformat a square or rectangular field into one or more long slits. Its advantages over previous designs of image slicers and other methods using fibres, lenses or narrow-band filtering are discussed, mainly: Large field, high spatial resolution, large number of spectral resolution elements, high transmission, and the small size of the instrument. The system is also easy to cool and is then well suited for infrared spectroscopy.

The proposed design is a new type of image slicer in which the original two-dimensional image is sliced into narrow sub-images that are re-imaged side by side to form one or more long one-dimensional images at the spectrograph input.

The flexibility of the concept at the base of this new design is highlighted through the description of 6 different slicer designs. Two of these are for future instruments now at the design phase: The UIST slicer (UKIRT) and the GNIRS slicer (GEMINI); three others are studies for possible future slicers on CGS4 (UKIRT), GMOS (GEMINI) and especially NGST for which the conceptual design of the ESA-NGST Integral Field Multi-Object Spectrograph (IFMOS) is now finished; the last is a laboratory slicer that has been built at Durham University and is now under study. These designs show how easily our slicer can be added to an existing instrument, how it can be incorporated to the slit wheel of future instruments, and how multi-slit reformatting permits a much larger field of view.

Keywords: Integral field spectroscopy, large telescope instrumentation.

1. Introduction

Integral field spectroscopy provides a spectrum simultaneously for each spatial sample of an extended, two-dimensional field. Although the data cube, with axis labelled by x-position, y-position and wavelength, can be obtained by longslit spectroscopy where the slit position is stepped across the target or by scanning Fabry-Perot systems, these techniques are less efficient because they require a number of separate exposures to be taken. This considerably reduces the amount of light that would be detected by a single exposure in the same total integration

time. The same is true for scanning Fourier transform systems although the advantage of integral field systems diminishes when their detector is limited by the dark current. Also, with these scanning techniques, any intensity variations with time degrade the quality of the data.

The scientific advantages of integral field systems for studies of extended objects are obvious but it must also be remembered that spectroscopy of single unresolved objects or integrated spectroscopy of extended objects also benefit because light from the full extend of the image is sampled simultaneously without the need of a wide slit which would degrade the spectral resolution. The product of spectral resolution and throughput is then larger with an integral field system than with a single slit.

As this volume show, different methods of integral field spectroscopy have been developed. Most, if not all, of them can be classified into 4 groups:

In what is probably the older group, the two-dimensional field is optically divided into small images that are re-imaged at the entrance focal surface of a spectrograph (entrance slit or entrance mask for multi-slit spectrographs). One is the method proposed for the Gemini Multi-Object Spectrographs (GMOS) in Durham (Davies et al. 1997; Allington-Smith et al. 2000), in which the original field is divided into small images by an array of microlenses, each feeding a fibre. The fibre outputs are positioned side by side at the entrance slit of a spectrograph where another set of microlenses changes the focal ratio to match it to the spectrograph. A simpler technique uses bare fibres (Arribas et al. 2000) without microlens arrays at the input and output. Another technique uses an array of microlenses to split the field into an array of micropupils (Miller et al. 2000; Ferruit et al. 2000); each of these pupil images is a source on the entrance focal surface of a spectrograph. These 3 techniques have in common to cut the field in both directions of space into individual subfields where the spatial information is lost, each subfield then giving an individual spectrum.

Another group of techniques gives a series of images of the same field each at a different wavelength. In one technique, many dichroic mirrors are placed in the beam, splitting it into many beams with different bandwidth each ending on a different detector (Furusawa et al. 1998). Another technique uses a spectrograph with a two-dimensional image at the input, that is a very large slit. A slicing device in the spectrum image surface sample many narrow bandwidths and sends each sampled beam back into the spectrograph to reconstruct the white light images. At the output, the beams give a series of two-dimensional images of the field that appear side by side on the detector, each with a different bandwidth. These techniques have in common cutting the spectrum into narrow bandwidths instead of cutting the field.

Image slicing (Content 1997; Murphy et al. 2000; Krabbe, et al. 1997; Crampton et al. 2000) makes another group. The original image is sliced in narrow but long images the same length than the field. As for the other techniques that cut the field, the slices are re-imaged side by side to form the image at the entrance focal plane of a spectrograph. This technique cut the field in one direction of space only instead of two; the spatial information is then maintained in the other direction, as with a normal slit.

Finally, efforts to develop a no-slicing-or-cutting-device that can directly give a measurement of the photon energy are now showing significant results.

Superconducting (STJ) photon counting cameras (Perryman et al. 2000; Cabrera et al. 2000) are now sensitive enough to read the photon energy with a precision of a few percents in the visible. Supercooled photographic plates (Keller et al. 2000) may give the ultimate in integral field spectroscopy in the future, with thousands of spectral elements and millions of spatial elements.

2. Description of the design

Figure 1 shows a diagram of the design. The light from the telescope is first captured by a series of mirrors that slow the focal ratio to get the right image size on the slicing mirror. The image of the source is then formed on the slicing mirror, which is made of long and narrow mirrors (the slices) with different tilts. Each slice beam is reflected in a different direction onto another mirror, that we usually call a “pupil mirror” or an “imaging mirror” on which an image of the telescope stop appears. Each pupil mirror re-images a slice of the original field on the entrance slit of the spectrograph. The slice images are side by side with little space between them and form together a very long image on the spectrograph entrance focal surface. Field optics are used to re-image the telescope stop at the correct position in the spectrograph.

This design must be considered as the basis of a new design family. Many variations are possible. For example, In Figure 1, the slicing mirror has some power to re-image the telescope stop on the imaging mirrors; this is not necessary, the pre-optics that are used to change the magnification can also be used to put the telescope stop images at the correct position. The slicing mirror can then be made of flat sub-mirrors. The field optics are made of lenses, but usually they are made of 45 deg spherical mirrors (Fig. 3, top). The images of the slices on the spectrograph slit have slightly different magnifications; this can be corrected by moving the imaging mirrors so that the ratio of the distances imaging-mirror-to-slicing-mirror and imaging-mirror-to-image-of-slice be the same for all slices. The ratio must not be near the value one for this to be possible. The pupil images do not need to be exactly on the imaging mirrors, only near them in order for the beam size to be sufficiently small on them, etc.

3. Advantages over other methods

The methods that cut the spectrum, giving many images of the field at different wavelengths, suffer from a small number of spectral pixels, usually less than 20. The same is true for the camera that measures the photon energy. The maximum spectral resolution is also quite low unless there are significant transmission losses and/or a severely restricted range of accessible wavelengths. Image slicing and the other field cutting methods have all the advantages that can give a spectrograph, with very low to very high spectral resolutions, a large range of accessible wavelengths and a spectral length in the thousands of pixels.

The novel technique of supercooled photographic plates suffers from a very low signal to noise ratio and a small bandwidth.

The methods that cut the field with fibres and lenses suffer from limitations due to their cutting of the field in both directions of space instead of one only with a slicer, and from the losses of the spatial information in each spatial

Figure 1. Diagram of the new design of image slicer. In this example, the focal plane is sliced in 3 subfields that are reimaged side by side at the spectrograph slit.

Figure 2. Preliminary design of the slicer for the Gemini Near Infrared Spectrograph. The light of the central slice is shown.

Figure 3. The new design of the Gemini NIRS low resolution slicer system (top); the slit slide of the Gemini NIRS (centre); the image quality on the slicing mirror of the Gemini NIRS low resolution slicer system (bottom) with the size of the slices on the right.

subfield. Also, microlens arrays and fibres are components of low optical quality. We will now look at each advantage of the slicer separately.

3.1. Maximum packing of the spectra

The slicer gives the maximum packing of the spectra on the detector. Since the spatial information is conserved in the direction parallel to the slit, each pixel column on the detector gives a spectrum. To have 2 pixels per FWHM in the spectral direction with the other field cutting methods, it is necessary to have 2 pixels per spatial element. Clear space between the spectra is also necessary, for example, in the method using an array of microlenses to give an array of telescope stop images at the spectrograph input, a space is needed between adjacent spectra to avoid contamination. A normal number of free pixels would be 2 if a small contamination is acceptable. There are then 4 pixels per spatial elements altogether, which is 4 times fewer spectra than with a slicer. With the method using bare fibres, the simple fact that the fibre core is smaller than the fibres themselves makes it impossible for the fibre output images to touch each other, and the images are circular, not square as with the slicer, giving a larger size for the same surface area of the spatial element. The method that uses lenses and fibres also needs some space between spectra to avoid vignetting by the output microlenses.

In order to put each spectrum on one pixel instead of 4, the spectrograph must be massively oversized. If the density of lines on the grating is fixed for the maximum spectral resolution (increasing the density reduces the throughput), the pupil size in the spectrograph, so the optics, must be made 4 times larger while the focal length of the camera remains the same, making it 4 times faster; the collimator focal ratio would remain the same, giving a focal length 4 times longer. A voluntary degradation of the image in the spectral direction would give the 2-pixel width. All of this would considerably increase the cost, the size and the weight of the spectrograph.

3.2. Smaller diffraction in pupil plane

With the development of adaptive optics and space astronomy, there is a need for a spatial resolution smaller than the diffraction limit. Unfortunately, this generates a huge amount of diffraction in the pupil plane of the spectrograph. With 2 spatial elements per diffraction PSF FWHM, a spectrograph stop diameter 3 times larger than the geometric pupil image would still vignett more than a quarter of the light. Another way to describe this is to say that there is a huge focal ratio degradation due to diffraction. Since the methods using lenses and fibres cut the field in both directions the diffraction is also in both directions. The slicer, however, has a significant diffraction effect only in the direction perpendicular to the slices, the size of the slices in the parallel direction being at least an order of magnitude larger than the slice width. The focal ratio of the spectrograph would then need to be increased in one direction only; for example the grating would need to be longer but no larger.

However, there is a trick that can be used with the slicer to get around this problem. Since we want the same spatial resolution in the direction perpendicular to the slice than in the parallel direction, we would need slices one pixel wide only. However, we want the slices to be at least 2 pixel wide to have a correctly

sampled spectrum. A way to get both at the same time is to use toroidal optics in the slicer system to get a different magnification in the 2 directions. For the best image quality, toroidal optics are necessary anyway. If we magnify the slice images by a factor of 2 in the spectral direction, a one pixel wide slice gives a 2-pixel wide image on the detector. Inversely, this reduces by a factor of 2 the pupil size in the spectral direction, which is the direction of the huge diffraction! This trick then gives us everything at the same time: The same spatial sampling in both directions, an image 2-pixel wide in the spectral direction and a spectrograph stop size not much different from if there was no diffraction. Altogether, for diffraction PSF, the surface area of the spectrograph pupil image due to field cutting diffraction is at least 6 times smaller with a slicer, and still more light is contained in that area.

The slicer system itself does not have any significant vignetting. Theoretically, light could be vignetted by the pupil mirrors (Fig. 3, top) since the slices generate a large diffraction effect there. However, those mirrors can easily be oversized in the direction of diffraction to a size many times the geometrical pupil image. In the slit plane, the slit can be larger than the slice image, only cutting the stray light. A slit width equal to 2 widths of the slice image seems reasonable.

The effect of diffraction due to field cutting is even worse than it seems for the other methods. With bare fibres, the core diameter of the fibres is smaller than the distance between fibre cores; if there are 2 fibres per diffraction PSF FWHM, then there are more than 2 fibre core diameters per FWHM, increasing the width of the diffracted beam at the output. With the method using lenses and fibres, both the input and the output microlens arrays will diffract the light, although the latter has a smaller effect. With the method with lenses only, the wings of the diffracted beams cause additional contamination between spectra.

3.3. Other advantages

The discussion in section 3.1 does not include the degradation inherent to the methods using fibres and/or lenses. For example fibres suffer from errors in the angles at which the rays go out (the so called focal ratio degradation) due to the multiple reflections in the fibres. This results in a loss of light through vignetting or in an increase in the image size or in an oversizing of the spectrograph. Another problem is the limited quality of microlens arrays. Tests made at Durham (Lee et al. 1998) have shown large amount of diffuse light from the microlenses, especially with the larger one used in the lenses-only method. Values between 20% and 50% were found. This has similar effects than the degradation in the fibres: to reduce the vignetted light, one must accept a larger output image, which means a lower spectral resolution and a lower number of spectra, or must oversize the spectrograph.

The total transmission of the slicer system is very good, especially in the infrared. A normal system has only 6-reflections. With 98% reflection per surface, the total transmission would be greater than 88%. Even better reflection coefficients can be obtained in the infrared with gold coatings. Our designs of integral field units with fibres and lenses have total transmission in the range 50% to 80% depending of the design. In the ultra-violet, a slicer would needs special reflection coatings to be added in order to have a good total transmission.

A better solution for the UV may be a hybrid system with 3 transmissions and 2 reflections. In fact, a completely transmissive system is possible, the slicing mirror being made of prisms instead of mirrors (Crampton et al. 2000). Except for the ultra-violet, which may need a dedicated slicer, a slicer system can cover a very large range of wavelengths since it is all reflective.

In the infrared, the slicer system can easily be cooled to cryogenic temperature. This has already been done with a different slicer design (Krabbe et al. 1997). In future all reflective spectrographs made of aluminium, an aluminium slicer system made by diamond turning would give the adequate image quality at reasonable cost. From our discussion with Precision Optic Engineering, the cost for a multiple-mirror made of 20 slices is around \$10,000. Since there are 3 multiple-mirrors in a slicer, the total cost is around \$30,000 for the "special optics."

Fibre systems are limited in the number of spatial elements of resolution by the difficulty to manipulate the resulting "spaghetti." Although there are projects for systems with up to 4000 fibres, we believe from our experience that even a 1000-fibre system will be significantly difficult to achieve. A slicer is limited only by the number of pixels along the slit.

One may think that an advantage of the methods with fibres and/or lenses is the possibility to put many spectra in the spectral direction, for example by having many slits with a fibre system. However, this is also possible with the new slicer design. A 2 slit system is quite simple and a system with 9 slits has been designed for the ESA-NGST IFMOS slicer system (section 5.6).

4. Advantages over previous slicer designs

Slicers have been first developed for increasing the spectral resolution of unresolved sources without the losses of light due to undersized slit widths. In these systems, each slice gives an image with a smaller width than the original image. All spectra would be added together to obtain the spectrum of the source.

The main previous design of a slicer for integral field spectroscopy (Krabbe et al. 1997) used flat mirrors for the slicing mirror and the pupil mirrors (Fig. 4, bottom left). Most of the slicers described in this volume use this design, which cause different problems. First, the size of the beam on the imaging mirrors is much larger since it is equal to the pupil size plus the slice length instead of the pupil size alone in the new design. The larger beams force the imaging mirrors to be much larger, which significantly limits the number of slices and increases the tilt of the sub-mirrors in the slicing mirror. This increased tilt translates into a larger defocus at the edge of the slices. The slit image is also virtual instead of real. Unless a spectrograph is built around the slicer, additional large optics will be necessary to re-image the whole virtual image on the spectrograph slit.

But the most important limitation is the 1:1 magnification between the slicing mirror and the slit. With the new design, the slice images can be demagnified by the imaging mirrors, leaving a slit considerably smaller. For example, the high resolution spectrograph of the ESA-NGST IFMOS has a demagnification of about 20 giving a slit 75 mm long instead of 1500 mm. All of this makes the previous design considerably larger than the new design, which is typically 1 slit length wide by 1 to 3 slit lengths long by a few centimetres thick. This makes

Figure 4. The design of the Cooled Grating Spectrograph 4 slicer system (top); The most common design of slicer system in use today (bottom left); Diagram of the solution for the vignetting problem of the CGS4 slicer system (bottom right) because the field mirrors are not in the focal plane: The pupil of each slice is nearer to the focal plane than the global pupil of the slit; this removes the vignetting at the edges of the slices while maintaining all of the beams inside the spectrograph entrance pupil since the output focal ratio of the slicer system is slower than the spectrograph.

it easy to insert a slicer system of the new design inside a future spectrograph or in front of an existing spectrograph.

5. Description of the design

The new slicer design is particularly flexible. A slicer system can be added to a spectrograph in many ways. Since we started to study this design 3.5 years ago, 2 projects of slicer have been started and we are studying 3 more for possible future projects. We also built a slicer for laboratory test and another one is presently in the manufacturing process. Those 5 designs illustrate well how adaptable the design is. They shows that the slicer system can be placed before the slit wheel, or after, that it can be placed before the spectrograph even if the focal plane of the telescope is inside the spectrograph; It also shows that a slicer system can place the slice images on 2 slits or more, significantly increasing the field of view. Table 1 shows the characteristics of a few slicers. The slicer names in boldface are planned instruments at the design phase while the slicer names in *italic* are studies for possible future instruments. Since my preceding paper on the subject (Content 1998), these slicer designs have evolved. This paper gives an update of these designs.

Table 1. Some slicers on the line and under study

Instrument	field (arcsec ²)	spatial pixel (arcsec)	number of elements	
			spatial (pixel)	spectral (FWHM)
UKIRT-CGS4	21	0.4	132	256
UKIRT-UIST	29	0.12 x 0.24	1000	512
GNIRS	1.6	0.04	1000	512
	15	0.15	1000	512
GMOS	94	0.07 x 0.14	9100	1500
ESA-NGST IFMOS	10	0.05	4000	2000
	1860	0.19	54000	210

5.1. The laboratory proof of concept

As part of our effort to develop this concept, we had a slicing mirror manufactured for laboratory tests. The mirror is made of 20 aluminium slices 15 mm long and 0.5 mm large. The slices were diamond turned with angles from 13.8 to 15 degrees. This permitted to test if our method was correct for the maximum angle that could be necessary in a slicer system. The slices were machined spherical, which gives better result than a flat mirror since the tilt is obtained by placing the mirror to be off-axis on the diamond turning tool instead of tilting it. This permits to machine a high precision reference that gives a higher angular precision. The power on the slices also helps in the optical design to reduce the aberrations. The small difference between angles was chosen to have

most images forming on the same detector when the slicing mirror is uniformly illuminated.

Following this, a new laboratory slicer system is now in the manufacturing phase. It will not only include a slicing mirror but also a segmented pupil mirror, which function is to reimage the slices on a slit. It follows the basic design of the ESA-NGST IFMOS slicer system (section 5.6) and will then constitute a proof of concept for this part of the instrument.

5.2. The slicer system of the Gemini Near InfraRed Spectrograph

The design of the spectrograph (Elias et al. 1998) was already well advanced when we started studying this slicer. The available free space was already highly limited. Two positions were possible, at the telescope focus and at the slit wheel. Using the small size of the optics at best, a design was made that uses the slit wheel, the "decker" wheel under it and the space between the decker wheel and the next mirror on the optical axis (Fig. 2). When the slicer is in use, the image on the focal plane is captured by a pickoff mirror and re-imaged on the slicing mirror where it is cut in 18 slices (we have now increased that number to 21 slices for the low resolution and 26 for the high resolution). All beams are sent to a flat mirror used to increase the distance to the imaging mirrors, then each beam hits its corresponding imaging mirror (here called pupil mirrors) which is near the pupil image. The imaging mirrors each re-image a slice on its corresponding field mirror at the slit plane. The system is made of 3 parts: The slit block, the decker block and the fixed mirrors. When used in no-slicing mode, "normal" slit and decker modules are on the optical axis. The light goes through between the fixed mirrors, which are unused. When a slicer is needed, its slit and decker modules are rotated on the optical axis. This system makes it very easy to add the slicer to the spectrograph design since no new moving systems need to be added. Two slicer units will be available, each for a different spatial resolution (Table 1).

Since then however, the design of the Gemini NIRS has gone through a complete review and the slit and decker wheels have been replaced by a slit and a decker "slides," which move back and forth instead of rotating. This makes it possible to put all the optics of each slicer system in one block attached to the slit slide. Not only are the alignments easier, but the system can be made smaller and with a higher image quality since the angles between incident and reflected beams are smaller. Figure 3 (top) shows the new fore-optics of the low resolution. The system is now much smaller (compare with Figure 2). Figure 3 (centre) shows a general top view of the slide: the slits for normal use of the spectrograph are in the middle while the slicer systems are on each side. One selects a slit or a slicer by moving the slide (and the decker slide) in the left-right direction. The image quality on the slit with the old design is typically 80% ray tracing EE in 0.3 pixel for the low resolution and in 1.1 pixels for the high resolution. However, the image quality will be significantly better with the new design, for example the image quality of the low resolution fore-optics is very high (Figure 3, bottom). Each slicer system fits in a box of approximately 60 x 60 x 150 mm.

Ray tracing simulations have shown that the precision in the alignment of the slide is not critical since it must have a high degree of repeatability anyway.

The high spatial resolution is for use with full adaptive optic compensation while the low resolution, which has a larger field of view, is for tip-tilt corrected images. An interesting characteristic of this design is that all the slicer optics are positioned after the slit plane.

5.3. The UIST slicer

UIST (Ramsay Howat et al. 1998) is a future spectrograph for the United-Kingdom InfraRed Telescope (UKIRT) that includes a slicer system fully integrated. This slicer is the result of a collaboration between Durham and Edinburgh. A complete description of this slicer design is given in (Ramsay Howat et al. 2000) in this volume. The slicer has been planned from the beginning so a space has been left specifically for it as opposed to the Gemini NIRS slicer that has to be placed where some space remains. The spatial resolution (Table 1) has been chosen to use at best the new tip-tilt secondary on the telescope while maximising the field of view. As for the new design of the Gemini NIRS slicer, all optics are attached to the slit changing mechanism, in this case a slit wheel instead of a slit slide.

In an effort to reduce the complexity of the system, a design has been developed where the slit mirrors, which are needed to position the pupil images from each pupil mirror at the right place in the spectrograph, have been replaced by a unique flat mirror. This reduces the total number of individual mirrors by 25%, which should reduce the total cost, but it comes with some disadvantages due to the reduction in flexibility of the design. All mirrors have also been made spherical or flat. This spherical shape and the new design forced the slicer system to be larger in order to reduce the aberrations. One disadvantage is the need to have 2 slits instead of one, each with their mirror system from the slicer to the slits, which increases the complexity of the system, reduces the bandwidth common to all spectra and increases the problem of contamination between adjacent slice images. This is necessary because there is not enough space to put the pupil mirrors side by side. Another is the loss of the ability to correct the slice angular errors using the slit mirrors, since there are no individual slit mirrors anymore. With slit mirrors, one can measure in the laboratory the tip-tilt error of each slice and add an inverse tip-tilt on the slit mirrors where the error is much easier to correct. Without slit mirrors, the machining precision of the slices in the slicing mirror must be much higher. Another is the aspect ratio of 2 for the spatial resolution (Table 1) due to the use of spherical mirrors, which makes it impossible to have a different magnification in the spatial and spectral directions. Aberrations are also larger in the focal and pupil images. There is also at least one more reflection along the path.

However a careful analysis shows that most if not all of these effects are small, typically a few percents. As for the size of the slicer system, since our systems are small anyway, even doubling the size gives a system small compared to the size of the spectrograph. For other instruments that would use this design, the disadvantages increase with the slit length and the "smallness" of the available space.

An important difference between this design and the GNIRS design is that the slicer optics are positioned before (or after but near) the focal plane instead of after.

5.4. The CGS4 slicer

The fourth Cooled Grating Spectrograph (CGS4) is the present spectrograph on UKIRT. We already have built an integral field unit of the type fibres plus microlenses for this spectrograph (Allington-Smith et al. 2000). A new slicer has been investigated (Fig. 4, top) which would nearly double the number of spatial resolution elements, increase the spatial resolution and have a considerably lower emissivity for use in the K band. To do this, the slicer system will be cooled. It will also have a transmission at least 1.7 times higher. The major difficulty with the design was that no additional hardware could be added to the spectrograph. The slicer system must then be positioned in front of the spectrograph. An additional difficulty was that the focal plane of the telescope is inside the spectrograph, 102 mm in the back of the window. The slit mirrors could then not be placed at their normal position in the focal plane. Another difficulty is that, to avoid stray light between slices, a stop must be present before the slicing mirror. This was the case for the GNIRS and UIST slicers since these spectrographs have a stop before the slit plane where the slicer systems are positioned, but it is not the case here since the slicer is in front of the spectrograph.

The problem of the stop before the slicer was resolved by having an image of the telescope stop on one of the pre-optic mirrors (Fig. 4, top). This mirror has a diameter exactly the size of the pupil image (1 mm). This is only possible by giving some power to the pickoff mirror since it is the only mirror between the telescope stop and its image, which increases the alignment difficulty. A larger pupil image and a flat pickoff mirror are possible if 2 more mirrors are added to the design. This would slightly reduce the transmission.

The problem of the inaccessibility of the focal plane inside the spectrograph could be solved by adding 2 large optics after the slit mirrors that would re-image the whole slit at the focal plane inside the spectrograph. Apart from the fact that this would add 2 reflections, then slightly reducing the transmission, it also goes against one of the reason of the new design, which is to avoid these large optics. Since this slicer must be a demonstrator of the design, those large optics must be avoided. The problem that happens when the line of field mirrors is placed outside the spectrograph far from the focal plane (115 mm in this design) is a defocus that enlarges the beam length of the slices on these mirrors. At the edge of the slices, part of the light is vignetted due to this defocus. Fortunately, the scientific needs came to the rescue. The astronomers were craving more for a high spatial resolution than for a large field of view. By magnifying the original image by a factor of 1.55, the pixels have a width of 0.4" instead of 0.62", which takes advantage of the new tip-tilt corrector. At the same time, it slows the focal ratio in the same proportion. It is then possible to use a trick to remove the vignetting at the ends of the slices (Fig. 4, bottom right). By placing the pupil of each slice nearer to the focal plane, the beams at the edges of each slice are pushed back onto their field mirror. Because the focal ratio is slower than the spectrograph, all the light still goes through the spectrograph stop without vignetting even with this "defocusing" of the pupil image. The global pupil image of the whole slit remains the spectrograph entrance pupil, which is at the position of the secondary mirror of the telescope (a virtual pupil image here).

Altogether, the optics fit in a box 36 mm thick only by one slit length by a little more than 2 slit lengths. The optics will be placed in a dewar with 2 widows, one on the telescope side and one on the spectrograph side.

5.5. A slicer design for the Gemini Multi-Object Spectrographs

The Gemini Multi-Object Spectrograph (GMOS; Davies et al. 1997) is another example of a design already well on its way to completion when a study for a slicer system is started. The available space for a slicer system is extremely small. This space is 60 mm thick only before the spectrograph slit plane and 1.2 slit lengths long only in the spectral direction. Still, the new design of slicer is so small that there is enough space to put two slicers facing each other (Fig. 5, top). This gives an additional flexibility to the observer: If one of the two slits is masked, the surface area of the field of view would be 2 times smaller but the spectral length would be more than 2 times longer. Since the observer has the choice, many cases where 2 images would have been necessary to cover the needed spectrum length will be done with one image only. In this design, the 2 pickoff mirrors, one for each slicer, behave as a first slicing mirror that cut the field in 2 beams. To avoid any defocus, the common edge of these 2 mirrors is right in the focal plane. The rest of the design consists in 2 standard slicer systems.

Since the last time this design has been shown in a conference (Content 1998), an important improvement has been added. The old design had slices only 0.12 mm wide, which is difficult to manufacture. The new design has slices 0.50 mm wide, the same size than in the MPE slicing mirror (Krabbe et al. 1997). This is obtained by adding a long flat mirror the length of the slit to increase the path length, as in the old design of the GNIRS slicer, and by reducing the distance between the pupil mirrors and the slit. Altogether, this permits the much larger demagnification between the slicer and the slit.

A comparison of the slicer system with the instrument using fibres and lenses that we have designed for GMOS (Allington-Smith et al. 2000) shows a multiplex advantage of approximately 10 for the slicer system. Larger values are obtained when more complex formulas are used to calculate this advantage. A complete description of the old design and its performances is available in (Content 1997).

5.6. The European Space Agency Next Generation Space Telescope Integral Field Multi Object Spectrograph Slicer System

The new design of slicer is very small and has a large multiplex advantage over the other methods. For a space telescope, all of these advantages are even larger, especially at longer wavelengths. First, the available space and the acceptable weight for the spectrograph are highly limited while the field of view and the number of spectral elements of resolution must be as large as possible. Second, the diffraction in the pupil plane (which corresponds to a focal ratio degradation) due to the field cutting becomes critical if a good sampling is to be achieved in the focal plane with diffraction limited PSFs. This is why the Durham slicer system was chosen by the European Space Agency (ESA) for the Integral Field Multi Object Spectrograph (IFMOS; Lefevre, O., et al 2000) that it is proposing for the Next Generation Space Telescope (NGST; Seery & Smith 1998).

The ESA-NGST IFMOS is made of 2 spectrographs, one for high spatial and spectral resolution, the other for low resolution but a large field of view. A full optical and mechanical design has now been done, part of it after the conference so that it is slightly different from what has been presented, but the general design remains the same. Figure 5 (bottom) shows the basic principle of the instrument while Table 2 shows some of the characteristics of the spectrographs.

Together with another planed instrument, the high resolution camera, it can be seen as a system giving the choice between field of view and spectral length (in pixels) and resolution: The larger the field, the smaller the spectral length and the spectral resolution with a choice of small, medium and large. The total weight of the instrument with the present design is 150 kg, but this must be considered a maximum.

Table 2. Characteristics of the spectrographs and comparison with a planed camera

	High resolution	Low resolution	Camera (not part of IFMOS)
Size of field of view (")	2.6 x 3.8	40 x 46	240 x 240
Spectral length (pixel)	4100	420	1
Spectral resolution	3000	160	5
			(broadband filters)
Number of exposures to cover 1.25 to 5 microns	2	1	7
Spatial resolution (")	0.05	0.19	0.03
Number of slicing mirrors	1	6	0
Number of slices/slicing mirror	52	72	0

The high resolution spectrograph The primary goal of this spectrograph is to give a good spatial sampling at the diffraction limit. The chosen pixel size of 0.05" is slightly larger than the diffraction limited resolution ($1.22 \lambda/D$) at the minimum wavelength of 1.25 microns.

The light from the telescope focal plane is first imaged on the slicing mirror by a set of fore-optics, then sliced into 52 slices that are reimaged onto a unique slit at the entrance field of the spectrograph. This spectrograph works in double-pass to minimise the weight. It has 2 gratings that cover the range 1.25 to 2.5 microns and 2.5 to 5 microns. This turret with 2 gratings (and 2 filters) is the only mechanical piece of the whole instrument apart from the piezo-electric focus adjustments. Figure 6 (centre and bottom) shows the image quality on the slicing mirror and on the slit (in this case for a typical slice). The image on the slicing mirror is of very high quality, the PSF being negligible compared to the slice width. The image of the slicing mirror on the detector gives 80% ray tracing Encircled Energy (EE) into one pixel width on about 80% of the pixels, the worst PSF giving 80% EE in 2.1 pixels. This is however a pessimistic result since the calculations are very difficult with conventional ray tracing programs because they don't calculate the diffraction by the slices (only the diffraction by the exit pupil is calculated with the ZEMAX program that we use). A pessimistic approximation had then to be used.

Figure 5. Design of the Gemini Multi-Object Spectrograph slicer system (top), one of the 2 slicer systems is shown while only the beginning and the end of the other are visible; principle of the ESA-NGST IFMOS (bottom), which has a low resolution (LR) and a high resolution (HR) spectrographs.

Figure 6: Comparison between the ray tracing pupil images and the stop of the high and low resolution spectrographs (top); PSF on the slicing mirrors of the HR spectrograph (centre) and on the slit (bottom), the boxes are the size of the slice width (2 pixels on the detector).

Figure 6 (top left) shows the ray tracing pupil image and the effective stop of the spectrograph. Both are highly elliptical instead of circular. The ray tracing pupil image is 2 times larger than wide because the magnification is 2 times larger in the spectral direction in order to have 2 pixels/FWHM on the detector while having square spatial pixels on the sky. The stop of the spectrograph is elliptical because it is easier to correct the aberrations in one direction than in the other with a spectrograph using mirrors. The two are "cleverly" combined to compensate for the huge diffraction by the slices. In the direction of diffraction, the stop is then 4.6 times larger than the ray tracing pupil image ($F/4$ instead of $F/18.6$). Even so, there are still 11% of the diffracted light vignetted by the stop at 5 microns while 3% are lost at 1.25 microns.

The low resolution spectrograph The primary goal of this spectrograph is to give a large field of view so that it can be used not only as an integral field spectrograph with an observed unique large source but also as a multi-object spectrograph when a large number of small sources are observed. This can be achieved only by accepting a lower spatial and spectral resolution. By combining a fast focal ratio for the cameras with the distribution of the field between a few spectrographs, a field of $40''$ by $46''$ was achieved while maintaining a relatively high spatial resolution of $0.19''$.

This spectrograph is made of 3 subspectrographs. The light in the focal plane is splitted in 6 components by 6 pickoff mirrors, each feeding the fore-optics of a slicing mirror. These slicing mirrors are grouped in 3 pairs each feeding a spectrograph. Each spectrograph has 2 channels each with a prism and a camera. A dichroic following the collimator splits the light at 2.5 microns between the 2 channels. These spectrographs then cover the full band from 1.25 to 5 microns at once. To cover a large field of view, the slicing mirrors of each spectrograph distribute the slice images onto 9 slits instead of one in the other instruments described in this paper. This increases the field area by a factor of 9 but is possible only by reducing the spectrum length to less than $1/9$ of the detector size so the spectra from adjacent slits don't overlap. This is why the spectral resolution must be small, here only 160. Since the spectrograph covers the whole bandwidth at once, there are no mechanical pieces in this spectrograph except the piezo-electric focusing system.

As in the high resolution spectrograph, the elliptical shape of the spectrograph stop and of the ray tracing pupil image are used at best (figure 6, top right). In this case the diffraction by the slices is smaller due to their larger size ($0.19''$ instead of $0.05''$) and the ray tracing beam in the camera is faster. This permits to have a stop only 1.25 time larger than the ray tracing pupil image. Still, 20% of the light is vignetted at 5 microns while only 4% is vignetted at 1.25 microns.

Figure 7 (top and centre) shows the image quality on the slicing mirror and on the slit. The image quality on the slicing mirror is quite good while the quality of the slice images on the slit is excellent with 100% ray tracing EE in 1 pixel on nearly 100% of the field.

Figure 7 (bottom) shows how the 4 spectrographs are positioned with respect to each other and the field of view. They are integrated in a way that minimises the total weight, for example the components of the high resolution spectrograph use the same supports than components of the low resolution spec-

Figure 7. PSF on the slicing mirrors of the LR spectrograph (top) with boxes the size of the slices, and on the slit (centre) with boxes the size of one pixel; Diagram of the spectrograph slicer systems (bottom).

trograph. An important difference with the other slicer systems is also visible on this figure (bottom right): On the other instruments, the light arrives on the slit at an angle of about 90 degrees with the optical axis of the spectrograph, but here the high density of slits makes it difficult to do so. There is just not enough space. The solution was to have the light entering roughly parallel to the optical axis and to position the segmented pupil mirrors between the slits. At the same time, this reduces the angle between incident and reflected beams, which reduces the aberrations and the difficulty of the design.

6. Conclusion

The new design of slicer has a large multiplex advantage over other methods. The most significant advantages are the maximum packing of the spectra on the detector without the massive oversizing of the spectrograph necessary with the other field cutting methods, a much smaller diffraction due to the field cutting in one direction only, and a small size compared to previous image slicers. Compared with the methods that cut the spectrum, the slicer has the main advantage of a much larger number of spectral pixels. Other advantages are its high transmission, large range of wavelength and how easy it can be cooled to cryogenic temperatures for the infrared. The flexibility of the design has been illustrated by 5 designs. The slicer can be positioned before or after the slit wheel, or before the spectrograph as a completely independent module. It can disperse the slice images over 2 slits or more for larger field of view. Finally, it seems to be the perfect match for the Next Generation Space Telescope.

My thanks to the peoples that have actively supported the development of the new slicer at Durham in the Astronomical Instrumentation Group especially J.R. Allington-Smith which has helped producing this paper, at Edinburgh especially the group designing the UIST spectrograph, and in Tucson especially the group designing the Gemini NIRS. This work was supported by Durham University and the Particle Physics & Astronomy Research Council. I acknowledge the support of the European Space Agency for supporting the IFMOS work and thank our partners in the IFMOS consortium.

References

- Allington-Smith, J. R., et al. 2000, this volume.
- Arribas, S., et al. 2000, this volume.
- Davies, R., et al. 1997, Proc. SPIE 2871, 1099-1106.
- Cabrera, B., et al. 2000, this volume
- Content, R., 1997, Proc. SPIE 2871, 1295-1305.
- Content, R., 1998, Proc. SPIE 3354, 187-200.
- Crampton, D., et al. 2000, this volume.
- Davies, R., et al. 2000, this volume.
- Elias, J., et al. 1998, Proc. SPIE 3354, 555-565.
- Ferruit, P., et al, et al. 2000, this volume.
- Furusawa, H., et al. 1998, Proc. SPIE 3355, 646-657.
- Keller, C., et al. 2000, this volume.

- Krabbe, A., et al. 1997, Proc. SPIE 2871, 1179-1186.
Lee, D., et al. 1998, Proc. SPIE 3355, 810-820.
Lefevre, O., et al. 2000, this volume.
Miller, B., et al. 2000, this volume.
Murphy, T., et al. 2000, this volume.
Perryman, M., et al. 2000, this volume.
Ramsay Howat, S. R., et al. 1998, Proc. SPIE 3354, 456-467.
Ramsay Howat, S. R., et al. 2000, this volume.
Seery, B. D., & Smith, E. P., 1998, Proc. SPIE 3356, 2-13.

Focussing Image Slicers: Refractive and Reflective

E.H. Richardson

EHR Optical Systems, 1871 Elmhurst Pl., Victoria, V8N 1R1, Canada

A. Moore

CREO, 3700 Gilmore Way, Vancouver, V5G 4M1, Canada

T. Tillemann

NOAO, 950 N. Cherry Ave., Tucson, AZ, 85726-6732

David Crampton

National Research Council of Canada, Victoria, V8X 4M6, Canada

Abstract. A number of design options for image slicers for NGST and Gemini are being investigated. These image slicers are all of the focussing type and both refractive and reflective solutions are being explored. One such device, an image slicer that focuses 10 slices on a spectrograph slit is now in operation at the McMath Solar telescope. It consists of three lenslet arrays, and additionally acts as a focal reducer and provides correction for astigmatism of the telescope. A combined refractive and reflective slicer designed for use on NGST delivers near-diffraction limited images for up to 40 slices.

1. Introduction

Image slicers were initially designed and used to reformat the seeing disk of stars into a shape that could efficiently be introduced into the slit of a spectrograph. The first of these, designed by Bowen (1938), reflected nearly-focussed starlight parallel to the slit and then reflected slices of it into the slit using thin mirror facets whose width equalled the slit width. An inherent disadvantage of this type of slicer is that only one slice is in focus at the slit, although it, and the more elegant Walraven image slicer design, were still useful for their original purpose. In practice, the non-uniform illumination of the spectrograph collimator frequently led to substantial problems and they were not extensively used. The superpositioning image slicers (Richardson, 1968) were designed for use with a single-slit spectrograph in which photographic plates were used as detectors. Each slice of the blurry star is elongated and then superimposed on the slit providing deliberate scrambling (for more accurate radial velocities and more uniform illumination along the slit). In this way, the image slicer increased slit transmission and improved the spectrophotometric accuracy, especially with

non-linear detectors such as photographic emulsions. A more detailed discussion of these image slicers is given by Richardson, Fletcher & Grundmann (1984).

Now that superb spatial resolution can be achieved through use of adaptive optics or by telescopes in space, image slicers are being considered to reformat the more-or-less circular images of resolved targets such as galaxies into long slit-shaped apertures in order to obtain 2-D spectral information or to study kinematics of the whole object at once. Of course, to preserve the spatial information, these new designs (e.g., Diego 1994; Content 1997) must ensure that the slices are in proper focus. We are examining both all-reflective and all-refractive types of these focussing image slicers. In the following sections we describe two specific examples of these Richardson Focussing Image Slicers (RFIS).

2. RFIS 2.1: McMath Focussing Image Slicer

As mentioned above, a disadvantage of the Bowen-Walraven type of image slicer is that only one slice is in perfect focus on the slit, the others being out of focus because the light travels a larger (or smaller) distance before reaching the slit. In 1994, one of us (EHR) designed a stacking type image slicer for Drew Potter of Lockheed/NASA, Houston for use on the McMath solar telescope spectrograph. The basic goal was to slice the planet Mercury into 10 slices and then stack these along the 20 mm long slit, demagnified by a factor of 2, with each slice in focus in order to preserve the spatial resolution of planetary features along the slit.

The detailed specifications were:

- slice the 10'' disk of the planet Mercury into ten 4mm $\times$ 0.4mm slices
- demagnify the slices by a factor of 2 (from f/54 to f/27)
- correct for astigmatism of telescope
- maintain focus of slices along the slit
- conform to mechanical constraints of the existing Bowen/Walraven image slicer

The optical design, by EHR Optical Systems, comprises 3 sets of lenslet arrays, 30 lenslets total, all BK7 glass. The lens fabrication was carried out by Lumonics (formerly Interoptics) Ottawa, and the mechanical assembly was designed by A. Moore. An overall layout and a photograph of the completed slicer are shown in Figure 1. The components were successfully aligned and installed in 1996. Performance tests are ongoing, but it apparently works as expected.

3. Richardson Focussing Image Slicers designed for NGST

3.1. RFIS 3.1

An all-refractive image slicer, similar to McMath design described above, was designed for NGST. In this case, however, the design is considerably simpler since it doesn't have to incorporate focal reduction nor to compensate for telescope

Figure 1. Left: The overall mechanical layout of RFIS 2.1. Three sets of lenslets arrays are used to slice up the image and refocus the slices into the slit. A beam splitter above the entrance to the device is used to deflect some light to the guide camera. Right: A schematic diagram of RFIS 3.2. Light from the telescope enters the glass block from the upper left, reflects off the first mirrorlet array, returns to the top and reflects off another (long) array, and then is redirected by two nearly right angle prisms. It then passes through an array of field lenses just before the slit and into the spectrograph. The longest dimension is 126mm.

aberration. A ten-slice image slicer with an entrance aperture of $3'' \times 1''$ (ten $0''.1$ slices) is straightforward and larger ones appear possible.

3.2. RFIS 3.2

A combined refractive and reflective design appears to offer more advantages. This design, shown on the right in Figure 1, uses concave back-surface mirrors (the incident and reflected light beams are inside glass) and has only two air-glass surfaces (one at the $2''.0 \times 3''.0$ aperture, the other at the $65''$ long slit). The powered optics consist of the first mirrorlet array, the reimaging mirrorlet array and the field lenslet array located at the exit slit. Two near-right angle internally reflecting prisms redirect light to the fixed slit location. All of the optical surfaces are spherical or flat.

Features of the prototype design:

- Slices a $2'' \times 3''$ object at NGST focus into 20 $0''.1$ slices
- Slices form a slit $65''$ (68mm) long times $0''.1$ wide. Each slice is offset from the next by $0''.12$ along the slit to preserve spatial information at the ends of the slices.

- Input is offset laterally from output by 43", but location of the telescope focus/spectrograph input focus is preserved
- RMS image diameters are typically 46 microns (0".05) i.e., nearly diffraction-limited. The worst image (outermost) is 57 microns or 0".06
- Very stable alignment since mirrors and lenses are bonded to a block of glass of the same material. For the NGST core wavelength range (1 - 5 μm), CaF_2 would be used.
- Good focus maintained over entire (1 - 5 μm) wavelength range
- Very high throughput, very low light loss and very low scattered light since there are only two air/glass surfaces and since all other reflective surfaces are internal.
- Design compensates for non-telecentricity of NGST telescope design.
- Very compact (126mm by about 12mm wide)

This design can be extended to provide either longer slices or to provide more slices of the same length. If one wished to sample the diffraction-limited images at 2 μm with two pixels, then the nominal 4K detector would allow a total slit length equivalent to 100". Hence a slightly larger format is possible with the strawman detector. Obviously, if one chose to accept lower spatial resolution then the field could be substantially enlarged. Another option to increase the field size (while retaining the superb resolution) would be to increase the length of the detector mosaic.

The McMath image slicer demonstrates that such an image slicer could be easily achieved in practice. The lenslets for the McMath device were made by fabricating two lenses for each stage which were then cut up (every second lenslet was destroyed by the saw cuts, so two lenses were required) to produce the lenslets. Subsequent alignment of the lenslets and of the image slicer posed no difficulties. Hence, the RFIS 3.2 design concept appears to provide an excellent image slicer for NGST scientific applications.

We are grateful to Murray Fletcher for assistance with the diagram of RFIS 3.2 which is shown in Figure 1.

References

- Bowen, I.S. 1938, Ap. J., 88, 113
 Content, R. 1997, SPIE 2871, 1295
 Diego, F. 1994, SPIE 2198, 525
 Richardson, E.H. 1968, JRASC, 62, 313
 Richardson, E.H., Fletcher, J.M., & Grundmann, W.A. 1984, Proc. IAU Colloq. No. 79: Very Large Telescopes, their Instrumentation and Programs, Garching, 469

3-D Spectroscopy with Fibers: Extraction Methods and their Limitations

T.Becker, M.M.Roth, J.Schmoll

*Astrophysikalisches Institut Potsdam
An der Sternwarte 16, D-14482 Potsdam, Germany*

Abstract. In preparation for PMAS, the Potsdam Multi Aperture Spectrograph, which is currently under construction at the Astrophysikalisches Institut Potsdam (Roth et al. 1998), real data from existing fiber coupled 3D instruments have been analyzed in order to investigate the influence of overlap of spectra perpendicular to the direction of dispersion on the quality of different extraction methods.

1. Introduction

A single spectrum for a given optical fiber spectrograph can be imagined as the superposition of an infinite number of monochromatic fibre images (MFI) on the detector along the direction of dispersion. The details of the 2-dimensional profile depend on the fiber diameter, optical scale, point spread function of the spectrograph, detector sampling, etc. To a crude approximation, we can think of a 2-dimensional Gaussian. The purpose of the *spectra extraction* is to break down this superposition and to derive a 1-dimensional spectrum which is independent of the profile. An exact algorithm therefore has to fit the MFI to all possible positions along the strip, which in fact is not realistic.

A more obvious idea is to reduce the 2D-problem to one dimension, simply collecting the signal along all possible cuts through the spectrum in crossdispersion direction. The result is a 1-dimensional spectrum for each fiber which is already exact for the continuum representing a convolution of the real line with the reduced 1-dimensional instrumental profile in the direction of dispersion. The situation becomes some more complicated when there is crosstalk between neighbouring spectra. For a good extraction method one has to decompose the overlap contribution assuming the *slit function*, which is the profile of the cut, for each fibre. Profile fitting is not a problem in principle. However, it is a matter of computational effort, being relatively sensitive against variations of the MFI.

2. Profile Variances

In reality the MFI is not exactly a Gaussian and not constant, but rather a function of the fiber, the position on the detector and of time, caused by a changing transmission characteristics of the fibers. The spatial variation of the MFI must be determined for a good profile fitting algorithm. If the spectra are well separated, the slit function can be determined directly from the exposure itself, so temporal variability is not a problem. Otherwise it is very important to measure the MFI under unchanged conditions.

We tested the quality of the MFI using a run of 9 nights in September 1998 with MPFS at the Selentschuk 6m Telescope and discovered a profile variation of 1.4% with time in the peak related to the total profile intensity. We also detected a variance from fiber to fiber of 3% compared to a variance of 6% that was measured with data from a run with INTEGRAL/WYFFOS at the WHT in Las Palmas in December 1998. Longscale variations over the field of view of the spectrograph amount to 2% in the peak of the profiles in the case of WYFFOS, and 1% with MPFS.

For non-gaussian MFI's, it can be shown that even if the MFI is exactly known, profile fitting algorithms introduce errors, because the slit function depends on the shape of the spectrum itself and reacts very sensitively to sharp spectral features.

Slit functions of 256 fibers with MPFS (top left) compared with 219 WYFFOS fibers (bottom left) plotted one above the other. Profile variations for MPFS (top right) and WYFFOS (bottom right) fibers: *uppermost curves*: total variance of the slit function (solid) and variation from fiber to fiber (broken); *lower curves*: longscale variance (broken) and uncertainty of the profile determination because of photon noise (dashed).

3. Extraction of Well Separated Spectra

At WYFFOS we found the situation of relatively well separated spectra with an overlap of only 13% of the maximum between to spectra. The picture shows the slit function of two neighbouring spectra.

Standard Extraction: This method is simply summing the signals along the profile without any fitting, working sufficiently only in the case of well separated spectra. The advantage is the independance of the underlying profile. The variance is composed of the photon noise, CCD read out noise (RON), which has its greatest impact for low signals and, if not corrected for, the pixel-to-pixel variations of the CCD response which leads to a constant Signal Noise Ratio (SNR) at higher intensities.

Weighted Extractions: With the knowledge of the slit function each participant pixel independant of the others gives rise about the total signal of the cut, where the pixels in the wings carry more uncertain information than those in the middle, especially in noisy exposures. An optimal extraction algorithm combines all this information to a weighted sum, the weights corresponding to the certainty of the signal information (Horne 1986). For low signals this weighted extraction improves the quality with respect to the standard extraction method. For higher intensities the optimal weighted extraction gives the same result as the standard extraction if CCD-flatfielding is done, and can with low uncertainty of the MFI even improve the relation to the standard extraction if a CCD-flatfielding is missing. For well separated spectra it is possible to determine the profile with a great certainty directly from the image.

4. Extraction with crosstalk

At MPFS we found a really extreme situation: 256 spectra are placed on a 1x1K-Chip – the distance between neighbouring spectra being not more than 4 pixels at a FWHM of 4 pixels. The result is an overlap of 67% of the peak intensity in the middle between the spectra.

Centre Extraction without Fitting: Without profile fitting, one can ignore the overlap and collect pixel signals from the centre of the profile out to the wings as far as possible taking into account that one misses the wings and that a part of the sum comes from the neighbour. The total flux is preserved, but sharp signals like emission or absorption lines are merged with the neighbours. This leads to a loss of spatial resolution and the detection of objects in a bright background will be spoiled. As an advantage, no profile determination is required, uncertainties occur only because of a variation of the overlap caused by a profile variation.

Exact Centre Extraction: Based on the extraction method above, it is possible to remove the systematic error if one considers what is left in the wings of each profile. However the elimination of the error introduces a large profile dependance and essentially increases the uncertainty of the extracted results.

Iterative extraction: This method works like simultaneously fitting a known profile to each spectrum. With an approximation of the total signal intensity of each spectrum the individual pixel intensities can be subdivided into neighbour contribution. The standard extraction of these results improves the afore given approximation of the total signal and provides a new input for the next iteration step. The quality of this method is better than the others for overlapping spectra, but does hardly surpass a good extraction method in the well separated situation. Though overlap can help to reduce the influence of the RON for very low signals, but the the quality of extracted spectra is very sensitive against profile uncertainties for high signals.

Different extraction methods under the influence of Flat Variations (FV), Profile Variations (PV) and RON. *left:* SNR for the Standard extraction under different CCD Flat conditions. The dashed curve indicates the quality limit given by the photon noise. *upper:* Weighted extractions without crosstalk compared with the standard extraction. *lower:* Extraction methods with crosstalk again compared to the standard extraction without overlap.

References

- Horne, K., 1986, ASP, 98, 609
 Roth, M.M., Bauer, S.-M., Dionies, F., Fechner, T., Hahn, T., Laux, U., Nickel, U., Popow, E., Schmoll, J., Wolter, D., SPIE Conf. Ser., 3555

SIMON: The Infrared Multi-Object Spectrometer of the Observatoire du Mont Mégantic

René Doyon, Daniel Nadeau, and Philippe Vallée

*Département de Physique, Université de Montréal and Observatoire du
Mont Mégantic, C.P. 6128, Succ. Centre-ville, Montréal, Qc, H3C 3J7,
Canada.*

Abstract. SIMON is a near-infrared (0.8–2.5 μm) 1024x1024 camera/spectrometer currently under development at the Université de Montréal. The instrument will be used on the 3.6 m Canada-France-Hawaii telescope (CFHT) and the 1.6 m telescope of the Observatoire du Mont Mégantic (OMM). The instrument features a separate cryogenic slit wheel unit that can be easily removed and warmed up rapidly for easy replacement of multi-object masks. This paper presents the optical/mechanical design and expected performance of the instrument.

1. Introduction

The availability of large-format infrared (IR) detectors has opened the way to the development of efficient three-dimensional spectrographs. Virtually all IR instrumentation development plans for many telescopes require their instruments to have either (if not both) multi-object and integral field spectroscopy capabilities. One of the projects at the OMM “Laboratoire d’Astrophysique Expérimentale” is the development of a near-IR (0.8-2.5 μm) multi-object spectrometer for use on the OMM 1.6 m telescope and CFHT, a project funded by grants from the Natural Sciences and Engineering Research Council of Canada. The design of the instrument (SIMON: “*Spectromètre Infrarouge de Montréal*”) is presented in this paper.

2. General characteristics of the Instrument

The basic characteristics of SIMON are summarized in Table 1. The instrument houses a HgCdTe 1024x1024 detector sensitive between 0.8 μm and 2.5 μm and cryogenic (~ 80 K) reimaging optics. It has both imaging and spectroscopic capabilities and the required flexibility to be used on several telescopes (OMM 1.6 m, CFHT 3.6 m, Gemini) and configurations (direct f/8 or adaptive optics (AO) foci). A unique feature of SIMON is a separate focal plane cryostat used for rapid exchange of cold multi-object masks.

Table 1. Instrument Characteristics

Detector format:	HgCdTe 1024×1024			
Wavelength range:	0.8-2.5 μm			
Nominal slit width:	4 pixels			
Low resolution mode:	R~400, 1.17 to 2.35 μm (1st order)			
Medium resolution mode:	R~1500 covering, J,H or K (R is doubled with a 2-pixel slit)			
	OMM	CFHT	AOB	GEMINI
Pixel scale ("):	0.46	0.20	0.08	0.045

3. Optical Design

A layout of the instrument, as mounted on the OMM, is presented in Figure 1. The f/8 beam enters the focal plane cryostat through a CaF_2 window and then crosses the CaF_2 window of the detector cryostat towards the collimator (BaF_2 -IRG2 doublet). The beam then passes through a ten-position filter wheel followed by the Lyot wheel and the grism wheel. The Lyot wheel is mounted with several cold stops of different dimensions and locations for accomodating different telescope configurations. The nominal cold stop diameter at f/8 is 20 mm. The collimator is coupled with a f/5.3 quadruplet camera (BaF_2 -IRG2- BaF_2 -IRG2) providing pixel scales of 0.2" on the f/8 cassegrain focus of CFHT (0.08" with AO), 0.46" on the OMM (0.18" with AO) and 0.045" on GEMINI.

The optics is achromatic from 1.1 μm and 2.4 μm allowing the same focus set-up to be used for *J*, *H* and *K*, a crucial feature for acquiring good quality spectra over a wide spectral range (see below). Spot diagrams are shown in Figure 2. The nominal 80% encircled energy diameter is <0.18" and <0.30" when tolerances are taken into account. The geometrical distortion is only 0.1% in the corner of the array. With coating, the average throughput of the optics (lenses and windows) in imaging mode is expected to be 80-85%.

Spectroscopy is obtained by inserting a grism in the collimated beam with an appropriate slit in the focal plane cryostat. Two resolution modes are currently planned. The low-resolution grism, a 125 l/mm grating replicated onto a 33.5° IRGN6 prism, provides a 1st order spectrum from 1.17 μm to 2.35 μm with an average resolving power of 400 (4-pixel slit). An appropriate blocking filter must be used for this mode. The medium-resolution grism is a 130 l/mm grating ruled directly onto a 37.9° KRS-5 prism providing a resolving power of 1500 at *J*, (5th order), *H* (4th order) and *K* (3rd order) with a full wavelength coverage in each band. The low-resolution grism can also be used in combination with standard *JHK* filters for multi-object spectroscopy. For instance, used with the *K* filter, the low-resolution grism yields *K*-band spectra spanning 1/3 of a row on the array.

Figure 1. Layout of SIMON. The instrument is normally oriented with the optical axis pointing vertically.

4. Mechanical Design

The instrument comprises two cryostats: the detector cryostat and the focal plane cryostat (see Figure 1) both cooled with liquid nitrogen. The former includes all the optics and the three cryogenic mechanisms (filter, Lyot and grism wheels) operated with cryogenic motors. The main structure of the detector cryostat is a section of an aluminium cylinder constituting both the cryogen vessel and the cold surface. It is surrounded with a vapor-cooled radiation shield. Both the shield and the cold surface are attached at both ends of the cryostats by 3 G-10 tabs making the whole structure very stiff. The hold time of the detector cryostat is ~ 50 hours.

The focal plane cryostat (FPC) is smaller and meant to be cycled frequently. It features a toric cryogen vessel, with a hold time of 25 hours at room temperature, mounted with an 11-position slit wheel driven by a cryogenic motor. One open position is used for imaging, three for long slit masks (2, 3 and 4 pixels width) and 7 positions for custom multi-object masks. The FPC has its own vacuum valve and getter and it is equipped with heating fingers mounted on the cold surface to reduce the warm-up time. Once warm, the vacuum of the FPC is broken and the unit is easily removed by removing one of the sides of the spacer box. The slit wheel has separate mounting cells for each mask that can be easily removed/replaced from the bottom hole of the FPC. The cool-down/warm-up time should each be less than 3 hours. We expect that less than 8 hours will be

Figure 2. Polychromatic (1.1, 1.65 and 2.4 μm) spot diagram for the center, edge and corner of the array (on CFHT f/8). Each box is the size of a pixel.

needed to warm-up the FPC, service it and bring it back to operation. Mask replacement can take place during the day and will not affect observing time.

5. Expected Performance

The throughput of SIMON, from the top of the atmosphere to the detector, is predicted to be 35% in imaging and 20% in spectroscopy. On CFHT (without AO), this translates into a 3σ 1 hour sensitivity of $K = 21.9 \text{ mag/arcsec}^2$ in broad band imaging and $K = 19.2 \text{ mag/arcsec}^2$ in spectroscopy at $R = 500$. The optimum slit width should be about the same size as the seeing disk (3-4 pixels) without AO. With AO, the slit can be reduced to 2 pixels with little penalty on transmission to reach a resolving power of ~ 3000 . At such a resolution, 80% of the pixels of an H band spectrum are free of bright OH lines. Thus, SIMON combined with an AO system will have a very good software OH suppression mode. For this reason, SIMON represents a particularly good option for AO multi-object NIR spectroscopy on GEMINI.

The flexibility of SIMON to be operated on several telescopes will open a wide range of scientific programs: follow-up spectroscopic studies of very low mass star candidates from the DENIS and 2MASS survey, study of the initial mass function in young embedded star forming regions, high spatial resolution stellar kinematics and stellar population studies in the Galactic center, super star clusters studies of starburst galaxies, AGNs, studies of distant galaxy clusters, quasars and gravitational lenses.

6. Project Status

SIMON has entered the construction phase. All optical components will be in hand by the end of 1999. Prototypes for the various mechanisms have been built and tested. First light on the OMM is expected during the spring of 2000.

Tunable Gratings: Imaging the Universe in 3-D with Volume-Phase Holographic Gratings

S. C. Barden and J. B. Williams

*National Optical Astronomy Observatories¹, 950 N. Cherry Ave.,
Tucson, AZ 85719*

J. A. Arns and W. S. Colburn

Kaiser Optical Systems, Inc., 371 Parkland Plaza, Ann Arbor, MI 48106

Abstract. Volume-phase holographic (VPH) grating technology can enable innovative capabilities for integral field spectrographs beyond that achievable with classical diffraction gratings. VPH gratings can be tuned to maximize diffraction efficiency at the wavelength of interest and complex grating structures can be fabricated to allow the simultaneous observation of different spectral regions. The encapsulated nature of these gratings also allow them to be more robust and easier to handle and clean than classical gratings.

A brief review of the physics and performance for VPH gratings is given along with a presentation of some application concepts. In particular, discussion focuses on the performance of a dual grating that diffracts $H\alpha$ and $H\beta$ to the same angle. Such a grating may be of interest for IFU applications.

1. Introduction

Technologies that are viable for integral field spectroscopy are rapidly evolving and revolutionizing the design of astronomical instruments. One such area is that of diffraction grating technology. Recent advances have been made in grating immersion and grating mosaics, but these devices are still based upon the classical, surface-relief grating. An alternative diffraction grating has recently become available through advances in holographic technology. These gratings, called volume-phase holographic (VPH) gratings, show superior performance over surface-relief gratings in many applications.

Barden, Arns, and Colburn (1998) present an overview of this technology and its potential for astronomical use. We review some of that discussion and present discussion focused on VPH gratings of particular interest in IFU applications.

¹Operated by the Association of Universities for Research in Astronomy, Inc. (AURA) under cooperative agreement with the National Science Foundation.

Figure 1. VPH grating structures. A. Littrow transmission grating. B. Non-Littrow transmission grating. C. Non-dispersive reflection grating (notch filter). D. Dispersive reflection grating.

2. VPH Grating Physics

Instead of diffracting light by surface structure, as in either diamond-ruled or holographically generated surface-relief gratings, VPH gratings diffract light by refractive index modulations within the volume of a thin layer of gelatin material. Light is diffracted at angles corresponding to the classical grating equation

$$m\nu\lambda = \sin(\alpha) + \sin(\beta)$$

where ν is the line frequency of the index modulation as it intersects the surface of the grating volume, (α and β are the incident and diffracted angles, respectively; m is the diffraction order; λ is the wavelength of light). A surface-relief grating “blazes” the diffracted energy by having a tilted facet acting like a mirror. VPH gratings, on the other hand, distribute the diffracted energy according to the Bragg condition

$$m\nu_g\lambda = n_g 2 \sin(\alpha_g)$$

in which ν_g is the frequency of the index modulation within the grating volume itself, n_g is the average refractive index of the grating material (1.5 for the gelatin), and α_g is the Bragg angle within the grating. Figure 1 shows the relationship of α, β, ν , and $\Lambda = 1/\nu_g$ for the different possible VPH grating structures.

The efficiency at which light is diffracted when it satisfies the Bragg condition is a function of both the grating depth (d) and the contrast provided by the index modulation (Δn). In general, rigorous coupled wave analysis (RCWA) (Magnusson and Gaylord, 1977 and Gaylord and Mōharam, 1985) or modal analysis (Burckhardt, 1966 and Magnusson and Gaylord, 1978) is required to properly model the diffraction efficiency for a VPH grating. Kogelnik (1969) developed a set of scaling relations that provide an estimate for the diffraction

efficiency of a VPH grating. The diffraction efficiency for the two planes of polarization for a transmission grating can be estimated by the following:

$$\eta_s = \sin^2 \left[\frac{\pi \cdot \Delta n \cdot d}{\lambda \cdot \cos(\alpha_g)} \right]$$

and

$$\eta_p = \eta_s \cdot \cos(\alpha_g + \beta_g)$$

in which α_g and β_g are the angles of incidence and diffraction within the grating volume. In general, the p plane of polarization efficiency is not that significantly different from that for the s plane as long as α_g and β_g do not add up to a total of 90 degrees. In that case, the grating becomes a perfect polarization beam splitter in which only one of the polarization planes is diffracted and the other is passed straight through.

Light that meets or nearly satisfies the Bragg condition can be efficiently diffracted for appropriate grating parameters. Both the angle of incidence and the wavelength of light are factors which determine whether or not the Bragg condition is met. The angular Bragg envelope, the diffraction efficiency as a function of incident angle for a fixed wavelength, scales as

$$\Delta\alpha_{FWHM} \propto \frac{\Lambda}{d}$$

and the wavelength or spectral Bragg envelope, the diffraction efficiency as a function of wavelength for a fixed angle of incidence, scales approximately as

$$\Delta\lambda_{FWHM} \propto \frac{\Lambda}{d} \cot(\alpha_g).$$

3. VPH Grating Structure

The VPH gratings discussed in this paper are constructed with Dichromated Gelatin (DCG), a material that has been extensively used in holographic devices over the past several decades. Although there are other materials available for the recording of fringe structures, DCG provides superior diffraction efficiencies and has great clarity and excellent transmittance between 300 and 2700 nm. The index modulations are generated by first photo-sensitizing the gelatin, exposing it to the fringe pattern produced by the properly aligned holographic exposure system, swelling the gelatin with water, and then rapidly drying it in a bath of alcohol. In this bath, the regions that were exposed to light collapse to a different density than those regions which were not as intensely exposed. The resultant variations in density translate directly to a refractive index modulation.

Since DCG is very hygroscopic, it must be encapsulated and sealed from exposure to any moisture or humidity. Properly sealed, DCG can have optimum performance over very long lifetimes (at least 20 years). The encapsulated grating can be handled much more easily than surface-relief gratings that are more easily damaged by handling and environmental conditions. The substrates of VPH gratings can be made of any optically suitable material such as BK7 or Fused Silica. High performance anti-reflection coatings can be applied to the

outer surfaces of the grating substrates to further enhance the performance of the VPH grating.

VPH gratings can have line densities ranging from about 300 to 6000 1/mm. The grating depth can vary from 4 to 30 μm . The average refractive index for DCG is about 1.5 and the index modulation can range from 0.02 to about 0.10. Gratings have routinely been fabricated to function in the optical and near-infrared wavelengths.

Grating size is another trait in which VPH gratings are potentially superior to classical, diamond-ruled, surface-relief gratings. Surface gratings are currently limited at about the 200 mm size due to wear on the diamond head making the ruling and to limitations in the grating replication process due to the surface area of such a large grating. Larger surface-relief gratings must be made from mosaicing together a set of smaller gratings. Although not impossible, such mosaics are expensive and delicate. VPH grating technology (and to some extent, holographically generated surface-relief gratings) is limited by the size of the holographic exposure beam, the power of the exposing laser, the stability of the exposure system against vibration and drifts, and the capacity of the DCG processing equipment. DCG holograms of greater than 400 mm have been fabricated and efforts are underway to produce gratings with apertures of 1 meter in diameter. There are also techniques in which DCG gratings might be mosaiced into even larger grating formats, either during the exposure process or by replication techniques. At present, the VPH gratings produced by Kaiser Optical Systems, Inc. (KOSI) are limited in size (75 by 100 mm) by their holographic exposure system dedicated to the fabrication of diffraction gratings. Efforts are currently underway to upgrade their collimators to a size that could expose gratings as large as 300 mm.

4. Measured Grating Performance

Barden, Arns, and Colburn (1998) presented the analysis of a 600 1/mm VPH grating with a peak efficiency of 80% at 700 nm that provided good performance at other wavelengths and diffraction orders when the grating was tuned to the appropriate angles. Funding from the National Science Foundation was received for the fabrication and analysis of an additional eight VPH gratings. The following gratings are under such evaluation:

- 300 1/mm transmission grating designed for optimal performance at 1064 nm.
- 1200 1/mm transmission grating with peak efficiency at 532 nm.
- 2400 1/mm transmission grating with optimal efficiency at 532 nm.
- 2400 1/mm transmission grating with peak diffraction efficiency at 1064 nm. This grating requires prism substrates to extract the otherwise evanescent first order diffraction at the design wavelength.
- 4800 1/mm transmission grating with prism substrates and optimal performance at 532 nm.

- 1200 l/mm reflection grating at 532 nm.
- 1200/1620 l/mm transmission multiplex grating for optimal performance at both $H\alpha$ and $H\beta$. Both wavelengths are diffracted to the same angle.
- 300 l/mm transmission grating with 532 nm in the tenth order of diffraction. This is an attempt to make an echelle-like VPH grating.

As Barden, Arns, and Colburn (1998) pointed out, VPH gratings can be tuned to peak up the efficiency at the desired wavelength of observation. This allows the grating to be used outside the spectral bandwidth of the grating at the design wavelength. Figure 2 demonstrates this for the 300 l/mm grating. In effect, the blaze envelope can be adjusted by tilting the grating. We call the envelope defined by the tunable peak efficiency the “super blaze”.

Figure 3 shows the diffraction performance of the 1200 l/mm grating at a specific grating angle along with the “super blaze”. This particular grating gives a peak efficiency of 87% for unpolarized light at 500 nm, nearly a factor of two in comparison to a comparable 1200 l/mm surface-relief grating.

Note that the spectrograph camera must also be adjusted in angle to re-center the wavelength that the grating has been tuned to. The camera should either be articulated to directly adjust for the diffraction angle that the grating is tuned for, or a set of prisms could be used to modify the angular deviation back to a fixed camera angle in a similar fashion to classical grisms.

5. VPH Gratings of Particular Interest for IFU Spectrographs

Integral field unit (IFU) spectrographs may be able to take advantage of particular aspects in VPH grating technology that may not be specifically suited for more conventional spectroscopic applications. Two possible VPH configurations are

- VPH multiplex gratings in which two different spectral regions are diffracted to the same angle.
- VPH reflection gratings which can have extremely high diffraction efficiency over a narrow bandwidth while the undiffracted light passes through the grating undispersed.

5.1. Multiplexed Gratings

Kaiser Optical Systems, Inc. can produce VPH gratings containing a stack of more than one grating in a single holographic unit (Owen, Battey, Pelletier, and Slater, 1995). Since the spectral bandwidth of a VPH grating can be controlled by either the depth of the grating or the contrast of the index modulation, one can design a grating pair in which one grating diffracts one spectral region while having minimal diffraction efficiency for another spectral window and the second grating diffracts the second region while being invisible to the first. One can then manipulate these two spectral regions somewhat independently. One possibility is to diffract both regions to the same angle where they could both be simultaneously imaged onto the spectrograph detector. Spatial separation

Figure 2. Diffracted energy profiles in unpolarized light for the 300 l/mm VPH grating at three different grating angles.

Figure 3. Measured diffraction efficiency for the 1200 l/mm VPH grating in unpolarized light at a grating angle of 19 degrees. The “super blaze” shows the tunable range of this grating.

of the two regions can be achieved by simply having the fringe planes of the two gratings slightly rotated with respect to each other. The zeroth order of diffraction for both gratings would overlap each other, but higher diffraction orders would be well separated. This type of multiplex grating performs the near equivalent of a multi-order grating with cross-dispersion, but without the need for the wavelength regions to be integral multiples of each other or the need for an additional cross-dispersing prism or grating. An IFU instrument could be designed to utilize such a grating to provide simultaneous spectral imaging in two spectral bands.

For the NSF study, a multiplex grating was designed to diffract $H\alpha$ (656 nm) and $H\beta$ (486 nm) to a diffraction angle of about 23 degrees. A 1200 l/mm grating diffracts 656 nm while the 486 nm spectral region is diffracted by a 1620 l/mm grating component. The bandwidths for the gratings were designed so that at a 23 degree angle of incidence (or diffraction), each would have a minimum in their diffraction efficiencies at the complementary spectral region. Spectral separation is achieved by rotating the gratings by about 5 degrees with respect to each other.

The tunable nature of VPH gratings allows this multiplex grating to be tuned to match the redshifted wavelengths of a particular astronomical target. Adequate diffraction efficiency can be achieved with this grating up to a redshift

of about 0.5 (where $H\alpha$ is shifted to 985 nm and $H\beta$ to 729 nm) by illuminating the grating at 36 degrees. The forbidden lines of Oxygen at 500.7 nm ([OIII]) and 372.7 ([OII]) have a similar wavelength ratio (1.34) to that of $H\alpha$ and $H\beta$. Tuning the grating to an angle of 17.5 degrees shifts the diffraction energy envelope to those wavelengths.

Figure 4 and figure 5 show the measured efficiency for the $H\alpha/H\beta$ grating in unpolarized light and as a function of grating angle. The 1200 l/mm com-

Figure 4. The measured efficiency for the 1200 l/mm component ($H\alpha$ channel) of the multiplex grating in unpolarized light at grating angles of 17, 23, and 33 degrees. This grating achieves optimal peak efficiency (93%), however the spectral bandwidth minimum falls redward of the design goal so does not quite align with the diffraction efficiency peak of the 1620 l/mm component.

ponent has a peak efficiency of 93% at 656 nm! Both grating components were fabricated with peak efficiencies near that predicted by RCWA. However, the spectral bandwidth for the 1200 l/mm grating turned out to be slightly narrower than that desired. The efficiency minimum for that grating does not align with the wavelength for peak efficiency in the 1620 l/mm grating. This has the effect of lowering the effective efficiency of the 1620 l/mm grating by about 15% at 486 nm.

On-sky testing was carried out for this grating by using a fiber optic feed on the 2.1-meter telescope at Kitt Peak National Observatory. The 30 meter long fiber subtended 2.5 arc-seconds on the sky and fed into the QDS (quick and dirty spectrograph). The QDS consisted of three achromatic lenses, the VPH grating, and a 2048 by 2048 CCD detector (T2KB). The detector was binned into a 512 by 512 format in order to better sample the image formed by the

Figure 5. The measured efficiency for the 1620 l/mm component ($H\beta$ channel) of the multiplex grating in unpolarized light at grating angles of 17, 23, and 33 degrees. Light inadvertently diffracted by the 1200 l/mm component lowers the efficiency performance of the 1620 l/mm component by about 15%.

spectrograph. The seeing and sky conditions were quite good during one night of the five night run, with seeing measuring at about 1.2 arc-seconds. Observation of a spectrophotometric standard yielded total detected quantum efficiencies (inclusive of telescope, fiber, seeing, spectrograph, grating, and detector) of 29% for the 1200 l/mm grating at $H\alpha$ and 18% for the 1620 l/mm grating component at $H\beta$! The spectrum of a blue compact galaxy (B=17) with a redshift of 0.0217 is displayed in figure 6. Obvious features are $H\alpha$, [SII], $H\beta$, and the [OIII] doublet.

5.2. Reflection Gratings

A VPH reflection grating has the potential for achieving diffraction efficiencies of nearly 100%. Unfortunately, the spectral bandwidths for these gratings tend to be quite narrow with widths of only a few hundred Å. Many applications in IFU spectroscopy, however, are often concentrated on specific spectral features that don't require wide spectral bandwidths. As such, VPH reflection gratings may be ideal dispersing elements for such IFU systems. An additional benefit of a VPH grating is the fact that the wavelengths that are not diffracted (ie. fall outside the spectral bandwidth) will pass through the grating unaffected. This minimizes the need for bandwidth limiting filters and allows the possibility

Figure 6. Spectrum of an 18th magnitude, blue compact galaxy. The CCD image of the spectrum is displayed on the left. The extracted spectra on the right have had the sky spectrum subtracted. Note the slight tilt between the two spectra in the CCD image. The spectra would intersect each other at the location of zeroth order.

of utilizing the undiffracted light for simultaneous imaging or dispersion by an additional spectroscopic channel in the instrument.

A 1200 l/mm VPH reflection grating has been fabricated through the NSF effort. The grating was designed to have optimal diffraction at 532 nm, corresponding to a laser wavelength that allowed monitoring of the gratings performance during the grating processing. Figure 7 shows the theoretical diffraction performance for the NSF grating.

The grating has been fabricated, but at the time of writing this paper, it has not been fully evaluated for its actual diffraction efficiency. Figure 8 shows the transmittance of the actual grating when the grating is tilted by 4 degree. Except for material absorption losses of a few percent, the diffraction efficiency can be estimated by the inverse of this plot. The wider bandwidth of the actual grating compared to the predicted curve is probably a result of a "warping" of the fringe structure during the processing of the grating. Intentional warping of the fringes is a way of increasing the spectral bandwidth of VPH reflection gratings.

6. Concluding Comments

VPH gratings show excellent efficiencies and provide unique diffraction properties that are either difficult or impossible to achieve with the classical, surface-

Figure 7. Plot showing the RCWA predicted diffraction efficiency for the VPH reflection grating according to the initial design parameters for that grating. Note that the efficiency peaks out at nearly 100%.

relief gratings. The high efficiencies, multiplex capability, and narrow bandwidths can all be exploited by integral field spectrographs. As implied by the title for this paper, it will indeed be possible to do 3-dimensional imaging of the Universe with volume-phase holographic gratings.

This work was funded under Cooperative Agreement AST-9613615 awarded by the National Science Foundation. Gratitude goes to Mark Benson for his dedication and efforts in constructing the VPH gratings at KOSI. Appreciation goes to Skip Andree at NOAO for his support in setting up the QDS instrument for the on-sky testing of the gratings.

References

- Barden, S. C., Arns, J. A., and Colburn, W. S. 1998, *Proc. SPIE*, 3355, 866
- Burckhardt, C. B. 1966, *J. Opt. Soc. Amer.*, 56, 1502
- Gaylord, T. K. and Moharam, M. G. 1985, *Proc. IEEE*, 73, 894
- Kogelnik, H. 1969, *The Bell System Technical Journal*, 48, 2909
- Magnusson, R. and Gaylord, T. K. 1977, *J. Opt. Soc. Amer.*, 67, 1165
- Magnusson, R. and Gaylord, T. K. 1978, *J. Opt. Soc. Amer.*, 68, 1777
- Owen, H., Battey, D. E., Pelletier, M. J., and Slater, J. B. 1995, *Proc. SPIE*, 2406, 260

Figure 8. The measured transmittance in unpolarized light for the 1200 l/mm VPH reflection grating. The diffraction efficiency can be estimated by taking the inverse of this plot.

Development of Grisms & Immersion Gratings for the Spectrographs of the Subaru Telescope

N. Ebizuka¹

*Space Technology Section, Space Communications Division,
Communications Research Laboratory, Koganei, Tokyo 184-8795
JAPAN*

H. Kobayashi & Y. Hirahara

Department of Earth and Planetary Science, Nagoya University

M. Wakaki

Department of Electro-Photo-Optics Engineering, Tokai University

K. Kawaguchi, T. Sasaki & M. Iye

National Astronomical Observatory

Abstract. A solid grism of *Cytop*TM for CIAO and a germanium immersion grating for the prototype IRHS have been successfully fabricated by using diamond turning and grinding method, respectively. Solid grisms of lithium niobate or ZnS for the highest dispersion grism of FOCAS and the integral field mode grism of CIAO are under development.

1. Introduction

To perform spectroscopic observations, several types of grisms will be employed in astronomical cameras, which are to be mounted on with the 8.2m Subaru telescope on Mauna Kea, Hawaii (Kaifu 1998, Iye 1998, Ebizuka et al. 1998A). As shown in Table 1, the Faint Object Camera And Spectrograph (FOCAS) uses six grisms and the Coronagraphic Imager with Adaptive Optics (CIAO, Tamura et al. 1998) uses three grisms for example.

The mid-InfraRed High dispersion Spectrograph (IRHS; tentative name), which is under design study as a second generation instrument for the Subaru telescope, is aiming at a resolving power; $\lambda/\Delta\lambda = 200,000$ at $10\mu\text{m}$, in order to observe vibrational transitions of molecules in dark clouds for instance. IRHS will employ a germanium immersion grating as the dispersing element.

We describe the fabrication processes and present the performance of the gratings developed.

¹Fellow of Telecommunications Advancement Organization, E-mail: ebizuka@crl.go.jp

Table 1. Spectroscopic capabilities of Subaru instruments.

Name (Focus)	Detectors (pix.×pix.×pc.)	Wavelength [μm]	Dispersing elements	R=λ/Δλ	Remarks
Suprime-Cam (Pri.)	Si-CCD (4K×2K×10)	0.35 - 1.0	20 Narrow Band Filters	25	
FOCAS (Cas.)	Si-CCD (4K×2K×2)	0.35 - 1.0	4 Grisms 2 Echelle Grisms 1 Direct Vision Prism	100 - 2,000 2,000 - 10,000 50 - 100	Multi Slit, Long Slit Multi Slit, CD ^a
CIAO (Cas.)	InSb (1K×1K×1)	0.9 - 2.7 (5.0)	2 Grisms (<i>Cytop</i> TM) 1 Grism (LN)	300 - 1,200 300 - 1,200	Imaging Mode Integral Field Mode
IRCS (Cas.)	InSb (1K×1K×2)	0.9 - 5.0	3 Grisms 1 Echelle Grating	600 - 2,000 20,000	Spectroscopic Mode Spectroscopic Mode
COMICS (Cas.)	Si:As (320×240×7)	8.0 - 22	1 Reflection Grating (1 Echelle Grating)	2,000 (10,000)	Spectroscopic Mode
HDS (Opt.Nath.)	Si-CCD (4K×2K×2)	0.35 - 1.1	1 Echelle Grating	100,000	+ 2 CDGratings (Blue, Red)
OHS (IR Nath.)	HgCdTe (1K×1K×1)	0.9 - 2.5	2 Reflection Gratings 3 Grisms 1 Direct Vision Prism	(5,600) (5,600) 50 - 100	OH Suppression Long Slit Slit Less
IRHS (IR Nath.)	Si:As (512×412×2)	8 - 22	2 Immersion Gratings	25,000 - 200,000	+ 2 CDGratings

^aCross Disperser

2. Development of grisms and immersion gratings

2.1. Grisms for CIAO

A solid grism of *Cytop*TM (fluorocarbon polymers, ASAHI GLASS Co.Ltd.) for CIAO has been successfully fabricated by using an ultra-high-precision 3D profile grinding/turning machine and a single crystal diamond tool. *Cytop*TM has a refractive index of 1.32 at 1.6μm and is transparent from 0.2 to 2.6μm.

Measured efficiencies are 65% for the second order (2.4μm), 70% for the third (1.6μm) and fourth (1.2μm) orders. In case of ideal groove shape, the diffraction efficiency should be 100% for the third order and 0% for the adjacent order at 1.6μm. However 10% of the light which ought to be third order is distributed to the second and third orders at 1.6μm. This implies that the useful facets are effecting about 90% of the ideal width.

2.2. Immersion gratings for IRHS

Since refractive index is 4.0 for infrared wavelengths from 1.6 to 20μm and the angular dispersion is proportional to the refractive index of the medium, a germanium immersion grating has four times higher angular dispersion than a conventional reflection grating of the same size. This means, by using a germanium immersion grating, length of optics are reduced to one fourth as compared with one that employs a conventional grating of the same angular dispersion, so that total volume of a spectrograph can be dramatically reduced.

A prototype germanium immersion grating, which is designed for a spectrograph with a resolving power of 250,000 at 2.2μm (Figure 1), was successfully fabricated by using a ultra-high-precision 3D profile grinding/turning machine and electrolytic in-process dressing (ELID) method for diamond grinding (Ohmori et al. 1992). As shown in figure 1 (right), the groove shape has an

Figure 1. Germanium immersion grating for the prototype IRHS (left) and cross section of the grating grooves (right). The immersion grating has a size of 30 (W) $\times$ 30 (H) $\times$ 72 (L) in mm, vertex angle of 68.75 degree and the groove density of 10 grooves/mm.

inevitable deviation from the ideal saw-tooth shape, due to the abrasion of the diamond grinding tool.

As shown in figure 2, the full width half maximum (FWHM) of diffraction images of the optical flat and immersion grating are 4.0 and 4.5 pixels, respectively. The FWHM difference of the immersion grating from the optical flat implies the presence of a large scale phase error and deviation of groove interval resulting from the inaccuracy of the fabricated machine. The wavelength for 633 nm in air is the same as that for $2.6\mu\text{m}$ inside the germanium because the refractive index of the medium is 4.1 times that in air. We assume that resolving power is deteriorated from 10 to 15% at $2.6\mu\text{m}$ while it is trivial at $10\mu\text{m}$.

Although the fabricated immersion grating for use at $2.2\mu\text{m}$ should be improved to achieve better shape of grooves and surface roughness of facet, the results of the trial fabrication suggest that a germanium immersion grating for use at $10\mu\text{m}$ can be realized by means of this fabrication method.

2.3. High dispersion grisms for FOCAS and CIAO

The highest dispersion grism of FOCAS and integral field mode grism of CIAO would employ a high index material, lithium niobate (LN) or ZnS for examples, as the prism substrate because the dispersion is proportional to the optical pass difference generated by a grism but the size of the grism is limited by a holder. Optical resin is not employed into the grisms because interface of substrate and resin will exceed a critical angle of incidence. In this reason the grisms should be solid.

A prototype high dispersion grism was made by attaching an LN prism with a blazed transmission grating formed on a Z-cut LN wafer (Ebizuka et al. 1998B). The blazed transmission grating was fabricated by an oblique ion etching method (Aoyagi et al. 1976). However it is difficult to fabricate deep grooves, exceed $1\mu\text{m}$ in depth, by means of ion etching. We are planning several test fabrication runs of deep grooves for high index materials by means of plasma etching.

Figure 2. Cross section of diffraction images of the immersion grating and optical flat at 633nm (He-Ne laser). Linear scale (left) and logarithmic scale (right).

The authors wish to thank staff members at the Institute of Physical and Chemical Research (RIKEN) for advice on fabricating gratings and their help in using the diamond grinding/turning facility. The work was funded in part by a Grant-in-Aid for Scientific Research (09559018) from the Ministry of Science, Education, Sports and Culture of Japan.

References

- N. Kaifu, "Subaru telescope", Advanced Technology Optical/IR Telescopes VI, ed. Larry M. Stepp, 1998, Proc. SPIE 3352, 14-22.
- M. Iye, "Overview of Subaru instrumentation", Optical Astronomical Instrumentation, ed. S. D'Odorico, 1998, Proc. SPIE, 3355, 344-353.
- N. Ebizuka, M. Iye, T. Sasaki and M. Wakaki, "Development of High dispersion grisms and immersion gratings for spectrographs of Subaru Telescope", Optical Astronomical Instrumentation, ed. S. D'Odorico, 1998, Proc. SPIE, 3355, 409-416.
- M. Tamura, H. Suto, H. Takami, Y. Ito, N. Ebizuka, K. Murakawa, N. Kaifu, N. Takato, N. Takeyama, M. Iye, K. Chikami, and Y. Oasa, "Coronagraphic imager with adaptive optics (CIAO) for the Subaru 8-m Telescope", Infrared Astronomical Instrumentation, ed. A.M. Fowler, 1998, Proc. SPIE, 3354, 845-852.
- H. Ohmori, "Electrolytic In-Process Dressing (ELID) Grinding", 1992, Int. J. Jpn. Soc. Prec. Eng., 26, 273-278.
- N. Ebizuka, M. Iye and T. Sasaki, "Optically Anisotropic Crystalline Grisms for Astronomical Spectrographs", 1998, Appl. Opt. 37, 1236-1242.
- Y. Aoyagi and S. Namba, "Blazed ion-etched holographic gratings", 1976, Optica Acta 23, 701-707.

Adaptive Optics High Resolution Spectroscopy: Present Status and Future Direction

J. Ge, D. Ciarlo, P. Kuzmenko, C. Alcock, B. Macintosh, K. Cook, C. Max

Lawrence Livermore National Lab, Livermore, CA 94551

R. Angel, N. Woolf, M. Lloyd-Hart

Steward Observatory, Univ. of Arizona, Tucson, AZ 85721

J. Najita

National Optical Astronomy Observatories, Tucson, AZ 85726

Abstract.

High resolution spectroscopy experiments with visible adaptive optics telescopes at Starfire Optical Range and Mt. Wilson have demonstrated that spectral resolution can be routinely improved by a factor of ~ 10 over the seeing-limited case at visible wavelengths. With large CCDs now available, a very wide wavelength range can be covered in a single exposure.

In the near future, most large ground-based telescopes will be equipped with powerful adaptive optics systems. Most of these systems are aimed primarily at diffraction-limited operation in the near IR. An exciting new opportunity will thus open up for high resolution IR spectroscopy. Immersion echelle gratings with much coarser grooves being developed by us at LLNL will play a critical role in achieving high spectral resolution with a compact and low cost IR cryogenic cooled spectrograph and simultaneous large wavelength coverage on relatively small IR detectors.

The combination provided by adaptive optics of very high resolution and simultaneous large wavelength coverage allows much improved sensitivity in the search for extra-solar planets through precision Doppler techniques and in the detection of newly formed giant planets through their influence on the protoplanetary accretion disk. It also opens the possibility of studying stellar activity through the simultaneous observation of a large number of stellar absorption line profiles for the first time.

1. Introduction

Very high-resolution optical and infrared spectroscopy is one of the most exciting new fields to be explored in astronomy. Its major applications include search for newly formed extra-solar planets (Carr & Najita 1997) and extra-solar planets

around old stars (Marcy & Butler 1998), and the study of the structure, physics and chemistry of proto-planetary and planet forming circumstellar disks. It also offers unprecedented sensitivity for study of stellar magnetic fields, interstellar and circumstellar medium, stellar abundance and isotopic abundance.

Despite the tremendous potential of very high resolution optical and infrared spectroscopy to provide major breakthroughs in astronomy, it was very difficult to be achieved because of lack of efficient techniques (Ge 1998). Adaptive optics correcting telescope wavefront errors caused by the atmospheric turbulence and restoring diffraction-limited images however has the potential for very high resolution spectroscopy in an efficient way for the first time. For a well corrected telescope, the spectral resolution can be potentially increased by a factor of D/r_0 , where r_0 is the atmospheric coherence length and D is the telescope aperture size. At a good seeing site, $D/r_0 \sim 10$ in the near IR for a 8m telescope and in the optical for a 2m telescope.

Current adaptive optics systems being developed will provide high order wavefront correction in the IR for 8m class telescopes. The same adaptive optics systems will be able to correct wavefront errors in the optical for 2m class telescopes. Therefore, spectral resolution can be potentially boosted by a factor of 10 in the IR for large size telescopes and in the optical for medium size telescopes. Current state-of-the-art high spectral resolution in the optical and IR is $R \sim 30,000$. Therefore, the next generation high spectral resolution range with adaptive optics will be $R \sim 300,000$.

Optical adaptive optics spectroscopy experiments with a prototype spectrograph at the SOR 1.5m and Mt. Wilson 100inch telescopes demonstrated a spectral resolution of $R \sim 250,000$ in the optical (Ge et al. 1996a,b; 1998). This resolution is a factor of ten times higher than that provided by the same spectrograph operated under seeing conditions. The spectrograph efficiency is about a few percent, similar to seeing limited spectrographs with $R \sim 30,000$ (Suntzeff 1995).

In this paper, we report the progress on a newly developed adaptive optics optimized high resolution optical spectrograph. We also report the progress on developing silicon immersion echelle gratings for IR high resolution spectroscopy. Because IR spectrographs have to be cooled to cryogenic temperatures to minimize thermal background, spectrograph size is the major constraint factor for the design of IR high resolution spectrographs. Special technique is required to reduce the spectrograph size. The most promising technique is to apply immersion gratings made of silicon to provide a factor of 3.4 increased spectral dispersion over conventional reflective gratings. Therefore, the total cryogenic chamber size for housing the spectrograph can be substantially reduced while the resolution is maintained, and the total cost is greatly reduced.

2. An optimized adaptive optics cross-dispersed optical echelle spectrograph with $R = 200,000$

This spectrograph is designed to be coupled with the adaptive optics corrected beams through a near single-mode fiber to provide a spectral resolution of $R = 200,000$ and a large wavelength coverage. Construction and initial test of the spectrograph was finished at Steward Observatory in November, 1998. The

spectrograph test demonstrates its excellent performance, especially some results are better than expected.

The spectrograph focal length is 1.9m and focal ratio is $f/16$. It is arranged in a near Littrow configuration. A novel Maksutov-type design is adopted for the spectrograph, i.e. all surfaces are spherical, to significantly reduce the total construction cost (Ge 1998). The spectrograph is fed with a near single mode fused silica fiber with 10 μm core diameter. A fiber coupler converts $f/4$ beams from the near single mode fiber to $f/16$ spectrograph beams. A virtual image of the fiber core mimics the spectrograph entrance slit. The spectrograph collimator and camera mirror is a 10 inch diameter sphere coated with protected silver to increase its reflectivity. A Maksutov plate is used to correct coma and astigmatism. The main disperser of a Milton Roy R2 echelle with $250 \times 125 \text{ mm}^2$ ruled area and the cross-disperser of a 21 deg BK7 prism are installed in a box for protection. The prism is applied in double path to increase the spectrograph cross-dispersion. Multi-layer anti-reflection coatings are used to the Prism and Maksutov plate to provide more than 98% transmission per surface over 0.35 to 1.0 μm .

About 90 cross-dispersed orders from 0.5 to 1.0 μm , or a total of 1600 $\text{\AA}$ wavelengths, are covered by a $1\text{k} \times 1\text{k}$ CCD array with 24 μm pixel size from Apogee Inc. in a single exposure. Order separations are quite homogeneous over this wavelength range. For instance, the minimum measured order separation is 240 μm and the maximum is 320 μm . Cross-dispersed orders are well separated from each other. Stray light level between these orders are less than 2% of the peak flux of the corresponding orders. A new $2\text{k} \times 4\text{k}$ Loral CCD camera with 15 μm pixel size being developed at Steward Observatory is planned to use for scientific observations, which will cover more than 100 cross-dispersed orders and 4000 $\text{\AA}$ in one exposure.

Figure 1 demonstrates the spectral profile of a HeNe laser at 6328 $\text{\AA}$ with this spectrograph. The FWHM is 38 μm (1.6 Apogee CCD pixels), corresponding to a spectral resolution of $R = 190,000$ at the laser frequency. Because of the undersampling of the spectral resolution element with the Apogee CCD, a higher spectral resolution can be potentially achieved with this spectrograph with the finer pixel size CCD array.

3. Silicon immersion echelle developments for very high resolution IR spectroscopy

We are developing silicon immersion gratings in the Microtechnology center at LLNL. New combination of techniques to overcome previous problems and improve immersion grating quality are being and will be applied to make them useful for spectroscopic applications (see Kuzmenko & Ciarlo 1998, 1994 for technical details). The new techniques include

- Implementation of an endpoint detection system on the plasma etcher to monitor completion of the nitride etching to significantly reduce grating wavefront errors.
- Selective plasma etch processes to further reduce overetching and nonuniformity, decreasing the wavefront errors

Figure 1. Spectral profile of a HeNe laser at 6238 Å.

- Ultrasonic agitation during the KOH chemical etching to break free the hydrogen bubbles and reduce the roughness of the etched grating facets
- Deposition by maniscus coating and/or electrodeposition for high resolution printing on thick, heavy silicon disks

So far we have installed a monochromator on our plasma etch chamber as an endpoint detection system. This monochromator looks at the nitrogen emission line and shuts down the etcher when this line disappears during the etching of silicon nitride. We have also obtained C_3F_8 gas for the selective etching of silicon nitride over silicon. A number of experiments have been performed on thin wafers and we have improved our ability to selectively etch the chemical etch mask without damaging the super flat underlying silicon surface. Figure 2 shows etched testing gratings with new techniques. The grating quality is being evaluated. In addition, several thick silicon disks, 100 mm diameter and 4 cm thick have been x-ray crystal orientated and polished and are ready to be etched with new techniques.

A prototype IR immersion echelle spectrograph has been designed. A silicon grating with $100 \times 50 \text{ mm}^2$ etched surface will provide $R \sim 200,000$ in the near IR. The spectrograph size is about 1m long, 0.5m wide and 0.5m high. The spectrograph is expected to be constructed within a year.

The fiber-fed AO high resolution spectrograph is supported by NASA grant through NAG5-7200 and NSF grant through AST9731176. The silicon immersion grating development is supported by LLNL Laboratory Directed Research and Development funds. The work is performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract W-7405-Eng-48.

Figure 2. $10 \times 10 \text{ mm}^2$ testing gratings are etched on a silicon wafer for testing new techniques.

References

- Carr, J., & Najita, J. 1997, ASP conf ser. 133 "Science with the NGST", ed. EP Smith and A Koratkar, p. 163
- Ge, J., Angel, J.R.F., & Shelton, C. 1998, Proc. SPIE, 3355, 253
- Ge, J. 1998, Ph.D. Dissertation, The University of Arizona
- Ge, J., et al. 1996, in *Adaptive Optics*, Vol. 13, OSA Technical Digest Series (Optical Society of America, Washington DC), 122
- Ge, J., et al. 1996, *BAAS*, 28, 905
- Marcy, G.W. & Butler, R.P. 1998, *ARA&A*, 36, 57
- Suntzeff, N.B. 1995, *PASP*, 107, 990
- Kuzmenko, P.J., Chiarlo, D.R., & Stevens, C.G. 1994, Proc. SPIE, 2266, 566
- Kuzmenko, P.J., & Chiarlo, D.R. 1998, Proc. SPIE, 3354, 357

Slit Mask Exchange in a Cryogenic IR Multi-Object Spectrograph

P. R. Hastings and G. S. Wright

*UK Astronomy Technology Centre, Royal Observatory, Blackford Hill,
Edinburgh, EH9 3HJ, UK*

Abstract. The use of multi-slit masks in the IR offers the same observational advantages as in the visible. The major challenge, particularly with large instruments, is to move new masks into a cooled instrument and remove those no longer needed.

1. Introduction

Multi-object spectroscopy requires light from a large number of objects, placed anywhere in the available field of view, to be dispersed. Two of the techniques to achieve this in optical spectrographs are slitlet arrays and multi-slit masks at a focal plane. Slitlet arrays tend to be mechanically complex, have a fixed length of straight slitlet in the spatial direction and are confined to a single slitlet per row of detector pixels in the dispersion direction. Multi-slit masks are mechanically simpler at the focal plane, can have slits tailored to specific objects and can have more than one slit per row of pixels in the dispersion direction. Masks are thus the most promising technology for providing multi-object spectroscopy in the IR, where the usual cryogenic constraints apply:

- The masks have to be cold — from 80K to 120K for near-IR (K band and beyond) work.
- Sufficient masks for at least one night's observing must be held within the instrument.
- The mask deployment mechanism must work reliably in cryogenic conditions, must place a selected mask with sufficient accuracy and hold it securely.
- Masks must be removed from, and placed in, the spectrograph without warming the instrument significantly.

2. Cold masks

The working mask must be kept sufficiently cold. The main heat load on the mask is incident radiation falling on its outer side, which is not transmitted by the slits. The most effective way of reducing incident radiation is to place the mask at a re-imaged focal plane after a cold stop and all necessary filtration.

Figure 1. Mask Storage and Exchange.

The mask is thus exposed to radiation which is as restricted as possible in both wavelength and extent. The re-imaging system also changes the plate scale to reduce the physical size of the mask, and all its associated hardware. Heatsinking of the mask requires it to be held against a cold surface. Whether this can be done by its deployment mechanism or whether a clamp will be needed is under investigation. The mask must also behave sufficiently predictably to keep astrometric accuracy between cutting slits at 300K and operation at 80K.

2.1. Mask dimensions

Even without knowledge of the instrument or its layout, a mask size can still be estimated. For a 2048×2048 array of $18\mu\text{m}$ pixels, an f-ratio at the detector focal plane, $F_d = f/2$, and an f-ratio at the telescope focal plane, $F_t = f/15$, then the f-ratio at the re-imaged focal plane is $F_r = \sqrt{F_d \times F_t} = f/5.5$. The size of the mask is then $2048 \times 18\mu\text{m} \times F_r/F_d = 101\text{mm}$ square, and a 2 pixel slit is $99\mu\text{m}$ wide. These sizes fall comfortably within the normal realms of fabrication and handling. Existing techniques (Stilburn, 1998) of mask cutting by laser machining have already shown accuracies which are consistent with likely tolerances — a slit-to-slit positional accuracy (non-cumulative) of 10% of a slit width and a straightness of slit of 5%.

3. Mask exchange and selection

It is probable that two mechanisms will be needed — one to select a particular mask and one to move the mask to and from storage. Mask selection can be most readily achieved by storing the masks in a linear cassette as shown in Figure 1. Storage in a carousel is effectively precluded by the size of hole in the vacuum vessel needed to change it. Masks are selected by moving the cassette axially to

place the mask in a position where it can be moved to the focal plane by the mask changer.

Either a linear or a rotary mechanism can move the mask from its cassette to the focal plane or vice-versa. A linear "pick-and-place" mechanism allows the masks to be packed closely in their cassette but is more complex than a rotary mechanism in both construction and operation. A rotary mechanism requires the cassette to have intervening walls between each mask and is therefore bulkier, but allows "permanent" masks for long-slit spectroscopy to be fitted in the wheel which swings the mask from the cassette to the focal plane. Axial movement of the cassette puts the required mask in the plane of the wheel. A half rotation of the wheel both moves the incoming mask from the cassette to the focal plane and puts the currently deployed mask in the vacated slot in the cassette. The mask locates by spring pressure against fixed abutments in the slot in the wheel and is held similarly against the back wall of its housing to provide cooling and control of stray-light.

4. Exchanging the mask cassette

This, of course, is the major technical challenge — a cassette full of masks has to be removed from the cryostat, and another one substituted, possibly daily. The frequency with which this is done is an operational decision depending, of course, on the capacity of the cassette but a cassette to hold significantly more than a few night's supply would be inconveniently large. The exchange procedure outlined below starts with there being no cassette in the cryostat.

4.1. Inserting a mask cassette

An additional piece of equipment is needed — a cassette handling Dewar, shown top right in Figure 2. This ends in a gate valve - which can be sealed directly to an airlock on the vacuum vessel of the spectrograph. The airlock has a pumping port and valve on it for evacuation. Two Dewars will probably be needed, one wherever masks are cut, to carry the next set of masks, and the other at the telescope ready to receive the used masks.

After loading and evacuation of the Dewar the cassette and its masks are pre-cooled with LN₂. In operation, the Dewar is attached to the cryostat airlock, which is pumped out and sealed, and both gate valves are opened. The pre-cooled cassette is then slid into the cryostat, pushing open a sprung flap on the radiation shield and engaging with the linear drive for the cassette. The cassette is then driven into the cryostat to its forward endstop and the sprung flap in the radiation shield closes behind it. Both gate valves are closed, the airlock is opened to atmosphere and the Dewar is removed.

4.2. Removing a mask cassette

As shown in Figure 2, an empty (but cold) Dewar is first attached to the airlock which is pumped out and sealed. The cassette is then driven to its fully-forward position. A storage box for the cassette is pushed in far enough to open the sprung flap in the radiation shield and disengage the outer endstop. The cassette is then driven to the outer end of its drive where it disengages. A

Figure 2. Exchanging the Mask Cassette.

threaded shaft is attached to the end of the cassette which is then drawn into the storage box. Finally, the box and its cassette are withdrawn beyond the gate valve on the handling Dewar, the gate valves are closed, the airlock is brought up to atmosphere and the Dewar is detached.

5. Conclusion

The preceding description and illustrations suggest that a practicable handling method for cold multi-slit masks can be designed. As ever, the challenge will be in the detail design of the mechanisms, the focal plane and the handling Dewars, and these are currently the subject of a development study for Gemini.

We would like to thank our colleagues at the UK-ATC and the University of Durham for useful discussions about the practicalities of cold slit mask exchange.

References

Stilburn, J. 1998 Private communication on the Gemini MOS mask cutter

An Image Slicer for the UKIRT Imager Spectrometer

Martyn Wells, Peter Hastings and Suzanne Ramsay Howat

*UK Astronomy Technology Centre, Royal Observatory, Blackford Hill,
Edinburgh EH9 3HJ, U.K.*

1. Introduction

As infrared arrays increase in size, the possibility of providing a powerful three-dimensional spectroscopic option is becoming a reality. 3D spectroscopy is now seen as essential for general purpose, common user spectrometers on large telescopes. This introduces the requirement for deployable integral field units which are compact in design. We present the design of an integral field unit for the UKIRT Imager spectrometer (UIST). This IFU is based on a reflective image slicer rather than a fibre bundle, principally because of the requirement that the UIST IFU should work cryogenically. The UIST IFU design develops ideas used in the 3D instrument (Krabbe et al 1995) and those presented by Content (1997). The generic optical requirements of a reflective IFU are to: reformat a small area of the sky into narrow slices; reimage the slices onto the focal plane of a spectrometer such that the slices do not overlap in the non-dispersion direction; ensure that all images have a common pupil. In addition to these considerations, the UIST design aims to minimise the number of components and their complexity and to allow tight control of scattered light and cross-talk between slices without impacting the scientific value of the IFU.

2. UIST

UIST has been designed to exploit the spatial resolution achieved with the UK Infrared Telescope (UKIRT) following the installation of a tip-tilt secondary mirror. The 20-percentile seeing with UKIRT at K is measured to be 0.25arcsec. The basic scientific functions of the instrument are: 1-5 micron wavelength coverage with an Aladdin array; 0.12arcsec, 0.06arcsec cold exchangeable cameras and pupil imaging; broad and narrow band filters for imaging; imaging polarimetry; JHKLM spectroscopy, complete coverage of each atmospheric window with $R \sim 1500$; JHK spectroscopy with $R \sim 3000-4000$; HK cross-dispersed spectroscopy with $R \sim 1500$; spectropolarimetry.

The UIST optics are refractive with the exception of a cold image rotator and two flat fold mirrors, which are diamond machined from aluminium. The optical components are mounted mechanically to a central optical bench. The mechanical layout is organised as a number of modules, one of which contains the slit wheel. The IFU is designed to be mounted in the slit module and deployed as required. More details of the UIST design can be found in Ramsay Howat et al. (1998).

Figure 1. The optical layout of the image slicer.

3. Optical design of the UIST Integral Field Unit

The slicer width is chosen to match two UIST pixels (0.24arcsec) to sample the spectrum. Spatial sampling is at the detector for the non-dispersion direction and at the slicer for the dispersion direction. Therefore the sampling is 0.12arcsec in the non-dispersion direction and 0.24arcsec in the dispersion direction. The UIST design has 18 slices, giving a field of view of $6.8\text{arcsec} \times 4.3\text{arcsec}$ on the sky. When deployed, a fold mirror directs the beam into the IFU (Figure 1). The slit focal plane is located just ahead of a spherical f-converter mirror which provides an f-220 beam. This focal ratio is chosen to give a slice width of 0.95mm . The re-imaged focal plane is formed at the 18 slicer mirrors. Each of these is $0.95\text{mm} \times 27\text{mm}$ in size with a radius of curvature of 191mm . They will be machined from a common spherical surface and displaced to provide separation of the optical beams. Two sets of slices pass either side of the output slit. In the output focal plane area, two sets of spherical re-imaging mirrors ($2.4\text{mm} \times 1.2\text{mm}$ aperture) produce an f-10 beam at the focal plane. A roof mirror redirects light to the slit mask and the rest of the UIST optics. More detail on the focal plane unit is shown in Figure 2. The beam from the third fold mirror passes through a mask at the re-imaged pupil plane which provides baffling for stray light. The nine pairs of re-imaging mirrors are cut on the sides of small aluminium blocks and mounted from mechanical reference surfaces. The re-imaged slices are re-directed onto a slit mask which provides control of the cross-talk between slices on the array. The mask slices are undersized relative to the image slicer mirrors. The output from the IFU is two “staggered” slits which are separated by 36 pixels on the array, giving a small wavelength shift of each alternate slice relative to the adjacent slice on the sky.

The optical system has been optimised for use in the K window ($2.2\mu\text{m}$), where the best telescope images are expected. The final spots at the detector at this wavelength are shown in Figure 3. The optical system is diffraction limited across the whole detector in the non-dispersion direction. The reflective design

Figure 2. Mechanical details of the focal plane unit showing the staggered slits and the pupil mask, both of which help control scattered light and provide baffling.

will function over the complete 1-5 μ m range of the spectrometer. At J (1.25 μ m) and H (1.6 μ m) excellent performance is seen and diffraction effects are smaller than at K, where $\sim 20\%$ vignetted is expected. The L and M diffraction spots are larger than the slices, leading to significant ($\sim 50\%$) vignetting.

4. Mounting the IFU in the slit wheel.

Figure 4 shows a mechanical drawing of the way in which the IFU module is mounted on the UIST slit wheel. A variety of focal plane masks are available, including field stops for imaging, masks for polarimetry and a range of slit widths. The assembled and aligned UIST IFU module is mounted with the pick-off mirror assembly replacing one of these masks and is deployed simply by rotating the slit wheel. The staggered IFU slit appears as a long slit to the spectrometer optics, so that the UIST IFU may be used with any of the UIST grisms, giving two spectral resolution modes ($R \sim 1500, 3500$).

The authors gratefully acknowledge useful discussions with Robert Content and Alfred Krabbe during the development of this image slicer. Thanks also to Jason Cowan of the UK ATC Photolabs who produced the poster. The UIST project is funded by the UK Particle Physics and Astronomy Research Council.

References

- Content, R. 1997, Proc. SPIE Vol. 2871, p1295.
- Krabbe, A., Weitzel, L., Kroker, H., Tacconi-Garman, L.E., Cameron, M., Thatte, N., Samann, G., Boeker, T., Genzel, R. & Drapatz, S. (1995), Proc. SPIE Vol. 2475, p172.
- Ramsay Howat, S.K., Ettedgui-Atad, E., Bennett, R.J., Bridger, A., Content, R., Ellis, M., Hastings, P.R., Strachan, M., Wall, R. & Wells, M. 1998, Proc. SPIE Vol. 3354, p456.

Figure 3. Spot diagrams for the centre (above) and edge slices of the UIST image slicer. The circle indicates the 1st dark ring of the diffraction pattern.

Figure 4. The IFU mounted on the UIST slit wheel, 180degrees from its deployed position. The optic axis is indicated at the top of the slit module.

PMAS: An Integral Field Spectrophotometer for the Calar Alto Observatory 3.5m Telescope

M. M. Roth, W. Altmann, S.-M. Bauer, T. Becker, F. Dionies,
T. Fechner, T. Hahn, U. Laux, E. Popow, J. Schmoll, D. Wolter

*Astrophysikalisches Institut Potsdam, An der Sternwarte 16,
D-14482 Potsdam, Germany*

Abstract. PMAS, the Potsdam Multiaperture Spectrophotometer, is a new integral field spectrograph (IFS) with photometric capabilities, currently under development at the Astrophysical Institute Potsdam, Germany. The instrument is conceptually based on the lens-array / fiber type of IFS which not only provides the means of contiguous sampling of the telescope focal plane, but also employs the micropupil principle playing an essential role for accuracy and stability issues. Contrary to the concept of so-called "IFUs" which are plugged-in solutions for ordinary large field spectrographs, thus providing an additional mode of operation, PMAS is a dedicated integral field instrument with a specifically designed, high performance fiber spectrograph. We describe the basic layout of PMAS and illustrate the progress of the project.

PMAS, the Potsdam Multi-Aperture Spectrophotometer, is currently being designed and built at the Astrophysical Institute Potsdam, Germany, for use at the Calar Alto Observatory 3.5m Telescope. It is characterized by a modular design, consisting of a telescope module, an optical fiber module, the fiber spectrograph module, and the instrument control module. The replacement of only a few minor optical components in the re-imaging section of the telescope module allows for a simple change-over for operation at other telescopes. The principle of operation is of the lens-array / fiber coupled spectrograph type with low/medium resolution.

Of the various modules, the dedicated fiber spectrograph determines the overall performance in an important way. It employs large gratings and all-dioptic, apochromatic collimator and camera optics. Thanks to high-efficiency, broad-band AR coatings, transmission of the lenses is high and (without grating) nearly constant at 80% in the optical, dropping to about 60% at 350nm. The image quality according to design is excellent ($D80 \leq 20\mu$) and well-balanced over the 12° field of the camera. The collimator has a pseudo-slit of 96mm length.

The first prototype will be operated with approx. 300 fibers, i.e. with a 16×16 element lens array plus separate sky fibers, filling about half of the design capacity of the spectrograph. A planned upgrade with fibers of smaller diameter will allow up to 1024 spectra with full wavelength coverage across 4096 pixels.

The size and weight of the 200mm diameter collimator and camera tubes require special care for handling. The spectrograph housing is designed with sufficient mechanical stability in order to avoid flexure to the best possible de-

Figure 1. CAD view of PMAS instrument as seen from above the telescope flange. The compact device to the front-right is the fiber spectrograph, with the CCD dewar attached to it. A second dewar, oriented at a right angle to the spectrograph is mounted to the telescope module and allows for A&G, monitoring of field stars for differential photometry, and other special functions. The boxes to the rear-left contain the electronics and the CCD data acquisition SPARC workstation.

gree. FEA simulations of the aluminum cast housing indicate an image motion of less than $2.5\mu\text{m}$ when the instrument is tilted to airmass 2, and rotated in azimuth about 360° .

A more detailed description of the design is given in Roth et al. 1997, 1998, and Schmoll et al. 1998. The following photographs are intended to illustrate the progress being made with the manufacture of some of the subsystems.

The PMAS project is supported by the German Verbundforschung under grants 05 3PA4141 and 05 AL9BA1, and by DFG under HA1850/10.

References

Roth, M.M., Seydack, M., Bauer, S., Laux, U. 1997, SPIE 2871, 1235

Figure 2. Pre-assembled lens front sections in Zeiss workshop, front right: collimator, rear left: camera.

Figure 3. Complete camera optics in protective handling frame. Right: collimator tube section.

Figure 4. Fiber spectrograph aluminum cast housing: grating cell

Roth, M.M., Laux, U. 1998, in *Fiber Optics in Astronomy III*, ASP Conf. Ser. 152, eds. Santiago Arribas, Evencio Mediavilla, Fred Watson, p.168

Roth, M.M., Bauer, S.M., Dionies, F., Fechner, T., Hahn, T., Laux, U., Nickel, U., Popow, E., Schmoll, J., Wolter, D. 1998, in *Opt. Astr. Instrumentation*, ed. S. D'Odorico, SPIE3355, 798

Schmoll, J., Popow, E., Roth, M.M. 1998, in *Fiber Optics in Astronomy III*, ASP Conf. Ser. 152, eds. Santiago Arribas, Evencio Mediavilla, Fred Watson

An OH Airglow Suppression Spectrograph with Multi-Object Feeder

T. Maihara and F. Iwamuro

Department of Physics, Kyoto University, Kyoto 606-8502, Japan

Abstract. The design and expected performance of the unique infrared instrument OHS (OH airglow suppression spectrograph) for the Subaru Telescope is presented. The system consists of a large spectroscopic filter to eliminate most of OH lines and of a cooled spectrograph/camera (CISCO) to get low-to-medium resolution spectra as well as photometric intensity. It is also planned to provide with multi-object capability by the use of 16 manipulatable fiber probes. OHS is scheduled to be installed on the Nasmyth platform in late 1999, and to be commissioned in late 2000. The scientific justifications of the instrument are mentioned in relation with the deep infrared survey project performed by CISCO mounted on the Cassegrain port of Subaru Telescope in the early test phase.

1. Introduction

Recently built, or planned 8 - 10 m class telescopes are supposed to tackle the quest for early epoch galaxy formation, or even for the formation process of first generation stars after the Big Bang. In fact, efforts to search for the highest redshift galaxies have already broken records of redshift values at a fairly rapid pace. In order to get information in the epoch between $z=5$ and 10, or in the even higher redshift region, near-infrared observations are obviously most important. However, the capability of observing extremely faint objects in the infrared bands longer than $1\ \mu\text{m}$ would not easily be secured due to severe constraints caused by intense background emission of the OH airglow in the 1 to 2 micron region and also by the thermal radiation at wavelength longer than $2\ \mu\text{m}$.

In this report, we describe an effective method of elimination of OH airglow lines in the 1 to $2\ \mu\text{m}$ region, present a specific layout of the real instrument in the making, and show expected performance of the newly designed OH suppression spectrograph for Subaru Telescope (OHS).

2. Scientific Justifications

Theoretical studies treating the formation process of the so-called Population III stars have been carried on, giving suggestions on the initial star formation epoch at around $z=10$ to 20. Although these are not established understandings, we may agree with the view that before forming observable primordial galaxies, the pregalactic gas could have been, more or less, contaminated by a small

abundance of metals. Under such conditions, the global galaxy formation and build-up as well as the young quasar production would have occurred in the $z=5$ to 10 period, which are the virtual targets for the big telescopes together with advanced instruments in preparation.

The Hubble Deep Field (HDF) image certainly manifests that galaxies are smaller in size and mass at early epochs of the Universe. This is interpreted in the way that primitive galaxies are intrinsically small, and they have got larger as they coalesce each other possibly multiple times during the cosmological time scale. Small compact galaxies in the high- z region would be the realistic target objects of our infrared observational instruments. Assuming the mass of a galaxy as $5 \times 10^9 M_{\odot}$, we present synthetic SEDs of model galaxies with various ages after initial burst star formation assuming an exponential-type ($\tau = 0.1 \text{ Gyr}$) burst model. As seen from this figure, it is necessary for instruments to have sensitivity level of the order of $0.1 \mu\text{Jy}$ level in the 1 to 2 μm region.

Figure 1. Time dependences of SEDs of a model galaxy with a mass of $5 \times 10^9 M_{\odot}$. The location of the galaxy is assumed to be at $z=7$ (thick lines), and $z=10$ (thin lines). The filled squares represent limiting flux levels of CISCO in z, J, H and K' bands under conditions of SN=10, seeing size=0.4" in FWHM, and accumulated integration=6hrs. The filled circles are plotted similarly of FOCAS, the optical spectrograph/camera for Subaru Telescope.

A key project in the test phase of Subaru Telescope with a Cassegrain near-infrared camera (CISCO: Cooled Infrared Spectrograph/Camera for OHS, see Motohara et al. 1998) has been proposed, which is called the Subaru Deep Field

(SDF) survey. The limiting magnitudes or flux levels in the J and K' bands supposedly achieved by the SDF survey are shown in Fig.1, as compared with model galaxies. Crucial objects to be targeted by OHS may be provided in the course of this project, if accumulated integration time of some ten hours or so could be secured. The estimation of OHS performance will be given in subsection 3.2.

3. Instrument Concept

3.1. Optical Layout

The basic concept of the OH airglow suppression spectrograph is described by Miahara et al. (1993, 1994), and test results using a prototype OHS is reported by Iwamuro et al. (1994). The optical layout of the OHS for Subaru Telescope under construction is shown in Fig. 2. Here, the incident beam to OHS passing through the "SLIT" hits the "Collimator Mirror", dispersed by the "Grating", and then is focused by the big "Sp. Camera Mirror" onto the "Mask Mirror". The OH line mask of a thin metal plate is placed on the surface of the "Mask Mirror". The mask is made of a number of thin blackened stripes as many as more than 200, thereby reducing background by a factor of about 40.

Figure 2. The Subaru OHS layout. The beam train is described in the text.

The designed slit width is 0.5mm , corresponding to $1''$, so that the resolving power of the spectral image on the "Mask Mirror" (having a dimension of about $560\text{mm} \times 30\text{mm}$) is about 5500. Although the theoretical line width on the mirror is 0.25mm , we have individual absorbing lines for masking widened by a factor of 1.4 to ensure more complete suppression, resulting in area loss of a little more than 25 % by 260 line masks. Note that both the J and H band spectra are focused on the mask mirror by the use of 4-th and 3-rd orders of the grating, respectively.

The substantial improvements of the present model of the Subaru-OHS are, first, the usage of a cooled spectrograph instead of the warm prism-based

spectrometer in the prototype OHS, and, second, the exploitation of a 16-object fiber-based light guide placed in between the Nasmyth focus and the incident slit of OHS. We also intend to have a new OH-line mask which can remove more than 220 lines compared with the 130 in the prototype one, thereby reducing the background by a factor of about 40 in the H-band. The wavelengths of OH lines and O_2 lines in the J-band are referred to results of observations made by Ramsey et al.(1992) and Maihara et al.(1993). The optical throughput of the rather complicated OHS optics is estimated to be at least 40 % or more. We have measured efficiency of CISCO including the sky and telescope which is a little less than 30 % in the H-band, and is about 17 % in the J-band, including the detector's quantum efficiency. Now we estimate the overall system throughput including the transmissions of sky and telescope, as well as detector characteristics, to be about 12 % or so in the H-band.

3.2. OHS/CISCO Performance

During the course of the latest test observations spanning from February to May 1999, we have found that the seeing size ranges from 0.3" to 0.6" in most nights except for bad nights exceeding 0.6" so that the average is about 0.45" in FWHM. We have undertaken the very deep survey (SDF) under this superb condition using CISCO. The limiting magnitudes of the CISCO plus OHS system when attached to Subaru telescope are shown in Table 1.

Table 1. Limiting magnitudes of OHS observations.

OBSERVATION MODE	INTEGRATION TIME (HOURS)	J-BAND	LIMITING MAG H-BAND	K'-B AND
CISCO(SDF Survey)	16	24.5	—	23.5
OHS (Spectral, R=400)	4	24.5	23.5	—
OHS (Photometry)	4	25.5	24.5	—

^aActual seeing size in the SDF survey is in the range from 0.3" to 0.5".

^bLimiting magnitudes in the table are defined by 10σ .

These are estimated on the basis of the results of the prototype OHS and the recent CISCO observations, which are basically the same as presented in Fig.1. Different integration times are assumed as tabulated in Table 1 for CISCO imaging, OHS spectroscopic and photometric observations, respectively. These are defined as limiting sensitivity for continuum emission, but we could detect fainter objects possessing emission lines.

The total number of galaxies is more than 5×10^5 per square degree at this limiting magnitude in the K' band, and therefore, the number of objects should be about 1000 galaxies or more within the $2' \times 2'$ field of the SDF survey, compared with the number of objects found in the HDF image of similar sky coverage which is about 3500 as listed in the original catalog presented by Williams et al.(1996). Note that if 10σ is taken, the total number of objects in HDF is about 1800.

3.3. Multi-Object Feeder Option

As a result of deep infrared/optical surveys, a fairly large number of intriguing objects will be obtained, and therefore OHS with multi-object capability should become a powerful spectrometer. We have designed a fiber-probe assembly in which we have 16 pairs of low-OH silica fibers with special tapers on both sides. Here the taper is defined as a diameter variance of 300-to-100 μm per 3 m at the entrance, and as 100-to-500 μm per 1m at the output terminal edge (see Fig. 3). This particular specification of fiber tapers has been determined from test results of several type of fibers.

Figure 3. The cross sectional view of adopted tapered fibers (upper). The entrance of fibers is organized to 16 individually manipulatable pairs on the Nasmyth focal plane and then transmitted to a linear format output at the OHS slit (lower).

The assembly consists of 16 manipulators driven by individual XY-stages on which a pair of tapered fibers are mounted. Each pair of fibers has two entrances of 0.6" in effective diameter intended to measure a target object and a blank sky, simultaneously (see the lower diagram in Fig. 3). We have measured light intensity distribution in the output beam using a single side-tapered fiber (N.A.=0.12) which has tapered as 400 μm - to - 100 μ per 2m and has a 100 μm core on the other side. When an F/6 beam is illuminated on the 100 μm core surface, the output beam is well concentrated ($\sim 80\%$) within the F/12 beam. The multiple XY stage mounting these fiber pairs uses ball screw type positioners driven by small stepping motors. Just outside of the Nasmyth field of a little smaller than 4 arcmin in diameter we put four stages. Each stage cover each quadrant of the field. In this way, we can choose up to 16 objects for OHS spectroscopic or photometric measurements. The paired configuration of fiber probes is intended to switch object and sky alternatively to secure better sky subtraction.

The OHS/CISCO system described here is scheduled to be tested on the Nasmyth port of Subaru Telescope in the end of 1999. Then the fiber multi-object feeder is installed and tested in early 2000, and the full operation will be started in the late half of the year just after the planned commission phase.

The authors would like to thank colleagues in Kyoto University participating in the OHS development, who are S. Oya, K. Motohara, H. Terada, M. Goto, R. Hata, and T. Taguchi. We should acknowledge Yu Nakajima and N. Kikuchi for their continuous efforts in the OHS design and fabrication. Developmental works of the OHS project are supported by Subaru Telescope staff and sponsored by National Astronomical Observatory, Japan.

References

- Iwamuro, F., Maihara, T., Oya, S., Tsukamoto, H., Hall, D.N.B., Cowie, L.L., Tokunaga, A.T., & Pickles, A.J. 1994, PASJ, 46, 515
- Maihara, T., Iwamuro, F., Hall, D.N.B., Cowie, L.L., Tokunaga, A.T., & Pickles, A.J. 1993, PASP, 105, 940
- Maihara, T., Iwamuro, F., Hall, D.N.B., Cowie, L.L., Tokunaga, A.T., & Pickles, A.J. 1993, Proc. SPIE, 1946, 581
- Maihara, T., Iwamuro, F., Oya, S., Tsukamoto, H., Hall, D.N.B., Cowie, L.L., Tokunaga, A.T., & Pickles, A.J. 1994, Proc. SPIE, 2198, 194
- Motohara, K., Maihara, T., Iwamuro, F., Oya, S., Imanishi, M., Terada, H., Goto, M., Iwai, J., Tanabe, H., Tsukamoto, H., and Sekiguchi, K. 1998, Proc. SPIE, 3354, 659
- Ramsey, S.K., Mountain, C.M., Geballe, T.R. 1992, MNRAS, 259, 71
- Williams, R.E., Blacker, B., Dickinson, M., Dixon, W., Van Dyke, H.C., Fruchter, A.S., Giavalisco, M., Gilliland, R.L., Heyer, I., Katsanis, R., Levay, Z., Lucas, R.A., McElroy, D.B., Petro, L., Postman, M., Adorf, H.-M., Hook, R.N. 1996, AJ, 112, 1335

Recent Developments in Fabry-Perot Interferometers

K.A.R.B. Pietraszewski

*Queensgate Instruments Ltd, Willoughby Road, Bracknell, Berkshire,
RG12 8FB, England*

Abstract.

The Fabry-Perot etalon has advanced dramatically since Perot and Fabry built their first instrument at the end of the 19th Century. The interferometer, that until recently was described as a laboratory curiosity in one university text book, can now be operated at wavelengths ranging from < 2000 to $> 15 \mu\text{m}$ with apertures ranging from $< 0.5\text{mm}$ to 150mm . The observing methods developed for Astronomy have also been used in environmental monitoring and analysis and Raman spectroscopy.

The technology of optical glass polishing, closed loop capacitance micrometry and piezo electric actuation has also been successfully utilized at low temperatures. Furthermore these techniques allow etalon substrates with clear apertures of 150mm to be polished to a flatness of at least $\lambda/200$, to be coated with very broad or special dielectric coatings, and to have a minimum $1\mu\text{m}$ air gap between the coated surfaces. This opens up their use further in tunable filters and higher resolution etalons for 3D imaging, spectroscopy and photometry on ever-larger telescopes.

At the other end of the aperture spectrum recent advances in the miniaturization and control of fibre coupled, small aperture, $< 0.5\text{mm}$, FPs at Queensgate Instruments have increasingly found a market in the telecommunications industry for channel analysis and selection.

1. Introduction

Fabry-Perot interferometers are increasingly being recognized as a powerful tool for spectroscopy, both in Astronomy and other areas such as Telecommunications and environmental monitoring. Their high efficiency and large throughput make them ideal for low light level applications. The versatility of their construction makes it possible to choose their resolving power almost at will for the job in hand. The use of a variety of substrate materials and dielectric coating types covers a very broad spectral range. The inherent mechanical strength of the Queensgate ET Series II etalon construction can be exploited in interferometers whose working apertures range from 28mm to 150mm and larger making it possible to match these etalons to any size of telescope. The closed loop servo control of the cavity spacing and plate parallelism coupled with the mechanical rigidity and robustness of the glass assembly make these etalons unsurpassed for stability with typical drifts of $0 \pm 50 \text{ pm K}^{-1}$ and cavity spacing noise of $10 \text{ pm Hz}^{-1/2}$ rms.

2. The Queensgate Fabry-Perot (FP): A Brief Historical Background

The first engineered Fabry-Perot was built by Jobin in 1899 and described by Fabry (1923) using semi-silvered mirrors, the transmission function having been described mathematically by George Airy in 1831. Chabbal (1958) and Jacquinet (1954) described the high etendue properties of the FP. Other methods of scanning FPs such as tilt tuning, whether mechanically or photographically, pressure scanning whereby the refractive index within the cavity was changed rather than the plate separation were attempted with varying success and utility.

By the time Queensgate Instruments came into being in the late 1970s piezo electric ceramic actuators built into optically contacted etalon assemblies had been attempted with limited success. By incorporating capacitance sensors and a highly stable feedback control system Queensgate were able to start the development of a world beating Fabry-Perot interferometer that eliminated the hysteresis and creep effects and maintained both the parallelism and separation of the plates. See Hicks et al 1984.

The traditional market for these etalons for over 20 years has been in Astronomy but in recent years the etalons have been built into LIDAR instruments for environmental monitoring, into Raman devices for the study of surfaces and surface contamination in, for example, semiconductor manufacture, in gas analysis and in fibre optic telecommunications.

3. The Queensgate Fabry-Perot Control System

Using capacitance micrometers as the sensor in a feedback loop affords great precision in the measurement of changes in separation. The electrodes are attached to the mirrors and monitor their separation directly. There is zero hysteresis in the measurement, which is also very linear. The mechanical configuration ensures low drift and no power is dissipated by the capacitor so that the mirrors are not subject to internally generated temperature fluctuations. The sensitivity is such that changes in separation of the capacitors down to 3-4pm can be detected. This is considerably smaller than the 106pm diameter of a Hydrogen atom.

The control system is a null system based on the comparison of the sensing capacitor to a reference or other capacitor. Changes in spacing are generated by changing the balance between the capacitor by means of offsets generated within the controller.

The controller has very low noise by design, as stated above, and imparts a jitter to the mirrors of amplitude $<0.1\text{nm}$ rms. The thermal stability of the controller imparts a drift of $<0.05\text{nm/K}$ to the cavity spacing and tuning between cavity spacing settings within the range of the piezo actuators has a response time of $<1\text{msec}$.

The long-term drift of the system at a constant temperature, pressure and humidity is of the order of $\lambda/10,000$ per hour, $\lambda/3000$ per day will be better than $\lambda/200$ indefinitely.

4. The Queensgate Fabry-Perot Etalon

4.1. Mirror flatness

Advances in polishing techniques have allowed the surface figure of etalons to be made flat over a much larger proportion of the available area of the mirror blank. Edge effects are now confined to within 1mm of the edge of the mirror itself. Mirrors are individually polished to a flatness of $\lambda/50$ or better and then matched to each other to give a transmitted wavefront error of $\lambda/200$ or better for the interferometer. The flatness is defined at a wavelength of 632.8nm, that of the HeNe red line. Departures from flatness are generally spherical.

A flatness of $\lambda/200$ represents a maximum difference in cavity spacing of about 3nm and for an etalon mirror of diameter 70mm, polished to $\lambda/200$ the radius of curvature of the surface is approximately 50km.

With care it is possible to polish the mirrors such that they are matched to $\lambda/300$ or better, however, such levels of flatness are difficult to maintain through the coating process and would be effected by temperature gradients during operation.

The surface roughness of the polished surfaces of the mirrors is typically 3Å peak to peak. This level of roughness does not degrade the performance of the working etalon.

The surface flatness and roughness will depend upon the material being used, which in turn is dictated by the wavelength region of interest. The values quoted here are for etalons made of fused silica and transparent crystals used in the visible.

Fabry-Perot etalons have traditionally been used in well-collimated beams and therefore flat plates have been used to maximize the finesse and transmission of the cavity. Matching the curvature of etalon mirrors to faster beams in an attempt to reduce the number of optical elements within an instrument is possible although not practical. Apart from the difficulty of matching two curved surfaces to $\lambda/200$ (and the measurement problems raised) such an etalon would be of very limited use. The two mirrors would need to have a common centre of curvature. The etalon would have to be aligned so that the incident beam also shared this centre of curvature. Scanning of the cavity spacing would not be possible beyond a small fraction of a fringe as this would introduce a 'wedge' between the plates by mismatching the centres of curvature.

4.2. Cavity Spacing

The separation of the coated etalon mirrors is set by the length of the three piezo actuators which hold the two mirrors rigidly apart. The lengths of these actuators are polished to be well within $1\mu\text{m}$ of each other so that corrections for parallelism are minimal. It is possible to polish these actuators such that the absolute separation of the mirrors is as low as $2\text{--}3\mu\text{m}$. The risk of damaging the delicate, coated surfaces makes it imprudent to attempt smaller spacings bearing in mind that the size of the average dust particle falls within this region. At these cavity spacing levels there is also a risk driving the mirrors together under servo control as the piezo actuators have a nominal range of $\pm 1.8\mu\text{m}$.

Tunable filters, such as the Taurus Tunable Filters (TTFs), have a cavity scan range in excess of $10\mu\text{m}$ and great care must be taken in the manufacture

of these etalons to avoid building a self-damaging etalon (Bland-Hawthorn & Jones 1998*a, b*). In recent years, several such etalons have been made for Dr. J. Bland-Hawthorn which have a large cavity scan range and a very small minimum cavity spacing giving great utility in 3d imaging spectroscopy.

It is not necessary to limit the cavity spacing to a few microns. It is possible to build up a very large cavity separation using spacer rings as supports for the piezo actuators and capacitance micrometers. Such techniques have been used recently to manufacture etalons with spacings of the order of 10cm. The very high resolving power of such etalons has been utilized in the testing of semiconductor laser diodes emitting very narrow lines.

4.3. Coatings

The coating deposited onto the etalon mirrors defines both the usable wavelength range and the finesse of the etalon. This together with the cavity spacing determines the resolving power of the assembled etalon. These parameters are always very carefully chosen to best match the requirements of any given application.

The pressure to produce versatile and cost effective etalons that have a long-term usefulness has resulted in the evolution of coating designs towards increased wavelength coverage.

Increasing the coating bandwidth usually requires an increase in the number of layers of dielectric material that must be deposited on the substrates. Very broad coatings often cause distortions to the surface figure beyond that which is expected for simpler coating designs with the result that the transmitted wavefront error is degraded. Such degradation may not always be significant, however, for low light level applications where the flatness of the plates determines the overall efficiency of the instrument it is essential that the plates remain as flat as possible through the coating process.

In order to produce the best possible flatness we have developed a technique whereby the plates are polished, coated and mapped, in effect calibrating the coating induced stresses. The etalon is then re-polished and left with a corrected flatness that will compensate for the stresses of the coating.

Broadband coatings with many layers also have more variation of reflectivity across the wavelength band. This is a natural consequence of there being many layers within such coatings that give rise to inter-layer interference effects. Coating designers are being constantly pushed towards reducing these wavelength dependent reflectivity variations by refining the designs and improving the control of the deposition of the layers.

When used in low cavity spacing etalons broadband coatings display a rapidly changing phase change with wavelength on reflection. This gives rise to the effect that the cavity spacing appears significantly bigger than expected from the physical separation of the mirrors. =20

Figure 1 shows a typical simple coating and a typical broadband coating. The quarter wave stack design produces a much more reproducible overall shape than the broadband coating.

4.4. Special Coatings

Broadband coatings with a nominally constant reflectivity over the wavelength range will exhibit an increasing finesse with wavelength. This is due to the

Figure 1. Quarter wave coating with 90% reflection at 532nm (dotted line); broadband coating (solid line) with $85\% \pm 3\%$ reflection, 550nm to 850nm.

effective flatness of the plates scaling with wavelength provided that the reflectivity is not the limiting defect. This is not always desirable and a coating with constant finesse may be more appropriate. It is possible to design coatings that have a falling or rising average reflectivity with wavelength. The variations in reflectivity remain in these designs. Queensgate have made several such coatings.

When using etalons with very low cavity spacings it is desirable to minimise the effect of the variation in reflective phase across the wavelength range in order that the minimum cavity spacing is achieved throughout the range. Coatings can be designed to minimise the variation of the phase change on reflection. Queensgate have some experience with this approach, however, the coating reflectivity variations can become uncontrolled resulting in large variations across the band despite the low reflective phase variations.

The ability to deposit individual layers within a coating to within 1 to 2nm of a given thickness can be used to produce etalons with two (or more) regions with different cavity spacings on the same substrate. The difference in cavity spacing between areas can be 'tuned' to spectral feature such as the red and blue sides of an absorption line. Incidentally, this tolerance is sufficient to give the observed run to run variability of broadband coatings, however it is used to great effect in the manufacture of narrowband, multi-cavity interference filters.

Coatings with low absorption and scattering are available that are suitable for applications where very high finesse, >200 , are required. These coatings are deposited by the Ion Beam Sputtering technique at elevated temperatures. Low loss coatings deform large etalon plates and are only suitable for very small aperture applications such as in fibre optically coupled etalons where the effective aperture is $<0.5\text{mm}$ in diameter.

4.5. Aperture size

The aperture of an etalon is determined by the application. Imaging systems on large telescopes viewing large fields require large aperture etalons. Queensgate

have recently manufactured etalons with working apertures of 150mm. There is scope to increase this further to 200mm. The polishing techniques exist that will cope with this aperture. The main obstacle in producing ultra-large aperture etalons is likely to be cost. The large size of the substrates will require special production runs to produce the blanks that will then take a long time to polish. Coating chambers large enough to accommodate these sizes exist but would need extensive calibration and to ensure coating uniformity. It is possible to test large etalons for flatness by a variety of means to a higher precision than is possible with smaller etalons. Extending to beyond 200mm is an exciting prospect.

5. Evolution from the FP

The techniques of actuation and position sensing developed for QI etalons have been adapted for use in Nanopositioning and NanoMechanisms. Sub-nanometre positioning precession is easily possible and mechanisms that move $100\mu\text{m}$ that have negligible out of plane movement have been developed. These mechanisms are used in the testing of computer disc drives and in the manufacture of semi-conductors increasing process efficiency considerably. The most exciting aspects of Nanopositioning have involved the development of digital controllers with 21-bit precision. These controllers give access to the feedback loop parameters allowing rapid scanning, variation of response time according to demand and the control of noise and bandwidth. Digital controllers of this type would be of great benefit to the operation of FP etalons.

6. Cryogenic and Space etalons

Queensgate Instruments have made cryogenic etalons and space systems using etalon technology (Reay and Pietraszewski 1995). These activities are very exciting but require huge resources, which have, as yet, not materialized.

The three-channel etalon controller, piezo actuators and capacitance sensors configured as in an ET Series II etalon are at present driving the moving mirror in the WINDII Michelson instrument on the UARS satellite and since launch have generated in excess of 25 million images. Etalon mirrors manufactured by Queensgate are also observing the Sun in the LASCO instrument on SOHO.

References

- Fabry Ch. 1923, *Les Applications des Lumineuses Editions de la Revue d'Optique theorique et instrumentale* (Paris).
- Bland-Hawthorn J., Jones, D.H. 1998a, *PASA*, 15, 44
- Bland-Hawthorn J., Jones, D.H. 1998b, *In Optical Astronomical Instrumentation*, Kona, Hawaii, SPIE Proceedings, vol. 3355, in press
- Chabbal R. 1958, *J. Phys. Rad.* 19, 295
- Jacquinet P. 1954, *J. Opt.Soc.Am* 44(10), 76
- Hicks T.R., Reay N.K. and Atherton P.D., 1984, *J Phys E: Sci Instrum* 17, 49-55
- Reay N.K. and Pietraszewski K.A.R.B., 1994, *ASP Conference Series*, Vol 71, 1995

OSIRIS: A Tunable Filter Spectrograph for the GTC 10m Telescope

J. Cepa

*Instituto de Astrofísica de Canarias, E-38200 La Laguna, Tenerife,
Spain*

*Departamento de Astrofísica, Facultad de Física, Universidad de La
Laguna, E-38071 La Laguna, Tenerife, Spain*

J. Bland-Hawthorn

*Anglo-Australian Observatory, P.O. Box 296 (167 Vimiera Road),
Epping, NSW 2121, Australia*

J.J. González

*Instituto de Astronomía-UNAM, Apdo. Postal 70-264, México D.F.,
Mexico*

and the OSIRIS consortium

Abstract. Grantecan (GTC), a Spanish company created for the design and construction of a 10m telescope in the Canary Islands, recently approved the development of a tunable imaging spectrograph (OSIRIS: Optical System for Imaging and low/intermediate Resolution Integrated Spectroscopy) as the first light instrument for the GTC. This Keck-style telescope, now fully funded by the Spanish government, is to be built on La Palma with First Light anticipated in 2003. OSIRIS is a fully transmissive spectrograph which combines low to moderate resolution spectroscopy ($R=500-5000$) and tunable imaging ($R=100-3000$) over an $8'$ field (goal: $11.3'$). An off-axis, multi-coated, reflecting collimator gives the instrument high throughput over a very wide spectral range (300-1600nm) as well as a robust, compact design. The instrument is expected to work at both the Nasmyth and Cassegrain $f/17$ foci, and incorporates the latest developments in tunable filters and volume phase holographic gratings. The design includes three cameras to optimize performance over three wavelength intervals (in anticipated order of commissioning): (I) 365-1100 nm, (II) 300-450 nm, (III) 1000-1600 nm.

1. Introduction

GTC, a 10m segmented mirror telescope was fully funded and approved in March 1998 by the spanish government. GTC will be operative in the Observatorio del Roque de Los Muchachos (La Palma island, Spain) late in 2003. The civil work

of the building will begin by October 1999, and Schott Glass has already been contracted as the supplier of the blanks for the primary mirror.

An Announcement of Opportunity (AO) for the first generation of instruments was issued in July 15, 1998 with deadline December 15, 1998. Six instruments answered the AO and, in March 11, 1999, after a Conceptual Design Review, OSIRIS was selected as the GTC first light optical instrument. OSIRIS has entered now into the Preliminary Design, with a Preliminary Design Review (PDR) scheduled early in 2000 and a first light on telescope for instrument commissioning by August 2003.

OSIRIS will be the first instrument belonging to the new class of tunable spectrographs (e.g. Barden, Arns & Colburn 1998). This new class of instrument implements the last advances in Volume Phase Holographic Gratings (VPHs) and tunable imaging (using tunable filters, TFs) combined with charge shuffling. An articulated camera allows one to select the desired VPH order while a tiltable wheel allows one to tune the desired wavelength within the same VPH order. The high efficiency of VPHs and their similar behaviour with the s and p polarization modes, combined with the high OSIRIS high throughput optical design, will provide the GTC with a unique tool for very efficient multi-object spectroscopy. Also, the combination of charge shuffling with multiobject spectroscopy allows to observe a larger number of targets. Moreover, tunable imaging facilities using tunable filters (TFs) allow us to optimize the observation of very faint emission line objects well beyond other existing facilities, as have already been proved in 4m class telescopes. Finally, OSIRIS will provide a high number of very interesting extensions in the near future: polarimetry (imaging and spectroscopy) and narrow band imaging up to $1.65\mu\text{m}$ (using a NIR TF plus a new camera) among others. This powerful capabilities, combined with the relatively simple design and low weight, will provide the GTC with the first light instrument needed for a full exploitation of its capabilities.

2. The OSIRIS Consortium

The OSIRIS Consortium is constituted by the Instituto de Astrofísica de Canarias (IAC), the Instituto de Astronomía of the Universidad Nacional Autónoma de México (IA-UNAM) and the Universidad de La Laguna (ULL). The following researchers compose the Instrument Definition Team (IDT): G. Hill, E. Alfaro, I. González, M. Sánchez, J. Casares, H. Castañeda and F. Prada. The following engineers have been in charge of the conceptual design: F. Cobos and C. Tejada (IA-UNAM, optics) and C. Militello (ULL, mechanics).

IA-UNAM will be in charge of the optical design and optical construction. The AAO is charged with the design and characterization of TFs, VPHs and their coatings, and the implementation of charge shuffling, among other tasks.

3. OSIRIS observing modes, user requirements and throughput

The following tables show OSIRIS observing modes, throughput, efficiencies, limiting magnitudes and user requirements, respectively.

OSIRIS has been designed to fit within the Cassegrain envelope of the GTC (2 meter diameter per two meter length), which is more restrictive than the

Observing Mode	Charge Shuffling	Telescope Nodding	Fast Photometry
Direct Imaging	N	N	Y
Tunable Imaging	Y	N	Y
Long Slit Spectroscopy	Y	Y	Y
Multiobject Spectroscopy	Y	Y	N
Fast Photometry	Y	N	—

Table 1. OSIRIS observing modes available on first light. Capabilities for charge shuffling, telescope switching and fast photometry are specified for each mode. After first light OSIRIS capabilities will be extended to allow imaging (conventional and tunable) and spectroscopy polarimetry.

Wavelength (Å)	OSIRIS Efficiency
3650	0.23
4000	0.46
5000	0.56
6200	0.56
10600	0.08

Table 2. OSIRIS efficiencies. For this calculations, on axis thickness for each lens is used. Collimator reflectivity is for fresh deposited aluminium. On the air-glass interfaces, blue coating MgF_2 + SolGel is assumed. Detector response and telescope efficiency have been included.

OSIRIS Monochromatic efficiencies				
U	B	V	R	I
0.15	0.31	0.38	0.38	0.26

Table 3. OSIRIS net monochromatic efficiencies on telescope including telescope, atmosphere, optics, filters and detector.

Exp. time (s)	Limiting magnitudes				
	U	B	V	R	I
100	24.3	25.9	25.5	25.4	24.6
3600	26.2	27.8	27.4	27.3	26.5

Focus	Nasmyth, Cassegrain
FOV	8' in diameter (goal 11.3')
Pixel size	$\sim 0.12''$
Wavelength Range	0.365-1.1 μm
Image quality	80% EE < 0.25" (goal 0.2")
Detector	Two buttable 4k $\times$ 2k of 15 μm MIT/LL CCID-20 blue enhanced
Conventional filters	18 positions (goal 27) broad band, plus order sorters
Tunable Filters	Blue and Red TFs: 0.30-0.45 μm and 0.45-1.00 μm
Charge Shuffling	
Fast Shutter	Fast photometry & spectroscopy
Resolutions	500, 2500 (goal 5000) using VPHs
Long slits	Different fixed widths
Slit mask wheel	20 masks (goal 30)
Multiobject Spectroscopy	20-200 targets (at R=500) Conventional Spectroscopy and Microslits with beam switching, respectively

Table 5. OSIRIS user requirements

Nasmyth platform envelope. Then OSIRIS can be used in both foci. The total weight will not exceed one ton.

The optical design is fully transmissive and consists in one off-axis parabolic mirror as collimator and a camera. At the pupil and before there are three tiltable filter wheels (four is a goal) with 10 positions each. One wheel is located in the pupil and contains TFs and VPHs. Another two (goal 3) before that contain broad band and order sorting filters. Wheels can be tilted to minimize TF ghosts or to tune VPHs to the desired wavelength. The camera plus detector is articulated and can rotate with respect to the pupil to select the desired VPH order. With this system and immersed VPHs a resolution up to 5000 is achievable.

The aperture (mask) wheel is located on the opposite side of the articulated camera and can accommodate up to 20 masks (goal 30).

References

Barden, S.C., Arns J.A., and Colburn, W.S. 1998, SPIE, 3355, 866

Multi-notch and Multi-band Interference Coatings

S. Cianci

*Chatterton Astronomy Department, School of Physics, University of
Sydney, NSW 2006, Australia*

J. Bland-Hawthorn

*Anglo-Australian Observatory, PO Box 296, Epping, NSW 2121,
Australia*

J. O'Byrne

*Chatterton Astronomy Department, School of Physics, University of
Sydney, NSW 2006, Australia*

Abstract. Astronomical instrumentation generally makes use of filter components to isolate discrete spectral bands. Many of these components use thin film interference coatings to create the filtering properties. Improved thin film deposition and monitoring techniques provide greater control over the spectral output of the filter, allowing for more complex designs for specific applications. One multi-band filter design, with pass-bands at $H\alpha$ and $H\beta$ that will allow for high spectral resolution imaging at these two widely spaced emission lines, is discussed here. It will also make use of multi-notch anti-reflection coatings. This filter will be made at the University of Sydney, in conjunction with Francis Lord Optics.

1. Introduction

In recent years, there have been major advances across a broad front in the application of thin film, interference coatings. Technological improvements in deposition and monitoring techniques permit much better control of 'out of band' interference. This allows for more sophisticated interference designs, including multi-band transmission filters and multi-notch, anti-reflection coatings. There are many aspects of astronomy where these new types of filters can be applied. We discuss here the use of multi-layer thin films for one specific application of multi-band filters, and the considerations in the filter design process.

2. Multi-layer thin film coatings as interference filters

In many commercial applications, multi-layer thin film coatings consist of alternating layers of dielectric materials with high and low refractive index deposited onto a glass substrate. For optimal performance, the materials used are chosen such that the geometric mean refractive index of the thin film components is

comparable to the substrate index. Dielectric materials (e.g. MgF_2 , ZnS , SiO_2 , TiO_2) are often used as they have a very low absorptance.

Conventional thin film coatings use 'quarter wave stacks' of dielectric materials, i.e. the optical thickness of each layer is a quarter of the nominal wavelength at which the filter will operate (Baumeister & Pincus, 1970). At each thin-film interface, the partial transmission and reflection of light creates multiple beams with different relative phases, that interfere to produce a wavelength-dependent pattern of high transmittance and reflectance. More sophisticated interference filters can be designed using 'non-quarter wave' coatings, where the thickness of each layer is adjusted to allow for much steeper cut-offs to the pass- or stop-bands.

For many years, thin film coatings have been deposited using ion-assisted deposition techniques (George, 1992, Chapter 3). Low-energy ions are used to bombard the growing film, e.g. during electron beam evaporation deposition (Martin *et al*, 1983). This minimises variation in the refractive index of film, as the ions increase the packing density of the thin film, reducing the adsorption of moisture after removal from the deposition chamber. Ion-assisted deposition, therefore, increases the optical stability of thin films, thus improving their performance for multi-layer structures. Improvements in optical monitoring and in situ ellipsometry (McFall *et al*, 1996) are also important, providing more accurate control of the film thickness.

3. $\text{H}\alpha/\text{H}\beta$ double-passband filter

Conventionally, tunable filters require an order-sorting filter to block the neighbouring orders. If rapid switching between two bands is required in narrowband imaging, a single broad blocking filter spanning both bands is usually necessary. This forces the tunable filter to space its orders widely, causing its resolution to be several times broader than optimal.

With multi-band filters, however, this is no longer true. An example, soon to be manufactured, is a blocking filter with high transmission passbands at $\text{H}\alpha$ (656.3 nm) and $\text{H}\beta$ (486.1 nm), as shown in Figure 1. Used in tandem with a tunable filter (e.g. the Taurus Tunable Filter (TTF) at the Anglo-Australian Telescope) and charge shuffling techniques (see Bland-Hawthorn & Jones 1998*a, b*), this blocking filter allows for differential narrowband imaging between $\text{H}\alpha$ and $\text{H}\beta$ using the one optical path.

This was designed using *TFCalc* software (version 3.2, Software Spectra), by adjusting the thicknesses of each layer to give a filter profile closely matching the desired profile. The passbands are approximately 300Å in width, with the centroid of each band shifted 5nm to the red from the nominal emission line wavelengths. This slight shift allows for a small redshift in the objects to be observed, with the possibility of tilting the filter to observe further blueward.

The thin film will be comprised of 31 alternating layers of TiO_2 and SiO_2 , chosen for their 'high' and 'low' refractive index values compared with glass. Near-IR transmission is rejected by depositing a 15-layer quarter-wave stack of TiO_2 and SiO_2 and a one-eighth layer of SiO_2 (at a reference wavelength of 875nm) on the opposite BK7 glass substrate face. The inclusion of Schott filter glass GG435 will assist in rejecting the near-UV transmission through the filter.

Figure 1. Transmission profile of a 31-layer $\text{TiO}_2/\text{SiO}_2$ coating on a BK7 glass substrate, with GG435 glass for blocking near-UV transmission and a quarter-wave stack for near-IR rejection.

The $\text{H}\alpha/\text{H}\beta$ filter will also include a multi-notch AR coating design, i.e. a coating that provides very low reflectance at $\text{H}\alpha$ and $\text{H}\beta$. The AR coating will be important in decreasing the effect of ghost reflections from the detector. For this filter, a 5-layer AR coating (using TiO_2 , SiO_2 and MgF_2) can be applied to the 'back' side of the device, facing the detector. This will decrease the reflectance of the filter over the two passbands and increase the overall transmittance by $\sim 2\%$.

Differential imaging between $\text{H}\alpha$ and $\text{H}\beta$ will be possible with this filter, by alternately tuning the TTF between the two emission lines, as shown in Figure 2. The TTF spans a wide range of interference orders ($m = 4 - 40$). Thus, it is always possible to choose a free spectral range such that the periodic response (shown in (b)) places the two widely spaced H-bands in (a), isolated by the double passband filter, side by side within a single order of the TTF scan range (see (c)). It therefore becomes possible to obtain accurate differential measurements through shuffle imaging between two widely spaced, narrow spectral bands. This is a major step for differential imaging of galactic nebulae since observations with the highest contrast and photometric integrity must be obtained in bandpasses which are well matched to the observed spread in emission.

There are many advantages to such differential imaging: (i) charge shuffling averages out variations in the response of the atmosphere, telescope and the detector; (ii) non-photometric conditions can be used as the ratio of the two lines is normally required; (iii) systematic errors like flatfield problems and telescope oscillation due to wind shake are greatly reduced; (iv) due to the small systematic errors, the differenced quantity often has a truly poissonian form, i.e. proper noise integrity, for physical measurement with an associated proper error.

The $\text{H}\alpha/\text{H}\beta$ filter will enable us to obtain very accurate measurements of the local dust extinction throughout spiral galaxies, and to correct the local star formation rates for dust using the well known Balmer decrement for HII regions. There is some evidence that the brighter HII regions are less dust extinguished than the fainter regions which can be easily tested (de Jong 1996). Broadband photometry, coupled with the $\text{H}\alpha$ - $\text{H}\beta$ images, should allow us to interpret blue

Figure 2. Two widely spaced bands (a) can be placed side by side in convolution space (c) when convolved with the periodic profile of the TTF, (b).

gradients in spiral disks as arising from gradients in age, metallicity or dust fraction.

There are many other applications for which multi-band interference filters will be extremely useful: e.g. [O II]372nm/[O III]500nm to test if luminous ionization regions are density bounded (e.g. McCall et al 1985); He II 468nm / Hα for investigations of the hardness of ionizing spectral shapes; [S II]/[S III]906nm and [N II]658nm/[N II]575nm for studying the electron temperature of warm, ionized gas; [N II]/[O II] to investigate metal abundances.

References

- Baumeister, P. & Pincus, G. 1970, *Scientific American*, 223, 59
- Bland-Hawthorn, J. & Jones, D.H. 1998a, *PASA*, 15, 44
- Bland-Hawthorn, J. & Jones, D.H. 1998b, *SPIE*, 3355, 855
- George, J. 1992, *Preparation of Thin Films*, Marcel Dekker, USA
- de Jong, R.S. 1996, *A&A*, 313, 377
- Martin, P.J., Macleod, H.A., Netterfield, R.P., Pacey, C.G. & Sainty, W.G. 1983, *Applied Optics*, 22, 178
- McCall, L.M., Rybski, P.M. & Shields, G.A. 1985, *ApJS*, 57, 1
- McFall, W.D., McKenzie, D.R. & Netterfield, R.P. 1996, *Surface Coatings & Tech.*, 81, 72

The GSFC Lyot Filter for Night Time Astronomy

Povilas Palunas¹, Bruce E. Woodgate, Anthony C. Danks & Daniel A. Klingsmith

*Laboratory for Astronomy & Solar Physics, Code 681, NASA/Goddard
Space Flight Center, Greenbelt, MD 20771*

Richard B. Dunn

*National Solar Observatory/Sacramento Peak, P.O. Box 62, Sunspot,
NM 88349*

Abstract. We are developing a tunable narrow-band Lyot filter for night time astronomy. The Lyot filter is an interferometer which acts by recombining beams of light for which a relative phase shift is induced along the fast- and slow-axes of a birefringent crystal. The filter has been fabricated by Dunn & Ramsey and is based on a wide-field tunable design by Title & Rosenberg. The use of high transmission (97%) polarizers and index matching oil between the elements should allow 35% transmission. The aperture is 10cm in diameter and the wide-field elements remain monochromatic in f/3 beam. In this configuration the filter achieves a 400x throughput gain over Fabry-Perot interferometers. The filter has tunable range between 4200 and 7300Å and a resolution of 4-8ÅFWHM. These properties make it an ideal instrument for detection of extended emission line objects. The filter is currently being integrated with a 2048x2048 STIS-SITe CCD camera. We plan to use the instrument in two modes: at prime focus of a large aperture telescope providing a field of view of 12' (4m f/3 telescope) and behind 300mm focal length f/2.5 lens as a "parking-lot" surveyor of Galactic emission providing a field of view of 8 degrees.

1. Introduction

The Lyot filter is an interference device which consists of a combination of polarizers and birefringent crystals. The great advantage of the Lyot filter is that it can be phase corrected over a large range of acceptance angles and thus achieve very high etendue (area x solid angle). Historically, Lyot filters have had very poor transmission and were used primarily for Solar astronomy. Lockheed has recently developed polarizers with up to 97% transmission in the parallel plane that have made it possible to build a Lyot intended for night time astronomy.

¹NASA/NRC Resident Research Associate

Figure 1. A cascade of simple elements, and the system response.

2. The Lyot principle

In a simple birefringent element an entrance polarizer acts as a beam splitter, a birefringent crystal acts as a retarder and an exit polarizer acts as a beam combiner. The entrance polarizer splits the incident beam into two equal amplitude parts along the ordinary (o) and extraordinary (e) axes of the birefringent crystal. The birefringent crystal introduces a relative phase shift between the o and e beams $\delta\lambda = d(\mu_o - \mu_e)$ where d is the thickness of the birefringent crystal and μ_o and μ_e are the refractive indices along the o and e axes. The exit Polarizer recombines the o and e beams: constructively, if $d = n\lambda$ and destructively if $d = (n+1/2)\lambda$, where n is the order. A simple element can be designed to transmit at a given wavelength by using a birefringent crystal of suitable thickness. The finesse of a simple element, however, is quite poor.

The finesse can be increased by making a cascade of simple elements (Fig. 1). The thickness of the birefringent crystal in each succeeding element is twice as thick as that of the last. The thickest elements interfere in the highest order and have the highest resolution. The thinner elements act as blockers for the thicker elements.

Wide-field narrow-band imaging with interference filters is generally hampered by low etendue (area $\times$ solid angle) because off-axis rays suffer a different phase delay than on-axis rays. The great advantage of the Lyot birefringent filter is that it can be phase corrected over a large range of acceptance angles. In a simple birefringent element rays inclined toward the e -axis have a larger phase delay than on-axis rays, while rays inclined toward the o -axis have a smaller phase delay than on-axis rays. In the Lyot design the off-axis rays are phase corrected by using two birefringent crystals rotated 90° with respect to each other. The extra phase delay along the e -axis is effectively canceled by the shorter phase delay along the o -axis. A half-wave plate, which rotates the

Figure 2. A phase-corrected tunable element.

polarization of the beam by 90° , is inserted between the crystals to prevent the overall delay from being canceled (Fig. 2).

The beam emerging from the birefringent crystal is elliptically polarized with ellipticity, ϵ , varying continuously with wavelength, λ : $\epsilon = \tan \pi \delta \lambda / \lambda$. If $\delta \lambda = n \lambda$ then $\epsilon = 0$ and the emerging beam is linearly polarized parallel to the polarizer axes and if $d \lambda = (n + 1/2) \lambda$ then $\epsilon = \infty$ and the beam is linearly polarized perpendicular to the polarizer axes. A quarter-wave plate with its axis aligned along the major-axis of an elliptically polarized beam, i.e. parallel to the axis of the entrance polarizer, converts the elliptically polarized beam into a linearly polarized beam with an orientation $\theta = \arctan \epsilon = \pi \delta \lambda / \lambda$. An element with a achromatic quarter-wave plate can therefore be tuned away from the design wavelength by rotating the exit polarizer by an angle θ . The range over which the element can be tuned depends on the achromaticity of the quarter-wave plate.

3. The GSFC Lyot Filter Design

The GSFC Lyot filter design has eight elements (Fig. 3). The five highest order elements are phase corrected for wide-field performance. The lower part of the filter uses subtractive elements to avoid extremely thin retarders. The complete stack is immersed in index matching oil to reduce reflections. The polarizers are bleached and have 95-98% transmission. The filter should achieve a total transmission of 35% including the 50% polarization loss. The filter is designed to transmit at $H\alpha$ with a resolution of $8\text{\AA}$, and tuned from $4200\text{-}7300\text{\AA}$ with $4\text{-}8\text{\AA}$ resolution. The 10cm aperture accepts an $f/3$ beam giving it $400\times$ the etendue of a Fabry-Perot! A minimal number of order sorting blockers are required to cover the complete wavelength range of the Lyot filter.

The anti-backlash transmission drives each optical element in the ratio 2:1. The transmission is driven by a computer controlled stepper motor. The Lyot

Figure 3. The GSFC Lyot design.

filter is currently being integrated with a SITe STIS technology 2048x2048 CCD with 21 μ m pixels. The filter and the CCD are controlled from a Sun Workstation.

4. Summary

The GSFC Lyot filter provides a high-throughput, wide-field, narrow-band imaging capability for night-time astronomy. It has up to 400 times the etendue of a Fabry-Perot and thus much higher sensitivity for detecting extended line emission. We plan to use the filter in two modes: At prime focus of a large aperture telescope to detect low flux levels. At a 4m telescope with an f/3 prime focus the instrument covers a 12 arcmin square area. As a "parking-lot" surveyor to detect very extended galactic emission. We have a 300mm focal length f/2.5 lens which will provide a field-of view of 8 degrees.

This work was funded by by NASA grant NRA-98-03-UVG-011

References

- Lyot, B. 1944, Ann. Astroph. 7, 31
- Title, A.M. & Rosenberg, W.J. 1981, Opt. Eng., 20, 815
- Title, A.M. & Rosenberg, W.J. 1979, App. Opt., 18, 3443

Conference Summary

Bruce E. Woodgate

*Code 681, NASA/Goddard Space Flight Center, Greenbelt, MD 20771,
USA*

Abstract. This conference demonstrated the coming of age of 3-D astronomy as a standard tool, rather than as a technological niche. In particular, several integral field spectrometers are now in use, and many more are now in development for the new ground-based 6-10 meter class telescopes, and in study for space use, and are described in the papers of this conference. The astronomical roles of various forms of 3-D technologies are summarized.

1. Introduction

Until recently, 3D instruments occupied a technological development niche, and were used primarily by their builders. An exception in recent years has been multi-object spectrographs (MOS), used for redshift surveys of galaxies. The large investment in 6-10 meter telescopes coming on line has generated the recognition of a need to maximize the information from these telescopes, and triggered the need for a more substantial investment in their instruments. Together with enabling technological developments such as in fiber optics, lenslet arrays, micromirrors and microapertures, this is bringing 3D instrumentation into the mainstream of astronomy. The combination of 3D methods with the dramatically improved angular resolution being newly realized with adaptive optics is especially powerful. This also necessitates the spatial scales of the instruments be selectable for use with adaptive optics when conditions allow, or without it when they do not.

While 3D most obviously suggests two angular and one wavelength dimension, those of linear and circular polarization and time are also possible. 3D instruments may be separated into two classes; true 3D energy-resolving detectors, and 3D information projected onto 2D array detectors. We may divide the latter category into 2 sub-categories: 3D projected simultaneously onto 2D, such as Multi-Object Spectrographs (MOS), Integral Field Spectrographs (IFS), and wide field or slitless spectrographs; and 3D obtained by sequentially scanning 2D images across the 3rd dimension, such as Imaging Fourier Transform Spectrographs (IFTS), tuneable filters including Fabry-Perot, birefringent filters such as Lyot, Solc, etc., and acousto-optical tuneable filters (AOTF), and long slit spectra.

2. Energy-resolving detectors

True 3D detectors, imaging arrays where each pixel has energy resolution, are being used at x-ray energies in space instruments on ASTRO-D, and soon on CXO (AXAF), XMM and ASTRO-E (Canizares). The first three use CCDs, read out frequently so that only one photon/pixel is recorded each readout time. The x-ray energy is converted into electron-hole pairs (2.36 eV each), so that the charge read out is proportional to the photon energy. The readout time for a storage device such as a CCD therefore limits the counting rate per pixel, not a problem for most astronomical x-ray source fluxes. The maturity of CCD technology allows for large array sizes. The x-ray calorimeter (Moseley et. al., 1984) on ASTRO-E provides better spectral resolution than a CCD because the x-ray energy is more finely divided into a larger number of lower energy pieces, in this case phonons of a few meV energy each. The statistical accuracy of measuring the larger number of secondary events is higher. Each much lower energy event must be counted clearly above the thermal noise of the detector, so it must be operated at a very low temperature, about 0.1K.

Similarly, energy resolving detectors in the UV and visible, with lower energy photons to divide, require similar or lower temperatures to operate. Two technologies are being pursued, Superconducting Tunnel Junctions (STJ) (Perryman and Peacock, Labov) and Transition Edge Sensors (TES) (Cabrera). In STJs, a photon is absorbed in a superconductor, creating excess charge carriers (quasiparticles) which increase the subgap current of an attached tunnel junction. The integral of the increase in current, the charge, is proportional to the photon energy. The current may be measured by MOSFETs or by single electron transistors (SETs). In TESs, heating due to the photon energy is measured, using the sharp change in resistance through the superconducting transition, and maintaining equilibrium through voltage bias feedback, with readout of the bias changes using a SQUID.

The fabrication of both one and two spatial dimension arrays is at the stage of a few to a few tens of pixels. The investment in resources and time required to make these competitive with the projection of 3D information on large arrays with two spatial dimensions will be large, and it would be helpful in sustaining this effort during development to identify areas of uniqueness for the early small arrays.

Since these detectors count photons promptly, even small arrays can be used for fast timing studies with modest spectral resolution, such as for pulsars and cataclysmic variables. Recently demonstrations have been performed measuring the pulse profile of the Crab pulsar (Perryman et. al., 1999, Cabrera).

Important features of these detectors, in addition to their 3D capability, is their high QE and zero read noise. When substantial linear arrays with good spectral resolving power are developed, these features will make them highly desirable as cross-dispersers for echelle spectrographs, for example to measure UV absorption line spectra of faint QSOs.

Eventually, if large 2 spatial dimension arrays can be made, the large scale structure as a function of redshift could be investigated using spectra of multiple faint QSOs in the field.

3. 3D projected onto 2D

3.1. Simultaneous 3D projections

Multi-Object Spectrographs (MOS) Spectroscopy of multiple objects has been available on several ground-based telescopes for several years, and is being improved in coverage and throughput. They are of two general types: fiber optics with a moveable front end send the light from a target field into the long slit of a spectrograph; or a multi-aperture plate is made at a wide field spectrograph entrance. In the latter case the spectra are not aligned, but are dispersed from the imaged positions of the targets, but they generally have more throughput than the fiber method. The throughput of the fibers can be improved by using microlens arrays to feed the fiber inputs. Two very powerful spectroscopic surveys emphasizing galaxy and QSO redshifts are the Sloan survey described by Szalay and the 2dF survey described by Boyle.

For use in space, the movement of many fibers to feed the slit of a spectrograph would be difficult. It would also be impractical to fabricate an aperture mask for each target field. For NGST, two studies are underway, a programmable micro-mirror array and a micro-aperture array, to image the field of selected objects into a spectrograph, reported on at this conference by MacKenty and Stiavelli, and by Moseley, Fettig, Kuttyrev et. al. respectively.

MOS require prior imaging to position the apertures and to select targets of interest. They are used for spectroscopic surveys of fields with known positions but unknown redshifts, velocity structures and spectral character, and have been used most notably for mapping the large scale structure by measuring the redshifts of many galaxies at once, and for measuring the masses of clusters of galaxies via cluster dynamics.

Integral Field Spectrographs (IFS) The most notable area of change in spectroscopic methods demonstrated at the conference was the proliferation of IFS in operation and in planning, particularly to take advantage of the new 6 - 10 meter class telescopes and adaptive optics.

IFS convert a contiguous 2D area into a line array or into a sparse 2D array at the input to a spectrograph, by several methods:

A microlens array can produce very small images of each pixel which are magnified and projected so that spectra of each pixel can be interleaved between them, as in the TIGER instrument and its derivatives such as OASIS (Bacon, Ferruit).

A fiber optic array can be redistributed to produce a line array at its back end, and it can be preceded by a matched microlens array to improve throughput by focussing the light into the center of each fiber (Thatte, R. Martin, Prieto). Repositioning of multiple subarrays by moving separate imaging fiber bundles can extend this technique to allow several targets each to be examined in 3D, creating multi-IFS units (Wright, Taylor, Monnet, Allington-Smith). The use of fiber IFS for coronagraphy was described by Arribas.

A reflective image slicer using a set of small tilted mirrors relayed to a second set of tilted mirrors can convert a 2D array into a line array (Krabbe, Genzel, Ramsay Howat). This method has been adapted for space use in a design for NGST discussed at the conference by LeFevre. Larger areas for surveys can be

built up by sequential mosaicing of the fields. Instruments using far-IR image slicers for FIRST and SOFIA were described by Poglitsch.

IFS are particularly powerful for the examination of the spatial and spectral structure of one or a few extended objects. Examples include the dynamics of galaxies and the surroundings of their nuclei, gravitational lens arclets, QSO environments, starburst galaxies, young stellar objects and protoplanetary disks.

Slitless Spectrographs One of the oldest methods of 3D spectroscopy is that of using a wide aperture, typically the whole detector, most often implemented using an objective prism or grism for wide field spectroscopic surveys. The spectra of objects can overlap those of other objects in the dispersion direction, and so this method is most useful for relatively dilute fields, where overlaps are infrequent.

On HST several instruments (FOC, STIS, NICMOS) have used and are using wide field spectroscopic modes. UV wide field spectroscopic modes of the GALEX Explorer space program were described by C. Martin.

3.2. Scanning Imagers

Imaging Fourier Transform Spectrographs (IFTS) IFTS instruments are described at this conference by Bennett, who also provides comparisons with other 3D instruments. FTS designs for NGST were described by Graham and by Mailard. They use 2-beam interferometry, where the spectrum for each field pixel is created by scanning in path difference between the 2 beams. For each pixel the scans are fourier transformed to produce the spectrum at that field point. Each interferometrically combined output beam uses half the input light, but the other half can be recovered by using a second detector in the other output beam.

IFTS have the advantage that they observe all the imaging and spectral elements within the field and bandpass for all the observing time. This is advantageous for multiband imaging for spectral resolving powers where the sky noise is the dominant background, and the whole image is required without prior target selection, for example for broad surveys with unknown positions and unknown redshifts. It would also be competitive where spectra of a large number (>1000) of bright objects over a wide field are required. For higher resolution spectra, each spectral element suffers the sky noise from the whole bandpass, so the limiting sensitivity for a spectrum of a point source cannot compete with a dispersed spectrum. In addition, when used with a non-zero read noise detector such as a CCD or IR imager, each scan position must be read out separately, so the accumulated read noise can be dominant for high spectral resolutions.

Tuneable Filters Sequential images are obtained by scanning tuneable filters in wavelength. For this reason they are most competitive when all the pixels of a wide field are required but relatively few spectral pixels are needed, such as for velocity scanning of a spectral feature, emission line images, or differential images at selectable wavelengths.

Fabry-Perot Imagers Fabry-Perot imagers became a practical astronomical tool with the advent of stable etalons by application of a capacitance AC bridge

servo to stabilize piezo-electric drives by Queensgate Ltd., (Atherton et al, 1981). Their current status was described by Pietraszewski.

A current form of the Fabry-Perot was described by Bland-Hawthorn in his instrumental review. It is called the Taurus Tunable Filter, and can chop multiple times between several spectral bands without incurring multiple read noise. It achieves very low background differential imaging, minimizing systematic errors, by moving images around on the CCD detector in a process called charge shuffling. Its use in surveying for gravitationally lensed high redshift emission line galaxies was described by Hewett, for radio galaxies by Tadhunter, for starburst winds by Veilleux, and for QSO surroundings by Shopbell.

Birefringent Filters In the Lyot design for birefringent filters the beam is polarized, then its plane of polarization is rotated by a birefringent crystal, and then the beam is passed through another polarizer. The wavelength dependence of the birefringence creates a sinusoidal transmission function in wavelength. This element is followed by similar elements thinner by factors 2, 4, etc, each element removing neighboring transmitted wavelengths from that desired, until the next resulting transmitted wavelength can be removed by a blocking filter. Wavelength tuning is accomplished by rotating the elements. This method has been used in solar physics for several decades for very narrow band imaging. A wide-fielding method developed by Title and Rosenberg (1979) uses the hyperbolic fringes as a function of field angle, where the wavelength shift is oppositely blue-shifted or red-shifted in orthogonal field directions. By halving the birefringent crystals and rotating one half with respect to the other by 90 degrees, nearly equal and opposite wavelength shifts are induced for each field angle. This widens the monochromatic field by a factor 20 - 30 compared to a Fabry-Perot. Combined with a method of bleaching high contrast polarizers to obtain high transmissivity, a useful night-time tuneable monochromatic wide-field imager can be produced, as described by Palunas et. al.

Acousto-Optical Tuneable Filters (AOTF) Fast electronic wavelength tuneability for monochromatic imagers can be produced by AOTFs. A refractive index grating pattern is induced in a crystal by a strong radio-frequency beam producing a standing wave. The optical beam is diffracted by the grating by an angle depending on the wavelength. The diffraction angle can be tuned by changing the radio frequency. The acceptance angle for a given monochromaticity can be larger than that of a Fabry-Perot, but the beam diameters are generally smaller, and crystal heating by the radio beam can limit their stability.

Long Slit Spectrographs It is equally valid to obtain the 3rd dimension by using a long slit spectrograph and scanning it in the slit width direction to create a data cube. Since this is usually an operational mode of standard long slit spectrographs rather than special instrumentation, it was not highlighted at the conference. It is most advantageous when detailed spectral information is required, but relatively sparse information is needed in the second spatial dimension. This has been used many times, for example to map the velocity structure of the Crab nebula, and is being used on HST/STIS to map the velocity structure around massive black holes in galactic nuclei, and velocities of stars in the cores of globular clusters.

4. Polarization

Linear polarization occurs directly by synchrotron radiation from directed beams in magnetic fields, as in AGN jets or pulsars, or as anisotropically scattered light. An important example of scattered light is from obscured AGNs off the circumnuclear gas in the galaxy. Observing the AGN spectrum in polarized light allows it to be separated from the ambient narrow-line emission region gas. This led to the paradigm of classes of AGNs being distinguished by our direction of view to the nucleus relative to the obscuring torus (Antonucci and Miller, 1985). Similarly, information on the amount of material and its size and orientation in protoplanetary disks may be obtained from linear polarization.

Circular polarization from Zeeman splitting enables magnetic fields to be measured, and would be useful, for example, in investigating the mechanism of angular momentum exchange between stars in formation, their protoplanetary disks and polar jets.

5. Comparisons - Areas of Strengths

The Figure shows areas of parameter space where each technique has its best application for observing faint objects, using the NGST focal plane as an example:

MOS can obtain high sensitivity spectra of objects from a large area, but are limited to one or a few objects in each row in the dispersion direction, depending on the length of each spectrum.

IFS can obtain high sensitivity spectra of all the pixels in their fields of view, but would require multiple observations to cover a similar field to a MOS.

IPTS can obtain spectra of all the objects in a large field of view, but have competitive sensitivity only at low spectral resolving powers.

Numerical comparisons between several of the 3D methods were presented at the conference by Bennett.

6. Special Instruments and Techniques

A number of special projects and techniques were described which do not easily fall into the categories above, but contributed significantly to the sense of excitement at the meeting.

The Overwhelmingly Large Telescope (OWL), a concept for a 50-m ground-based telescope, was presented by Monnet. The 3 Degree Telescope, a design for a 6-m survey telescope with an etendue (area x solid angle) of $230 \text{ m}^2 \text{ degree}^2$, over 10 times larger than the Sloan survey, was presented by Angel. The Tera-el, a very unusual concept for a detector/filter with 4×10^{11} elements, a spectral hole-burning device using a dye-doped polymer film, was presented by Keller. A dispersive wide angle interferometry concept for measuring the radial velocity changes in a star due to planetary motion to $<1 \text{ m/s}$, was presented by Erskine. A scheme for interferometry between 2 geosynchronous earth satellites to measure biological spectral indicators on earth-like planets around other stars was presented by Abrams.

Instrument Phase Space - NGST example

Figure 1. Areas of advantage for 3D instruments.

References

- Antonucci, R.R.J. and Miller, J.S., 1985, *Ap.J.*, 297, 621
 Atherton, P.D., Reay, N.K. and Hicks, T.R., 1981, *Opt. Eng.* 20, 806
 Moseley, S. H., Mather, J. and McCammon, D., 1984, *J. Appl. Phys.*, 56, 1257
 Perryman et. al., 1999, *Astron. and Astrophys.*, 346, L30
 Title, A.M. and Rosenberg, W.J., 1979, *App. Opt.*, 18, 3443