Title: RADIATION HYDRODYNAMICS IN STELLAR ATMOSPHERES

Author(s): DIMITRI MIHALAS
LOS ALAMOS NATIONAL LABORATORY
X-3

Submitted to: INTERNATIONAL MEETING ON MODELING STELLAR ATMOSPHERES TO BE HELD IN TIEBIGEN, GERMANY, APRIL 2002
1 Introduction

I am supposed to provide this morning’s entertainment by talking about radiation hydrodynamics in stellar atmospheres. But I must offer two caveats. First, my last paper on stellar atmospheres was written 20 years ago, in 1981; so I am a kind of a fossil, left over from an earlier era. Indeed, my motive in coming to this conference is to learn from you about the incredible progress made in this field over the past 20 years. I have spent that time working on radiation hydrodynamics, mostly for Laboratory applications; perhaps that will be a useful contribution to this meeting.

Second, I warn you that it is tempting for an elderly scientist (I am now 62, one of the oldest people here) talking about a favorite field, to reminisce. Those of you who know me personally can affirm I can resist almost anything, ... except temptation. So I shall yield to it gracefully, but briefly.

I came to Caltech in 1959 thinking that I might work on stellar interiors and stellar evolution. Two things made me change my mind: 1) I realized that stellar-interiors modeling results in only two numbers connected to the real world, a star’s “radius” and luminosity, and even then the theoretician’s numbers must be converted to observable quantities using models of stellar atmospheres. In contrast, the spectrum of a star contains a wealth of data about its physical structure and composition, just pleading for interpretation. 2) Fred Hoyle was at Caltech in the winter of 1960/61, and gave us a discouraging account of his fruitless efforts to compute the structure of highly-evolved stars by the old technique of finding a unique fit between core and envelope models. It was a complicated process, and it wasn’t working. Ironically, I heard about Louis Henyey’s new two-point boundary-value method for making stellar models at the Berkeley IAU meeting the following summer. His work was an adaptation, to astrophysics, of methodology from nuclear-weapons codes, about which Henyey learned while consulting at Lawrence Livermore Laboratory. But the die was already cast for me.

I started my thesis work on the atmospheres of O-stars almost exactly 40 years ago, in 1961. Caltech was then a hotbed for work on stellar atmospheres, and the significance of the results to broader astrophysical questions. Jesse Greenstein’s spectroscopic stellar-abundance project was in full swing, and producing data of
relevance to ideas of nucleosynthesis in stars being developed by Willie Fowler, Hoyle, and Goeff and Margaret Burbidge, next door in Kellogg Laboratory. We had a constant stream of visitors, including luminaries like Albrecht Unsöld and Cecilia Payne-Gaposchkin, and brilliant newcomers like Roger and Guisa Cayrel, Wal Sargent, Leonard Searle, and George Wallerstein. Bev Oke was leading in the field of photoelectric spectrophotometry, and had completed a two-channel photoelectric scanner for line-profile work at the 100" coude. At Mount Wilson, Armin Deutsch was working on peculiar A-stars, Paul Merrill on the spectra of evolved giants, and Bob Kraft on the spectra of Cepheids. Also, Bob Christy had begun his monumental computational work on RR Lyr pulsations, using techniques he had mastered at Los Alamos during the Manhattan Project.

In the autumn of 1961, Caltech got access to an IBM 7090 computer at JPL. Crucial work had been done by Gene Avrett, Max Krook, & Steve Strom at Harvard on computing non-grey model atmospheres in radiative equilibrium. It was clear to me then that one had both the hardware and software (thanks to John Backus's group at IBM having developed FORTRAN) to attempt a computational analysis of stellar atmospheres. If one were to repeat my two-year thesis work today, using modern algorithms and a good workstation, the entire set of computations could be done in a few minutes! For the next 18 years I worked on problems of non-LTE modeling and line-formation in static and expanding atmospheres with my colleagues and friends Larry Auer, David Hummer, and Paul Kunasz. By then, I felt that, given only the computational techniques I knew, and the inadequate atomic database needed, we had reached a dead end.

So I changed fields. In that connection, let me say that in forty years of teaching, I have told every class of students that "One of the more progressive forces in astrophysics is called the funeral". It is when the Old Man releases his death grip on a field, and new blood comes pouring in with the energy and inventiveness of youth, that there is suddenly rapid, and previously unimaginable, new progress. Happily, I did not have to follow my dictum to its ultimate conclusion to verify its truth; I simply got out of the way, and you have done it for me.

2 Radiation Hydrodynamics

I have often told colleagues, that radiation hydrodynamics becomes important when radiation dominates the energy or momentum content, or the energy or momentum transport in the composite radiating fluid comprising radiation and material particles. An estimate of when radiation is important in a flow is the ratio $R$ of the material internal energy density to the radiation energy density. For a perfect gas and equilibrium radiation, one finds:

$$R = \frac{3k}{2a_R} \left( \frac{N}{T^3} \right) \sim 2.8 \times 10^{-2} \left( \frac{N}{T^3} \right)$$

(1)
where \( k \) is Boltzmann’s constant, and \( a_R \) is the radiation constant. Obviously radiation is most important at high temperatures and/or low densities. One finds that \( R \approx 1 \) when \( T_{\text{keV}} \approx 2 \rho^{1/3} \), where \( T_{\text{keV}} \) is the temperature in measured in kilovolts (~\( 1.2 \times 10^7 \text{K} \)). This formula shows that radiation must be important in the dynamics of any object from which one is observing X-rays, characteristic of temperatures of 1 - 100 KeV (not to mention gamma rays!), even at “normal” densities (i.e. 1 - 10 gm/cm\(^3\)). One notes that radiation is dynamically important at the center of the Sun, and becomes more important in the cores of more massive stars which have higher temperatures and lower densities at their centers. Moreover, radiation can also act to change the excitation and ionization state of the material, often in ways far from equilibrium. As a result, other important mechanisms can come into play. Some examples of these different regimes are discussed below.

2.1 Laboratory vs. astrophysical radiation hydrodynamics

While there are a number of common themes in both Laboratory and astrophysical radiation hydrodynamics, there are also important differences, and it pays to be aware of them. First, in Laboratory phenomena, a central problem is often geometry, because very different materials (e.g. low-Z and high-Z) can be within optical proximity (i.e. one photon mean-free-path) of one another. In such cases it may be unrealistic to use 1-D, or even 2-D calculations, and imperative to attempt full 3-D simulations. These efforts require an immense amount of computation and sophisticated algorithms. Such is the goal of the ASCI (Accelerated Strategic Computing Initiative) at both Los Alamos and Livermore. By the end of 2001, Los Alamos will have a massively-parallel machine capable of 30 TeraOps, and in fiscal, it expects to get a 100 TeraOps machine; Livermore will get comparable machines. Second, the physical properties, e.g. opacity and EOS, of the exotic materials in some Lab work may not be known with great precision for the extreme conditions encountered experimentally. Third, time-scales in experiments of interest can be extraordinarily short, and it may be necessary to track radiation fronts traveling almost at the speed of light. On the other hand, one can sometimes make simplifications in the treatment of radiation transport. Typically, densities are so high that collision rates dominate, and LTE is a good approximation. Also, in some problems, there are no open boundaries, so it is possible to use the diffusion approximation. A penetrating discussion of such problems can be found in the brilliant book by Zel’dovich & Raizer [133].

The situation in astrophysical problems of interest to this group is almost the polar opposite of the Laboratory regime. Here one needs to know the full distribution function for photons that emerge from a star, and are received by a distant observer. The presence of an open boundary is decisive in determining that distribution function. To obtain it, one must deal with the strongly nonequilibrium conditions in low-density layers having large gradients of physical properties, which
lie between the almost-perfect-equilibrium stellar interior, and the empty blackness of interstellar space. This is where spectrum-formation in both lines and continua takes place. In these regions one can have no expectation that LTE is valid, a fact borne out by calculation and observation. The diffusion approximation is worthless here, and one must solve the transfer equation, coupled to rate equations for all mechanisms that can populate/depopulate the energy levels giving rise to observed spectral features.

For stellar atmospheres one assumes 1-D planar or spherical geometry because one generally has no information otherwise, at least for single stars. The material in stars is composed of quantum-mechanically simple light elements and ions; nonetheless, one may have to deal with millions of lines in the low ions of elements up through iron, and perhaps hundreds of millions of molecular lines. In time-variable stellar atmospheres, one must explicitly account for time-variation of the integrated moments of the radiation field (i.e. total radiant energy density, flux, and radiation pressure), which are the dynamical properties that determine its energy- and momentum-exchange with the material. But almost always the photon flight-time, $t_p \sim L/c$ ($L$ is a characteristic size of the atmosphere, and $c$ is the speed of light), is orders of magnitude shorter than the hydrodynamic time-scale, $t_f \sim \ell/v$ ($\ell$ is a characteristic length, say the optical scale-height, and $v$ is a typical fluid velocity in the atmosphere). Therefore, the radiation field adjusts "instantaneously" to the state of the material at a given time step in a hydrodynamical calculation. Thus one can calculate the angle-frequency distribution of the emergent spectrum by using snapshots of the radiation at the given time. Exceptions to this generality are: 1) if an R-type ionization front is present [1], or 2) in a resonance line, well-approximated by an effective-two-level-atom with a small effective thermalization parameter $\epsilon$, whose source function at the surface depends on photons that were created/destroyed over an entire thermalization depth $\Lambda$, which take a time $t_d \sim (\Lambda/\ell)^2(\ell/c)$ to diffuse to the surface. In such cases it might be necessary to explore the consequences of a time-dependent solution.

Another crucial difference between Laboratory and astrophysical radiation hydrodynamics is the absence or presence of gravity. In Laboratory flows the experimental apparatus is too small for gravitational gradients to be significant. An exception is if one is modeling the fireball from an intense (nuclear or not) explosion; here gravity comes into play in determining the shape and internal dynamics of the fireball. In contrast, gravity almost always plays a basic role in astrophysical radiation hydrodynamics.

Failure to note these differences has led to miscommunication between the astrophysical and the aerodynamical/physical-hydrodynamics communities. An example is the puzzlement of both groups in I.A.U. Colloquia 12 and 28 [124, 125]. The physicists did not grasp that the astronomers have a transparent upstream boundary condition, because usually they usually deal with problems [133], in which all photons emitted at the viscous shock surface are re-absorbed upstream,
their energy is thermalized into the material, and carried back downstream into
the shock. This simple fact changes the shock morphology drastically.

2.2 Examples of astrophysical radiation hydrodynamics

I list below some astrophysical phenomena where radiation hydrodynamics is
of central importance. The first four are "legitimate" subjects for discussion at
this conference; the others really have other venues.

- **Convection**

  Convection results when a large luminous flux encounters a steep increase
  of opacity in gravitationally stratified material, and energy transport by hy-
  drodynamic motion is more efficient than photon transport. In stars, con-
 vection produces multidimensional flows. If photons can freely move through
  convective "cells", then one must deal with multidimensional radiation trans-
  port; such problems will be discussed later in this meeting. In Laboratory
  flows where gravity is irrelevant, opaque material exposed to "hot" radiation
  experiences *ablation* instead of convection. Moreover, Laboratory flows are
  generally too short-lived for even a single convective turn-over time, though
  small-scale *turbulence* produced by shear flows and various instabilities may
  occur.

- **Stellar winds**

  Radiation-driven stellar winds, discussed later in this conference, are
  unique because they result from photon *momentum* deposition, rather than
  energy-deposition. This regime is unusual because if the photons were in
  equilibrium with the material, they would have much lower momentum
  per unit energy, i.e. \((\hbar v/c)/(\hbar v) = 1/c\), than the material particles \(\sim
  (mv)/(mv^2) = 1/v\). Driving by radiation force on the material occurs be-
  cause the material is rarefied, but has strong scattering (resonance) lines
  that can remove momentum from the intense stellar radiation field, while
  not absorbing much energy. (In passing, let me mention a pet peeve: One
  should not say "radiation pressure drives stellar" winds. Radiation produces
  a force on the material. This force is related to a pressure gradient through
  a momentum-conservation equation.) Again, gravity plays a key role in the
determining the subsonic-to-supersonic transition of the flow. Most studies
of stellar winds assume spherical symmetry; but there is evidence that stellar
winds are unstable, and degenerate into complex flow patterns.

- **Stellar pulsation**

  Here the key mechanism, originally proposed by Eddington in the early
1900's, is that the opacity variation of ionizing/recombining stellar material,
in response to changes in temperature and density, can act as a *value* on the
radiation flux created at the star’s center. Thus the material becomes a kind of *thermodynamic engine* that drives pulsation. I shall talk further about pulsating stars below.

**Supernovae**

Radiation hydrodynamics is key to core-collapse supernovae. The collapse is triggered by endothermic photodissociation of iron nuclei in the energetically-inert core. This sink of internal energy leads to an abrupt gravitational contraction. Degenerate electrons are crushed into protons to form neutrons and neutrinos. The crash of the star's now-unsupported envelope onto the rigid neutron core produces a gigantic shock. Transport of neutrinos out of the shock bleeds off its energy, and it "fizzles". Apparently the envelope above the shock becomes unstable to convection, transporting energy efficiently upward. Once radiative transport again reigns, the high luminous flux creates a D-front [76], which drives an emergent shock. The observed phenomena result from photon transport in a mildly relativistic expanding, opaque, envelope. All the tools presented at this meeting can be used to model observed supernova spectra.

The following list contains objects outside of the range of phenomena appropriate to this conference, but which are nevertheless interesting.

**Novae**

Radiation can be important both dynamically, and in determining the excitation/ionization state, in novae. But novae are complicated because they occur in double stars where the geometry of the flow is quite complex.

**Protostars**

Radiation hydrodynamics in protostars enters in describing how essentially the entire luminosity from the object is generated in the intense accretion shock at the surface of abrupt transition from free-falling material to a hydrostatic envelope.

**Radiation-driven star formation**

In a large molecular or atomic interstellar cloud, formation of even a single O-star (say, via Jeans instability, triggered when a denser part of the cloud goes through a spiral-arm shock) can result in radiation trapped in the "hohlraum" provided by the optically-thick "walls" of the cloud. If strong enough, the radiation will drive ablation shocks into nearby density concentrations, thus producing radiatively-driven collapse of more stars, hence clusters of nearly coeval stars.
• **Radiative Acceleration of Interstellar Clouds**

  The Oort-Spitzer "rocket effect" [100] is used to explain high-velocity interstellar clouds in our Galaxy. Here luminous O-stars ionize material on one side of the cloud; the hot ionized material expands explosively; and the reaction force to this flow drives the cloud exactly as its hot exhaust drives a rocket. In Laboratory jargon one would say a Marschak wave [91] produces an ablation front driving a blowoff that accelerates the medium. Or, in the words of Freeman Kahn's seminal work [76] it can be described as a weak or critical D-type radiation-front, depending on the incident radiation flux. Kahn's categorization of radiation fronts also proves useful in describing the radiation hydrodynamics of pulsating stars.

• **Black Holes, AGNs, and Quasars**

  The radiation produced from dissipation of gravitational potential energy (gravity again!) of material falling into the accretion disks around black-hole remnants of stars, and massive black holes at the centers of active galactic nuclei and quasars, merits more detailed study, using full transport methods, and real material properties (nongrey opacities). Here one has transport in a curved spacetime. Such problems have been explored, but need further elucidation both from diagnostic and dynamical points of view.

• **The Big Bang**

  The most interesting recent development in astrophysics has been the beginnings of a physical understanding of the Big Bang. It is fascinating how physics at the smallest known space-time scales (in the regime of GUTs or string theories), determines physics at the largest observable scales (the Universe), and, possibly, vice versa. Let me remark that while back-of-the-envelope calculations of the behavior of radiation in the pre- and post-recombination eras abound, I know of no full transport calculation of the abrupt transition between these near-equilibrium and strongly nonequilibrium states. The techniques discussed at this meeting might make a significant contribution.

### 3 The Basic Pulsation Mechanism of Cepheids and RR Lyr Stars

Pulsating variable stars are of extreme importance for both astrophysical reasons (e.g. as distance indicators), and because they are "simple" systems where radiation-hydrodynamics plays a central role. Originally, Eddington proposed that Cepheid pulsation is driven by valving of the radiation flow in the hydrogen ionization zone (HIZ). Further analysis showed the HIZ is at too shallow a depth to
produce the necessary thermal work; in fact, it is dissipative, not driving. But in 1953 Zhevakin showed that adequate driving can be produced in the He$^+ \rightarrow$ He$^{++}$ ionization zone (HeIZ). His ideas were confirmed theoretically by John Cox & Charles Whitney, computationally by Cox and colleagues at Los Alamos [39] (but see also [31]), and by R. Stobie [122] in the UK, for Cepheids; and by Christy for RR Lyr stars [28, 30, 31, 29, 32, 33, 34, 35] at Caltech. The key to the process is that the opacity in both the HIZ and the HeIZ increases proportional to a high power of temperature when H is mostly neutral, He is mostly neutral, or He is mostly singly ionized. Once H is mostly ionized, or He is mostly doubly ionized, the opacity decreases with increasing temperature.

An illuminating analysis of the nature of this mechanism and the observed “phase delay” between the luminosity and velocity curves expected from linear adiabatic pulsation models was made by John Castor [22]. Let me offer instead only a rough intuitive physical “cartoon” of the pulsation cycle. Consider first what one knows from observations. The basic data are: a) spectra, b) photometric magnitudes (and colors), and, if one is lucky, c) spectrophotometry. From spectra, one can determine radial velocities, spectral classes, equivalent widths of “metallic” lines, and line profiles of strong absorption and emission lines, a vast amount of information for interpretation. By integration of the radial-velocity curve, one finds differences in the “radius” of the star as a function of phase (assuming the lines measured are formed at the same depth in the atmosphere, as a function of phase, which has to be calibrated by modeling). From analysis of line strengths one can infer chemical abundances in the atmosphere. From spectral classes one can infer $T_{\text{eff}}$ (the meaning of which may not be the same in a dynamical atmosphere as in a static one). From photometry and spectrophotometry (corrected for interstellar reddening, of course) one can infer variations of the luminous output of the star, and obtain independent estimates of the variation of $T_{\text{eff}}$ as a function of phase.

What emerges from analysis of the data is that maximum light corresponds to the phase: a) of maximum velocity of expansion in the star’s photosphere; b) when the star’s radius is expanding through its average radius; c) of maximum photospheric temperature; d) when, from modeling, it has its largest outward extension of the HIZ. And the phase of minimum light corresponds to when the star has its: a) maximum contraction velocity; b) average radius while contracting; c) minimum photospheric temperature, d) smallest outward extension of the HIZ.

We can understand these phenomena in terms of the behavior of the HIZ and the HeIZ. Start at the phase of maximum radius, when the expansion velocity is zero, and the luminosity is starting to head for its minimum value. This phase is analogous the exhaust stroke of a metaphorical “internal combustion engine”. Gravity starts pulling this extended, cool, moderately-opaque atmosphere inward. Material rains back in on the star, terminating in an accretion shock when it slams into the hydrostatically-stratified envelope. Contraction of the atmosphere
increases to maximum velocity while the star passes through its average radius. The atmospheric material becomes more opaque as its density increases, hence the optical depth at the outer edge of the HIZ increases, the hot radiation emerging from the HIZ is therefore more strongly absorbed, and the star reaches minimum luminosity. The increased optical depth of the layers overlying the HeIZ dams up the luminosity created at the stellar core, hence the local radiation energy density and temperature rises. At first, the increase in temperature increases the local opacity, but eventually the ionization balance shifts from He$^+$ to He$^{++}$, and the opacity then decreases with increasing temperature. At this point conditions are right for a classic weak D-type ionization front (or, in Lab terms, a Marschak wave) in which the hot radiation field from below eats its way into the opaque neutral material above. As a result of the increased temperature, and the larger number of particles resulting from ionization behind the front, the gas pressure there increases, and a shock is driven into the overlying cooler material, ahead of the radiation front. The radiation front can only diffuse outward, relatively slowly. This shock is the main shock in the cycle. This is the time when mechanical work (analogous to the power stroke of an engine) is done on the envelope, accelerating the atmosphere outward. The material moves faster and faster outward, reaching maximum speed just when the outer surface of the HIZ, in Lab language, burns out of the atmosphere. At that point, the dammed-up interior luminosity escapes, and one sees the star at maximum light. At burnout, the nature of the radiation front shifts from a D-type to an R-type [76], which moves outward at high velocity (in principle approaching the speed of light in transparent material). This phenomenon is familiar in Lab experiments; it also has been seen in computations of Cepheid models by Tom Adams & Castor [1], and by Toshihito Ishida & Mine Takeuti [74]. But new generation codes must be used to track this transition precisely. Beyond that, the shock-driven material lofts into the gravitational field until it reaches maximum height, and the cycle repeats.

This broad-brush picture has been verified by calculations by Norman Simon, Shashi Kanbur, and me [118], using a variant (TGRID) of the (DYN) code (see §5.1). What we found is that at maximum light, one sees the "naked H ionization front". And because opacity in the front rises rapidly with temperature, one always sees into optical depth $\frac{2}{3}$ at nearly the same temperature (see figures 3 and 4 in [118]). This result provides a good explanation for the result found decades ago by Art Code [36] (see the figure in his paper) that Cepheids have about the same spectral type at maximum light, independent of period. Further, we showed that spectral type at minimum light, as a function of period, depends mainly on the amplitude of the light curve (see figures 9 and 10 in [118]). Kanbur & Paul Philips have also analyzed the connection between the behavior of the HIZ and the light curve of RR Lyr stars [77, 78]. They show that here, because the H in the atmosphere is already ionized in the atmosphere, one sees the upper edge of the HIZ, hence material at a nearly unique temperature, at light minimum.
Most of what I just said pertains to continuum light, for which the behavior of ionization fronts (both HIZ and HeIZ) is fundamental, and hydrodynamic shocks are “incidental consequences” of this behavior. I think this picture is the best one to gain insight into the basic physics. Many other discussions have focused on the shocks themselves. Shocks are relevant mainly for line formation because spectral lines can be sufficiently optically thick to show significant radiation from the shock layer itself. Also, Doppler shifts can lead to the observed phenomena of grossly-distorted line profiles, perhaps including emission components, in strong lines, and line-doubling of weaker lines.

4 Early Pulsation Models for Cepheids and RR Lyr stars

4.1 Lagrangean models

In the 1960's and 1970's there were many studies of Cepheid and RR Lyr pulsation based on nonlinear Lagrangean pulsation codes. The results were able to: 1) match the observed position of the instability strip in the H-R diagram, 2) identify plausible mechanisms that determine the edges of this strip, 3) find higher modes (overtone pulsations) that account for the “beat Cepheids”, 4) explain the Hertzsprung sequence of light-curve morphology as a function of period (the “bump” Cepheids), 5) and model light- and velocity-curves “reasonably well”.

All of these studies dealt with stellar interior properties; in this era there are no studies of these stars' atmospheres worth serious consideration. Also, the last “success” cited above is meaningless. For example, examine the light curves in Figure 1 on page 247 and the radial-velocity curves in Figure 2 on page 248 of [37], predicted by seven independent models from this era. One sees that the light curves are in poor qualitative agreement at the level of ± 0.2 mag, and the velocity curves are not in quantitative agreement to better than ± 10 km/sec. “Agreement” here is a case of beauty being in the eye of the beholder. These results certainly would not warm the heart of an observer who spent many a cold night at the telescope determining magnitudes to ± 0.01 mag and velocities ± 2 km/sec!

The spurious features in the computed light- and velocity-curves result from the inability of a Lagrangean grid to resolve the steep gradients in temperature, density, and opacity that exist in the HIZ. What happens is that when an outward-burning flux of radiation passes the boundary of a computational zone, it pre-heats, and raises the opacity of, the entire next zone. That zone becomes artificially optically thick, radiates at a temperature lower than that characterizing the incident flux, hence the emergent luminosity decreases. Finally, at the instant when the front burns all the way through the zone one once again sees the correct luminosity. An extreme example of these phenomena can be found in a calculation of the light-curve from a model 1Mt nuclear fireball in Figure 1 of [134]. As discussed
below (§5.1), these problems can be overcome by use of implicit adaptive grids in high-quality radiation-hydrodynamics codes.

4.2 Convection

The work cited above assumed that radiation is the sole energy-transport mechanism in the envelopes of Cepheids and RR Lyr stars. But investigation showed that many radiative models had convectively unstable regions. Several authors, notably Robert Stellingwerf, pursued the effects of convective transport using extensions of conventional stellar mixing-length theory. Convective transport is dominant at the red edge of the instability strip in the H-R diagram, and important elsewhere. The literature is voluminous, and I shall not cite it, because this again is a problem for interiors theorists.

4.3 Opacities

Other important discrepancies between Cepheid models and observed, or other theoretical, data persisted. In particular, Cepheid masses inferred from pulsation theory were significantly smaller than the masses demanded by stellar-evolution theory for the positions of Cepheids in the H-R diagram. Clearly something was wrong! A wide variety of possibilities for resolving this discrepancy were offered. But it was the persistent high-quality sensitivity studies done by G. K. Andreasen, J. O. Petersen, and Simon that finally identified the problem: the Los Alamos "stellar opacities", produced by Art Cox and his coworkers, which were then widely in use, were seriously in error. Twenty-five years of work had essentially been wasted! These findings stimulated heroic efforts to compute accurate new opacities both by the Livermore OPAL project and the US-UK Opacity Project (OP). Extensive results have been published by both groups [70, 71, 72, 73, 104, 105, 117]. The new opacities resolved the mass-discrepancy problem for beat- and bump-Cepheids, and also have important implications for other phases of stellar evolution.

5 "New Generation" Models

Robert Buchler, Stellingwerf, & Géza Kovacs [12, 13, 84, 121] have written good discussions of the limitations of Lagrangean models, of the physics used in the codes of the 70's and 80's, and the need for a "new generation" of pulsation codes. However, their discussion is from the point of view of the stellar interior modeler, and centers on getting the interior pulsation properties right, while ignoring prime issues for atmospheric modeling, such as nonequilibrium phenomena. The two main issues that emerge in their discussion are the need for adaptive-grid techniques, and improved convection theories.
5.1 Adaptive grid techniques

Virtually all of the spurious features in the light- and velocity-curves from Lagrangean Cepheid models can be eliminated by use of adaptive grid methods. The first such code was the DYN code developed by Castor, Cecil Davis, & David Davison [23], and later used by other workers [2, 3, 74, 78, 78, 118]. The method produced a refined grid that could resolve the temperature structure of the HIZ in the stellar envelope. Although the HeIZ drives the pulsation, the steep temperature gradient at the outer edge of the HIZ is close enough to the surface that it dominates the emergent radiation. This code led to much smoother light-curves, which can meaningfully be compared to observations.

The authors of the DYN code exploited their physical insight into the particular problem of stellar pulsation. Newer codes use methods that resolve gradients of all relevant physical quantities (temperature, opacity, ionization state, ...) that can influence a flow. Most of this work did not focus on stellar pulsation. One represents the entire set of nonlinear partial-differential equations governing a radiating flow by difference equations on a (nonuniform) discrete grid of depth-points, whose positions are determined simultaneously with the flow variables themselves, in such a way as to yield optimum resolution of gradients of any relevant physical quantity. The resulting equations are made globally conservative by using control-volume methods on the adaptive grid, making use of the adaptive-grid expansion formula, and adaptive-grid transport theorem [93]. Examples of this procedure are given in [130]. In order to determine the positions of the grid points at the advanced time level, one needs a grid equation that determines the grid consistently with the physical structure that will be found at the advanced time.

The first robust method of this kind was developed Karl-Heins Winkler & Michael Norman in their code WH80s, described in detail in [128, 129]. Applications of this methodology to pulsating variables are described in [38]. A different grid equation, which has a satisfying geometric interpretation, and is simpler to implement, was developed by Ernst Dorfi and his coworkers [40, 41, 45]. Either choice of grid equation has to be solved iteratively along with the nonlinear hydrodynamic equations by a second-order-convergent Newton-Raphson procedure. Dorfi's grid equation has also been used in other codes, e.g. [14, 61]. The great advantage of the implicit-adaptive-grid methods is that they can detect ionization fronts and shocks, and track them on the grid automatically. In Lagrangean codes, these radiation-hydrodynamic features necessarily move through zones which leads to severe loss of accuracy.

5.2 Convection

In the past decade, attempts have been made to derive an astrophysical convection theory which has more physical content than the customary mixing-length theory, see, e.g., [18, 19, 62, 63, 65, 66, 120, 131, 132]. While it has been stated that these theories yield improved agreement of models with observed data, particularly near the red edge of the instability strip, where convection quenches pulsation,
they all contain: a) a host of constants whose values cannot be set a priori, and b) a number of theoretical assumptions and simplifications of unknown validity, and which can never be tested by experiment. After all, convection is a three-dimensional, turbulent flow, and attempts to "reduce it" to a one-dimensional formulation must be problematic at best! The details of the convection theory are important for Cepheids and RR Lyr stars because they have solar-like temperatures and extensive hydrogen and helium ionization zones. They are particularly important for Cepheids, which are supergiants, whose low envelope densities assure that convection, when it occurs, will be highly superadiabatic, hence directly susceptible to errors in the theory. At higher densities where convective efficiency is so high that it is practically adiabatic, the problem is much less severe. The optimistic view is that these theories can be "calibrated" by their application to a wide variety of stars. My own view, given the glacial rate of progress (if at all!) in basic turbulence theory, is that it is unlikely to happen in my lifetime.

5.3 Radiative models

A small grid of RR Lyr models, using Dorfi's implicit adaptive-grid technique was constructed and analyzed by Dorfi & Michael Feuchtinger (DF) [42, 43, 48, 49]. These models are probably the best radiative RR Lyr pulsation models yet constructed; they give a good general match to observed light-curves and photometric colors. Yet another adaptive-grid code has been developed by Buchler, Zoltán Kolláth, & Ariel Marom (BKM) [14], which uses mass as the grid-variable instead of radius; they made this choice so that their code could be coupled to existing Lagrangean codes to exploit efficient methods of obtaining periodic, full-amplitude, limit-cycles. They use the material total energy equation (including kinetic and gravitational energy) instead of the internal energy equation used by DF. BKM report they get better total energy conservation in this way. This is an issue of importance because the kinetic energy in the pulsation is five orders of magnitude smaller than the ambient internal and gravitational energies, hence small errors in those quantities could perturb or even swamp the pulsation mechanism. Nevertheless, even though the total energy continuously increases in DF's calculation, the pulsation properties remain unaltered over 5000 pulsation cycles. In the context of Laboratory calculations, BKM again report that use of the total energy equation gives superior results for the infamous Noh stagnating-shock problem; in contrast, a calculation of the same problem by Michael Gehmeyr, Bao-Lian Cheng, and me [60], using the internal energy equation, appears to give better results than BKM's. Further research is needed to resolve this important issue.

5.4 Convective models

Modern convective models of RR Lyr stars use nonlinear, nonlocal, time-dependent convection theories. Adaptive-grid models by Gehmeyr [57, 58, 59], using the convection theory in [62, 63] and full radiation transport, showed a secondary maximum on the rising part of the light-curve right after minimum radius, and convective quenching of the pulsation at the red edge of the instability.
strip. An extensive grid of RR Lyr models by G. Bono and his collaborators is reported and analyzed in [7, 8]. They used Stellingwerf's [120, 121] convection theory and the diffusion approximation for radiative transport. They present an atlas of light curves and positions of the blue and red edges of the instability strip for $M = 0.75, 0.65,$ and $0.59M_\odot$ for both fundamental and first-overtone pulsators. They derive relationships among the masses, luminosities, radii, and effective temperatures of the models, and discuss the behavior of the theoretical light-curve amplitude, which is larger than observed. Adaptive-grid models using a superior convection theory [131, 132] have been published by Feuchtinger [46, 47]. An important feature of this work is the use of a convective flux-limiter. This procedure removes the spike on the ascending branch of the light curve found in [59] and [7]. Agreement with: a) the periods predicted by linear theory is to within 2%, and b) the coefficients in Fourier decompositions of observed light curves is obtained up through fifth order. Both theory and observation show decreases in Fourier amplitude ratios and asymmetry factors, and increases in phase differences. Amplitudes of light curves are controlled by two parameters: the turbulent mixing-length, and the turbulent viscosity. In the absence of an ab initio determination of these parameters, values were used that gave reasonable luminosity amplitudes. Thus, despite the improvements made in pulsation models, work remains to be done; the critical discussion by Kovács & Kanbur [85] shows that many problems remain in fitting light- and velocity-curves in detail.

6 Atmospheric Modeling of Cepheids and RR Lyr Stars

6.1 Observations

For model-atmosphere work, the observational data are the gold standard for correctness. If the data are fit well, one can have confidence in the model; otherwise one must simply try again!

6.1.1 The continuum: colors, spectrophotometry, and line-blanketing

The literature on photometry of Cepheids and RR Lyr variables is huge (many hundreds of papers). Such work is fundamental to inference of: a) the luminosities of these stars; b) when combined with radial-velocity data, their radii (by several variants of the Baade-Wesselink technique); c) calibration of the P-L and P-L-C relations; and d) photometric estimates of metal abundances. Many papers on computing theoretical colors from models to compare with observations exist; see, e.g. [6, 20, 44]. In contrast, there is relatively little work on spectrophotometry of Cepheid and RR Lyr variables. I found only 10 papers: [75, 96, 95, 97, 98, 99, 102, 103, 106, 107]. There is a great need for more data of this kind because it provides a
more detailed comparison standard for dynamical models. A few papers give values of line-blocking coefficients measured directly from high-dispersion spectrograms.

6.1.2 The line spectrum

There are many papers on spectroscopy of Cepheids and RR Lyr stars, though far fewer than on photometry. Some of the most important ones are based on photographic spectra, with all of their calibration, nonlinearity, and signal-to-noise problems. Renewed efforts to obtain high photometric-quality results with high-dispersion CCD cameras will be welcome! A small sample of results emphasizing a) emission-line phenomena, which can arise when the atmosphere is shocked or is experiencing burnout of a radiation front, and b) high-quality measurements obtained with modern instrumentation, can be found in [5, 9, 10, 11, 15, 16, 25, 64, 86, 87, 101, 108, 109, 110, 111, 112, 113, 114, 115, 116, 126, 127]. Further information on line-doubling of weaker lines, and detailed profiles of strong lines, as a function of phase in the pulsation, would be of value.

6.2 Early "atmospheric models"

6.2.1 Hydrostatic models

Continuum fluxes from hydrostatic atmospheres for values of $T_{\text{eff}}$, $\log g$, and metal abundances appropriate for RR Lyr stars were published by Strom [123]. They allow for convection by mixing-length theory, and departures from LTE in the continuum, assuming the H-lines are in detailed balance. Comparison with spectrophotometry suggests that these models can be used when the hydrodynamical accelerations in the atmosphere are smaller than the acceleration of gravity; but dynamical models are to be preferred. LTE line profiles for lines of hydrogen and several other ions, using quasi-static radiation snapshots, and taking into account the velocity field from dynamical models of $\delta$ Cephei, can be found in [67].

6.2.2 Lagrangean models

Crude "dynamical model atmospheres" for pulsating stars have been constructed with Lagrangean codes having more zones near the surface. The first such models were made at Los Alamos by Charles Keller & Paul Mutschlecner [82, 83]. They accounted for nongrey opacities in 10 frequency groups, used the diffusion approximation to compute the radiation flux in optically thick zones, and the transport equation in higher, optically-thin zones. Their results are in qualitative agreement with observed colors when line-blanketing corrections were taken into account. However, they still show the same "bumpiness" of the light curve, described above, from under-resolution of the temperature profile of the HIZ. Similar calculations were made by Alan Karp [79, 80]. Dynamical model atmospheres for RR Lyra, allowing for radiative transport (in the Eddington approximation), were made by Stephen Hill [68], who also made a detailed analysis of shocks in
the atmosphere; later Hill did similar calculations for the Cepheid $\beta$ Doradus [69]. Velocity-induced asymmetries in strong lines, computed from Lagrangean models, are discussed in [17, 81] for Cepheids; metallic line-doubling in RR Lyr is discussed in [55]; and the Van Hoof effect for metallic lines is discussed in [26, 92]. But all of this work is based on under-resolved models, and the assumption of LTE, which probably fails in the continuum, and certainly fails in the lines, particularly when in emission.

6.3 Reality check

Virtually all of the work described above was done from the point of view of the envelope/interior, with emphasis on the pulsation mechanism itself, and the determination a few basic properties. But if we are truly to model the atmospheres of these stars, then we must view matters quite differently from our "stellar pulsation" colleagues. I list here a few such heresies:

- **Thermodynamics is useless in the atmosphere**
  That is, one cannot use a tabular equation of state, caloric equation of state, or opacity, which depends on a couple of thermodynamic variables like temperature and density. Rather, one must specify the nonequilibrium populations of each atom/ion that contribute to the macroscopic fluid quantities; these depend on the (non-Planckian) radiation field, in principle over the entire spectrum, while, at the same time, determining that radiation field. This is a gigantic paradigm shift for "pulsation people", but we have been doing it since 1963 when the first NLTE atmospheres were constructed. In fact, calculations using LTE sometimes give ridiculous dynamical results. For example, both Hill's and A. Fokin's models of RR Lyr [54, 68] predict that the outer atmosphere is distended over $\sim 20$ scale-heights by the main shock. The radiative contributions to energy balance are very small at the resulting low densities, so the expansion is nearly adiabatic and leads to a dramatic cooling of the material, down to absurd values like 200K!

- **Stiff rate equations**
  The excitation/ionization state of the material is determined from kinetic rate equations, not statistical mechanical relations. In pulsating stars, neither the Eulerian nor Lagrangean (or, for that matter, adaptive) time-derivative of level populations is zero. But these variations (on a time-scale of order of a fraction of a pulsation period) are small compared to the rates of atomic depopulation/repopulation mechanisms. Thus the equations are numerically stiff, and one should probably formulate the kinetic equations in terms of net rates. Perhaps it will be useful to combine numerical and analytical integration techniques of these equations in order to get both accuracy and stability, as is done for the atmospheric sciences or chemically reacting flows.
• **We must calculate the interaction between nonequilibrium phenomena and the dynamics consistently.**
  In addition to accounting for nonequilibrium rate equations to determine the local state of the material, one must account for their nonlocal dynamical consequences on shock and radiation-front structures and propagation through radiation transport. One will need comprehensive and refined adaptive-grid techniques, equally effective for dynamics, kinetics, and transport. Both the propagation of a D-front, and its burnout and conversion to an R-front are fundamentally nonequilibrium phenomena.

• **Are there actually periodic limit cycles in the atmosphere?**
  One of the striking facts about Cepheid and RR Lyr pulsations is their remarkable regularity. By phasing light maxima from widely separated epochs, one can determine pulsation periods to many digits, and even track small secular changes in period, which contain clues about the star's interior evolution. Numerical techniques have been developed to obtain strictly-periodic models at limit cycles. But these methods may not be appropriate to an atmosphere. I have never believed there is any reason to suppose that the upper atmosphere, specifically where strong lines are formed, must necessarily vary in strict periodicity with the photosphere. Rather, it well can be quasi-periodic, with different physical structures during alternate (photospheric) cycles, or a complex variation within a single cycle. Recent observations of RR Lyr show “irregularities” in the strength and radial velocity of metallic lines from one pulsation cycle to the next [24] do exist.

  Our interiors/pulsation colleagues, notably Aikawa and Buchler, have written penetrating analyses using concepts of period-doubling, intermittency, resonances, and dynamical chaos to explain the behavior of Population II Cepheids such as the W Vir, RV Tau, and BL Her stars. But from the atmospheric point of view, keep in mind the work of Stein and Schwartz [119] who showed how perfectly-regular shock trains in a solar-like stratified atmosphere propagate as shocks when their period is less that the acoustic-cutoff period, but give rise to semi-regular beats when their period exceeds the cutoff period. If one accepts the position that we know almost nothing theoretically about the outer atmospheres of these stars at present, one should not be surprised by any “unexpected” dynamical behaviors.

• **Are “snapshots” of the radiation field adequate to compare with observations?**
  The crux of this question is whether or not it is sufficient to use the (nonequilibrium) state of the material, and its computed motion, to calculate the angle-frequency dependent radiation field using time-independent transport equations. I have argued above that it is, on the basis of time-scales. If this conclusion is true, it affords a great simplification because we can split the calculation into space, angle, and frequency-dependent parts, by introduc-
ing subsidiary "form factors" such as the well-known Eddington factor, and energy- and flux-spectra. This part of a dynamics nonequilibrium-spectral calculation will be, by far, the most computationally expensive. If the splitting is valid, then these calculations can be carried out in parallel on a massively-parallel computer.

7 The theoretical challenge

7.1 The continuum

Given the existence of voluminous accurate broad-band photometry and spectrophotometry of Cepheids and RR Lyr variables, it would be almost criminal not to compute accurate continua from new high-quality, full-transport, radiation-hydridynamic models of their atmospheres, including line-blanketing, for a comprehensive grid covering the relevant parts of the H-R Diagram. In such work one would need to allow for departures from LTE in H, and perhaps the first five levels of H (enough to get beyond the "opacity hole" of H near 1.6 microns). The H Lyman lines could be taken to be in detailed balance for such continuum calculations, but the subordinate H-lines must be treated in detail. Using the compendious atomic/molecular now available, opacity distribution functions for the whole spectrum can be constructed. Methods to account for line-shifts within the ODF formalism in moving media have been suggested [94].

For the RR Lyr and Population II Cepheids, we know the "metal" abundance is smaller by an order of magnitude or more, so my guess is that one could likely use LTE to account for line blanketing, at least as a first attempt. For the Cepheids, the metal abundances is about the same as it is for the Sun, and whether or not one needs to (or can!) allow for NLTE line-blanketing to get accurate continua is a question I leave for the experts here.

7.2 The line spectrum

Theoretical computation of the profiles of strong or chromospheric lines such as the H-lines (Lyman, Balmer, and Paschen), Ca II H & K, Mg II h & k, and He I λ10830, in pulsating stars is practically virgin territory. That is no surprise: for a the problem is forbidding in difficulty. For a truly consistent analysis, one would need to perform, at each time-step, a full NLTE calculation using model atoms that are complete enough to give all the important couplings, while allowing for the effects of Doppler shifts, and then account for the back-reaction of the resulting nonequilibrium state of the material ("EOS", internal energy, opacity) on the dynamics of the pulsation.

A partial step towards such an ultimate model has been made by Fokin and D. Gillet [53, 54, 56] who calculated hydrogen-line formation in dynamical models of W Vir, RR Lyr, and BL Her. The models were constructed with a Lagrangean
code using an explicit hydrodynamic step, followed by iterative implicit solution of the radiating-fluid energy-equation and radiation moment equations [52]. Fokin discusses how the comoving-frame line-transport algorithm is extended to the case of nonmonotonic velocity fields. Models at various phases in the pulsation cycle are then used to calculate H-line profiles for a three-level H-atom for W Vir, and a four-level H-atom for RR Lyr and BL Her, using an equivalent-two-level-atom method. While that approach is not as strongly convergent as more modern methods, it appears to work successfully because of desaturation of the lines by velocity shifts. A back-reaction of nonequilibrium populations in the H-atom on the pulsation dynamics is not taken into account.

For W Vir, the predicted light- and velocity-curves are in qualitative agreement with observations, including the well-known shock-induced break in the radial-velocity curve. Lyα and Lyβ have absorption profiles with a blue-shifted emission wing. Hα has a more complex behavior, ranging from Doppler-shifted pure absorption lines, to lines with a strong blue-shifted emission wing, distorted by an unshifted dark scattering core.

For RR Lyr, the theoretical light-curve agrees well with observation, and photospheric layers in the model show a simple periodic pulsation. In contrast, the upper atmospheric layers undergo alternating cycles having either extreme distention (~20 scale-heights), or a more compact structure with multiple shocks. No mass-loss is found. The main shock in RR Lyr occurs very high in the atmosphere (column mass density ~ 10^{-3} g/cm^2), as opposed to W Vir where it is much deeper (column mass density ~ 10^2 g/cm^2). These differences explain why the weak metallic lines show line-doubling in W Vir, but not in RR Lyr. The hydrogen Lyman lines are in strong emission, reaching maximum intensity at phase 0.0 (light maximum). Hα is generally in absorption, as observed, with a small emission feature near line center at certain phases, which is consistent with its level of formation lying below the main shock. The computed central intensity, Doppler shifts, and amplitude and phases of splitting of Hα are in qualitative agreement with high-quality observational data. The results for BL Her are intermediate between W Vir and RR Lyr.

Dimitar Sasselov & John Lester [110, 111] have constructed more sophisticated models to derive information about the chromospheric structure of ζ Gem and η Aql from an analysis of their observed He I λ 10830 profiles. They use a 1-D explicit, conservative, Godunov, Lagrangean hydrodynamics code, which is second-order in space and time, to solve local Riemann problems on the spatial grid at each timestep, in conjunction with the NLTE transport code MULTI [21]. The transport code is used to solve for the radiation field in model atoms of H, He I, He II, Ca II, and Mg II. They proceed in three steps: a) construct an initial semi-empirical hydrostatic atmosphere, b) use the static atmosphere as a starting condition for a hydrodynamic calculation which at first uses a rough line-cooling approximation until the pulsation reaches steady state, then c) use the full tran-
port code for all species until a stable pulsation is reached. For the ζ Gem model, the pulsations of the upper atmosphere are qualitatively quite different from those of the photosphere; the discordance is less for η Aql. The computed radial-velocity curve for ζ Gem is in good agreement with observations, as are the computed profiles for Mg II h, and He I λ 10830. They conclude that: a) in addition to heating by shocks, Cepheid chromospheres are heated by other sources (perhaps acoustic and/or MHD wave dissipation); b) Cepheids do not seem to have hot coronae; 3) the He I λ 10830 line gives evidence of a steady outflow of material from the upper layers of Cepheid chromospheres.

These pioneering investigations have already paid great rewards; much more intensive work in the future will ultimately allow us to construct a reliable theory of Cepheid atmospheres.

8 Can it be done?

The atmospheric phenomena in pulsating stars are, per se, mainly the result of nonequilibrium radiative and hydrodynamic effects. Yet the properties of an atmospheric model depend on its lower boundary condition, predicted by a full-envelope model, so convection does (alas!) influence the atmosphere. The atmospheric phenomena require a non-LTE theory including line-formation, coupled to dynamical models that resolve shocks and radiation fronts. In my opinion, a major challenge for future work is to find a way to achieve a seamless connection between both the atmosphere and the deeper envelope, using improved (e.g. time-dependent) convection theories, and full NLTE line-formation theory. As I said above, the rest is “mere work”.

The computational load will, of course, be heavy because one must construct perhaps scores of line-blanketed models for a single star over its pulsation period. Nevertheless, modern workstations with GBytes of RAM, scores of GBytes of disk storage, and clock speeds measured in GHz, are now available, and their cost is within reach for astrophysics departments or institutes, or even individual research groups. Under Unix or Linux operating systems one has automatic virtual-memory and multiprocessing capability. It is realistic to suppose that several cases could be run simultaneously, in background, on a given machine, 24 hours per day. Furthermore, massively-parallel machines are now here; one only needs to figure out how to get access to them.

It must be understood that a fully-implicit time-dependent dynamical computation is nothing more than a connected series of calculations, each of exactly the same type as a static atmosphere calculation, and we already know from work on static atmospheres that both our physical paradigm and computational techniques are sufficient. When we can truly model Cepheid and RR Lyr atmospheres accurately, it will be a crowning achievement of tremendous value to astrophysics. The only question is whether anyone will have the desire and will to do it. I wish
that I could turn my personal clock back forty years, and help make it happen; unfortunately it seems that would violate the physical principle of causality, the second law of thermodynamics, and laws of biophysics and biochemistry I cannot even name. I can only encourage the young, brave, and energetic, to take the flag, and carry it forward to victory.

This work was performed at Los Alamos National Laboratory, supported by the Department of Energy under Contract W-7405-ENG-36.

References


24


