CONSTRAINTS ON SOLAR MODELS FROM STELLAR EVOLUTION[1]

I. IBEN, JR.
The Pennsylvania State University
525 Davey Laboratory
University Park, PA 16802

ABSTRACT. This paper constitutes a brief and exceedingly elementary review of the consraints imposed on models of the solar interior by comparisons between theoretical models of evolving stars and the observed characteristics of stars, including the Sun.

1. Introduction

One of the dominant lessons to be learned from the science of stellar evolution is that, with a moderately simple set of approximations to the physics of matter and radiation under conditions expected in stellar interiors, one can account extremely well for the distribution in the Hertzsprung-Russell diagram of single stars in the field and in clusters which vary in heavy element composition by several orders of magnitude and range in age from 10^6yr to 10^{10}yr. The approximations include: spherical symmetry, no rotation, no acceleration in the pressure-gravity balance equation, initial homogeneity in composition, uniform mixing in convectively unstable regions, radiative and conductive flow in the diffusive approximation, no mixing in radiative regions, radiative opacity in the hydrogenic approximation, electron conductivity controlled by Coulomb scattering from randomly distributed nuclei, energy generation and particle transformation rates which are simple extrapolations of laboratory cross sections, the simplest of equations of state (radiation in a black body distribution, perfect gas for ions, perfect gas for free electrons with degeneracy allowed for, Saha equation with single term partition functions in regions of partial ionization, Coulomb forces taken crudely into account at high density if so desired), neutrino-loss rates from a well established theory of weak interactions, and an extremely simple Eddington-approximation surface boundary condition. The manner in which stars of

[1]Supported in part by the NSF grant AST 88-07773

G. Berthomieu and M. Cribier (eds.), Inside the Sun, 91–100.

https://doi.org/10.1017/S0252921100067841 Published online by Cambridge University Press

near solar mass and initial composition evolve from formation to the final cooling phase following the exhaustion of all exploitable fuels is a very solidly founded consequence of comparisons between the observations and models of evolving stars which have been constructed by using the enumerated approximations. Such stars (1) contract onto the core hydrogen-burning main-sequence band, (2) live within the main-sequence band until having exhausted approximately 10 percent of their hydrogen fuel, (3) leave the main-sequence band to become giants with electron degenerate cores, (4) ignite and burn helium in the core as clump stars, (5) become asymptotic giant branch stars after exhausting helium in the core, (6) eject a shell of material which is caused to fluoresce by emanation of energetic photons from the compact but luminous remnant, and (7) finally evolve into white dwarfs.

In short, there is every reason to believe that models of the Sun which are constructed by using the same input physics that successfully accounts for many of the observable features of single stars in clusters and in the field are reasonably accurate representations of the interior structure of the Sun.

2. The Sun as a Main-Sequence Star

From the fact that one sees emission and absorption lines in stellar spectra, one ascertains that matter at stellar surfaces is gaseous and that the distribution of radiated energy with respect to wavelength is very similar to that of a theoretical blackbody. From the location of the peak in the observed distribution, one may estimate a characteristic surface temperature. For the Sun, this characteristic temperature is of the order of 6000K. In the case of the Sun, one may also make use of the known angular diameter of the solar disk and of the mean temperature at the earth's surface to obtain another estimate of the temperature at the photosphere; this estimate is in agreement with the first one. Since energy flow requires a temperature gradient, one may infer that the temperature increases inward from the surface. Assuming that matter in the interior is gaseous, that the equation of state is that of a perfect gas, that the pressure gradient everywhere balances the gravitational acceleration, that the matter in the interior is of nearly homogeneous composition, and that the dominant element species are essentially completely ionized, one deduces that a star's central temperature is related to its mass M, radius R, mean molecular weight μ of constituent particles, and to fundamental constants (k = Boltzmann's constant, G = Newton's gravitational constant G, and M_H = mass of the hydrogen atom) by

$$kT \sim (GM/R)\, \mu M_H \sim 15 \times 10^6 K \times (M/R), \qquad (1)$$

where, on the far right, mass and radius are in solar units and μ has been chosen to accommodate the fact that by far the dominant species of element existing at stellar surfaces are hydrogen and helium. More

https://doi.org/10.1017/S0252921100067841 Published online by Cambridge University Press

sophisticated versions of this relationship may be constructed by taking into account the effect on the equation of state of radiation pressure, electron degeneracy, and Coulomb forces between charged particles, but, for stars of near solar mass, equation (1) turns out to be an excellent approximation.

From stars near enough to permit an estimate of distance by trigonometric parallax, one may translate the observed energy flux from the star into an estimate of the intrinsic bolometric luminosity L and use the expression

$$L = \sigma T_e^4 \times 4\pi R^2 \qquad (2)$$

to estimate the stellar radius R.

The vast majority of nearby stars (when expressed in terms of space density = stars pc^{-3}) lie in a very tightly constrained main-sequence band in the Hertzsprung-Russell diagram (log L versus log T_e). The Sun lies comfortably within this band. From nearby stars in binary systems, one finds that there is a tight correlation between luminosity and mass for main-sequence stars. In particular, for stars of the Sun's mass or larger,

$$L \sim M^4, \qquad (3)$$

where both L and M are in solar units. Along the main-sequence band mass and radius are related very roughly by

$$R \sim M^{0.7}. \qquad (4)$$

Having an estimate of the mean interior density, $\rho \sim M/(4\pi R^3/3)$, and a first estimate of interior temperatures, one may verify that the assumption of complete ionization for the dominant element species, hydrogen and helium, over most of the interior is thoroughly justified for main-sequence stars, including the Sun.

Equations (1), (3), and (4) imply that the local rate of energy generation in the interiors of main-sequence stars more massive than the Sun increases very steeply with increasing temperature. In fact, these equations show that, within main-sequence stars of mass between 1 and 10 times the mass of the Sun, the mean rate of energy generation varies as about the 14^{th} power of the central temperature. This dependence is in fair agreement with that expected for the CN cycle reactions, cross sections for which are measured in the laboratory; extrapolation to relevant stellar energies is quite straightforward.

For the Sun and less massive stars, the CN cycle reactions do not produce energy at a rate sufficient to account for the observations. However, there is strong theoretical and experimental basis for believing that the pp-reaction chains take up the slack. Unfortunately, the basic initiating reaction for these chains, the pp reaction itself, is not directly testable in the laboratory. Instead, one must rely on a calculation based on the theory of weak interactions normalized to fit with experimental results from β-decay experiments. Support for this reliance comes after constructing detailed main-sequence models which show that luminosity and mass are related in exactly the same way as in real main-sequence stars of the Sun's mass or less, viz.,

$$L \sim M^{2.2}. \qquad (5)$$

Further insight into the structure of main-sequence stars of mass as large or larger than the Sun is obtained by supposing that energy is carried outward to the surface from the region of energy production predominantly by radiative diffusion. The effective cross section between photons and ionized atoms at any point is summarized in the opacity κ, which is a function of the temperature T, density ρ, and composition. The main sources of opacity are scattering by free electrons ($\kappa_e \sim 0.2$-$0.35\ cm^2\ g^{-1}$), absorption by free electrons in the Coulomb potential of ionized nuclei, photoionization, and absorption between bound atomic levels (κ[free, bound-free]) $\propto \rho/T^{3.5}$). The photoionization and bound-bound contributions are strong functions of the abundances of heavy elements

The flux of energy carried by radiation is related to local conditions by

$$L(r)/4\pi r^2 = -(1/3)\ c\ \partial aT^4/\partial r\ 1/\kappa\rho, \qquad (6)$$

where c is the velocity of light and $a = 4\sigma/c$ is proportional to the Steffan-Boltzmann constant σ. Here, L(r) is the energy generated within a sphere of radius r and ρ and T are the local density and temperature, respectively. Combining equations (1) and (6) gives

$$L \sim A\ \mu^4\ M^3\ /\ \kappa, \qquad (7)$$

where κ is a "mean" interior opacity and A is a combination of fundamental constants G, σ, c, and k that is of the order of unity when total stellar luminosity L and stellar mass M are in solar units. Comparison between equations (7) and (3) shows that the mean opacity is of the order of unity and that all three opacity sources contribute comparably to the mean opacity. The value of mean κ deduced in this way bolsters the estimate from an analysis of spectral lines that hydrogen and helium are the dominant element species in main-sequence star interiors (Eddington thought that iron was one of the main constituents of stars, as in the earth).

In the outer layers of low mass model stars, including solar models, the opacity in regions where hydrogen and helium are ionized is so large, that energy is carried through these regions predominantly by convection. The lower the mass of the star, the larger is the fraction of the model star's mass contained in a convective envelope. Models less massive than about one-third of the Sun's mass are completely convective. It is this change in mode of energy flow with decreasing mass, in conjunction with the temperature dependence of the pp chains, that is responsible for the difference in the mass-luminosity relationship for stars more massive than the Sun (equation [3]) relative to that for stars less massive than the Sun (equation [5]).

3. The Sun as an Evolved Star

Incorporating the full panoply of physics just sketched, one may construct models whose observable properties change in consequence of nuclear transformations in the deep interior. These transformations bring about a reduction in the number of free particles per gram, and, since pressure is proportional to the number of free particles, a balance between pressure forces outward and gravitational forces inward can be maintained only by contraction of the region in which the transformations are occurring most rapidly. With contraction comes heating, and with heating comes an increase in the rate of nuclear reactions. The increased flux of radiation through outer layers forces these layers to expand. Thus, as hydrogen is converted into helium in the deep interior, the radius and luminosity of the model star increase. The model star evolves through the observational main-sequence band defined by the bulk of nearby stars. Once hydrogen is exhausted over about 10 percent of the mass of the model, the model evolves ever more rapidly into the region in the Hertzsprung-Russell diagram occupied by observed red giants (such as, e.g., Capella).

There are two clusters in the disk of our Galaxy which contain stars whose positions in the Hertzsprung-Russell diagram beautifully confirm the theoretical picture painted by the evolutionary models of stars of mass near the Sun's mass and of surface composition near that of the Sun (it must be emphasized that the confirmation preceded the models). The two clusters are M67 and NGC188. Each cluster is far enough away from us that the distances between stars in each cluster are quite small compared with the distance of the cluster; therefore, the relative bolometric luminosity of one star relative to another in each cluster is known quite accurately. In each cluster, all of the stars less luminous than the Sun lie in the Hertzsprung-Russell diagram within the main-sequence band defined by nearby stars. In each cluster there is a well populated giant branch extending to luminosities several orders of magnitude brighter than the Sun at surface temperatures much less than the Sun's, and there is a sprinkling of stars between the main sequence and the giant branch. The luminosity at which the main sequence "turns off" towards the giant branch is larger in M67 than in NGC188 by a factor of about 2.5. The comparison with models proceeds as follows.

From the luminosity of the "turnoff point" one may estimate (using, e.g., equation [3]) that the mass of a star leaving the main sequence in M67 is about $1.25M_{\odot}$ and that the mass of a star leaving the main sequence in NGC188 is about $1M_{\odot}$. Knowing from model calculations that departure from the main sequence occurs when hydrogen has been

https://doi.org/10.1017/S0252921100067841 Published online by Cambridge University Press

exhausted over 10 percent of the mass of the model, noting that the time scale for burning a mass δM of hydrogen at luminosity L is proportional to $\delta M/L$, and making use of equation (3), we infer that stars near turnoff in NGC188 are older than stars near turnoff in M67 by a factor of about two. Knowing that approximately 24 MeV are liberated when four protons are converted into an alpha particle, we can then infer that stars near turnoff in M67 are approximately 5×10^9 yr old and that stars near turnoff in NGC188 are near 10^{10} yr old. One may construct arguments to show that the timescale for star formation in a cluster is far shorter than these estimated cluster ages. Then, by evolving models of different masses for 5×10^9 yr and plotting their positions in the Hertzsprung-Russell diagram one may construct a synthetic model cluster locus which may be compared with the locus defined by stars in M67. The agreement is impressive, including a distinct gap in the density of stars near cluster turnoff corresponding to a phase of rapid contraction in the models when hydrogen is exhausted over a finite region about the center which was convectively mixed during the main-sequence phase. The agreement of model cluster loci for an assumed age of 10^{10} yr with the locus defined by stars in NGC188 is also impressive, including the absence of a gap in stellar density near cluster turnoff and a near constant luminosity of stars between cluster turnoff and the base of the giant branch.

The lesson for the Sun is that, at an age of 4.6×10^9 yr (as determined by radioactive dating using ^{238}U and ^{206}Pb), it is only halfway through the main-sequence phase and it does not possess a convective core.

4. Solar Models, Solar Neutrinos, and Solar Oscillations

Having argued that there is every reason to expect that solar models constructed with standard input physics are not grossly in error, it remains to address the persistent discrepancy between predictions of "standard" solar models and results of solar neutrino experiments. The relation between a predicted flux, central temperature, assumed composition parameters, and choices of relevant nuclear cross section factors may be approximated by

$$F_\nu \sim 0.91 \text{ SNU } y\, T_c^{6.7}\, S_{34}\, (S_{11}S_{33})^{-0.5}$$

$$\times\, (1 + 8.1\, y^{0.6}\, T_c^{13.6}\, S_{17}\, S_{e7}^{-1}), \qquad (8)$$

where F_ν is a predicted neutrino counting rate in the Davis Homestake mine experiment in units of 10^{-37} counts per ^{37}A atoms (= 1 SNU), T_c is the central temperature of the Sun in units of 15.23×10^6K , and S_{11}, S_{33}, S_{34}, S_{17}, and S_{e7} are center of mass cross section factors for the $p(p,e^+)d$, $^3He(^3He,2p)^4He$, $^3He(^4He,\gamma)^7Li$, $^7Be(p,\gamma)^8B$, and $^7Be(e^-,\nu)^7Li$ reactions in units of the most likely values for these cross sections. The parameter y is Y/0.25, where Y is the initial helium abundance by mass. The first term

https://doi.org/10.1017/S0252921100067841 Published online by Cambridge University Press

in expression (8) is due to neutrinos emitted in the reaction $^7Be(e^-,\nu)^7Li$ and the second term is due to neutrinos emitted in the $^8B(e^+,\nu)^8Be^*$ reaction. All other contributing reactions have been ignored. It is further the case that $T_c \propto \mu\, S_{11}^{-1/7} \propto y^{0.245}\, S_{11}^{-1/7}$ (see equation [1]), so that

$$F_\nu \sim 0.91 \text{ SNU } y^{2.65}\, S_{34}\, S_{33}^{-0.5}\, S_{11}^{-1.46}$$

$$\times\ (1 + 8.1\, y^{3.97}\, S_{17}\, S_{e7}^{-1}\, S_{11}^{-1.95}). \qquad (9)$$

Note that the abundances of elements heavier than helium do not appear explicitly in either equation (8) or (9). The reason for this lies in equation (7), from which it is clear that, for a fixed luminosity, a decrease in the estimated opacity (which is in turn related to the assumed abundances of elements heavier than helium) requires a concomitant decrease in the molecular weight (i.e., a decrease in the abundance of helium). The particular normalization in equations (8) and (9) follows from a specific assumed relationship between opacity and the abundances of elements heavier than helium.

The appropriate value of Y is not directly known from the observations. In the solar wind, the ratio of ^4He to ^1H is on the average about 0.05, or only half of what is found at the surfaces of main-sequence stars where accurate analyses can be made. Presumably, the acceleration mechanism which drives the solar wind preferentially accelerates protons, whose charge to mass ratio is twice that of alpha particles. Thus, the value of Y relevant to the interior is actually found by constructing theoretical solar models which are constrained to be of age 4.6×10^9 yr and to have an interior heavy element composition equal to that estimated from photospheric spectral line strengths. With best estimates of center of mass nuclear cross sections and of abundances of heavy elements, the abundance by mass of ^4He in solar models is Y ~ 0.25-0.26, only ~ 0.03 larger than the Y predicted by Big Bang models. With this value of Y, predicted solar neutrino fluxes are typically 6-8 SNU.

From equation (1) it is evident that a decrease in the initial abundance of ^4He means a decrease in the calculated central temperature and, from equation (8) this means a decrease in the predicted neutrino flux, all other things being equal. However, even reducing Y to the Big Bang value, predicted values of F_ν are significantly larger than the observed long-term mean of $F_\nu \sim 2.3$ SNU. In some quarters, this has been taken as evidence that the mass of at least one type of neutrino (the electron neutrino, the muon neutrino, or the postulated tau neutrino) is not zero and that oscillations between the (two or three types of) neutrinos are induced by the interaction between the electron neutrino and the electrons in the Sun as the composite neutrino makes its way outward through the Sun (the Mikheyev-Smirnov-Wolfenstein process). This could account for the discrepancy between the predicted counting rate and the observed average counting rate.

However, the counting rate from the Davis experiment is not constant, and has varied between essentially zero and about 4 SNU. Davis has suggested that the variation over the first two decades of the experiment is not necessarily stochastic and may be anticorrelated with the solar sunspot number, implying that the Sun's magnetic field may modify the characteristics of the electron neutrino. A neutrino magnetic moment is inconsistent with Majorana theory, except in the form of a "transition" magnetic moment; but theories are only theories until confirmed by experiment. It is significant that the first estimate of a solar neutrino flux from the Kamiokande experiment is consistent with the current counting rate from the Davis experiment, viz., ~ 4 SNU. If the two experiments continue to give consistent results and the variation with the solar cycle persists, there is a strong possibility that it is the physics of the neutrino and not a shortcoming of solar and stellar models that is at the bottom of the mystery.

Another component of standard stellar physics which explains the properties of a large class of stars and which is also relevant to the Sun is the theory of stellar pulsations. By allowing for dynamic motions in the linear approximation, it may be shown that, throughout most of the HR diagram, models are stable against acoustical pulsation in radial as well as non-radial modes. There are also two well defined boundaries to the red of which models are unstable to pulsation in either the fundamental radial mode or in the first overtone radial mode; the blue edges of these boundaries coincide precisely with the blue edge of a narrow strip in the HR diagram defined by large amplitude classical variables -- Cepheids, W Virginis stars, BL Her and RR Lyrae stars, delta Scuti stars, Am variables, and the white dwarf ZZ Ceti variables. The Sun does not lie within this strip and it is interesting to note that, as of this date, there do not appear to be any substantiated indications that the Sun is pulsating in radial modes (the fundamental radial mode should have a period between 55 and 65 min, and the first overtone should have a period of about 40 min). On the other hand, the Sun does oscillate in non radial modes at frequencies near 3 mHz; continuous excitation by coupling between these modes and turbulent convective motions (which show persistence time-scales of the order of 5-20 min) appears to be necessary to maintain the oscillations. An analysis of frequency splittings allows one to learn something about the variation of angular velocity with depth and current interpretations of the observations suggest that the Sun does not rotate significantly more rapidly in the interior than at the surface. Arguments about the existence of exotic particles such as WIMPs and about the accuracy of opacity estimates based on comparison between the observed mode structure and the mode structure given by detailed solar models remain unconvincing.

5. What the Sun Teaches us about Other Stars

Thus far, an attempt has been made to show that a fairly simple theory of single star evolution gives a good account of the observed characteristics of real stars and that this implies that theoretical models of the Sun's structure are probably a very good first approximation.

The impression one might form is that rotation, including rotationally induced mixing of various sorts, magnetic fields, surface winds, and the surface activity which occupies the attention of most of the solar physics community are minor perturbations from the standpoint of the overall, global evolution of most stars, including the Sun.

However, when one looks more closely at stars other than the Sun, there are a number of situations which can most easily be understood only by invoking physical processes which have been left out in the simple picture, but which the solar experience tells us must be playing a role in the evolution of real stars, a goodly number of which are rotating far more rapidly than the Sun. In the Sun, rotation, magnetic fields, convection, acoustic and Alfen-wave energy transfer all play a role in establishing a corona and supporting a wind which, although it currently carries away mass at an insignificant rate, carries away angular momentum at a rate which, when used as a normalization in a somewhat heuristic theory of magnetic braking, predicts very high rates of angular momentum loss for rapidly rotating stars which also sport convective envelopes. Direct evidence for this type of magnetic braking comes from a comparison of typical $< v \sin \underline{i} >$ values for stars of near solar mass in the Pleiades (of age about 10^8 yr) with $< v \sin \underline{i} >$ values for solar mass stars in the Hyades (of age about 10^9 yr) and with the Sun's equatorial rotation velocity. To a reasonable approximation $< v \sin \underline{i} > \propto (\text{age})^{-1/2}$, as first pointed out by Skumanitch. Several quantitative indicators of magnetic activity follow this same relationship and one can therefore construct a heuristic algorithm for estimating the rate of angular momentum loss by a "magnetic stellar wind" (MSW) from solar type stars as a function of rotation rate. This algorithm has proved very useful in understanding the evolution of close binary systems in which at least one of the components is cool enough to have a deep convective envelope and shows evidence for surface magnetic activity. The MSW carries away angular momentum from the cool star, but tidal torques keep it spinning at a rate close to the orbital frequency; thus, orbital angular momentum decreases, forcing the stars to come closer together. This mechanism appears to be a major factor in forcing R CVn stars to evolve into Roche-lobe contact, following which the system is transformed into an Algol binary. It appears also to be a major driver of mass transfer in cataclysmic variables and in Algols and to be a mechanism whereby components in contact systems such as W UMa stars are forced to merge into single stars. This latter process offers an explanation for the presence of blue stragglers in old clusters which also contain contact binaries. In summary, a process not normally included in the simple theory of single star evolution plays a central role in the evolution of some binary stars.

Finally, it must be confessed that the simple theory has demonstrable shortcomings and that, in several cases, these are not minor shortcomings. There are mass-loss processes other than a MSW which are not included in the simple theory and which are known observationally to play an exceedingly important role in the evolution of single stars much brighter than the Sun. Mass loss from initially very massive main-sequence stars can be strong enough to convert such stars into Wolf-Rayet stars, which reveal at their surfaces material that has been highly processed by nuclear burning in an earlier phase of evolution. The mechanism for mass loss is possibly the transfer of photon momentum to atoms by resonance absorption from the continuum. Mass loss from bright and cool AGB stars terminates the AGB phase, converting the AGB star into a compact remnant which emits photons that cause the ejected material to fluoresce as a planetary nebula. The mechanism for mass loss in this case is possibly a hydrodynamic instability which occurs following a helium shell flash.

Mixing processes not included in the simple theory are known to occur. For example, the abundances of CNO elements and of light elements such as lithium and beryllium at the surfaces of low mass red giants are not quantitatively in accord with the expectations of the simple theory and suggest some form of extra mixing through formally radiative regions during both the main-sequence phase and the red giant stage. The lack of an extended giant branch in the Hyades cluster, suggests that stars only 1.6 times the mass of the Sun do not develop electron-degenerate cores before they ignite helium at their centers, in contradiction with the simple theory which suggests that stars of mass less than about $2M_0$ should develop such a core. The discrepancy can be resolved if it is supposed that matter within the formal convective core of stars of mass larger than $\sim 1.2M_0$ is mixed outward beyond the edge of the formal core by some form of convective overshoot, thus forcing the star to mimic a more massive star (as modelled by the simple theory which does not take convective overshoot into account).

In summary, although the simple theory of stellar evolution does an excellent job of describing qualitatively and often quantitatively many aspects of the global evolution of stars, there are situations in which the theory is inadequate. Despite the known shortcomings, it is the view of this reviewer that these shortcomings are not such as to invalidate the first-order picture of solar evolution which the simple theory provides.