NOVA MODELS AND THEIR PROBLEMS

H.-C. Thomas
Max-Planck-Institut für Physik und Astrophysik,
Föhringer Ring 6, 8000 München 40, FRG.

Introduction:

Very detailed spectroscopic and photometric observations of Novae are available today. Unfortunately they do not directly tell us what suddenly makes a star a million times brighter, sometimes even transforming the familiar constellations on the sky. Twenty-five years after its publication, Mestel's (1952) metaphor of a gigantic hydrogen bomb seems to be most widely accepted, although he at that time applied his model to a Supernova outburst. Later Giannone and Weigert (1967) as well as Rose (1968), Saslaw (1968) and others have computed hydrostatic models of such an event, using Kraft's (1963) binary model of a late type main sequence star and a white dwarf. Detailed hydrodynamic computations were carried out by Starrfield and his collaborators (for references see Sparks, Starrfield, and Truran, 1977) and recently by Prialnik, Shara, and Shaviv (A & A 62, 339, 1978).

In this lecture I will try to answer two questions: first, what refinements have been added to Mestel's original idea and second, how far can we trust the details of todays models? It will not be a review on Nova theory, since more competent astronomers than myself are preparing such an article for Annual Reviews of Astronomy and Astrophysics. Therefore I shall feel free to pick out several points of interest rather than seek completeness.

Nova models ...

Degenerate equilibrium configurations have played a dominant role in Nova models since Milne's (1931) idea that the Nova outburst is caused by the collapse of a normal star into the white dwarf stage. From the observed minimum in the light curve of Nova Herculis 1934 around May 1st, 1935 Grotrian (1937) deduced a radius of $1.7 \cdot 10^9$ cm for the central object, which certainly is characteristic of a white dwarf. However, this argument is fallacious, because if dust has obscured the source, the deduced bolometric luminosity and therefore also the radius computed from it will be much too low. Hoyle (1947) has put forward the first non-spherical Nova model: rotational instability is the mechanism for mass ejection during the collapse toward the white dwarf stage. But then Mestel (1952) showed that there is an alternative to the collapse as an energy source, namely unstable nuclear burning in a shell of accreted hydrogen on top of a white dwarf. Since this instability is fundamental for today's Nova models, let me remind you of an analytical model for a thermal instability (Kippenhahn, Thomas, and Weigert, 1966; Thomas, 1967). There a quantity A is defined as

$$\frac{1}{A} = \frac{c_P T}{\nu\varepsilon}\left(1 - \frac{\delta\nabla_{ad}}{\alpha - \frac{3}{4}\frac{r_s^3}{r_s^3 - r_c^3}}\right) . \qquad (1)$$

The underlying stellar model consists of a degenerate core of fixed radius r_c, a shell with nuclear energy generation on top of it extending to radius r_s, and an envelope. Then $A > 0$ is a necessary condition for a thermal instability to develop in the shell source. For the case of a thin shell $r_s - r_c << r_s$ and one obtains

$$\frac{1}{A} \approx \frac{c_P T}{\nu\varepsilon}\,\frac{4}{3}\,\delta\nabla_{ad}\,\frac{r_s^3 - r_c^3}{r_s^3} \qquad (2)$$

which is always positive. This is the well-known shell-source instability (Schwarzschild and Härm, 1965; Weigert, 1966). For a thick shell ($r_c << r_s$) or central burning ($r_c = 0$) A is positive if the electron gas is degenerate ($\delta << 1$):

$$\frac{1}{A} \approx \frac{c_P T}{\nu\varepsilon}\left(1 - \delta\,\frac{\nabla_{ad}}{\alpha - 3/4}\right) . \qquad (3)$$

Complications due to energy transport out of the shell have been neglected in this simplified treatment, but it is possible to include this in the manner of Kippenhahn et al. (1966). The instability comes about because the stellar matter cannot respond to an additional energy input with sufficent expansion, but rather has to heat up, thereby increasing the nuclear energy generation.

Mestel (1952) came to the conclusion that in a rather hot, luminous white dwarf the accreted hydrogen will be burned immediately, the star being able to radiate all the additional energy into space, while for a cool white dwarf (below $\approx 10^7$ K) hydrogen will not burn but will instead cool down. Only if the temperature is slightly above the burning temperature will conditions be favorable for an instability to develope. This general picture has to be modified since, depending on the accretion rate, compressional heating for a cool white dwarf may raise the temperature sufficiently so that nuclear burning can start (Giannone and Weigert, 1967). As has been shown by Kraft (1963) the most likely cause of accretion is Roche lobe overflow in a close binary system. If the companion of the white dwarf is a low mass main sequence star, then its nuclear time scale will determine the mass transfer rate. That requires the mass ratio of the white dwarf to its main sequence companion to be larger than 0.7 (for details see Ritter, 1976), since otherwise the distance between the two stars will shrink and the mass transfer rate will increase to much higher values. Whether this mass transfer is a steady flow of matter or can develope an instability as proposed by Bath (1975) is an open question. One may wonder however, why main sequence stars in dwarf Novae follow this instability while those in ordinary Novae do not.

Apart from these instabilities it is still difficult to obtain an estimate of the mass transfer rate from the value of the secondary component's mass, since this may not be a typical main sequence star but one which has lost part of its envelope in earlier phases of mass transfer. It could therefore evolve on a nuclear time scale quite different from that of a main sequence star of the same mass, a time scale which depends on the amount of hydrogen already converted to helium in its interior. Another uncertainty is introduced by angular momentum loss via gravitational radiation (Faulkner, 1971), which cannot yet be determined with sufficient accuracy (Ehlers et al., 1976).

Another problem is the determination of the mass of the hydrogen-rich envelope at the beginning of the outburst. Mestel (1952) concluded that one tenth of a solar mass would be sufficient to maintain degeneracy long enough until a high temperature has been reached during the instability. Starrfield, Sparks, and Truran (1974) have reduced this to 10^{-3} solar masses for a one solar mass white dwarf, because their hydrodynamic computations do not result in ejection of material for higher envelope masses. One way to get an estimate for the mass in the envelope at least for recurrent Novae is to turn around an argument used by Schatzman (1965) by asking what is the maximum mass which will come back to thermal equilibrium in between two outbursts. That requires the Kelvin-Helmholtz time scale of the envelope to be less than or about 100 years. With a typical prenova luminosity of less than 100 solar luminosities and a temperature near the ignition temperature of hydrogen ($\approx 10^7$ K) one obtains (see also Rose, 1968)

$$M_{env} \leq \frac{\tau_{KH}\ L}{c_v T} \leq 9\cdot 10^{-4}\ M_{\odot}, \qquad (4)$$

a result which also fits the numbers for two ordinary Novae obtained by Gallagher and Holm (1974). In recurrent Novae and probably also in ordinary Novae only part of this envelope will be blown into space during an outburst, because recurrent Novae would have to accrete mass at a rate of 10^{-5} solar masses per year to replenish the envelope, while currently estimated transfer rates are of the order of 10^{-8} solar masses per year or less.

Another question raised already by Mestel (1952) is, whether or not matter can escape from the rather deep potential well of the white dwarf. For his envelope mass of 0.1 solar masses and an energy generation rate of $5\cdot 10^{17}$ erg g^{-1} sec^{-1} at $2\cdot 10^{8}$ K Mestel (1952) estimated expansion velocities of the order of 10^{10} cm sec^{-1} as compared to the escape velocity of $5\cdot 10^{8}$ cm sec^{-1} for a one solar mass white dwarf. However this estimate was based on an extrapolation of the energy generation rate, which is unjustified, since the β^{+}-decay of the CNO isotopes limits the energy generation rate to approximately $2\cdot 10^{14}$ erg g^{-1} sec^{-1} (Fowler, 1966). So instead one can ask what rate of energy generation is required to reach escape velocity (see Starrfield, Sparks, and Truran, 1974). During peak energy generation the burning region will expand nearly isothermally and the energy radiated will be negligible. Energy balance therefore requires

$$\varepsilon_H M_H = -c_P T \nabla_{ad} \frac{\dot{P}}{P} M_{ej}\ , \qquad (5)$$

where ε_H is the energy generation rate due to hydrogen burning, M_H the amount of material in the burning region and M_{ej} the amount of material brought to escape velocity. Neglecting the kinetic energy one can use the equation for hydrostatic equilibrium to obtain the speed of expansion:

$$\dot{P} = -\frac{GM}{R^2}\rho\,\dot{R} \qquad (6)$$

or inserting the escape velocity for $\dot{R}$:

$$\dot{P} = -\frac{(GM)^{3/2}}{R^{5/2}}\rho\ . \qquad (7)$$

Inserting equation (7) into equation (5) together with the equation of state

$$P = \frac{1}{\beta}\frac{\Re}{\mu}\rho\ T \tag{8}$$

gives

$$\varepsilon_H = \frac{(GM)^{3/2}}{R^{5/2}}\ \frac{c_P}{\Re/\mu}\ \beta\ \nabla_{ad}\ \frac{M_{ej}}{M_H}\ . \tag{9}$$

For $c_P = \frac{5}{2}\frac{\Re}{\mu}$, $\nabla_{ad} = 0.4$, $M = 1M_\odot$, $R = 10^9$ cm this is

$$\varepsilon_H = 4.8\cdot10^{16}\ \beta\ \frac{M_{ej}}{M_H}\ \mathrm{erg\ g^{-1}\ sec^{-1}}\ . \tag{10}$$

So we see that unless the envelope is already dominated by radiation pressure ($\beta \ll 1$) or the amount of matter ejected is much smaller than that producing the nuclear energy, the rates required are too high for normal chemical compositions, a result which is confirmed by the hydrodynamic models of Starrfield, Sparks, and Truran (1974). This led them to the conclusion, that the burning must take place in surroundings enriched in CNO nuclei, because the rate is directly proportional to the amount of these nuclei. However this seems difficult to achieve even if matter is accreted by a carbon-oxygen white dwarf. Colvin et al. (1977) have studied the question of convective mixing and found that accretion quenches any convection quickly, because compressional heating raises the effective temperature of the white dwarf. So only a very thin layer at the bottom of the envelope will be enriched in carbon and oxygen, if the original luminosity of the white dwarf was as low as $4\cdot10^{-4}$ solar luminosities and the mass accretion rate only 10^{-12} solar masses per year. Since it will already take a white dwarf of one solar mass $2\cdot10^{8}$ years to cool down to 10^{-3} solar luminosities (Savedoff, Van Horn, and Vila, 1969), it may never reach this stage between two outbursts. The other way out of the dilemma, namely radiation driven ejection, was studied by Sparks, Starrfield, and Truran (1977) recently, but it requires a white dwarf mass relatively close to the Chandrasekhar limit and may therefore not be applicable to Novae like DQ Herculis, although mass estimates for this system are rather uncertain (Robinson, 1977).

To summarize this section it is fair to say that Mestel's (1952) model for the cause of the outburst is still valid, despite the fact that he did not suggest it for Nova explosions. But our knowledge about the range of parameters necessary to produce an outburst in a spherically symmetric model has been greatly improved.

... and their problems:

Paczyński (1971) emphasized that since in cataclysmic variables accretion onto the white dwarf takes place out of a surrounding disk it cannot be spherically symmetric. Durisen (1977) summarized the problems connected with this type of accretion. He noted that the non-spherical geometry may provide an explanation for the equatorial rings and polar caps observed in Nova shells (see for instance Grotrian, 1937; Hutchings, 1972; Fehrenbach and Andrillat, 1977) and could result in carbon enrichment of the hydrogen burning region. He also estimated that dissipation of rotational energy could influence the thermal equilibrium of the white dwarf. Kippenhahn and Thomas (1977) studied in detail the effect of shear instabilities on white dwarf models prior to the outburst. It is important to realize that the accreted matter has two properties tied together, namely because of its hydrogen content it has a lower molecular weight than the original white dwarf material and it possesses angular momentum. In writing down the Richardson criterion for stability of shear flow the stabilizing effect of a gradient in molecular weight has to be included. It is then possible to compute the distribution of accreted matter in horizontal and vertical directions which is marginally stable. If accretion is slow then turbulence will keep the actual gradient in angular velocity and molecular weight close to the marginal one. A sequence of marginal distributions with increasing amount of accreted material can then be converted into a time sequence, if the accretion rate is given.

Let us now compare the results of Kippenhahn and Thomas (1977) with the parameters deduced from spherical models. Because of its angular momentum the accreted material cannot spread over the surface of the white dwarf. Instead, it forms a belt which penetrates deeper into the star, therefore reaching high temperatures much faster. After accretion of only $M_{acc} = 10^{-5}\ M_{\odot}$ onto a one solar mass white dwarf the temperature at the bottom of the belt reaches 30 million degrees. Nuclear burning will determine the time scale of evolution, when the instability developes faster than the time scale dictated by the accretion. At $M_{acc} = 10^{-5}\ M_{\odot}$ and for an accretion rate of $\dot{M} = 10^{-9}\ M_{\odot}$/year one has

$$\tau_{acc} = \frac{M_{acc}}{\dot{M}} = 10^{4} \text{ years.}$$

If the original white dwarf has evolved from a double shell source red giant, as has been argued by Ritter (1976), then the belt will most likely have penetrated the carbon-oxygen core and nuclear reactions will start in an environment consisting mainly of C^{12} and O^{16} with a small admixture of hydrogen and helium. Another consequence of the belt is that part of the kinetic energy of rotation is brought into the interior of the white dwarf and released there instead of possibly being radiated away at the surface in some unobservable wavelength band. For an accretion rate of 10^{-9} $M_{\odot}$/year this amounts to several tenths of a solar luminosity which will prevent the white dwarf from cooling down below this value. Since this luminosity is produced in a toroidal volume, circulations will be started which distribute the energy over the whole sphere (assuming for the moment that the belt does not disturb the equipotential surfaces of the original white dwarf to a large degree). However the resulting circulation velocities are so small, that mixing is insignificant. The same argument holds for the nuclear burning so that as long as hydrostatic equilibrium can be maintained and mixing stays insignificant, the instability will not differ much from the spherical case. Depending on the depth of penetration into the carbon-oxygen core it might however produce much larger enrichments of carbon, nitrogen, and oxygen in the Nova shell than in the spherical case. In conclusion of this section one might say that non-spherical models give some new answers but also introduce a lot more uncertainty into the theory of Nova outbursts. This is especially valid if one wants to consider the recurrancy of a Nova outburst for the structure of pre-outburst models. So one looks for answers from the observations.

What observations should tell:

In view of the problems with Nova models one would like to learn as much as possible from observations about masses, mass transfer rates, recurrance times, ejected masses, bolometric luminosities, radii, velocities of rotation, secular changes of orbital period and more such basic parameters of the systems. However, what we can get is mainly information about spectra which gives many details of the structure of the ejected shell but can hardly be traced back to the white dwarf structure before the outburst. How little is known can be seen from the reviews by Warner (1975) and Robinson (1977), so it may still take some time before Nova theories rest on solid ground. In some cases the magnetic field might be strong enough to modify accretion onto the white dwarf (Angel and Landstreet, 1971), a point which almost certainly is important for transient X-ray sources or X-ray Novae as they are sometimes called. Whether single stars may undergo a Nova outburst (Starrfield et al., 1976; Truran et al., 1977) is an interesting question, but Nova Cygni 1975 does not seem to provide a good example to proof this (Ferland, 1977).

I have carefully avoided to discuss models for dwarf Novae because without more detailed understanding of the problems of non-spherical accretion I feel unable to decide between a nuclear origin of the outburst and one which is caused by variations in the mass transfer rate as proposed by Bath (1975). In this connection the recurrent Nova T Coronae Borealis might be a case for Bath's suggested instability (see Plavec, 1973; Harmanec, 1974; Webbink, 1976) and I wonder whether this is the only Nova which cannot be pressed into the scheme of the common Nova models as discussed in this article.

References

Angel, J.R.P., Landstreet, J.D.: 1971, Astrophys. J. 164, L15
Bath, G.T.: 1975, Monthly Notices Roy. Astr. Soc. 171, 311
Colvin, J.D., Van Horn, H.M., Starrfield, S.G., Truran, J.W.: 1977 Astrophys. J. 212, 791
Durisen, R.H.: 1977, Astrophys. J. 213, 145
Ehlers, J., Rosenblum, A., Goldberg, J.N., Havas, P.: 1976, Astrophys. J. 208, L77
Faulkner, J.: 1971, Astrophys. J. 170, L99
Fehrenbach Ch., Andrillat, Y.: 1976, Astron. Astrophys. 52, 123
Ferland, G.J.: 1977, Astrophys. J. 212, L21
Fowler, W.A.: 1966, in: High Energy Astrophysics, Ed. L. Gratton, Academic Press, New York, p. 328
Gallagher, J.S., Holm, A.V.: 1974, Astrophys. J. 189, L123
Giannone, P., Weigert, A.: 1967, Z. Astrophys. 67, 41
Grotrian, W.: 1937, Z. Astrophys. 13, 215
Harmanec, P.: 1973, in: Late Stages of Stellar Evolution, Ed. R.J. Tayler, D. Reidel Publ. Co., Dordrecht, p. 195
Hoyle, F.: 1947, Monthly Notices Roy. Astr. Soc. 107, 231
Hutchings, J.B.: 1972, Monthly Notices Roy. Astr. Soc. 158, 177
Kippenhahn, R., Thomas, H.-C.: 1977, Astron. Astrophys., in press
Kippenhahn, R., Thomas, H.-C., Weigert, A.: 1966, Z. Astrophys. 64, 373
Kraft, R.P.: 1963, Adv. Astron. Astrophys. 2, 43
Mestel, L.: 1952, Monthly Notices Roy. Astr. Soc. 112, 598
Milne, E.A.: 1931, Monthly Notices Roy. Astr. Soc. 91, 4
Paczyński, B.: 1971, Ann. Rev. Astron. Astrophys. 9, 183
Plavec, M.: 1973, in: Extended Atmospheres and Circumstellar Matter in Spectroscopic Binary Systems, Ed. A.H. Batten, D. Reidel Publ. Co., Dordrecht, p. 244
Ritter, H.: 1976, Monthly Notices Roy. Astr. Soc. 175, 279
Robinson, E.L.: 1976, Ann. Rev. Astron. Astrophys. 14, 119
Rose, W.K.: 1968, Astrophys. J. 152, 245
Saslaw, W.C.: 1968, Monthly Notices Roy. Astr. Soc. 138, 337
Savedoff, M.P., Van Horn, H.M., Vila, S.C.: 1969, Astrophys. J. 155, 221

Schatzman, E.: 1965, in: Stellar Structure, Ed. L.H. Aller D.B. McLaughlin, Univ. Chicago Press, Chicago, p. 331

Schwarzschild, M., Härm, R.: 1965, Astrophys. J. 142, 855

Sparks, W.M., Starrfield, S., Truran, J.W.: 1977, preprint

Starrfield, S., Sparks, W.M., Truran, J.W.: 1974, Astrophys. J. Suppl. 28, 247

Starrfield, S., Truran, J.W., Gallagher, J.S., Sparks, W.M., Strittmatter, P., Van Horn, H.M.: 1976, Astrophys. J. 208, L23

Thomas, H.-C.: 1967, Z. Astrophys. 67, 420

Truran, J.W., Starrfield, S.G., Strittmatter, P., Wyatt, S.P., Sparks, W.M.: 1977, Astrophys. J. 211, 539

Warner, B.: 1976, in: Structure and Evolution of Close Binary Systems, Ed. P. Eggleton, S. Mitton, J. Whelan, D. Reidel Publ. Co., Dordrecht, p. 85

Webbink, R.F.: 1976, Nature 262, 271

Weigert, A.: 1966, Z. Astrophys. 64, 395

DISCUSSION of paper by H.C. THOMAS:

SHAVIV: 1. I object to the argument based on rates of energy production. The energy can be supplied on one time scale and ejected on another.

2. I do not understand how you apply the Richardson criteria to a turbulent disk. Moreover, the disk does not have a form confined to the equator of the white dwarf. The disk will be unstable, blows-up and engulfs the white dwarf.

3. The high temperatures you obtain are very promising because they affect the time for the nuclear runaway and hence make it possible to get runaway on such a wide range of M_{core}.

H.C. THOMAS: 1. I agree. For an outburst model it is the rate which is important, which is what I am talking about. This does not apply to something like continuous ejection.

2. The Richardson criterion is applied to the interior of the white dwarf, not to the disk. Standard disk models give a thickness smaller than white dwarf radii and our white dwarf models do not produce an instability in the disk.

3. This seems likely, although one has to carry computations through the instability, to verify it.

KIPPENHAHN: I think there is observational evidence that in novae and in dwarf novae systems the white dwarfs are not completely imbedded in a blown-up disk.

R.N. THOMAS: The instability in the belt will blow-up the disk; what happens later is still an open question, but the disk may engulf the star. Secondly, in old novae the disk is not seen. What you do see sometimes is a spot. Hence nothing will prevent the turbulent disk from surrounding the white dwarf.

SMAK: What is the source of energy to account for the $1L_\odot$ luminosity of the $1M_\odot$ white dwarf after it has accreted $10^{-5}\ M_\odot$?

H.C. THOMAS: The luminosity of the original white dwarf is a free parameter;we choose a $1L_\odot$ white dwarf from a computed cooling sequence, so the source is thermal energy. More accurately, the luminosity will be 10% higher, which is a contribution from the marginal layer, the source in this case being rotational energy of the accreted material.

DE LOORE: Is there really no danger that the outer layers of the accreting white dwarf swell up during accretion? Possibly the accretion could then be stopped for some time, and afterwards start again? Another point: can the infalling matter not produce shock waves accompanied by extra heating effects?

H.C. THOMAS: 1. After accretion of $10^{-5}M_\odot$ the marginal layer still rotates slowly, so the equipotential surfaces (or some generalization of these) practically are spheres and the star remains at the white dwarf radius.

2. This will happen at the surface and result in some additional radiation, but has no consequences for the interior solution of the marginal layer, because the angular momentum is still there.