The Physical Nature of $\eta$ Carinae

Kris Davidson

Astronomy Department, University of Minnesota. Minneapolis, MN 55455

Abstract. Regarding quantitative theory, this object is surprisingly unexplored despite a wealth of modern observations; we do not truly understand its behavior on any timescale from a week to a millennium. The rotational and thermal structures may be far from equilibrium. Here I sketch a number of unsolved problems, potentially significant for stellar astrophysics in general, concerning its interior structure.

1. A Peculiar Situation

If $\eta$ Car were merely our best-studied very massive star, then it would be useful but not particularly fascinating. Its true significance, rather, lies in the fact that it is currently in an Interesting State. The Great Eruption around 1840 indicated some instability that we still don’t understand.\textsuperscript{1} The thermal and rotational structure, still recovering from that outburst, probably departs from equilibrium with characteristic timescales of 3 to 300 years. Being unique among observable objects, both the eruption and the recovery process should be immensely valuable for theory. Meanwhile the mass-loss rate currently exceeds a “normal” very-massive-star value by a factor of 100 or so, but may have been even larger before 1940; two or three episodes of rapid change have been observed during the post-eruption recovery process; some undiagnosed phenomenon gives the stellar wind a little kick once every 5.5 years; and various puzzles that we’ll ignore here occur in the diffuse ejecta.

In other words, $\eta$ Car is one of the great astrophysical counter-paradigms, having repeatedly exhibited phenomena that theory failed to predict. In terms of significance and breadth of work to be done, it amounts to a topic, functionally equivalent to a class of objects, not merely one object – and, so far, it’s been a topic with an unnaturally low ratio of theory to data. One can easily list numerous interesting puzzles that have scarcely been explored at all.

In that connection let me express an old-fashioned personal attitude justified by experience and by the history of astrophysics, but which the average astronomer today appears to regard as semi-heretical. Why, for example, has so little theoretical effort has been devoted to an object as unique and significant as $\eta$ Car? We once listed twenty specific problems that must be solved before any-

\textsuperscript{1}Wherever no specific reference is cited here, see lists of papers cited in Davidson (2000), Smith et al. (2003), and especially in Davidson & Humphreys (1997). This is a review of the physical situation, not of the literature.
one can understand this subject (Davidson & Humphreys 1997); observations have partially answered some of them, but theorists have simply failed to do their part. I fear that current astrophysical research is being channeled by three unacknowledged bad habits: (1) Researchers (and even some advisory panels) prefer to collect routine-quality data on many objects, rather than excellent information on the most significant cases. This attitude leads to oversimplified topics (spherically symmetric cows, as an old joke said), and crucial details can go undetected. (2) Most astronomers also seem to favor objects that are already more or less understood! During the past two decades some genuine mysteries have been ignored and almost forgotten. (3) Meanwhile most astrophysical “theorists” today avoid any problem that can’t be explored with their favorite computer codes. Acting in combination, those three factors have impeded work on some topics, while, equally unfortunate, they tend to make astronomy more bland, doctrinal, and generally less exciting than it should be.

Theorists certainly appear to have been intimidated by η Car’s unique combination of unsolved puzzles and abundant data – even though, philosophically and historically, this is precisely the sort of case that often leads to novel insights. I’ll refrain from quoting Bacon, Eddington, Feynman, et al. in support of this claim; but Shaviv’s (2000) paper on the Eddington limit, which employed η Car as a hint toward a more general end, may serve to illustrate my point.

Now to business. Before we get into theory, note that the usual estimate of η Car’s luminosity, \( L \approx 5 \times 10^6 \ L_\odot \), is more robust than most values quoted for very massive stars. During the past 20 years we’ve seen a succession of other objects advertised as “most luminous star,” but each later suffered demotion. Typically one had to guess the UV flux of a hot star by extrapolating from its much weaker visual-wavelength or even IR brightness. For η Car, however, we measure the luminosity simply by integrating the observed brightness re-radiated by dust at all IR wavelengths, with only minor corrections. If the result is seriously wrong, then it’s most likely an underestimate because some visual-wavelength light may escape through gaps in the ejecta. If Eta is a binary system, the primary star almost certainly accounts for more than 80% of the total luminosity (see next section below). Thus, anyone who wishes to be quite prudent – at least by the standards of this field – can safely assume \( L > 3 \times 10^6 \ L_\odot \) for the primary star. Here I’ll suppose that the true value is at least 30 percent larger, as seems probable. No other individual star has yet been proven to be more luminous. No doubt our Galaxy has at least a dozen objects, maybe several dozen, that would rank higher if we could observe them; but few are likely to be in η Car’s post-eruption structural state.

Unfortunately we have only broad limits for the star’s present-day mass, and its current evolutionary state is debatable. Since helium is perceptibly over-abundant in the ejecta, we can be sure that it has evolved; and, given the uncertainties, it may even be close to the end of its lifetime. (For the usual LBV reasons, η Car cannot become a red supergiant. Moreover, always remember that the opaque wind hides the star itself.) The observed luminosity suggests a zero-age mass above 150 \( M_\odot \), the star has obviously lost an appreciable amount, and the classical Eddington limit requires a present-day value \( M > 90 M_\odot \). We usually opine “maybe 120 or 130 \( M_\odot \),” a plausible compromise, but this doesn’t show whether \( L/M \) is within 10% or 20% of the classical Eddington limit.
2. The 5.5-year Period and the Binary Question

Most recent observations of η Car have focussed on a 5.5-year spectroscopic cycle, which may (or may not) be crucial for understanding the star’s long-term behavior. Before 1990 various observers reported “spectroscopic events,” when high-excitation emission disappeared from both the star and nearby ejecta for a few weeks or months (Gaviola 1953; Thackeray 1967; Rodgers & Searle 1967; Viotti 1968; Zanella et al. 1984). The hard X-ray flux crashes nearly to zero at the same time (e.g., Ishibashi et al. 1999; Pittard & Corcoran 2002). Always interesting as potential clues, the spectroscopic events moved to center stage when Damineli (1996), having seen another example in 1992, noticed that they recur with a 5.5-year period. Thanks to that discovery, the next two instances near 1998.0 and 2003.5 were observed fairly well.

Subsequent discussion followed a somewhat peculiar line. Zanella et al. (1984) had proposed a mass-ejection interpretation (see below) that still seems consistent with the observations; yet most authors have preferred a later, quite different idea that doesn’t work as well! Based on the behavior of hard X-rays (less than 0.01% of the radiation output) and just a few of the visual-to-infrared emission lines, several people asserted that η Car must be an eclipsing, colliding-wind binary (see, e.g., Pittard et al. 1998). Concentrating on just the X-rays, they regarded other colliding-wind binaries as role models although the resemblances are imperfect. In the form described by Pittard and Corcoran (2002), that wasn’t a bad hypothesis in 1997 – although a shell ejection à la Zanella et al. can produce much the same X-ray behavior. Anyway, since 1997 the eclipse idea has been expressed often and confidently, as though no alternatives exist. Hence this sober warning for non-specialists: In fact the binary eclipse scenario has not been proven; it produced some failed predictions but no strikingly successful ones; it must be bent and trimmed aggressively in order to match observations (see the elephant parable, Davidson 2002); and, in my opinion, it is far less interesting than rival hypotheses that fit the data as well or better. How is this pertinent to the internal structure of η Car? – Let me answer that later.

A hot companion star helps in three ways: (1) It explains the 5.5-year period, which has recurred several times with precision better than a percent; (2) the observed 2-to-10-keV thermal X-rays can be produced by colliding winds; and (3) a relatively hot star is useful as a source of far-ultraviolet radiation to ionize some ejecta a few hundred AU away. Therefore, let’s assume here that a companion star exists and somehow causes the spectroscopic event cycle.2

In that case we do have a few constraints on the parameters. Since each spectroscopic event is brief, almost everyone agrees that the orbit must be eccentric, $e > 0.6$, and an “event” occurs near periastron. The Homunculus Nebula’s

---

2Caveat: We should bear in mind that the binary hypothesis has not been proven, even though some authors assert that it has been; a particular class of single-star model may conceivably explain all the observations mentioned here (Davidson 1999, 2002; Smith et al. 2003). A possible single-star periodicity near 5.5 years will arise near the end of Section 3 in this discussion; while there is evidence that the primary star, contrary to expectations, can produce the high ejection speeds required to explain the X-rays (Dufour 1989; Weis et al. 2001; Smith & Morse 2004; Davidson 1999). A binary scenario is easier to work with, but isn’t the Only Game in Town.
orientation strongly suggests an orbit plane inclined near $i \sim 45^\circ$, far from edge-on. In order to produce the observed 2–10 keV X-ray spectrum via wind-wind shocks, the hypothetical companion must have a wind speed around 3000 km s$^{-1}$; *ergo*, it’s hot, most likely a mid-O-type star. But the ionizing-UV luminosity is only a few percent of $\eta$ Car’s total luminosity, since [Ne III] and similar emission lines in the circumstellar region are fairly weak; therefore the secondary star can’t exceed 50 $M_\odot$ or so. Several arguments, which we’ll skip here, make a W-R star or other evolved object very unlikely. Altogether – just an educated guess – a 40 $M_\odot$ main-sequence companion star seems reasonable and the periastron separation is most likely 3 to 5 AU. So far as I know, no one has yet assessed the long-term orbit evolution implied by mass loss.

The awkwardness lies in supposing that a spectroscopic event involves merely an eclipse of the hot companion star, causing a temporary decrease of ionization on our side of the wind. Some sort of eclipse may occur, but the frequency and exuberance of claims to have “confirmed” this interpretation should evoke skepticism. In fact the late-1997 X-ray crash did not match the form of eclipse that had been predicted (see, e.g., Fig. 1 in Ishibashi et al. 1999), and it more closely resembled the Zanella et al. mass-ejection scenario; but most eclipse enthusiasts, readjusting their parameters after the event, nevertheless claimed success. Incontrovertible evidence for the eclipse idea repeatedly appeared from 1997 to 2001 (see, e.g., Damineli et al. 1997, 2000) – but was soon controverted (e.g., Davidson 1997, Davidson et al. 2000).

As noted above, Zanella et al. (1984) proposed their shell-ejection scenario long before we knew of the 5.5-year period. Higher-than-normal gas densities around the star would temporarily quench the emergent UV radiation, leading to the observed type of spectroscopic changes. Today we can modify their suggestion as follows. (1) Such an event may be triggered by the close approach of a secondary star at a distance of the order of 3 AU. Some particular surface instability may be required.$^3$ (2) Most of the ionizing UV radiation probably comes from the hot companion; this scarcely affects the basic idea, since the ejected material quickly surrounds both stars, and the emission lines that mark the event originate at larger radii. (3) The X-ray emitting shocks, located fairly close to the secondary star, are suddenly disrupted by the increase in ambient density, and later recover. Indeed one expects the observed pattern of X-ray behavior; the fit is arguably easier than for an eclipse model. (4) The “ejection” process may really be a latitude-dependent disturbance in the wind, e.g., perhaps the higher-density polar wind briefly spreads to lower latitudes (Smith et al. 2003). Qualitatively, at least, this type of interpretation seems consistent with all existing data.

Independent of what happens during an event, the X-ray behavior during most of the 5.5-year cycle suggests that the orbit’s major axis is roughly perpendicular to our line of sight (Ishibashi 2001; Davidson 2002). Consistent with this interpretation, certain features in our HST/STIS data may hint that the secondary star was moving away from us at the time of the 2003.5 event, which,

$^3$Note that the star’s rotation rate is probably comparable to the angular rate of orbital motion near periastron, one or two radians per month.
in most scenarios, occurred near periastron. The point is that such an orbit
orientation disagrees with that which is assumed in most papers advocating an
eclipse model (e.g., Damineli et al. 2000; Pittard & Corcoran 2002). If the re-
vised picture is correct, then the secondary star and associated shock fronts are,
indeed, eclipsed by the primary wind a few weeks after periastron passage –
i.e., after the hypothetical disturbance has made the eclipse largely moot; see
remarks in Davidson (2002).

The competing explanations have very different consequences for the sub-
ject. If a mass ejection or wind disturbance scenario is correct, then it tells us
something about the primary star, perhaps involving a surface instability. If, on
the other hand, the spectroscopic events are superficial eclipse-like affairs, then
they’re far less significant. Most evidence, though not conclusive, seems to fit the
former alternative better. The most disappointing feature of the debate so far is
the fact that no one has attempted to develop a real model for the spectroscopic
behavior. The ideas mentioned above, as well as one or two variants, have been
qualitative hypotheses – mostly words and sketches, not calculations. Spheri-
cally symmetric wind codes are clearly unsuitable, for more than one reason.
The problem deserves far more serious theoretical effort than it has gotten.

3. Inside the Star

The biggest questions about η Car involve its Great Eruption of 1837–1858
and subsequent recovery. Obvious initial question: Did the instability arise near
the surface, or deep within the interior? The star’s suggestive closeness to the
Eddington limit, and the bipolar form of the ejecta, favor a surface phenomenon.
On the other hand, the eventual ejecta must have occupied most of the star’s
pre-eruption volume, suggesting to some a deeper cause (Guzik et al. 1999).

With the proviso that this question hasn’t been settled, let’s assume that
the eruption began near the surface. We call this a “geyser” model, because the
instability must have migrated downward through the stellar mass layers in order
to eject the observed amount of material.4 – Or perhaps we should say the inner
layers expanded outward as the eruption continued for several hundred times
the dynamical timescale. Anyway, for a wide range of ejected mass values, the
eventual ejecta extended down to roughly \( r \sim 0.5 R_\star \) in the pre-outburst star.
Most likely the basic instability occurred at temperatures of several hundred
thousand K (see Glatzel 2005 and refs. cited there) or, as a second choice, cooler
than 35000 K (the elusive modified Eddington limit, Humphreys & Davidson
1994; cf. Shaviv 2001). Rotation may play an important role but doesn’t change
the qualitative reasoning, and the main driving force, presumably, was contin-
uum radiation. Here we’re concerned mainly with consequences of the eruption,
not the details of what caused it.

The energy budget obviously depends on how much mass was ejected. In
the past we’ve tried to estimate the mass of the Homunculus ejecta-nebula from
its thermal IR flux: calculate the amount of dust required to radiate that much,
multiply by some plausible gas/dust ratio, and get a result of several \( M_\odot \). The

4 “Geyser” connotes this physical analogy and has nothing to do with any plume or jet.
logic is somewhat inconsistent, because, remember, the IR flux was also used to measure the total luminosity – so we shouldn’t be surprised if either quantity has been underestimated. Including cooler dust seen at longer wavelengths, Morris et al. (1999) and Smith (2005) suggest a mass in the range 6−15 $M_\odot$ for the Homunculus, or even larger.

Budgetary concerns impel me to urge restraint, however. If, for instance, 8 $M_\odot$ of material was ejected, then we require —

- Emergent radiative energy $\sim 10^{49}$ ergs (observed during the event);
- Emergent kinetic energy $\sim 2 \times 10^{49}$ ergs (r.m.s. speed $\sim 500$ km s$^{-1}$);
- Original binding energy of ejecta $\sim 2 \times 10^{49}$ to $6 \times 10^{49}$ ergs;
- Total energy of the eruption was “of the order of” $7 \times 10^{49}$ ergs.

Let’s not worry about the questionable assumptions used here; in the absence of a real model, at least these numbers give some impression of what we’re dealing with. If far more than 8 $M_\odot$ was ejected as the authors cited above suggest, then emergent radiation carried only a small fraction of the total energy. That’s OK for an interior (bottom-up) outburst, but seems unpalatable for an instability operating near the surface where photons can wriggle out along favorable paths (cf. Shaviv 2000). Thus I’m skeptical of ejecta masses above 10 $M_\odot$ or so. The far-IR dust estimates are often portrayed as robust lower limits, but that isn’t necessarily the case. They depend on grain properties – conductivity, imaginary part of the refractive index – which are not securely known. (See also Davidson & Smith 2000). The old guess $\sim 3 M_\odot$ was almost certainly an underestimate for the reason noted above, but values below 10 $M_\odot$ remain credible.

The Great Eruption must have seriously deranged the star’s thermal structure. Imagine a thought-experiment sketched in Fig. 1, where some quantity resembling temperature is plotted as a function of mass zone. Curve 1 is the pre-outburst state, mostly convective although radiation may carry much of the energy flux. Suppose that some mechanism rapidly ejects the outer layers. If the ejection time greatly exceeds the dynamical timescale (as in Eta’s Great Eruption), then the remaining layers quickly expand and remain close to hydrostatic equilibrium. However, curve 2 shows a kink in the temperature-like function immediately after the event, because we assume that the outermost layers provided most of the eruption energy. This kink or irregularity in the temperature (or entropy?) structure then migrates inward on a local thermal timescale, curve 3.
Since the gradient is larger than in an equilibrium model, “thermal timescale” may mean something faster than one would normally expect; although, I’m not sure of this, not having a real model. The core temperature may return to normal later, on a thermal timescale for the entire star. One obvious point of all this: The stellar envelope remains out of thermal equilibrium for an interesting length of time. If the ejection event were instantaneous, I suppose the surviving layers would be greatly disordered for a timescale of the order of 3 years, substantially disturbed on a 30-year timescale, and perceptibly out of equilibrium on a 300-year timescale. The emergent luminosity may evolve on these timescales, depending on the energy and/or entropy distribution at any given time. Someone please attempt a few realistic models of the thermal recovery!

(As hinted above, they won’t necessarily require much knowledge about the instability mechanism. One simple artifice would be to remove some outer mass and energy so the star resembles stage 2 in Fig. 1, and see what happens. André Maeder may comment at this meeting regarding the recovery timescale.)

Now here are two more related questions for theorists. First, why did the Great Eruption stop at a certain point? If one does a series of numerical experiments as suggested above, removing progressive larger masses \(\Delta M\), does some essential characteristic of the star change when \(\Delta M\) approaches 10 \(M_\odot\)? Second, will the recovery process lead to another Great Eruption, or was it a unique affair in the star’s evolution? Judging from \(\eta\) Car’s photometric history, its stability progressively decreased during the century before the outburst, and then was relatively stable after the event (with some exceptions, see below).

The ostentatiously bipolar Homunculus nebula suggests that the star’s rotational structure may be as important for this problem as its thermal structure. Since the outermost material not ejected in the Great Eruption must originally have been located around \(r \sim 0.5R_*\), and expanded outward as the outburst continued, the surface rotation rate should have been substantially reduced afterward. Later, presumably, it would have begun to speed up again as angular momentum diffused outward from the dense inner regions. I suppose the relevant timescale is somewhat smaller than \(R_*^2/\nu\), where \(\nu\) is the effective viscosity. Estimating the turbulent viscosity by setting the effective Reynolds number to 300 or so, we find a characteristic recovery timescale of the order of 100 \(\times\) (rotation period), i.e., seemingly less than 40 years. This depends on a flock of assumptions, though, and the details are a little bewildering – another question that needs genuine, industrial-strength modeling. Eta’s prolate wind (Smith et al. 2003; van Boekel et al. 2003) suggests that the present-day surface rotation speed is indeed substantial.

Still on the sub-topic of rotation, I’m a little surprised that the Great Eruption remained essentially bipolar as it progressed. Assuming that the ejection surface (sonic point?) did not shrink very much during the 10-or-20-year outburst, inner material must have continuously expanded outward toward that surface, retaining its specific angular momentum. Therefore one might have expected the surface rotation to slow as the eruption continued, causing the later ejecta to approach a spherical, not bipolar, form – contrary to modern observations of the Homunculus. One way to avoid that conclusion is to guess that the specific angular momentum increases inward in the star, with a considerable gradient – is this reasonable? If so, then rotation may have been essential during
the entire duration of the instability. We may even conjecture that a decrease in the surface rotation speed stopped the eruption at some critical mass layer.

Observationally, the post-eruption recovery has been fraught with incident, not just a boring trend. A substantial secondary eruption occurred in the 1890’s, and later the spectroscopic/photometric state shifted rapidly between 1937 and 1952 (see Humphreys et al. 1999; Humphreys 2005; Feast et al. 2001; de Vaucouleurs & Eggen 1952; Gaviola 1953; O’Connell 1956). Either the mass-loss rate or the latitude structure of the wind probably changed at that time. Another rapid alteration may currently be underway; the star’s apparent brightening accelerated around 1997 (Davidson et al. 1999; Martin & Koppelman 2004), and the 2003.5 spectroscopic event differed from that seen in 1997-98. Conceivably each occasion marked a time when the surface rotation rate had increased to a critical value. Thermal explanations are also worth considering, if the emergent luminosity varies as part of the recovery process. Is there some reason why these change-of-state episodes have occurred at roughly 50-year intervals?

The concept of recovery time may be pertinent to the 5.5-year cycle. If a “spectroscopic event” involves the sudden loss of, say, about $10^{-3} M_\odot$ of outer-layer material, then the surface rotation rate should temporarily decrease by about 2 percent. A very simple estimate of the subsequent spin-up time to return to the original speed, based on the same reasoning mentioned above in connection with the Great Eruption, is – surprise! – about 6 years. The thermal timescale for relevant layers may be similar. In a binary model, this means that one orbital period may be about the right amount of time for the star to become susceptible again. More interesting, if 5.5 years allows only partial recovery, then two successive spectroscopic events may differ from one another – i.e., the orbital period and recovery timescale can combine in a non-trivial way, with hysteresis effects. Potentially even more interesting than that, we can imagine a 5.5-year periodicity even if $\eta$ Car is a single star! Some instability may cause a mass ejection event, and then recurs after a rotational or thermal recovery timescale of 5.5 years, etc. There is no demonstrated reason why such a cycle cannot have an accurately consistent period (remember the periodic geysers in Yellowstone Park, surprisingly good thermal oscillators).

Finally, I was about to congratulate myself on the total absence of the dread word “magnetic” so far here, but candor requires the following remark. Astronomers traditionally think of stellar activity as a phenomenon for cooler stars, but it may occur in the most massive hot stars too. Close to the Eddington limit, with instabilities lurking near or just below the sonic-point radius, we should expect convection, maybe rotational turbulence, and related MHD phenomena which may even assist the mass outflow and might conceivably produce high-velocity streams in the wind. Non-spherical disturbances may arise, with latitude dependences and, I fear, additional timescales. Unfortunately a skeptic might reasonably say That Way Lies Madness, the situation is so complex.

In summary, $\eta$ Car is a repository of interesting, potentially significant, fairly well-posed theoretical problems that have received almost no serious effort yet. Since this is the only very massive star that we can study well in its post-eruption state, some of these questions may be essential for understanding the evolution of the most massive stars in general. Frankly, it seems a little bizarre that they have been so neglected.
I wish to acknowledge STScI’s support for the η Car HST Treasury Project, program numbers GO-9420 and GO-9973. The Space Telescope Science Institute is operated by AURA, Inc., under NASA Contract NAS5-26555.

References

Davidson, K. 1997, New Astr., 2, 397
Davidson, K., & Humphreys, R.M. 1997, ARA&A, 35, 1
Davidson, K. 1999, in ASP Conf. Ser. 179, Eta Carinae at the Millennium, ed. J.A. Morse et al., p. 304
Davidson, K., & Smith, N. 2000, Nature, 405, 532
de Vaucouleurs, G., & Eggen, O.J. 1952, PASP, 64, 185
Glatzel, W. 2005, in these proceedings
Humphreys, R.M., Davidson, K., & Smith, N. 1999, PASP, 111, 1124
Humphreys, R.M. 2005, in these proceedings
Ishibashi, K., Corcoran, M.F., Davidson, K., et al. 1999, in ASP Conf. Ser. 179, Eta Carinae at the Millennium, ed. J.A. Morse et al., p. 266
O’Connell, D.J.K. 1956, Vistas in Astronomy, 2, 1165
Smith, N. 2005, in these proceedings
Discussion

Walborn: 1. I don't know of any precedent for shell ejections with a rigorous period from a single star; however, rotating magnetic fields can produce rigorous periods as in the sun and Ap/B stars. 2. Could the lack of highly ionization features before 1948 be related to UV quenching by the dust now clearing? 3. The very sharp edges of the lobes vs. very ragged outer structure may be an important clue to the outburst physics.

Davidson: 1. Precedent, of course, has never worked very well for eta Car. About the solar cycle: A distant observer, in a certain frame of mind, would say Aha! That star has an 11 year companion! – And he/she would be right, i.e. Jupiter, which constitutes a sobering example for the philosophy of scientific inference! 2. Conceivable. On the other hand, some of the high-excitation gas would have been at r 300 AU then, too close for dust. 3. Yes! good point! the two lobes appear dramatically different from the sparse irregular debris that somehow got outside them. At first sight, one doesn't even notice bipolar/axial morphology in the outer stuff.