

# **FORMATION AND EVOLUTION OF BLACK HOLES IN THE GALAXY**

*Selected Papers with Commentary*

**World Scientific Series in 20th Century Physics** **Vol. 33**

**H. A. Bethe  
G. E. Brown  
C.-H. Lee**

*Editors*

World Scientific

**FORMATION AND EVOLUTION**

---

**OF**

---

**BLACK HOLES IN THE GALAXY**

*Selected Papers with Commentary*

**World Scientific Series in 20th Century Physics****Published**

- Vol. 1 Gauge Theories — Past and Future  
*edited by R. Akhoury, B. de Wit, and P. van Nieuwenhuizen*
- Vol. 2 Scientific Highlights in Memory of Léon Van Hove  
*edited by F. Nicodemi*
- Vol. 3 Selected Papers, with Commentary, of T. H. R. Skyrme  
*edited by G. E. Brown*
- Vol. 4 Salamleitschrift  
*edited by A. Ali, J. Ellis and S. Randjbar-Daemi*
- Vol. 5 Selected Papers of Abdus Salam (with Commentary)  
*edited by A. Ali, C. Isham, T. Kibble and Riazuddin*
- Vol. 6 Research on Particle Imaging Detectors  
*edited by G. Charpak*
- Vol. 7 A Career in Theoretical Physics  
*edited by P. W. Anderson*
- Vol. 8 Lepton Physics at CERN and Frascati  
*edited by N. Cabibbo*
- Vol. 9 Quantum Mechanics, High Energy Physics and Accelerators — Selected Papers of J. S. Bell (with Commentary)  
*edited by M. Bell, K. Gottfried and M. Veltman*
- Vol. 10 How We Learn; How We Remember: Toward an Understanding of Brain and Neural Systems — Selected Papers of Leon N. Cooper  
*edited by L. N. Cooper*
- Vol. 11 30 Years of the Landau Institute — Selected Papers  
*edited by I. M. Khalatnikov and V. P. Mineev*
- Vol. 12 Sir Nevill Mott — 65 Years in Physics  
*edited by N. Mott and A. S. Alexandrov*
- Vol. 13 Broken Symmetry — Selected Papers of Y. Nambu  
*edited by T. Eguchi and K. Nishijima*
- Vol. 14 Reflections on Experimental Science  
*edited by M. L. Perl*
- Vol. 15 Encounters in Magnetic Resonances — Selected Papers of Nicolaas Bloembergen (with Commentary)  
*edited by N. Bloembergen*
- Vol. 16 Encounters in Nonlinear Optics — Selected Papers of Nicolaas Bloembergen (with Commentary)  
*edited by N. Bloembergen*
- Vol. 17 The Collected Works of Lars Onsager (with Commentary)  
*edited by P. C. Hemmer, H. Holden and S. K. Ratke*
- Vol. 18 Selected Works of Hans A. Bethe (with Commentary)  
*edited by Hans A. Bethe*
- Vol. 19 Selected Scientific Papers of Sir Rudolf Peierls (with Commentary)  
*edited by R. H. Dalitz and R. Peierls*
- Vol. 20 The Origin of the Third Family  
*edited by O. Barnabei, L. Maiani, R. A. Ricci and F. R. Monaco*

- 
- Vol. 21 Spectroscopy with Coherent Radiation — Selected Papers of Norman F. Ramsey  
(with Commentary)  
*edited by N. F. Ramsey*
- Vol. 22 A Quest for Symmetry — Selected Works of Bunji Sakita  
*edited by K. Kikkawa, M. Virasoro and S. R. Wadia*
- Vol. 23 Selected Papers of Kun Huang (with Commentary)  
*edited by B.-F. Zhu*
- Vol. 24 Subnuclear Physics — The First 50 Years: Highlights from Erice to ELN  
*A. Zichichi; edited by O. Barnabei, P. Pupillo and F. Roversi Monaco*
- Vol. 25 The Creation of Quantum Chromodynamics and the Effective Energy  
*V. N. Gribov, G. 't Hooft, G. Veneziano and V. F. Weisskopf;*  
*edited by L. N. Lipatov*
- Vol. 26 A Quantum Legacy — Seminal Papers of Julian Schwinger  
*edited by K. A. Milton*
- Vol. 27 Selected Papers of Richard Feynman (with Commentary)  
*edited by L. M. Brown*
- Vol. 28 The Legacy of Léon Van Hove  
*edited by A. Giovannini*
- Vol. 29 Selected Works of Emil Wolf (with Commentary)  
*edited by E. Wolf*
- Vol. 30 Selected Papers of J. Robert Schrieffer — In Celebration of His 70th Birthday  
*edited by N. E. Bonesteel and L. P. Gor'kov*
- Vol. 31 From the Preshower to the New Technologies for Supercolliders — In Honour  
of Antonino Zichichi  
*edited by B. H. Wiik, A. Wagner and H. Wenninger*

*Forthcoming*

In Conclusion — A Collection of Summary Talks in High Energy Physics  
*edited by J. D. Bjorken*

Selected Papers of Kenneth G. Wilson  
*edited by M. E. Peskin and B. Baaquie*

This page is intentionally left blank

World Scientific Series in 20th Century Physics

**Vol. 33**

# **FORMATION AND EVOLUTION OF BLACK HOLES IN THE GALAXY**

*Selected Papers with Commentary*

*Editors*

**H. A. Bethe**

*Cornell University*

**G. E. Brown**

*State University of New York at Stony Brook*

**C.-H. Lee**

*Seoul National University*



**World Scientific**

New Jersey • London • Singapore • Hong Kong

*Published by*

**World Scientific Publishing Co. Pte. Ltd.**

**5 Toh Tuck Link, Singapore 596224**

**USA office:** Suite 202, 1060 Main Street, River Edge, NJ 07661

**UK office:** 57 Shelton Street, Covent Garden, London WC2H 9HE

**British Library Cataloguing-in-Publication Data**

A catalogue record for this book is available from the British Library.

The editors and the publisher would like to thank the following for their permission to reproduce the articles found in this volume:

**American Association for the Advancement of Science, American Astronomical Society, American Physical Society, Discover Magazine, Elsevier Science B.V., Scientific American, Wiley-VCH**

**Formation and Evolution of Black Holes in the Galaxy: Selected Papers with Commentary**

Copyright © 2003 by World Scientific Publishing Co. Pte. Ltd.

*All rights reserved. This book, or parts thereof, may not be reproduced in any form or by any means, electronic or mechanical, including photocopying, recording or any information storage and retrieval system now known or to be invented, without written permission from the Publisher.*

For photocopying of material in this volume, please pay a copying fee through the Copyright Clearance Center, Inc., 222 Rosewood Drive, Danvers, MA 01923, USA. In this case permission to photocopy is not required from the publisher.

ISBN 981-238-211-9

ISBN 981-238-250-X (pbk)

Printed in Singapore

We dedicate this collection of papers to Rose Bethe, who throughout their long evolution  
fed us, walked us, consoled us and cheered us on.

This page is intentionally left blank

---

## Contents

<i>Preface</i>	xi
Chapter 1 Equation of State in the Gravitational Collapse of Stars	1
Chapter 2 How a Supernova Explodes	51
Chapter 3 Accretion onto and Radiation from the Compact Objects Formed in SN 1987A	63
Chapter 4 A Scenario for a Large Number of Low-Mass Black Holes in the Galaxy	71
Chapter 5 Neutron Star Accretion and Binary Pulsar Formation	83
Chapter 6 How Collapsing Stars Might Hide Their Tracks in Black Holes	95
Chapter 7 Mystery of the Missing Star	101
Chapter 8 Observational Constraints on the Maximum Neutron Star Mass	109
Chapter 9 On the Formation of Low-Mass Black Holes in Massive Binary Stars	115
Chapter 10 The Evolution of Relativistic Binary Pulsars	125
Chapter 11 Supernova Explosions, Black Holes and Nucleon Stars	145
Chapter 12 Evolution of Binary Compact Objects That Merge	151
Chapter 13 Contribution of High-Mass Black Holes to Mergers of Compact Binaries	163
Chapter 14 The Formation of High-Mass Black Holes in Low-Mass X-Ray Binaries	175

---

Chapter 15 Evolution of Black Holes in the Galaxy	189
Chapter 16 The Blandford-Znajek Process as a Central Engine for a Gamma-Ray Burst	225
Chapter 17 A Theory of Gamma-Ray Bursts	261
Chapter 18 Hypercritical Advection-Dominated Accretion Flow	285
Chapter 19 Evolution of Neutron Star, Carbon-Oxygen White Dwarf Binaries	293
Chapter 20 Formation and Evolution of Black Hole X-Ray Transient Systems	307
Chapter 21 Formation of High-Mass X-Ray Black Hole Binaries	317
Chapter 22 Broad and Shifted Iron-Group Emission Lines in Gamma-Ray Bursts as Tests of the Hypernova Scenario	337
Chapter 23 Discovery of a Black Hole Mass-Period Correlation in Soft X-Ray Transients and Its Implication for Gamma-Ray Burst and Hypernova Mechanisms	349
Commentary on Appendices A-D	375
Appendix A Kaon Condensation in Dense Stellar Matter	379
Appendix B Kaon Production in Heavy-Ion Collisions and Maximum Mass of Neutron Stars	469
Appendix C $K^-/K^+$ Ratios in Relativistic Heavy-Ion Collisions	475
Appendix D Strangeness Equilibration at GSI Energies	485
<i>Bibliography</i>	505

---

## Preface

This book contains 23 papers on astrophysics, chiefly on compact objects, written over 23 years. In addition there is an Appendix on kaon condensation which we believe to be relevant to the equation of state in neutron stars, and to explain why black holes are formed at relatively low masses.

This work was begun in Copenhagen, then carried out intensively in Januaries in Kellogg Lab at Caltech, the National Theory Institute at Santa Barbara, the Astronomy Department at Santa Cruz and the National Institute of Nuclear Theory in Seattle. Of course collaboration continued by mail and telephone and more recently by fax the rest of each year. The first paper on the collapse of large stars and the last paper on the black-hole mass-period correlation in galactic black-hole binaries have observational support. Many of the papers in between and in the Appendix offer new scenarios to replace accepted wisdom. Observations will only be able to test these scenarios, as we discuss in the commentaries, at some time in the future. In some cases confirmation or rejection may be as much as a decade away. We should follow new observations with great interest. We hope that the reader will share with us some of the pleasure that we have had in carrying out the developments described in this collection.

This page is intentionally left blank

---

## Chapter 1

# Equation of State in the Gravitational Collapse of Stars

H.A. Bethe, G.E. Brown, J. Applegate and J. Lattimer  
Reprinted with permission from *Nuclear Physics*, A324 (1979) 487-533  
Copyright (1979) by Elsevier Science B.V.

### Commentary by G.E. Brown

April 1, 1978, my wife and I picked up the Bethes from the airport in Copenhagen and took them to the house where they were renting some rooms. Hans had sprained his ankle on Mount Pion in Turkey and I told him that that was a bad prognosis for his work on the pion-nucleon many-body problem in which each new higher-order term one calculated was as large as all of the lower-order terms put together. Hans asked, "What should we work on?" I replied, "Let's work out the theory of supernovae." Hans said that he didn't know anything about them. I replied that workers in the field had the nuclear physics all wrong and that I knew how to correct it. In fact, I'd written a research paper with Jim Lattimer and Ted Mazurek (Mazurek, Lattimer & Brown 1978) showing that the many excited states of nuclei had essentially been left out in works to date. But, I said, "I need an expert on explosions." Hans admitted that he knew something about them. (He has been head of theory in the atomic bomb program in the Los Alamos under Oppenheimer.) We delivered the Bethes to their apartment and then, before going home that afternoon from the Institute, I left on Hans' desk a computer printout of a large star that Stan Woosley had evolved up to the point of collapse.

I came in the next morning and went to see Hans. He said, "The entropy in the Fe core is very low, less than unity (in units of  $k_B$ ) per nucleon." "So what?" "That means that the Fe core will collapse without being held up until the Fe nuclei merge into nuclear matter, at nuclear matter density. There isn't enough entropy for them to break up," Hans said. But, I expostulated, "All supernova calculations on supercomputers since the war show that the collapse is held up at about 1/1000 of nuclear matter density." "They're wrong!" Hans was right. To stop the infall of matter a strong repulsive force is needed and this is only available at a density well above nuclear matter density.

In May of 1978 we had a topical meeting on the collapse of large stars and resulting explosions in Copenhagen, and everyone began tabulating entropy, for the first time. The point was that before my paper with Lattimer and Mazurek, no one had included sufficient excited states in nuclei to soak up the entropy without producing pressure.

Theorists kept inventing ways to increase the entropy up to  $\sim 1.5$  per nucleon, and Hans was willing to accommodate them, but unity seemed to me to be a very nice number. In the Commentary to Paper 21, the reader will see that the latest, best value is 0.76 from calculations of Alexander Heger.

Willy Fowler turned the initials BBAL into "Babble", and the paper retains this nickname in the community.

I gave a colloquium on our work in the spring of 1979 at Caltech. Feynman, from the front row, asked me whether unity was really a very low entropy. I asked him, "What do you think the entropy per nucleon is in the middle of the sun?" He didn't have even a guess. I said, "About 10." (It's actually more like 14; I'd left out the numerical 2.5 in the Sakur-Tetrode Law.)

Willy Fowler thanked me profusely for getting Hans back into the field of nuclear astrophysics. Hans and I thought we would clear up the explosion subsequent to collapse quickly. Alas, no satisfactory successful supernovae calculations exist to this day! But we're here and most of the material our body is composed of had to be formed in the collapse of stars, so we know that supernovae do work.

In any case, the Bethe-Brown collaboration in nuclear astrophysics began with "Babble". Jim Applegate became a professor in Columbia, Jim Lattimer in Stony Brook. From all points of view the paper was very successful, and I learned about entropy, taught it lovingly to my undergraduate class in thermal physics, ever since.

#### Reference

- (1) Mazurek, T.J., Lattimer, J.M., and Brown, G.E. (1978), *Astrophysical Journal* **229**, 713.

## EQUATION OF STATE IN THE GRAVITATIONAL COLLAPSE OF STARS

H. A. BETHE†

*The Niels Bohr Institute, DK-2100 Copenhagen Ø, Denmark*

and

G. E. BROWN††, J. APPLEGATE†† and J. M. LATTIMER‡

*NORDITA, DK-2100 Copenhagen Ø, Denmark*

Received 12 February 1979

**Abstract:** The equation of state in stellar collapse is derived from simple considerations, the crucial ingredient being that the entropy per nucleon remains small, of the order of unity (in units of  $k$ ), during the entire collapse. In the early regime,  $\rho \sim 10^{10}-10^{13} \text{ g/cm}^3$ , nuclei partially dissolve into  $\alpha$ -particles and neutrons; the  $\alpha$ -particles go back into the nuclei at higher densities. At the higher densities, nuclei are preserved right up to nuclear matter densities, at which point the nucleons are squeezed out of the nuclei. The low entropy per nucleon prevents the appearance of drip nucleons, which would add greatly to the net entropy.

We find that electrons are captured by nuclei, the capture on free protons being negligible in comparison. Carrying the difference of neutron and proton chemical potentials  $\mu_n - \mu_p$  in our capture equation forces the energy of the resulting neutrinos to be low. Nonetheless, neutrino trapping occurs at a density of about  $\rho \approx 10^{12} \text{ g/cm}^3$ . The fact that the ensuing development to higher densities is adiabatic makes our treatment in terms of entropy highly relevant.

The resulting equation of state has an adiabatic index of roughly  $\frac{4}{3}$  coming from the degenerate leptons, but lowered slightly by electrons changing into neutrinos and by the nuclei dissolving into  $\alpha$ -particles (although this latter process is reversed at the higher densities), right up to nuclear matter densities. At this point the equation of state suddenly stiffens, with  $\Gamma$  going up to  $\Gamma \approx 2.5$  and bounce at about three times nuclear matter density.

In the later stages of the collapse, only neutrinos of energy  $\leq 10 \text{ MeV}$  are able to get out into the photosphere, and these appear to be insufficient to blow off the mantle and envelope of the star. We do not carry our description into the region following the bounce, where a shock wave is presumably formed, and, therefore, we cannot answer the question as to whether the shock wave, in conjunction with neutrino transport, can dismantle the star, but a one-dimensional treatment shows the shock wave to be very promising in this respect.

### 1. Introduction

Massive stars ( $M \approx 8-10 M_\odot$ ) live for  $\sim 10^7 \text{ y}$ , developing peacefully as heavier and heavier nuclear fuels are burned. Burning continues up to  $^{56}\text{Fe}$ , the nucleus with

† Permanent address: Laboratory of Nuclear Studies, Cornell University, Ithaca, NY.

†† Also State University of New York, Stony Brook, NY. Supported in part by the US Department of Energy, under Contract No. EY-76-S-02-3001.

‡ Permanent address: Department of Astronomy, University of Illinois. Supported in part by the National Science Foundation under grant NSF AST 76-22673.

the highest binding energy per nucleon. Then, because no further nuclear burning can proceed to create pressure to hold the constituents of the star against gravitational collapse, the end of the star comes with a bang, implosion of the central part into a dense matter core taking a time of the order of 1 sec.

During or following this implosion, the mantle and envelope of the star is thought to be blown off by neutrino transport outward, or by the shock wave formed upon bounce of the central core or by a combination of these two mechanisms, leaving a dense central part which evolves quickly into a neutron star.

The equation of state of the central region of the star during the collapse is obviously important in determining the characteristics and outcome of the implosion. Although determination of the equation of state looks like a formidable problem, considerable simplification can be introduced by the recognition<sup>1,2)</sup> that, after a certain stage in the implosion, neutrinos will be trapped, so that the collapse is adiabatic after this time, and we shall exploit this here.

In earlier calculations<sup>3)</sup> neutrinos produced in neutronization of the core escaped. However, with the advent of weak neutral currents<sup>4,5)</sup>, it was realized<sup>6)</sup> that coherent scattering by nuclei would be effective in giving mean free paths for high-energy neutrinos which were small fractions of the dimensions of the core of the star. Thus, even though travelling between collisions with the speed of light, neutrinos are unable to diffuse out of the central region of the collapsing star in times comparable with collapse time. In a detailed calculation of hydrodynamic collapse, Arnett<sup>7)</sup> examined these questions and came to the conclusion that neutrinos are necessarily trapped for densities higher than  $\rho \approx 10^{12} \text{ g/cm}^3$  if neutrino cross sections are at all comparable with those given by current weak interaction theories<sup>4,5)</sup>.

We shall show in this paper that the crucial feature of the collapse in determining the equation of state is the low entropy involved. At no stage does the entropy, in units of  $k$ , exceed about 1.5 per particle. Since the available number of states goes as  $\exp S/k$ , where  $S$  is the entropy, this implies an astoundingly high order in the system, and this order persists throughout the collapse.

That there is a high order at the beginning of the collapse, which occurs following silicon burning, is understandable. Nucleons are completely in nuclei at this stage, so many ( $\sim 60$ ) nucleons must move in conjunction with each other. The electrons are already highly degenerate at these densities of  $\sim 5 \times 10^9 \text{ g/cm}^3$ , so that their entropy is low,  $\sim 1$  per electron†.

The common hydrodynamic kinetic infall motion of the material is ordered, so that an entropy increase can occur only through the neutronization of the protons. During the early stages of this, low-energy neutrinos can escape freely, and entropy increases. We show that this increase results from: (a) the holes in the electron distribution, which will later be filled in by Auger-like processes, and (b) the deexcitation of nuclear excited states, reached in the electron capture. Neither of

† We measure entropy in units of  $k$ .

these mechanisms gives a large increase in entropy, as we shall show by estimates, and the total increase in entropy *per nucleon* is particularly small, since only a small proportion of the initial electrons ( $\sim 25\%$ ) undergo capture. Thus, up to the stage of neutrino trapping, the entropy per nucleon does not exceed the order of 1.5. Once trapping sets in, the entropy changes only little, as we shall calculate.

Given the low entropy, nucleons have no choice but to remain in nuclei, since drip neutrons would carry a large entropy per particle,  $\sim 8$  at the point of trapping ( $\rho \approx 10^{12} \text{ g/cm}^3$ ), and, consequently, not more than a few percent of the nucleons can be in the drip phase. Nuclei therefore persist up to the point where they begin to touch; i.e., up to nearly nuclear matter density.

The low entropy means that the temperature is kept relatively low during the entire collapse,  $T \leq 6 \text{ MeV}$  right up to the point where nuclei touch. This is an alternative way of understanding that the nucleons are not boiled out of the nuclei on the way in, but only appear once the nuclei merge.

When the nuclei do merge into nuclear matter, the implosion carries this matter to densities of about three times nuclear-matter density, up to  $\rho \sim 8 \times 10^{14} \text{ g/cm}^3$ , before the point of velocity reversal is reached, at which point there is a bounce. This bounce results in the production of a shock wave, and we cannot extend our present description into the region of the shock, but we hope that by giving a simple and realistic equation of state up to the bounce we can remove much of the ambiguity in the initial conditions for the shock.

Our equation of state is exceedingly simple. In densities up to nuclear matter density, the pressure is determined almost completely by the relativistic, degenerate leptons and is consequently that of a relativistic gas of fermions with adiabatic index  $\Gamma$ , slightly less than  $\frac{4}{3}$ , the decrease from  $\frac{4}{3}$  coming chiefly from electron capture. At intermediate stages of the collapse, some nuclei break up into  $\alpha$ -particles and neutrons, and at later stages they recombine. These latter processes have small effects on the equation of state; such effects are well understood, and it would be straightforward to include them in a detailed collapse calculation, so we shall not describe them in detail.

The  $\Gamma \leq \frac{4}{3}$  equation of state for the central zones will persist up to the point where nuclei begin to "feel each other"; i.e., nearly up to nuclear matter densities  $\rho = 2.7 \times 10^{14} \text{ g/cm}^3$ . As the nuclei are squeezed together into nuclear matter, they initially form nuclear matter at normal density, for which the contribution to the pressure from the nucleons is zero. Because of the stiffness of nuclear matter, as it is compressed, the nucleons rapidly take over from the degenerate leptons in determining the pressure and the density cannot be increased very much further before bounce takes place. Throughout much of the region of nuclear matter compression, the increase in nuclear energy is determined by the compression modulus of nuclear matter. The effective  $\Gamma$  in the region above nuclear matter density is  $\sim 2.5$ .

We thus have a very simple and relatively unambiguous collapse scenario in which  $\Gamma$  is slightly less than  $\frac{4}{3}$  over the entire region of densities up to  $\rho \approx 2.7 \times 10^{14} \text{ g/cm}^3$

(nuclear matter density), and then is substantially larger,  $\sim 2.5$ , up to bounce. In a narrow region near nuclear matter density, the sharp edge in the curve for  $\Gamma$  following from the above will be rounded off by effects we do not consider in detail.

Bounce occurs because of the exceedingly stiff equation of state pertaining above nuclear matter density. The bounce is analogous to that of a very stiff spring.

Our description ends at the point of bounce, because we do not include the entropy-producing shock wave. We are able, however, to discuss neutrino diffusion in the absence of shock waves and we show in appendix C that the diffusion rate is too slow for the neutrinos by themselves to blow off the mantle and envelope of the star. A one-dimensional treatment of the shock wave shows it to be very promising in this respect, although a three-dimensional calculation would be necessary to establish this.

## 2. Electron fraction, energy and entropy

Collapse begins, following Si burning, when the mass of the core of the star exceeds the Chandrasekhar limit:  $\sim 1.2 M_{\odot}$ . We assume the core develops independently of the more massive mantle and envelope; this is reasonable to the extent that the characteristic time for Si burning is long compared with the time for collapse of the central core.

We begin from the initial central conditions of Arnett<sup>7</sup>:

$$\rho_c = 3.7 \times 10^9 \text{ g/cm}^3, \quad T_c = 8.0 \times 10^9 \text{ K} = 0.69 \text{ MeV}, \quad (1)$$

which are supposed to pertain following Si burning.

There is general agreement on the initial number of electrons per baryon  $Y_e^{(i)}$  and that the presence of the degenerate electron gas with initial chemical potential  $\mu_e = 6 \text{ MeV}$  pushes the nuclei toward neutron rich isotopes, such that

$$\hat{\mu} = \mu_n - \mu_p = \mu_e, \quad (1.1)$$

there being plenty of time to establish  $\beta$ -equilibrium. Arnett<sup>7</sup> arrived at  $Y_e^{(i)} = 0.41$  enforcing hydrostatic equilibrium at the end of Si burning. Weaver, Zimmerman and Woosley<sup>8</sup> calculate in detail the nuclear transformations during Si burning and conclude, as Arnett, that a lot of electron capture takes place already in the pre-supernova stage, especially because this stage (in their calculation) lasts fairly long, several days. They find that "the composition is dominated by very neutron-rich species such as  $^{48}\text{Ca}$ ,  $^{66}\text{Ni}$ ,  $^{50}\text{Ti}$ ,  $^{54}\text{Cr}$  near the center of the star" and gradually shifts to less neutron-rich species farther out. These species have, respectively,  $Y_e = 0.417, 0.424, 0.440$  and  $0.444$ , and, accordingly, the central electron abundance found by Weaver *et al.*<sup>8</sup> is  $Y_e^{(i)} = 0.427$  (for a star of  $15 M_{\odot}$ ). The favored nuclear species listed above mostly have at least one closed shell,  $N = 28$  (Ca, Ti),  $Z = 28$  (Ni) or  $Z = 20$  (Ca), which explains their high abundance in the stellar matter.

Our formula (4.4), which we shall give later as representing the best overall fit to  $\mu_n - \mu_p$  taking into account bulk, surface and Coulomb energies, would predict an initial  $Y_e$  between 0.41 and 0.42 (see table 1) but we shall show in appendix A, by comparison with binding energies of the Fe isotopes, that our bulk contribution to  $\mu_n - \mu_p$  appears to be a bit too small in this region (or the surface contribution is too large), so that  $Y_e^{(i)} \approx 0.42$  seems reasonable. We shall adopt this value. The precise  $Y_e^{(i)}$  we choose is important only for estimating the entropy increase during the process of electron capture and will not be critical for our results.

TABLE 1  
 $\hat{\mu} = \mu_n - \mu_p$  as function of  $Y_e$

$Y_e$	$\omega_{\text{surf}}$ eq. (4.1)	$\bar{A}$	$\hat{\mu} = \mu_n - \mu_p$ eq. (4.4)	Epstein's empirical $\hat{\mu}_e$ eq. (4.6)	$\frac{d\hat{\mu}}{d Y_e}$ (eq. (12.1))
0.48	18.1	69			-170
0.464	17.9	73	-2.7*	-0.6	
0.46	17.9	74	-2.0	0	-177
0.44	17.6	80	1.7	2.88	-184
0.42	17.2	86	5.6	5.76	-191
0.40	16.7	91	9.5	8.6	-198
0.38	16.1	98	13.7	11.5	-206
0.36	15.4	105	18.0	14.4	-214
0.34	14.6	111	22.4	17.3	-223
0.32	13.7	118	27.1	20.2	-232
0.30	12.8	125	31.7	23.0	-240
0.28	11.8	132	36.6	25.9	-250
0.26	10.7	137	41.6	28.8	-259
0.24	9.6	147	47.0	34.6	-269
0.22	8.5	154	52.4	37.4	-279
0.20	7.4	163	58.1		-289

The density  $\rho = 10^{12} \text{ g/cm}^3$  is assumed in evaluating  $\bar{A}$ , eq. (5).

\* This negative value is an indication that we should increase the bulk symmetry energy (or decrease the magnitude of the surface symmetry energy) as discussed in appendix A.

The initial entropy per nucleus for translational motion is given by

$$\frac{1}{N} \frac{S}{k} = \frac{5}{2} + \ln \left[ \frac{V}{N} \left( \frac{M k T}{2 \pi \hbar^2} \right)^{\frac{3}{2}} \right] = 16.70, \quad (2)$$

where  $N$  is the number of nuclei in volume  $V$ , and  $M$  the average mass, which we take to be that of  $^{56}\text{Fe}$ . To this we add the entropy of the excited states of the nucleus and the entropy due to the  $\alpha$  particles and neutrons present due to thermal dissociation of nuclei. At the temperature and density (1) the  $\alpha$ 's and neutrons

contribute 3.6 and the excited nuclear states contribute 4.8. Thus the initial entropy per nucleon, obtained by dividing by 56, is

$$(S/k)_{\text{per nucleon}} = 0.45. \quad (2.1)$$

For the electrons, the Fermi energy is initially  $\mu_e = 6.0 \text{ MeV} \approx 8.6 kT$ . Thus,

$$(S/k)_{\text{per electron}} = \pi^2 \frac{kT}{e_F} \approx 1.15, \quad (2.2)$$

and this gives a total entropy per nucleon of

$$(S/k)_{\text{total}} = 0.45 + 0.48 = 0.93. \quad (2.3)$$

Entropy will increase during the development of the core up to neutrino trapping ( $\rho \approx 10^{12} \text{ g/cm}^3$ ) through electron capture†. The electron capture increases entropy, firstly because it occurs in matter that is not in beta equilibrium, and secondly because electron captures lead to excited states of the daughter nuclei which decay then to the ground state, mainly by  $\gamma$ -emission. However, escaping neutrinos will carry off a large fraction of the entropy. To show this we begin with a result from thermodynamics

$$dE + P dV = \sum \mu_i dN_i + T dS = dQ, \quad (3)$$

where the sum goes over electrons, protons and neutrons (and neutrinos, if they are trapped). Here  $E$  and  $S$  are the total internal energy and entropy and  $N_i$  is the total number of particles of species  $i$ .  $Q$  is the total heat lost or gained during the collapse. In electron capture  $dN_n = -dN_e = -dN_p$  ( $= dN_\nu$ , if neutrinos are trapped). We are interested in the case in which neutrinos are not trapped and freely escape so that  $dQ$  will be  $dN_e$  times some fraction of the maximum energy available for electron capture,  $\Delta$ . In appendix F we show this fraction to be about  $\frac{2}{5}$ . As long as neutrinos are not trapped  $\Delta = \mu_e - \hat{\mu} - \Delta_n$ , where  $\Delta_n$  is the excitation energy of the daughter nucleus, shown in appendix B to be of order 3 MeV. Eq. (3) now becomes

$$T(dS)_{\text{per nucleon}} = -dY_e[\mu_e - \hat{\mu} - \frac{2}{5}\Delta], \quad (3.1)$$

† Entropy increase proceeds only through the weak interactions, since the various phases are in equilibrium under the strong interactions. Although relevant times for the important part of collapse are only milliseconds, this is a long time compared with equilibration times of the strong and electromagnetic interactions. Indeed, we shall find that even the weak interactions, which are a factor of  $\sim 10^{12}$  weaker than the strong interactions, equilibrate by densities of  $\rho = 10^{12} \text{ g/cm}^3$  to the extent that entropy production beyond this density in the collapse is small, so it is clear that we have many orders of magnitude in hand when we maintain that the system is in equilibrium as far as the strong interactions are concerned.

or

$$T(dS)_{\text{per nucleon}} = -dY_e \left[ \frac{2}{3}(\mu_e - \hat{\mu}) + \frac{3}{5}\Delta_n \right], \quad (3.2)$$

which clearly shows the two contributions to entropy generation. We calculate later (eq. (15.2)) that  $Y_e$  at trapping is about 0.31 so that  $\Delta Y_e \approx -0.11$ , and from eq. (14.3) we find  $\Delta$  to be about 6 MeV. The average temperature during this part of the collapse is  $T = 1.4$  MeV (see table 10). Thus  $0.11(9/1.4) = 0.71$  units of entropy per nucleon are generated by electron capture (the excitation energy adds to the amount the system is out of  $\beta$ -equilibrium), but  $0.11 \frac{3}{5}(6/1.4) = 0.28$  units of entropy are carried off by neutrinos. The difference, 0.43, is the total entropy per nucleon produced up to neutrino trapping. Adding this to (2.3) we would have a total entropy per nucleon of 1.36. We adopt the value of

$$(S/k)_{\text{per nucleon}} = 1-1.5, \quad (3.3)$$

as our estimate of the final entropy, although our arguments and results would be little changed were the entropy  $\sim 50\%$  higher. Detailed calculations of entropy production will be given later (see appendix F) and give numbers which agree with our estimates. In particular, table 9 gives results calculated under the assumption of no neutron drip, and the entropy increase there is shown to be smaller than our above estimates. In the more realistic calculations of table 10, which include neutron drip, the entropy for  $Y_e = 0.31$  is in line with our above estimate of  $S/k \sim 1.3$ , and even for  $\rho_{10} = 107.5$ ,  $S/k = 1.48$ , within the range of eq. (3.3).

The value of  $Y_e^{(0)}$  which was important in this calculation, is obtained from the difference between the chemical potentials of neutron and proton  $\hat{\mu} = \mu_n - \mu_p$ . We shall find in the next section that this  $\hat{\mu}$  stays close to the electron chemical potential  $\mu_e$ , which in turn depends mainly on the density of matter,  $\rho$ . We shall use the following expression for  $\hat{\mu}$ , taken from Baym, Bethe and Pethick<sup>9</sup> (BBP):

$$\hat{\mu} = \mu_n - \mu_p = -\frac{1}{A} (\partial W_N / \partial Y_e)_{\rho_N, V_N, u}, \quad (4)$$

where  $W_N$  is the total energy of a single nucleus:

$$W_N(Y_e, \rho_N, V_N, u) = W_{\text{bulk}} + \alpha(Y_e) V_N^{\frac{2}{3}} + \beta Y_e^2 \rho_N^2 V_N^{\frac{4}{3}} (1 - \frac{3}{2}u^{\frac{1}{3}} + \frac{1}{2}u). \quad (4.1)$$

These three terms are the nuclear bulk energy, surface energy and Coulomb energy respectively, and  $\rho_N$  is the density inside a nucleus,  $V_N$  is the nuclear volume and  $u$  is the fraction of the total volume occupied by nuclei so that  $A = \rho_N V_N$  and  $u = \rho/\rho_N$ . For  $\alpha$  we use the results of Ravenhall, Bennett and Pethick<sup>12</sup>) who calculated the surface energy as a function of  $Y_e$ . If the surface energy in MeV is expressed as  $\omega_{\text{surf}} A^{\frac{2}{3}}$  so that  $\alpha = \omega_{\text{surf}} \rho_N^{\frac{2}{3}}$ , a fit to ref.<sup>12</sup>) in the range  $0.25 \leq Y_e \leq 0.5$  gives

$$\omega_{\text{surf}} = 290 Y_e^2 (1 - Y_e)^2. \quad (4.2)$$

If the Coulomb energy in MeV is expressed as  $\omega_{\text{Coul}} A^{\frac{1}{3}}$ , we have, from BBP:

$$\omega_{\text{Coul}} = 0.75 Y_e^2 (1 - 0.236 \rho_{12}^{\frac{1}{3}} + 0.00194 \rho_{12}), \quad (4.3)$$

where  $\rho_{12}$  is the density in units of  $10^{12} \text{ g/cm}^3$ . Evaluating the derivative in eq. (4) gives

$$\hat{\mu} = 250(0.5 - Y_e) - \omega_{\text{surf}} A^{-\frac{1}{3}} \left( \frac{1}{Y_e} + \frac{2}{Y_e} \frac{1 - 2Y_e}{1 - Y_e} \right). \quad (4.4)$$

The first term comes from the bulk symmetry energy taken from BBP. The second term represents surface and Coulomb corrections. In obtaining eq. (4.4) we have used the BBP prescription that the nuclear surface energy be twice the Coulomb energy.

$$\omega_{\text{surf}} A^{\frac{1}{3}} = 2\omega_{\text{Coul}} A^{\frac{1}{3}}. \quad (4.5)$$

As we shall point out in appendix A, there is some ambiguity in determining  $\hat{\mu}$  empirically from nuclear masses, since surface symmetry energies can be changed in such a way as to compensate for rather large changes in the volume symmetry energy and the result can be rather different for  $\hat{\mu}$  in the region far from stability. An alternative expression for  $\hat{\mu}$  has been determined completely empirically by Epstein<sup>11</sup>). He examines the energies (mass excess) of observed, neutron-rich nuclei and compares them with their most stable isobars; by differentiation,  $\hat{\mu}$  is obtained. He finds that the nuclei most favorable for high neutronization (such as  $^{48}\text{Ca}$ ,  $^{66}\text{Ni}$ ,  $^{132}\text{Sn}$ ) can be fitted by

$$\hat{\mu} = 144 (0.46 - Y_e). \quad (4.6)$$

For  $Y_e \leq 0.35$  this gives somewhat smaller  $\hat{\mu}$ 's than we obtain from eq. (4.4). In table 1 we report values from Epstein's formula (4.6), as well as from eq. (4.4). Employing the generally smaller values from eq. (4.6) would not change our results appreciably.

We have found that the expression

$$\hat{\mu} = 207 (0.45 - Y_e)(1.32 - Y_e), \quad (4.7)$$

reproduces accurately results of our complete expression (4.4).

Formulae deduced from empirical binding energies may not be directly applicable to our problem. Whereas the nuclei most bound, because of shell effects and pairing, will be formed in the initial stages as long as the temperature is low; as soon as the temperature increases, nuclei will be in highly excited states, and there will be little influence from shell effects and pairing. For the temperature of 1.9 MeV relating to Arnett's<sup>7</sup> zone 13 which we often use as an example, the excitation energy of a nucleus with  $A = 56$  is  $U = 24$  MeV, computed from our later eq. (20.3). For this

excitation energy such a nucleus has a level density of  $2 \times 10^7$  levels per MeV. Thus, shell effects and pairing in the ground state will have little to do with the evolution of the collapse and the Fermi gas treatment of BBP [ref. <sup>9</sup>] and of ref. <sup>10</sup>) is appropriate.

A more detailed argument for this conclusion goes as follows: In nuclei with one or two closed shells, the shell closure lowers the ground state energy of the system through additional binding effects, but leaves the level density at any appreciable temperature at roughly the value appropriate for odd-odd nuclei. The excitation energy of the closed shell nucleus above the ground state would be higher, but then the additional binding of the ground state should be subtracted before calculating level densities. A similar discussion pertains to effects from pairing.

To conclude this section, we estimate the most likely atomic weight as a function of  $Y_e$ , which is needed to determine  $\hat{\mu}$  and the neutrino mean free path eq. (16). For this we simply use the BBP condition, eq. (4.5), which gives, after neglecting a small term,

$$A = 194(1 - Y_e)^2(1 - 0.236\rho_{12}^{1/2})^{-1}. \quad (5)$$

The resulting values of  $A$  are given in table 1 for  $\rho = 10^{12}$  g/cm<sup>3</sup>. It is seen that in the important range around  $Y_e = 0.3$ ,  $A$  is roughly twice as big as for Fe.

The neutron chemical potential can be found, following BBP eq. (2.10):

$$\mu_n = (\partial W_N / \partial A)_{Z, V_N, u}, \quad (5.1)$$

which we cast as

$$\mu_n = \frac{1}{V_N} \left( \frac{\partial W_N}{\partial \rho_N} \right)_{Y_e, V_N, u} - \frac{Y_e}{A} \left( \frac{\partial W_N}{\partial Y_e} \right)_{\rho_N, V_N, u}, \quad (5.2)$$

since  $W_N$  is an explicit function of  $V_N$ ,  $\rho_N$ ,  $Y_e$  and  $u$ . We obtain

$$\mu_n = -16 + 125(0.5 - Y_e) - 150(0.5 - Y_e)^2 - 2\omega_{\text{surf}} A^{-\frac{1}{3}} \frac{1 - 2Y_e}{1 - Y_e}, \quad (5.3)$$

where the first three terms come from the bulk energy taken from BBP. At zero temperature, neutron drip occurs when  $\mu_n = 0$ , or from eq. (5.3), when  $(Y_e)_{\text{drip}} \approx 0.30$ . This compares with BBP, who found  $(Y_e)_{\text{drip}} = 0.315$ . Whereas our value  $Y_e^{(0)} = 0.31$ , to be derived in the next section, is near  $(Y_e)_{\text{drip}}$ , so that neutrino trapping occurs in our calculations near the neutron drip, some neutron drip would in no way be catastrophic, since the leptons would still give the overwhelming contribution to the pressure below nuclear matter density. (Recall that the contribution of a free neutron to the pressure will be  $kT$ . For our maximum temperature of 6 MeV, before nuclei merge, this is small compared with the average contribution per electron of  $\frac{1}{4}\mu_e$ ,  $\mu_e$  being about 100 MeV in the region important for collapse. Thus, the number of drip neutrons could be comparable with the number of electrons with little modification in our results.) At finite temperatures, neutrons begin to drip at

larger  $Y_e$  values, however. To account for them, the  $Y_e$  of eqs. (4) through (5.3) should be replaced by  $x = Z/A$ , related to  $Y_e$  by

$$x = Y_e / (1 - X_n), \quad (5.4)$$

where  $X_n$  is the drip neutron mass fraction. For the low entropies of our calculations,  $X_n \leq 0.2$  at densities below  $10^{12} \text{ g/cm}^3$ . Thus drip neutrons have little effect on our calculation of  $Y_e^{(0)}$  (see appendix F).

### 3. Electron capture rate

We now discuss the rate of electron capture and "neutronization" of the nuclear matter. In the literature considerable complication is introduced involving strength functions for electron capture, networks of nuclei, etc. We believe that for general features these complications are unnecessary and that a description based on the shell model is both simpler and more accurate. Capture on free protons is assumed to play a dominant role in many earlier descriptions. Because of the paucity of these, (see table 4) they are actually unimportant, except possibly just in the beginning of collapse where the electron chemical potential of 6 MeV may be insufficient for the electrons to supply the excitation energy necessary in the daughter nucleus to find the main shell-model strength†.

In this section we shall use the Fermi gas model for the nuclei and give in appendix B the relation of the Fermi gas model to the shell model. The latter will vary from nucleus to nucleus; we give arguments that the Fermi gas model should reproduce average features.

The cross section for electron capture, with neutrino emission, on a proton at rest is ‡‡

$$\sigma = 1.18 \times 10^{-44} (\epsilon_\nu / m_e)^2 \text{ cm}^2. \quad (6)$$

Not all electrons and protons will be able to participate in the capture, however. Among the protons, only those near the maximum (Fermi) energy  $\mu_p$  will be effective; the number of these in an energy interval  $d\epsilon_p$  is

$$\frac{mk_f d\epsilon_p}{\frac{1}{3}k_f^3} \quad (\epsilon_p < \mu_p), \quad (6.1)$$

where  $m$  is the mass of a nucleon. The energy relation in the electron capture is

$$\epsilon_e + \epsilon_p = \epsilon_n + \epsilon_\nu. \quad (6.2)$$

† Even if free protons are responsible for the electron capture, they have to be replenished from nuclei. The resulting relation between  $\hat{\mu}$  and  $\mu_p$  will then be quite similar to (12.6) albeit with a different coefficient.

‡‡ Derived from the weak coupling constant  $g_V = 1.00 \times 10^{-5} \text{ m}^{-2}$ , where  $m$  is the nucleon mass, and the ratio of axial vector to vector coupling constants  $g_A/g_V = 1.25$ .

Assuming low temperature, in accord with the low entropy deduced in sect. 2, we must have

$$\epsilon_e = \mu_e - \Delta_e, \quad \epsilon_p = \mu_p - \Delta_p, \quad \epsilon_n = \mu_n + \Delta_n, \quad (6.3)$$

where the  $\mu$ 's are the chemical potentials (Fermi energies) and the  $\Delta$ 's are positive quantities. In particular,  $\Delta_n$  is the excitation energy in the daughter nucleus† which is needed to find the main shell-model strength for the  $\beta$ -transition. In appendix B, we show that

$$\Delta_n \approx 3 \text{ MeV}. \quad (6.4)$$

The other three quantities we combine, thus

$$\Delta = \Delta_e + \Delta_p + \epsilon_n, \quad (6.5)$$

and we also introduce

$$\hat{\mu} = \mu_n - \mu_p. \quad (6.6)$$

Then the energy relation (6.2) becomes

$$\mu_e = \hat{\mu} + \Delta + \Delta_n. \quad (6.7)$$

Using these notations, the energy-averaged proportion of the total number of electrons and protons which can participate, with inclusion of the energy factor  $\epsilon_\nu^2/m_e^2$  from eq. (6), is

$$X = \int_0^\Delta \frac{\epsilon_\nu^2}{m_e^2} d\epsilon_\nu \int_{\Delta - \epsilon_\nu}^{\Delta - \epsilon_n} \frac{d\epsilon_p}{k_t^2/3m} \frac{3\epsilon_e^2}{\mu_e^3}, \quad (7)$$

where, by our definition,

$$\epsilon_e = \mu_e - (\Delta - \Delta_p - \epsilon_\nu). \quad (7.1)$$

It will turn out that in general  $\Delta_e \ll \epsilon_e$ ; i.e., the electrons must come near the top of the Fermi distribution. We can set  $\epsilon_e \approx \mu_e$  and we easily obtain, upon integration

$$X = \frac{1}{4} \frac{\Delta^4}{m_e^2 (k_t^2/3m) \mu_e}. \quad (7.2)$$

Consequently, the time rate of change of the electron proportion is

$$\frac{dY_e}{dt} = -X\rho(\text{in g/cm}^3) Y_e N_0 c \sigma_0 Y_e. \quad (8)$$

† It could be argued that  $\Delta_n + \Delta_p$  is the excitation energy. In this case,  $\Delta$  in (15) etc. will be somewhat smaller.

where  $N_0$  is Avagadro's number ( $6 \times 10^{23}$ ). Here  $\rho Y_e N_0$  is just the number of protons/cm<sup>3</sup>. Using the cross section from eq. (6), we obtain

$$\frac{dY_e}{dt} = -2.12 X \rho_{10} Y_e^2 \text{ sec}^{-1}, \quad (8.1)$$

where  $\rho_{10}$  is the density measured in units of  $10^{10}$  g/cm<sup>3</sup>. In deriving this equation we have assumed that we are at sufficiently low energies that the neutrinos can escape freely; i.e., there is no blocking.

Since  $\rho$  occurs as a factor on the right hand side of (8.1) we would like to have an equation expressing  $Y_e$  as function of the density  $\rho$ ; to this end, we try to obtain  $\rho$  as a function of  $t$ . Epstein, Nørgaard and Bond<sup>14</sup>), in fitting Arnett's collapse calculation<sup>7</sup>) find‡

$$\frac{d(\ln \rho)}{dt} = 19 \rho_{10} \quad (9)$$

This formula gives a collapse rate which exceeds the free-fall rate slightly above  $\rho = 10^{12}$  g/cm<sup>3</sup>. We find this to be unreasonable because of the following argument.

Homologous collapse of a star in free fall gives a collapse rate

$$\frac{d(\ln \rho)}{dt} = 224_{10}^{\frac{1}{2}}. \quad (9.1)$$

This would be the correct rate of development were all of the leptons removed from a core of mass equal to the Chandrasekhar limit. However, at no point during the collapse have more than 26% of the leptons initially present disappeared (we take  $Y_e^{(0)} = 0.42$  and shall find  $Y_e^{(0)} = 0.31$ ), so the leptons offer considerable degeneracy pressure to resist collapse. In fact, since the pressure for a gas of relativistic degenerate Fermions is proportional to  $n^{\frac{5}{3}}$ , where  $n$  is the number density, a fraction

$$F = \frac{4}{3} \frac{Y_e^{(0)} - Y_e}{Y_e^{(0)}}, \quad (9.2)$$

of the lepton support pressure has been removed at the stage of collapse corresponding to a given  $Y_e$ . In our calculations,  $F$  never exceeds 35%, since our final  $Y_e$  is 0.31.

With retention of most of the leptons, the equation for homologous collapse of a core of radius  $R$  would read

$$\frac{d^2 R}{dt^2} = -\frac{GM}{R^2}(F), \quad (9.3)$$

† Of course, these protons are inside nuclei, as discussed earlier in this section.

‡ These authors give several values for the constant we write as 19 in eq. (9); their central value, used here, corresponds closely to our conditions. In an independent fit, before we knew of their work, we obtained an identical equation with 25 instead of 19.

where  $M$  is the mass of the core. Solution of this equation for  $F = 1$  and use of

$$\rho = M/\frac{4}{3}\pi R^3. \quad (9.4)$$

leads to the free-fall eq. (9.1). Eq. (9.3) would be difficult to solve in detail, because  $Y_e$ , and therefore  $F$ , is a function of  $t$ . However, if  $F$  is constant, the result is

$$\frac{d(\ln \rho)}{dt} = 224(\rho_{10}F)^{\frac{1}{2}}. \quad (9.5)$$

Since we have seen that  $F$  never exceeds 0.35, we take the average  $F$  to be 0.20 and get

$$\frac{d(\ln \rho)}{dt} = 100\rho_{10}^{\frac{1}{2}}. \quad (10)$$

This rate differs appreciably from that given in eq. (9).

In any case, it must be made clear just which density is to be used on the right-hand side of (10), so we go into a more detailed discussion of homologous collapse.

Collapse can be homologous only in the region where the sound velocity exceeds the infall velocity, since different parts of the core must be able to "communicate" with each other if they are to collapse together. From Arnett's <sup>7</sup>) fig. 5, this defines the radius of the homologous core as  $\sim 40$  km, roughly that of his zone 12. Taking  $r_k$  and  $u_k$  to be the radius and velocity of zone  $k$ , and

$$\frac{d(\ln \rho)}{dt} = -3 \frac{u_k}{r_k}, \quad (10.1)$$

to define a rate for this zone,  $k = 12$ , we find

$$\left( \frac{d(\ln \rho)}{dt} \right)_{\text{zone 12}} = 1166 \text{ sec}^{-1}. \quad (10.2)$$

On the other hand, for Arnett's zone 12,  $\rho_{10} = 135$ , so we would obtain  $1162 \text{ sec}^{-1}$  from eq. (10). The agreement is good, and serves to specify the density  $\rho$  in eq. (10) as the density at the outer edge of the homologous sphere. Taking zone 8, right in the middle of the core, we find  $-3u_k/r_k = 933 \text{ sec}^{-1}$ , not greatly different from the value for zone 12, indicating that the core really does collapse homologously.

Combining eqs. (8.1) and (10), we obtain

$$\frac{d Y_e}{d(\ln \rho)} = -\frac{2.12}{100} X \rho_{10}^{\frac{1}{2}} Y_e^2, \quad (11)$$

giving the development of  $Y_e$  with  $\rho$ .

The factor  $X$  depends sensitively on  $\Delta$ , eq. (7.2), i.e., on the difference  $\mu_e - \hat{\mu}$ , cf. (6.7). It is therefore desirable to convert (11) into an equation depending on the symmetry energy  $\hat{\mu} = \mu_n - \mu_p$  by writing  $Y_e$  in terms of  $\hat{\mu}$ . Combining eqs. (4.4) and

(5) and dropping the insignificant factor involving the density in the latter equation (we are interested now only in densities less than  $\sim 10^{12} \text{ g/cm}^3$ ), one has

$$\hat{\mu} = 250 (0.5 - Y_e) - 50 Y_e (1 - Y_e)^{\frac{1}{2}} (3 - 5 Y_e). \quad (12)$$

Taking the derivative,

$$\frac{d\hat{\mu}}{dY_e} = -250 - \frac{150}{(1 - Y_e)^{\frac{1}{2}}} \left( 1 - \frac{14}{3} Y_e + \frac{35}{9} Y_e^2 \right). \quad (12.1)$$

In table 1 we tabulate  $d\hat{\mu}/dY_e$  from this equation as a function of  $Y_e$ . One sees that  $d\hat{\mu}/dY_e$  is constant in the range  $0.2 < Y_e < 0.4$  to within 20%. Since we shall see that our results are affected by only the  $\frac{1}{4}$  power of this constant, we assume henceforth that  $d\hat{\mu}/dY_e$  follows the bulk behavior,

$$\frac{d\hat{\mu}}{dY_e} = -250. \quad (12.2)$$

Using this, we have

$$\frac{dY_e}{d(\ln \rho)} = -\frac{1}{250 \text{ MeV}} \frac{d\hat{\mu}}{d(\ln \rho)}. \quad (12.3)$$

Since  $\rho$  can be expressed in terms of  $Y_e$  and  $\mu_e$ ,

$$\rho = C\mu_e^3/Y_e, \quad (12.4)$$

with  $C$  a constant, and since it can be shown that the rate of change of  $\ln Y_e$  is at most 10% of that of  $\ln \mu_e^3$ , we have

$$\frac{d\hat{\mu}}{d(\ln \rho)} \approx \frac{\mu_e}{3} \frac{d\hat{\mu}}{d\mu_e} = -250 \text{ MeV} \frac{dY_e}{d(\ln \rho)}. \quad (12.5)$$

where we have used eq. (12.2) in the last step. Using eq. (11) we obtain, then,

$$\frac{d\hat{\mu}}{d\mu_e} = \frac{6.36}{100} \frac{250 \text{ MeV}}{\mu_e} X \rho_{10}^{\frac{1}{2}} Y_e^2. \quad (12.6)$$

Taking  $k_t^2/m = 80 \text{ MeV}$  in  $X$ , eq. (7.2), and expressing all energies in MeV, we have, finally,

$$\frac{d\hat{\mu}}{d\mu_e} = 0.57 \rho_{10}^{\frac{1}{2}} Y_e^2 \frac{\Delta^4}{\mu_e^2}. \quad (13)$$

This important equation gives the rate at which  $\hat{\mu} = \mu_n - \mu_p$  follows  $\mu_e$ . We note once more that, cf. (6.7),

$$\Delta = \mu_e - \hat{\mu} - \Delta_n, \quad (13.1)$$

where  $\Delta_n$  was previously shown to have the fixed value of 3 MeV.  $Y_e$  is expressible in terms of  $\hat{\mu}$ , eq. (12), and  $\rho$  in terms of  $\mu_e$ , eq. (12.4). Thus (13) is a differential

equation for  $\hat{\mu}$  in terms of  $\mu_e$ . It can be solved by direct numerical integration<sup>†</sup>. Regardless of the initial conditions on  $\mu_e$  and  $\hat{\mu}$ , the results are in agreement with the following simple consideration.

As the density, and hence  $\mu_e$ , increases, the right-hand side of (13) would become very large if  $\Delta$  were a constant fraction of  $\mu_e$ . The differential equation then would indicate that  $\hat{\mu}$  increases much faster than  $\mu_e$ , but this is impossible because of (13.1). Hence  $\Delta/\mu_e$  must decrease with increasing  $\mu_e$  (density), which means that  $\hat{\mu}$  follows  $\mu_e$  more and more closely. Therefore,

$$\frac{d\hat{\mu}}{d\mu_e} \rightarrow 1 \quad \text{for large } \mu_e. \quad (13.2)$$

Setting the left-hand side of (13) to one gives

$$\Delta = \frac{1.15}{\rho_{10}^{\frac{1}{3}}} \sqrt{\frac{\mu_e}{Y_e}}. \quad (14)$$

The factor  $\rho^{\frac{1}{3}}$  would not have occurred in the denominator had we used the collapse rate (9) rather than (10); its presence expresses the fact that with our slower collapse rate at high densities we are closer to equilibrium between electron capture and  $\beta$ -decay, so that  $\mu_n - \mu_p$  is closer to  $\mu_e$ .

If we insert in (14) the value of  $\mu_e$ ,

$$\mu_e = 11.1(\rho_{10} Y_e)^{\frac{1}{3}} \text{ MeV.} \quad (14.1)$$

we obtain

$$\Delta = 3.83 Y_e^{-\frac{1}{3}} \rho_{10}^{\frac{1}{3}}. \quad (14.2)$$

Using our standard value  $Y_e = 0.31$ ,

$$\Delta = 5.7 \rho_{10}^{\frac{1}{3}}. \quad (14.3)$$

The dependence on the density  $\rho_{10}$  is extremely weak; even at  $\rho = 10^{12} \text{ g/cm}^3$ ,  $\Delta = 6.9 \text{ MeV}$ . The total lag  $\mu_e - \hat{\mu}$  is  $\Delta + \Delta_n$ , eq. (13.1), and hence about 10 MeV. Going to a lower density, like  $\rho_{10} = 1$ , (14.3) still gives a large  $\Delta$ , so  $\mu_e - \hat{\mu}$  is quite large (see eq. (13.1)) which would mean that  $\hat{\mu}$  lags considerably behind  $\mu_e$  at low density. Our procedure is not applicable to low density, because: (a) the collapse rate is slower than eq. (10), and (b)  $d\hat{\mu}/d\mu_e < 1$ .

For zone 12 of Arnett<sup>7</sup>)  $\rho_{10} = 135$ ,  $\mu_e = 39.0$  and  $Y_e = 0.32$ . Using eq. (14),

$$\Delta = 6.88 \text{ MeV.} \quad (15)$$

From eq. (13.1) and using  $\Delta_n = 3 \text{ MeV}$  we find

$$\mu_n - \mu_p = 29.1 \text{ MeV}, \quad (15.1)$$

<sup>†</sup> We are indebted to R. Epstein for doing such an integration. We have independently repeated this integration; results are given in table 2.

TABLE 2  
Values of  $Y_e$  and  $\hat{\mu} = \mu_n - \mu_p$  as functions of density  $\rho$

$\log \rho$	$Y_e$	$\mu_e$ (MeV)	$\hat{\mu}$ (MeV)	$\frac{d\hat{\mu}}{d\mu_e}$	$\Delta = \mu_e - \hat{\mu} - \Delta_n$	$\Delta$ from eq. (14)
9.0	0.420	3.86	4.27	0.03	-3.41	4.65
9.5	0.420	5.66	4.33	0.03	-1.67	4.87
10.0	0.420	8.31	4.42	0.03	0.89	5.11
10.25	0.420	10.07	4.50	0.08	2.57	5.24
10.50	0.418	12.20	4.90	0.35	4.30	5.37
10.75	0.411	14.78	6.40	0.76	5.38	5.54
11.0	0.399	17.61	8.82	0.92	5.79	5.73
11.25	0.384	21.06	12.08	0.95	5.98	5.94
11.50	0.366	25.10	15.96	0.96	6.14	6.18
11.75	0.346	29.85	20.53	0.96	6.32	6.45
12.0	0.324	35.39	25.88	0.97	6.51	6.75
12.13*	0.311	38.58	28.97	0.97	6.61	6.94

Also tabulated are  $d\hat{\mu}/d\mu_e$  from eq. (13) and  $\Delta = \mu_e - \hat{\mu} - \Delta_n$ . In the last column, the asymptotic values of  $\Delta$  from eq. (14) are given. Remember that  $\Delta$  is also the maximum neutrino energy  $(\varepsilon_\nu)_{\max}$ , and  $\Delta_n$  is the assumed excitation energy of the daughter nucleus (see appendix B).

\* $\rho = 1.35 \times 10^{12} \text{ g/cm}^3$  (Arnett's zone 12).

for this zone. Using eq. (12), this can be translated into a  $Y_e$  consistent with our development,

$$Y_e = 0.31. \quad (15.2)$$

This is just the value of  $Y_e$  for this density obtained by numerical integration of eq. (11) (see table 2). To be consistent, we should use this  $Y_e$  back in eq. (14) and also correct  $\mu_e$ , to obtain

$$\mu_e = 38.5 \text{ MeV}, \quad \Delta = 6.92 \text{ MeV}. \quad (15.3)$$

As can be seen, there is considerable stability in these numbers.

After making the preceding arguments in this section, we checked their validity by direct computations. Eq. (11) can be solved straightforwardly by numerical integration. Following our earlier discussion, the initial value of  $Y_e$  is  $Y_e^{(i)} = 0.42$  at  $\rho = 3.7 \times 10^9 \text{ g/cm}^3$ . The quantity  $\Delta$  can be evaluated from eq. (6.7),  $\hat{\mu}$  being given as a function of  $Y_e$  by eq. (4.4) and  $\mu_e$  can be calculated from  $Y_e$  and  $\rho$  (see our eq. (14.1)). The excitation energy of the daughter nucleus was chosen to be  $\Delta_n = 3 \text{ MeV}$ , although this is probably too high a value initially. Our results are not very sensitive to this choice.

In table 2 we show results of the numerical integration, giving  $Y_e$  as a function of  $\rho$ . In deriving phase space factors (eq. (7)) we made approximations. In appendix F the derivation is carried out without these approximations; the  $Y_e$  as function of  $\rho$  which we shall show in table 9 differs only slightly from that obtained from our simple

equation (table 2). In neither of these calculations is neutron drip included and we shall see in appendix F that with its inclusion  $Y_e$  drops appreciably, being  $\sim 0.28$  at  $\rho = 10^{12} \text{ g/cm}^3$  rather than 0.31 found in table 2. It should be remembered that blocking due to some neutrino states being filled, in the density region  $\rho = 10^{11}$  to  $10^{12} \text{ g/cm}^3$  where neutrinos can no longer get out freely (see sect. 4.), has not been taken into account.

#### 4. Neutrino trapping

The neutrino mean free path has been calculated by Lamb and Pethick<sup>16)</sup> to be

$$\lambda_\nu \approx 1.0 \times 10^6 \text{ cm} \left( \frac{\rho}{10^{12} \text{ g/cm}^3} \right)^{-1} \left( \frac{1}{12} X_h \bar{A} + X_n \right)^{-1} \left( \frac{\varepsilon_\nu}{10 \text{ MeV}} \right)^{-2}. \quad (16)$$

Here  $X_h$  and  $X_n$  are the mass fractions in heavy nuclei and in neutrons, respectively. This result is based on Weinberg's theory<sup>4)</sup> which gives for the nucleonic neutral current

$$j^\mu = \bar{\psi}_n \frac{\tau_3}{2} \gamma^\mu (1 - \gamma_5) \psi_n - 2 \sin^2 \theta_W \bar{\psi}_n \left( \frac{1 + \tau_3}{2} \right) \gamma^\mu \psi_n, \quad (16.1)$$

where  $\psi_n$  is the nucleon wave function, and  $\theta_W$ , the Weinberg angle. Lamb and Pethick assumed equal numbers of neutrons and protons, and used  $\sin^2 \theta_W = 0.35$ , so the relevant part of  $j_3^\mu$  for them was

$$(j_3^\mu)_{coh} = -A \sin^2 \theta_W \bar{\psi}_n \gamma^\mu \psi_n = -0.35 A \bar{\psi}_n \gamma^\mu \psi_n. \quad (16.2)$$

In our case, for unequal numbers of neutrons and protons,  $-A \sin^2 \theta_W$  would be replaced by  $-\frac{1}{2}N + \frac{1}{2}(1 - 4 \sin^2 \theta_W)Z$ . On the other hand, the preferred value of  $\sin^2 \theta_W$  is now  $\sim \frac{1}{4}$ , so that, practically, only the first term  $-\frac{1}{2}N = -\frac{1}{2}(1 - Y_e)A$  contributes. For relevant values of  $Y_e \sim 0.30-0.36$ , this is little different from the value  $-0.35A$  used by Lamb and Pethick. More generally, we may replace  $\frac{1}{12}\bar{A}$  in (16) by

$$0.12 \bar{N}, \quad (16.3)$$

where  $\bar{N}$  is the average number of neutrons in the nucleus.

Neutrino trapping begins roughly in Arnett's zone 12 whose physical conditions we listed in eq. (15) ff. Table 1 indicates that for  $Y_e = 0.34$ , the average nucleus has  $A \approx 110$ , hence  $\bar{N} \approx 72$ . Taking  $X_h = 1$ , and substituting (16.3) in (16), we find for  $\varepsilon_\nu = 10 \text{ MeV}$ ,

$$\lambda_\nu = 0.85 \times 10^5 \text{ cm}. \quad (16.4)$$

The radius of zone 12 is  $4.27 \times 10^6 \text{ cm}$ , which is about 50 times  $\lambda_\nu$ . Thus we conclude that neutrinos in zone 12 will be effectively trapped.

In appendix C, we calculate the diffusion time for the neutrinos at a density of  $6 \times 10^{12} \text{ g/cm}^3$  and find it to be about 100 msec. The collapse time from  $\rho = 10^{12}$  to  $\rho = \infty$ , according to (10), is only 2 msec, so diffusion times are long compared to hydrodynamic ones, indicating effective trapping.

A more precise argument is given in appendix D. Again starting from diffusion theory, we calculate the drift velocity  $v_d$  of neutrinos relative to the nuclear matter; it is directed outward because of the gradient of the neutrino density. This matter moves with a velocity  $u$  relative to the center of the star; we have  $u < 0$  (inward motion). Hence the neutrino radial velocity is

$$\dot{r}_\nu = u + v_d. \quad (16.5)$$

But the core of the star contracts more or less homologously; hence at a given  $r$ , the material density  $\rho$  stays about constant with time. If  $\dot{r}_\nu > 0$ , the neutrinos move (on the average) to lower density and will therefore ultimately escape; if  $\dot{r}_\nu < 0$ , they move with time to higher density and are therefore trapped. The trapping limit,  $\dot{r}_\nu = 0$ , is shown in appendix D to occur at  $\rho < 10^{12}$ .

Once neutrino trapping sets in, the total number of leptons in a given material element remains constant,

$$Y_e + Y_\nu = Y_{e0}, \quad (17)$$

where  $Y_{e0}$  is the electron fraction at the beginning of trapping. This equation assumes that there is no diffusion of neutrinos, and we show in appendix C that, with the mean free path given by (16), the diffusion is indeed minimal in the short time available ( $< 2$  msec). The lower limit on the outer integral in eq. (7) now becomes  $\mu_\nu$ , so that

$$X = \frac{1}{4} \frac{(\Delta - \mu_\nu)^2 (\Delta^2 + 2\Delta\mu_\nu + 3\mu_\nu^2)}{m_e^2 (k_f^2 / 3m) \mu_e} \quad (17.1)$$

For  $\mu_\nu \rightarrow 0$ , this is identical with (7.2) so there is a smooth transition from the state in which neutrinos flow unimpeded, to the trapping regime. (17.1) may be inserted into (12.6) (which remains unchanged) giving,

$$\frac{d\hat{\mu}}{d\mu_e} = 0.57 \rho_{10}^{1/2} Y_e \frac{\Delta^2 + 2\Delta\mu_\nu + 3\mu_\nu^2}{\mu_e^2} (\Delta - \mu_\nu)^2. \quad (17.2)$$

In the region of trapping  $\mu_\nu$  will increase rapidly as  $\mu_e$  increases and whereas  $\hat{\mu}$  will continue to increase because of some further electron capture,  $d\hat{\mu}/d\mu_e$  will tend to become smaller. Because of the increase in the factors  $\rho_{10}^{1/2}$  and  $(\Delta^2 + 2\Delta\mu_\nu + 3\mu_\nu^2)/\mu_e^2$  in eq. (17.2) and the decrease in  $d\hat{\mu}/d\mu_e$ ,  $\Delta - \mu_\nu$  will quickly become small compared with  $\Delta_n$ , and both quantities will be negligible compared with the  $\mu$ 's. Hence in the trapping region, we soon reach the equilibrium condition

$$\mu_e - \mu_\nu = \hat{\mu} = \mu_n - \mu_p; \quad (18)$$

i.e., the rate of electron capture is fast enough to preserve statistical equilibrium.

The conditions (17) and (18), together with (4), determine  $Y_e$  and  $Y_\nu$  at every density  $\rho$ . Choosing  $\rho = 10^{14}$ , close to nuclear density, and  $Y_{et} = 0.31$ , we get

$$\begin{aligned} Y_\nu &= 0.04, & Y_e &= 0.27, \\ \mu_\nu &= 103, & \mu_e &= 155 \text{ MeV}, \\ \hat{\mu} &= 49 \text{ MeV}, \end{aligned} \quad (18.1)$$

in the trapping region. In this case  $dQ \approx 0$ , and application of eq. (3) gives

$$T(dS)_{\text{per nucleon}} = -dY_e(\mu_e - \hat{\mu} - \mu_\nu) \approx -dY_e \Delta_n, \quad (18.2)$$

since  $\Delta - \mu_\nu$  goes to zero beyond the trapping density, and the excitation energy is approximately equal to the amount the system is out of  $\beta$ -equilibrium. The temperature is now higher,  $> 3$  MeV (see table 4), so the entropy added is only  $0.04 \times \frac{3}{3} = 0.04$  per nucleon. Thus the recognition that neutrinos are trapped<sup>12)</sup> was very important for the simple description of the type we give.

### 5. Entropy of nuclei and of nuclear matter

The salient feature in the collapse is the low entropy per nucleon. Drip nucleons would have entropies of 5–10 per nucleon through most of the range of densities covered in the collapse, and the low entropy per nucleon keeps the nucleons mainly inside nuclei. To make quantitative estimates of the proportion of drip nucleons we should estimate the entropy of nucleons inside nuclei. This latter estimate will also tell us the temperature at which nuclear matter is formed when the nuclei squeeze together.

In the Fermi gas model, the nuclear level density at excitation energy  $U$  is given by<sup>17,18)</sup>:

$$\omega(U) = \sum_J (2J+1)\rho(U, J) = \frac{1}{12}\sqrt{\pi} \frac{e^{(4aU)^{1/2}}}{a^{\frac{1}{4}} U^{\frac{3}{4}}}, \quad (19)$$

where  $\rho(U, J)$  is the density of levels of angular momentum  $J$  at energy  $U$ . Here,

$$\frac{a}{A} = \frac{1}{4}\pi^2 \frac{1}{\epsilon_F} \approx 0.067 \text{ MeV}^{-1}. \quad (19.1)$$

We shall use the Fermi gas model as a guide, and later discuss deviations from it.

The partition function  $\Phi$  is obtained from

$$\Phi = \frac{1}{12}\sqrt{\pi}a \int_0^\infty \frac{e^{-U/kT + 2\sqrt{aU}}}{(aU)^{\frac{1}{4}}} dU. \quad (20)$$

Extending the lower limit of the integral to zero would cause difficulties, but it is known that the empirical level density at low excitation energies is substantially less

than that given by eq. (19), so one usually switches over<sup>18)</sup> to an exponential formula, unmodified by the denominator in (19). This switch has the effect of regularizing the integrand at our lower limit. Completing the square,

$$\begin{aligned}\Phi &= \frac{1}{12} \sqrt{\pi} a e^{akT} \int_{-\infty}^{\infty} \frac{e^{-(\sqrt{U/kT}-\sqrt{akT})^2}}{(aU)^{\frac{3}{4}}} dU \\ &= \frac{1}{6} \sqrt{\pi} (akT)^{-\frac{1}{4}} e^{akT} \int_{-\infty}^{\infty} \frac{e^{-(y-\sqrt{akT})^2}}{y^{\frac{3}{2}}} dy.\end{aligned}\quad (20.1)$$

Remembering that the integrand should be changed to something nonsingular for small  $y$ , it is clear that the main contribution to the integral comes from  $y \approx \sqrt{akT}$ . Therefore, we set the  $y^{\frac{3}{2}}$  in the denominator equal to  $(akT)^{\frac{3}{2}}$  and extend the integral from  $-\infty$  to  $\infty$ , obtaining

$$\Phi = \frac{1}{6} \pi \frac{e^{akT}}{akT}. \quad (20.2)$$

This differs slightly from the result of Mazurek *et al.*<sup>19)</sup>, who mistakenly used  $\rho(U, J)$ . (See eq. (19)) rather than  $\omega(u)$ , although their exponential factor was correct. For most of our purposes, the exponential factor  $\exp(akT)$  will be sufficient and, in fact, this is the only reliable part of the expression since the level density at low excitation energies is substantially changed by shell corrections and pairing effects.

Note that the average nuclear excitation energy will be given by the value of  $U$  which makes the exponent in the integrand zero in eq. (20.1), namely,

$$U = a(kT)^2. \quad (20.3)$$

Use of eq. (19.1) gives immediately the excitation energy per particle:

$$\frac{U}{A} = \frac{1}{4} \pi^2 \frac{(kT)^2}{\varepsilon_F}, \quad (20.4)$$

the well-known formula for a degenerate Fermi gas.

The free energy for nuclei is

$$F = -kT \ln \left[ \frac{V}{N} \left( \frac{MkT}{2\pi\hbar^2} \right)^{\frac{3}{2}} \Phi \right], \quad (21)$$

the contribution  $-kT \ln \Phi$  coming from the nuclear level density. As noted, at temperatures  $kT$  of a few MeV the leading term in  $\Phi$  is

$$\Phi \approx e^{akT}. \quad (21.1)$$

The entropy can be obtained from

$$S = -\frac{\partial F}{\partial T}, \quad (21.2)$$

so that the contribution to the entropy from  $\Phi$  of eq. (21) is

$$\frac{S}{k} = \frac{\partial}{\partial T} akT^2 = 2akT. \quad (21.3)$$

Thus, the entropy per nucleon from the level density is

$$\frac{S}{Ak} = \frac{2akT}{A} = \frac{1}{2}\pi^2 \frac{kT}{\epsilon_F}, \quad (21.4)$$

where we have used eq. (19.1). This is the well known entropy per nucleon appropriate for a degenerate Fermi gas of density corresponding to  $k_F = \sqrt{2m\epsilon_F}$ .

The level density parameter  $a$  appropriate for finite nuclei is, however, about twice the theoretical value given by eq. (19.1). A convincing demonstration of this is given in figs. 2-12 of Bohr and Mottelson<sup>20</sup>). Bohr and Mottelson attribute the larger value of  $a$  to the larger effective volume implied by the diffuse surface of the nucleus. It should be noted that the level density also depends upon the effective mass  $m^*$  at the Fermi surface, and the large effective masses found empirically<sup>21</sup>) and explained theoretically<sup>22</sup>) would also tend to increase the level density. In our derivation of the partition function, the low-energy excitations were handled roughly and our derivation is justified only for  $kT$  of the order of a few MeV.

## 6. Equation of state during collapse

In the early stages of the collapse, nuclei partially break up into  $\alpha$ -particles and neutrons<sup>23</sup>). This partial dissolution uses up energy and, therefore, lowers the pressure. This is one of the agencies lowering the effective adiabatic index of the material below  $\frac{4}{3}$ , the other one, in the early stage of the collapse, being the disappearance of leptons through electron capture and the subsequent emission of the neutrinos.

The partial dissolution of nuclei into  $\alpha$ -particles is well understood<sup>23</sup>), although it does not seem to be generally realized that the  $\alpha$ -particles go back into the nuclei at a later stage, a point we shall develop later. In order to bring out quantitatively the relation of this dissolution to entropy, we show in table 3 results for



at a density of  $\rho = 5.9 \times 10^{11} \text{ g/cm}^3$ . The density was chosen because it corresponds to Arnett's zone 13 the last zone before trapping occurs.

Our  $S/k$ , in each case, is the entropy per particle. The total entropy per nucleon is obtained in a straightforward way as

$$\left(\frac{S}{k}\right)_{\text{per nucleon}} = \frac{1}{56} \left[ (1-\delta) \left(\frac{S}{k}\right)_{\text{Fe}} + 13\delta \left(\frac{S}{k}\right)_{\alpha} + 4\delta \left(\frac{S}{k}\right)_n \right]. \quad (22.1)$$

where  $\delta$  is the fraction of  $^{56}\text{Fe}$  nuclei broken up.

TABLE 3

Compositions and temperatures for  $\rho = 5.9 \times 10^{11} \text{ g/cm}^3$ (a) Breakup of  $^{56}\text{Fe}$  ( $Y_e = 0.464$ ;  $Q = 124.4 \text{ MeV}$ ).

Proportion $\delta$ of Fe broken up	Entropy per Fe nucleus ( $S/k$ ) <sub>Fe</sub>	Entropy Per $\alpha$ particle ( $S/k$ ) <sub><math>\alpha</math></sub>	Entropy per neutron ( $S/k$ ) <sub>n</sub>	Total entropy per nucleon ( $S/k$ ) <sub>per nucleon</sub>	$kT$ (MeV)
0.1	12.76	8.43	8.22	0.46	1.38
0.2	13.05	7.92	7.71	0.66	1.55
0.3	13.28	7.60	7.39	0.85	1.65
0.4	13.50	7.39	7.18	1.04	1.73
0.5	13.74	7.22	7.01	1.21	1.80
0.6	14.04	7.11	6.90	1.39	1.88

(b) Breakup of  $A_H = 56$  with  $Y_e = 0.393$  ( $Q = 160$ ).

Proportion $\delta$ broken up	Entropy per nucleus	Entropy per $\alpha$ particle	Entropy per neutron	Total entropy per nucleon	$kT$ (MeV)
0.1	12.79	8.62	7.15	0.53	1.41
0.2	13.08	8.11	6.64	0.79	1.58
0.3	13.33	7.81	6.34	1.03	1.70
0.4	13.55	7.60	6.13	1.27	1.79
0.5	13.81	7.45	5.98	1.50	1.88
0.6	14.10	7.33	5.86	1.72	1.96

(c) Breakup of  $A_H = 56$  with  $Y_e = 0.357$  ( $Q = 163$ ) (column headings as given under (b))

0.1	12.73	8.66	6.80	0.55	1.35
0.2	13.03	8.16	6.30	0.84	1.53
0.3	13.28	7.85	6.00	1.10	1.65
0.4	13.51	7.66	5.80	1.35	1.74
0.5	13.77	7.51	5.65	1.60	1.83
0.6	14.06	7.39	5.53	1.84	1.91

\*Entropies in (b) and (c) should be increased, to take into account neutron evaporation from nuclei, before comparison with (a), as discussed in appendix E.

Calculations giving table 3 were carried out using the Saha equation:

$$\frac{n_{\alpha}^{13} n_n^4}{n_{\text{Fe}}} = \frac{1}{\Phi_{\text{Fe}}} \frac{2^{43}}{(56)^{\frac{1}{2}}} \left( \frac{m k T}{2 \pi \hbar^2} \right)^{24} e^{-Q/kT}, \quad (23)$$

where  $n_{\text{Fe}}$  is the number density for Fe,

$$n_{\text{Fe}} = N_{\text{Fe}} / V, \quad (23.1)$$

$\Phi_{Fe}$  is the partition function,  $n_\alpha$  and  $n_n$  are the number densities for  $\alpha$ -particles and neutrons,  $m$  is the nucleon mass and  $Q = 124.4$  MeV is the  $Q$ -value of the reaction. We assumed  $n_n = \frac{4}{13}n_\alpha$ , corresponding to the dissociation equation (22).

The above calculations are appropriate for  $Y_e = 0.464$ , but as nuclei become neutron rich, major changes occur in their breakup; especially their breakup into  $\alpha$ -particles is strongly inhibited.

In appendix E we discuss a model problem, the breakup of fictitious  $A_H = 56$  nuclei which have  $Y_e = 0.393$  and  $0.357$ , corresponding to 22 and 20 protons, respectively. This allows a direct comparison with the results for  $^{56}\text{Fe}$ , the only difference being the increased neutron richness. The results are given below those for  $^{56}\text{Fe}$  in table 3. From these results, we see that, for a given entropy, there is a smaller amount of breakup, and that the temperature is lower, as compared with the case of  $^{56}\text{Fe}$ , resulting from the larger number of neutrons which carry a higher entropy for the same amount of breakup and temperature.

Although useful for comparison with the  $^{56}\text{Fe}$  breakup, this calculation is unrealistic in at least two aspects: (i) The value of  $A$  will increase with decreasing  $Y_e$ , and this should be taken into account. (ii) Some drip neutrons are present, and these will inhibit the breakup into  $\alpha$ -particles and neutrons (law of mass action). We should, consequently, discuss a three-component system, consisting of nuclei,  $\alpha$ -particles and nucleons. Numerical results will be given in table 10.

Entropies in table 3 for  $Y_e = 0.393$  and  $0.357$  should be increased to take into account neutron drip. For  $Y_e = 0.357$  and  $\delta = 0.1$ , this increase is  $\sim 50\%$ . In the results shown in table 3 we also neglected the nuclear partition function  $\Phi$ . Inclusion of  $\Phi$  according to eq. (21.1) increases the breakup temperature  $kT$  by 0.13 MeV for  $Y_e = 0.357$  and  $\delta = 0.1$ , bringing the results of table 3 close to those of table 10. (For  $Y_e = 0.357$ , the corrected temperature in table 3 for  $\delta = 0.1$  is 1.48 MeV, to be compared with 1.51 MeV for  $\rho = 5.9 \times 10^{11} \text{ g/cm}^3$  from table 10.)

(In appendix E we show that the entropy for  $Y_e = 0.357$  and  $\delta = 0.1$  is about 1.33 per nucleon, to be compared with 1.31 from table 10.)

The relevant Saha equation for the dissociation of  $\alpha$ -particles into nucleons is

$$\frac{n_p^2 n_n^2}{n_\alpha} = 2 \left( \frac{mkT}{2\pi\hbar^2} \right)^{\frac{3}{2}} e^{-Q_\alpha/kT}, \quad (25)$$

where

$$Q_\alpha = 28.3 \text{ MeV}. \quad (25.1)$$

Proton and neutron densities are related by

$$\frac{n_p}{n_n} = e^{(\mu_p - \mu_n)/kT} = e^{-\hat{\mu}/kT}, \quad (25.2)$$

because  $\mu_p$  and  $\mu_n$  are the same for free nucleons and for those inside nuclei. We see from table 1 that

$$\hat{\mu} = Q_\alpha \quad \text{for } Y_e = 0.315, \quad (25.3)$$

$$2\hat{\mu} = Q_\alpha \quad \text{for } Y_e = 0.378. \quad (25.4)$$

Eq. (25) may now be written as

$$\frac{n_\alpha}{n_n} = \frac{1}{2} \left[ n_n \left( \frac{2\pi\hbar^2}{mkT} \right)^{\frac{3}{2}} \right]^3 e^{(Q_\alpha - 2\hat{\mu})/kT}. \quad (25.5)$$

The expression inside the bracket may be written

$$F = n_n \left( \frac{2\pi\hbar^2}{mkT} \right)^{\frac{3}{2}} = \frac{2.53\rho_{12}X_n}{T^{\frac{3}{2}}} \quad (25.6)$$

where  $T$  is in MeV. For  $X_n \approx 0.1$ ,  $\rho \approx 10^{13}$  g/cm<sup>3</sup> and  $T \approx 2$  MeV, this is of the order unity. (For lower density or higher temperature, it is even less.) The important factor in (25.5) is, therefore, the exponential, and this is  $\ll 1$  as soon as  $Y_e$  is substantially less than 0.378, eq. (25.4). This is true for most of the interesting region.

When  $Y_e = 0.315$ , eq. (25.3), we even have  $\hat{\mu} = Q_\alpha$  and we may then rewrite (25.5) as

$$\frac{n_\alpha}{n_p} = \frac{1}{2} F^3 e^{(Q_\alpha - \hat{\mu})/kT}. \quad (25.7)$$

Now the exponential is  $\sim 1$ , and thus, if also  $F \approx 1$ ,

$$n_\alpha \approx n_p. \quad (25.8)$$

This means that the  $\alpha$ 's are no more abundant than the very rare protons.

This makes it possible to use the theory of Lattimer and Ravenhall<sup>10)</sup> who have studied the equilibrium between neutron-rich nuclei and a "vapor" consisting of free neutrons and a few protons. This theory takes into account only volume and temperature effects; it leaves out surface, Coulomb and shell effects. The theory is based on a Skyrme interaction between the nucleons; this may somewhat exaggerate the effect of the repulsive core, and thus give too "hard" an equation of state at densities above nuclear density.

The properties of the mixture (including electrons) as a function of density has been calculated by one of us (J.L.) from the Lattimer-Ravenhall theory: The results for  $Y_e = 0.30$  are given in table 4, as a function of the entropy  $S$ , and of the density  $\rho_{13}$  (i.e. in units of  $10^{13}$  g/cm<sup>3</sup>). The temperature is given in MeV, then the percentage of nuclear matter (by mass) which has been evaporated in the form of single nucleons,  $X_n$ . The remainder,  $X_h = 1 - X_n$ , is in heavy nuclei. The next column,  $X_p$ , gives the evaporated protons as a percentage of total nuclear matter. Since the overall fraction of protons is 30%, essentially all of the protons form part of the heavy nuclei. The last column gives the baryon pressure,  $P_b$ , as a percentage of the electron pressure.

TABLE 4

Properties of the equilibrium mixture between neutron-rich nuclei and a free-neutron, free-proton vapor, calculated from the Lattimer-Ravenhall theory for  $Y_e = 0.3$

$S/k$ (per nucleon)	$\rho_{13}$	$T$ (MeV)	$X_n$ (%)	$X_p$ (%)	$P_v/P_e$ (%)
0.5	2.6	2	8	$4 \times 10^{-8}$	3
	10	3	3	$6 \times 10^{-6}$	1.5
1.0	0.33	2	18	$4 \times 10^{-5}$	8
	2.3	4	14	$5 \times 10^{-3}$	8
	4.7	5	11	0.015	7
	9.7	6	7	0.023	5
	1.5	0.27	24	0.006	20
2.0	2.1	6	21	0.27	17
	5.7	8	15	0.43	15
	9.2	9	10	0.37	11
	0.24	4	30	0.36	32
2.5	2.2	8	28	1.9	28
	4.6	10	24	2.1	24
	8.9	12	16	1.7	20

The value of  $S/k$  includes the electron entropy.

A striking result is the small value of  $X_n$ ; most of the nucleons are in heavy nuclei at all densities up to  $\rho = 10^{14}$ . The reason for the small  $X_n$  is the same as for the small number of  $\alpha$ -particles at  $\rho = 10^{14}$ : the entropy of the heavy nuclei, due to their excitation, is so large that essentially no entropy remains for dissociation – per nucleon, the entropy of the individual nucleons in the vapor is of course still higher. The internal excitation of the heavy nuclei is contained in the Lattimer-Ravenhall theory in the form of the motion of the Fermi gas of nucleons inside the heavy nucleus, precisely the method by which the level-density formula (19) was derived. It should be noted, however, that if we were to use the empirical level density parameter  $a$  in eq. (21.3), which is twice the liquid-drop value used to obtain (21.4) (see the discussion following that equation), we would find an entropy per nucleon of

$$S/k = \pi^2 \frac{kT}{\epsilon_F} \approx 0.25 kT, \quad (25.9)$$

and we would claim that this is the appropriate entropy per nucleon as long as the temperature is high enough to use the approximation (21.1) for  $\Phi$ , if the nuclei are separated. The above does not include the entropy of the electrons, and is, therefore, likely to be an underestimate. From this result (25.9) we see that for entropy  $S/k = 1.5$ ,  $kT$  cannot exceed 6 MeV, which is at variance with the results in table 4. The Lattimer-Ravenhall theory therefore ascribes too little entropy to the nucleus in the nuclei. Part of this is due to leaving out surface effects in their published work <sup>10</sup>).

After completion of that work, these authors found (private communication) that inclusion of entropy from the nuclear surface raised the total entropy per nucleon by  $\sim 50\%$ . This is in line with the discussion of Bohr and Mottelson<sup>20)</sup> who find the level density parameter  $a$  to be increased by 50% in going from the liquid drop model, which has a sharp surface, to the harmonic oscillator potential, which includes surface effects. This increase in  $a$  would be accompanied by the same percentage increase in  $S/k$  (see eq. (21.3)). We can thus understand roughly half the difference between the Lattimer-Ravenhall entropy for nuclei and that implied by eq. (25.9) for the empirical level density parameter. We look to the behavior of the nucleon effective mass in order to explain at least part of the remaining discrepancy. The Skyrme interaction used by Lattimer and Ravenhall<sup>10)</sup> had an effective mass  $m^*/m = 0.9$  while Jeukenne, Lejeune and Mahaux<sup>22)</sup> find that  $m^*$  peaks at larger values ( $\sim 1.2$ ) at an energy somewhat above  $\varepsilon_F$ , just in the region important for us. It is tempting to attribute a substantial part of the large empirical level parameter  $a$ , which is proportional to  $m^*$ , to the peaking of  $m^*$  in this region.

Because of the above considerations, we believe that in the real world, the figures listed under  $S/k = 1.0$  in table 4 would more appropriately refer to a larger entropy,  $S/k \approx 1.5$ , say.

The small fraction of mass in the vapor has the consequence that the nucleons contribute little to the pressure. As long as there is no degeneracy, the pressure is proportional to the number of particles, and that number is clearly small if most of the nucleons are in heavy nuclei. By contrast, in cold neutron-star models such as BBP, most of the nucleons are in the vapor, as free neutrons. The difference in our situation from that in BBP is a direct consequence of our relatively large proportion of protons,  $\sim 30\%$ . Moreover, in this case of BBP the density becomes high enough to lead to degeneracy, which in turn makes for a further increase in pressure. In our case, on the other hand, the pressure is almost entirely due to electrons and neutrinos, as long as the heavy nuclei exist. As long as this is the case, the adiabatic index is close to  $\frac{4}{3}$  and therefore gravitational collapse will continue. This point has been emphasized by Lamb *et al.*<sup>24)</sup> who have shown for reasonable values of the lepton number  $Y_e > 0.2$  and of the entropy  $S/k < 3$ , that in the density interval  $\rho = 10^{13}$  to  $10^{14} \text{ g/cm}^3$  the adiabatic index is less than  $\frac{4}{3}$  because: (a) nucleons are going back into nuclei and (b) the nucleon-nucleon potential is attractive here. The work of these authors shows that the entropy could be double, or the entropy increase three times, what we find here without substantially changing our picture.

For a given density, the fraction evaporated,  $X_n$ , is roughly proportional to the entropy,  $S$ . For given entropy,  $X_n$  decreases with increasing density above  $10^{13} \text{ g/cm}^3$  in accord with table 4. The fraction of evaporated protons is seen to be very small.

The most important result is the pressure due to baryons. This comes from the evaporated nucleons, and since their number is small, the baryon pressure is likewise.

We shall show in the next section that the baryon pressure must be considerably greater than the electron pressure to stop the gravitational collapse. At our low entropy,  $S/k = 1.0$  or  $1.5$ , the baryon pressure is only a small fraction of the electron pressure, and will therefore have hardly any influence on the rate of collapse.

Table 4 and the results of Lamb *et al.*<sup>24)</sup> therefore prove that the collapse cannot be stopped at any density below nuclear density. Once the latter density is reached, the nuclei merge into "one big nucleus" which we may consider as nuclear matter. Inside this, the nucleons move independently and exert pressure, as will be discussed in the next section.

### 7. Densities above nuclear density

At a density somewhat below that of normal nuclear matter, the nuclei will touch, and at about nuclear density  $\rho_0$ , they will merge completely. This transition, for zero temperature, is described in some detail in BBP, sect. 9. Once the nuclei have merged, we have a uniform medium of neutrons and protons to which the Lattimer-Ravenhall theory may be applied.

It was pointed out in the last section that neutrons do not substantially evaporate from nuclei below nuclear density. We see now that above that density, nuclei are squeezed together and change into individual nucleons by squeezing. These nucleons may now be considered as a degenerate Fermi gas with interactions. The interactions are such that at density  $\rho_0 = 0.16$  nucleon per  $\text{fm}^3$  (corresponding to  $2.7 \times 10^{14}$  g/cm<sup>3</sup>) the energy is a minimum, and hence the pressure is zero. Above  $\rho_0$ , the pressure increases rapidly because the kinetic (Fermi) energy dominates over the attractive interaction between the nucleons. The nucleon pressure is *not* due to the hard core in the nuclear interaction (as is sometimes claimed in the literature); the hard core becomes important only at densities of about  $5\rho_0$  or higher.

We calculate the nucleon gas pressure in two ways: (a) by using the empirical compression modulus  $K$  of nuclear matter and (b) from the Lattimer-Ravenhall theory. Finally, we shall use a combination:

(a) The compression modulus is defined as

$$K = 9 dp/d\rho. \quad (26)$$

Recently, it has been possible to determine  $K$  from the energy of the "breathing mode" of heavy nuclei; the result is<sup>25)</sup>

$$K = 220 \text{ MeV}. \quad (26.1)$$

The experiments were done on such nuclei as  $^{208}\text{Pb}$  which has  $Y_e = Z/A = 0.394$ ; we believe that the extrapolation to  $Y_e = 0.36$  or  $0.30$  is justified, especially since  $K$  is not expected to be a sensitive function of  $Y_e$ . The experiment (26.1) refers, of course, to temperature essentially zero, but it is easy to see (cf. table 5) that the pressure does

TABLE 5  
Baryon pressures, Lattimer-Ravenhall, for  $Y_e = 0.36$  and  $Y_e = 0.30$  in MeV/fm<sup>3</sup>

$Y_e = 0.36$		$P_b$	$T(\text{MeV})$		$S_e$
$\rho(\text{fm}^{-3})$	$S_b = 0.5$	1.0	2.0	$S_b = 1.0$	$S_b = 1.0$
0.16	0.16	0.68	2.78	7.4	0.113
0.20	2.55	3.39	6.76	8.6	0.121
0.24	7.03	8.28	13.3	9.7	0.129
0.32	24.6	26.9	36.5	11.8	0.142
0.40	57.3	61.3	77.2	13.7	0.153
0.48	110	116	140	15.5	0.162

$Y_e = 0.30$		$P_b$	$T(\text{MeV})$		$S_e$
$\rho(\text{fm}^{-3})$	$S_b = 0.5$	1.0	2.0	$S_b = 1.0$	$S_b = 1.0$
0.16	0.65	1.18	3.26	7.5	0.100
0.20	3.26	4.11	7.45	8.7	0.108
0.24	7.97	9.25	14.2	9.8	0.115
0.32	26.0	28.4	37.8	11.9	0.127
0.40	59.1	63.1	78.9	13.8	0.136
0.48	112	118	142	15.6	0.145

not depend much on temperature (entropy). We shall, in this section, measure pressures in the unit:

$$\text{MeV/fm}^3 = 1.6 \times 10^{33} \text{ dyne/cm}^2. \quad (26.2)$$

In this unit, then, the baryon pressure is

$$p_b = 24.5(\rho - \rho_0), \quad (26.3)$$

where the density is measured in the unit:

$$\text{nucleons/fm}^3 = 1.66 \times 10^{15} \text{ g/cm}^3. \quad (26.4)$$

For  $\rho = 2\rho_0 = 0.32$  nucleons/fm<sup>3</sup>, this gives  $p_b = 3.9$  MeV/fm<sup>3</sup>.

(b) The results of the Lattimer-Ravenhall (LR) theory are given in table 5, for proton fractions  $Y_e = 0.36$  and  $Y_e = 0.30$  and three different values of the entropy in baryons,  $S_b = 0.5, 1$  and  $2$ . The most striking feature is the rapid increase of pressure with density; extrapolating to  $S = 0$ , the data are well fitted by

$$p_b = 2.33[(\rho/\rho_0)^{3.5} - 1]. \quad (27)$$

This is substantially steeper than the result (26.3) derived from constant compres-

sibility; in fact, (27) yields the compression modulus

$$K = 9 \frac{dp_b}{d\rho} = 460 \left( \frac{\rho}{\rho_0} \right)^{2.5}. \quad (27.1)$$

We believe the increase of the compression modulus with density is more realistic than a constant  $K$ ; the increase comes about because the attractive nuclear forces become less important compared to the repulsive forces and the kinetic energy as the density increases. On the other hand, we feel that the empirical  $K = 220$  at  $\rho = \rho_0$  is more accurate than the 460 in (27.1), and we shall therefore multiply the LR pressures by  $\frac{220}{460} = \frac{11}{23}$  in the following calculations. We attribute the excessive compression modulus of 460 to the three-body term in the Skyrme interaction used by LR.

The pressure difference between the different entropies in table 5 is not great. As long as the temperature  $kT$  is small compared with the nucleon Fermi energy  $\epsilon_F$ , the effect on the pressure should be proportional to  $T^2$ , and since  $T$  is proportional to  $S_b$ , according to (21.3), it is also proportional to  $S_b^2$ . Hence we should have at any density,

$$p(S_b = 2) - p(S_b = 1) = 4[p(S_b = 1) - p(S_b = 0.5)]. \quad (27.2)$$

This is well fulfilled in table 5. Finally, the pressure does not appreciably depend on  $Y_e$ , e.g. we find at  $\rho = 0.48$  and  $S = 0.5$ :

$$p(Y_e = 0.3) - p(Y_e = 0.4) = 1.4 \text{ MeV/fm}^3 = 0.025p(Y_e = 0.4). \quad (27.3)$$

In table 5, we also list the temperature for  $S_b = 1$ ; it is in the neighborhood of 10 MeV. (Note that these temperatures are for nuclear matter, not nuclei. See the remarks at the end of sect. 6.) We also give the electron entropy, using (2.2). Both these quantities are proportional to  $S_b$ . The electron entropy adds 11 to 16% to  $S_b$ , so  $S_b = 1$  corresponds to  $S \approx 1.15$ , close to our initial conditions, see sect. 2. (It would have been more useful to calculate the pressures for constant total  $S = S_b + S_e$ , but our procedure saved a lot of work, and the ratio  $S_e/S_b$  is nearly constant in our density range.)

In table 6, we list  $\frac{11}{23}$  of the  $p_b$  calculated for  $S_b = 1$  (in accord with our argument above), to which we add the electron pressure  $p_e$  for the given  $Y_e$  (we assume we have only electrons; if 13% of these are converted into neutrinos, as (18.1) would imply, the combined pressure of the leptons would be slightly smaller for the same total lepton fraction,  $Y_{ei}$ ). The total pressure is well represented by

$$p = 3.6(\rho/\rho_0)^{2.5}, \quad (28)$$

although this formula is not quite as accurate as (27) for  $p_b$ . It is also seen that the baryon pressure exceeds the electron pressure above about  $\rho = 0.27$ . The equation (28) corresponds to an adiabatic index

$$\Gamma = 2.5. \quad (28.1)$$

TABLE 6  
Pressures in MeV/fm<sup>3</sup>, for  $S/k = 1$  and  $Y_e = 0.30$

$\rho$ (fm <sup>-3</sup> )	$P_e$	$P_b$	$P$
0.16	2.65	0.56	3.21
0.20	3.57	1.97	5.54
0.24	4.55	4.42	8.97
0.32	6.68	13.6	20.3
0.40	8.99	30.2	39.2
0.48	11.5	56.4	67.9

The energy per nucleon can be calculated by integrating (27) because (27) corresponds to a nucleon adiabat. Multiplying again (27) by  $\frac{11}{25}$ , we get

$$E_b = \int_{\rho_0} p \frac{dp}{\rho^2} = 6.9 \left[ 0.4 \left( \frac{\rho}{\rho_0} \right)^{2.5} + \left( \frac{\rho}{\rho_0} \right)^{-1} - 1.4 \right]. \quad (29)$$

From this we can calculate an approximate value of the density at which the collapse stops, as follows.

If we had the original number of electrons,  $Y_e^{(i)} = 0.42$  per nucleon, then the electron energy would at all times equal the negative gravitational energy; this energy is

$$\begin{aligned} -E_g = E_{e0} &= \frac{3}{4} \mu_e^{(i)} Y_e^{(i)} = \frac{3}{4} 111 (10^{-13} \rho_0 0.42)^{\frac{1}{3}} \left( \frac{\rho}{\rho_0} \right)^{\frac{1}{3}} 0.42 \\ &= 66 (\rho/\rho_0)^{\frac{1}{3}} \text{ MeV/nucleon}. \end{aligned} \quad (29.1)$$

Actually, we have only  $Y_e = 0.31$  electrons per nucleon, so that the electron energy is

$$E_e = (0.31/0.42)^{\frac{1}{3}} E_{e0} = 0.67 E_{e0}. \quad (29.2)$$

In order to stop the inward motion of the material we must have

$$E_b + E_e + E_g = 0, \quad (29.3)$$

and therefore

$$E_b = 22 (\rho/\rho_0)^{\frac{1}{3}}. \quad (29.4)$$

Setting this equal to (29) we find

$$\rho_s/\rho_0 = 2.9 \quad (30)$$

(subscript s for stopping). The inward motion stops only when the density reaches three times normal nuclear density, or  $0.8 \times 10^{15} \text{ g/cm}^3$ .

This estimate is very rough. First, we have calculated as if the entire core of the star had the same density; our result can at best apply to some average density, the central

density will be higher, that of the outer part lower. We have neglected, on the one hand, the thermal energy of the nuclear matter, which tends to use up some of the kinetic energy of the collapse. On the other hand, nuclei lose their surface and Coulomb energy upon merging, the transformation of a few of the electrons into neutrinos has lowered the pressure slightly, and the effect of general relativity would be to increase  $|E_s|$ , and therefore to increase the stopping density  $\rho_s$ .

The density  $\rho_s$  is sufficiently high to lead to a pion condensed phase of nuclear matter<sup>26)</sup>;  $\pi^-$  condensation is calculated to occur at twice nuclear matter density<sup>27)</sup> and  $\pi^0$  condensation might appear even earlier, given the large number of protons present. Pion condensation leads to a further softening of the equation of state, so we can foresee reaching densities of  $\sim 10^{15}$  g/cm<sup>3</sup> in the implosion.

The baryon energy increases rapidly near  $\rho_s$ , so the effect is almost as if the material were running into a brick wall. The "bounce" of the stellar material is very hard. We have not investigated the consequences of this on the shock reflected from the center of the star.

In order to get a rough idea on the consequences of this hard bounce, we have applied the Riemann method<sup>28)</sup> to a one-dimensional model for the same equation of state. This indicates that the outer part of the core acquires a substantial outward velocity after it is reached by the shock originating from the "brick wall"; in fact, for a large portion of it, this velocity exceeds the escape velocity (when we translate back to three dimensions). The "hard" equation of state is essential for this result. To obtain quantitative results it will be necessary to do numerical computations in three dimensions, and to treat the outgoing shock caused by the "brick wall" (the Riemann method treats only continuous changes of density and velocity, not shocks). However, on the basis of the one-dimensional model we are very hopeful that the hard equation of state will lead to ejection of material, and thus to the actual supernova phenomenon.

We would like to thank Richard Epstein for many useful discussions and Chris Pethick and Don Lamb for suggestions and criticism. We are grateful to Madeleine Soyeur for help with the calculations of electron capture. One of us (G.E.B.) would like to thank Gordon Baym for much basic tuition introducing him to supernova phenomena, and to Ted Mazurek for many interesting and helpful discussions. We would like to thank Willy Fowler for pointing out an error in our use of the nuclear level density in an earlier version of this paper.

#### Appendix A

##### NUCLEAR SYMMETRY ENERGIES

We need, in our work, an expression for  $\hat{\mu} = \mu_n - \mu_p$  which can be used for values of  $Y_e$  smaller than found in known nuclei, and we must be able to determine  $\hat{\mu}$  as a

function of  $A$ , so that a knowledge of individual contributions: bulk, surface and Coulomb, is necessary. Initially, we consider the iron isotopes, because  $^{56}\text{Fe}$  has minimum energy for  $Y_e = 0.464$ . Following eq. (4.5) we can eliminate the Coulomb energy in favor of the surface energy when near a minimum in the energy. In this way we arrived at eq. (4.4).

There is some ambiguity in fitting known masses, in that changes in the bulk symmetry energy can be traded off, to a great extent, against changes in the surface symmetry energy. It is important, therefore, to obtain one or the other independently from the fits to nuclear masses. The bulk symmetry energy is clearly the easier to calculate theoretically. Siemens<sup>29)</sup> obtained a value of 31 MeV for the symmetry energy from the Reid soft core potential, but after correction of an error in his program, this came down<sup>30)</sup> to 27 MeV. This compares well with the value of 28 MeV found, for this potential, by Bäckman<sup>31)</sup>. From the meson-theoretical potential of Jackson, Riska and Verwest<sup>32)</sup>, Bäckman finds 31 MeV. These numbers should be multiplied by 8 to obtain the coefficient in the bulk symmetry energy, and they are then not very far from the 250 MeV in eq. (4.4) taken from BBP.

We rely on Ravenhall, Bennett and Pethick<sup>12)</sup> for the dependence of  $\omega_{\text{surf}}$ , eq. (4.2) on  $Y_e$ .

For  $Y_e \sim \frac{1}{2}$ , we can compare the results of RBP with those of Myers and Swiatecki<sup>33)</sup> who found

$$\omega_{\text{surf}}(Y_e) = \omega_{\text{surf}}(0.5)[1 - (1 - 2Y_e)^2 C_{\text{symm}}]A^{\frac{2}{3}}, \quad (\text{A.1})$$

with  $C_{\text{symm}} = 1.79$  and  $\omega_{\text{surf}}(0.5) = 18.56$  MeV. Expansion of our expression (4.2) about  $Y_e = \frac{1}{2}$  gives  $C_{\text{symm}} = 2.00$ , 10% higher.

Mackie and Baym<sup>34)</sup> arrived at a functional form of the  $Y_e$  dependent surface energy, extrapolating from the situation where the density of drip neutrons is nearly equal to the density inside the nuclei. Expansion of their results about  $Y_e \approx \frac{1}{2}$  gives an expression of the form (A.1) with  $\omega_{\text{surf}}(0.5) = 17.6$  MeV and  $C_{\text{symm}}$  about half the value of Myers and Swiatecki<sup>33)</sup>, even though the value of 27 MeV found by Mackie and Baym for the bulk symmetry energy is close to that obtained by Myers and Swiatecki. Whereas their results for low densities of drip neutrons involves an extrapolation, it is important that they are able to obtain a good fit to the binding energies of medium and heavy ( $A \geq 40$ ) nuclei with their smaller surface symmetry energy.

From fig. 2 of Mackie and Baym, it can be seen that the surface correction to  $\partial W_N / \partial Y_e$ , the quantity determining  $\mu_n - \mu_p$  (see eq. (4)) is substantially smaller than in the results of Ravenhall, Bennett and Pethick<sup>12)</sup>, in the region  $0.3 < Y_e < 0.4$ , so that  $\hat{\mu}$  will be larger, for a given  $Y_e$ . The connection between  $\hat{\mu}$  and  $Y_e$  is particularly used in establishing the  $Y_e$  at neutrino trapping. (See eqs. (15.1) and (15.2).) If we use the expression of Mackie and Baym for the surface energy, we find  $Y_e = 0.33$  at trapping, which is substantially above neutron drip.

Whereas there is ambiguity in the surface symmetry energy, we believe the results of Ravenhall, Bennett and Pethick to give an upper limit for the surface corrections to  $\hat{\mu}$ , or a lower limit for the total  $\hat{\mu}$  in the region  $0.3 < Y_e < 0.4$  important for us.

Let us take  $\omega_{\text{surf}} = 18 \text{ MeV}$ ; there seems to be little ambiguity in this value. We then take the formula:

$$\hat{\mu} = C(0.50 - Y_e) - \frac{\omega_{\text{surf}}}{A^{\frac{1}{3}}} \left[ \frac{1}{Y_e} + \frac{2}{Y_e(1-Y_e)} \right], \quad (\text{A.2})$$

and determine  $C$  by fits to the known Fe isotopes, masses of which are taken from the Argonne Tables<sup>35</sup>). We find the results shown in table 7. For  $^{52}\text{Fe}$ ,  $Y_e = 0.5$ , and we cannot determine  $C$ . On the other hand, we can then determine  $\omega_{\text{surf}} = 18.0 \text{ MeV}$ , consistent with the value we have assumed.

TABLE 7  
 $\hat{\mu} = \mu_n - \mu_p$  for the Fe isotopes

Isotope <i>A</i>	<i>Y<sub>e</sub></i>	$\hat{\mu} = \mu_n - \mu_p$ from mass tables	Value of <i>C</i> implied by empirical $\hat{\mu}$ (MeV)	$\hat{\mu}$ from Epstein's formula $\mu = 144(0.46 - Y_e)$
50	0.520	-13.97	307	-8.64
52	0.500	-9.66		-5.76
54	0.4815	-4.52	366	-3.10
56	0.464	-0.99	330	-0.58
58	0.448	1.78	305	1.73
60	0.433	4.21	290	3.89

From table 7 it can be seen that the value  $C = 250 \text{ MeV}$  used in the text is, if anything, too small in the region of the Fe isotopes. As discussed above, we believe our surface energy correction to  $\hat{\mu} = \mu_n - \mu_p$  to be too large. Thus, our values of  $\hat{\mu}$  for the  $Y_e$ 's of interest are probably somewhat too small, and we view the value  $Y_e^{(l)} = 0.31$  at trapping to be a lower limit.

## Appendix B

### SHELL-MODEL DESCRIPTION OF ELECTRON CAPTURE

Our discussion in sect. 3 was in terms of the Fermi gas model, although we used an average excitation energy  $\Delta_n$  of 3 MeV for the daughter nucleus which we said would be derived in this appendix. In sect. 3 we found a maximum neutrino energy  $\Delta = 6.9 \text{ MeV}$ . Looking at the derivation there, one finds that protons in the Fermi sea to within a distance of  $\sim \frac{1}{3}\Delta$  will contribute. In the case of the Fermi sea, this means that a fraction

$$f = \frac{1}{3} \Delta \frac{k_f m / \pi^2}{2(\frac{4}{3}\pi k_f^3 / 2\pi)^3} = \frac{3}{5} \frac{\Delta}{k_f^2 / m}, \quad (\text{B.1})$$

will contribute. For  $\Delta = 6.9 \text{ MeV}$ ,  $k_F^2/m = 80 \text{ MeV}$ , a fraction  $f = 0.05$  is implied.

Now in the shell model,  $^{56}\text{Fe}$  has six  $f_{7/2}$  protons, all of which can easily participate. The main Fermi transition to the  $f_{7/2}$  neutron orbitals is blocked, because these are full, but the Gamow-Teller transition to the  $f_{5/2}$  neutron orbital is allowed. The Fermi transition involves  $g_V$ ; the Gamow-Teller  $g_A \sigma \tau_-$ , with  $\tau_- = \sqrt{\frac{1}{2}}(\tau_1 - i\tau_2)$ , and  $g_A \approx 1.25$ . The total transition strength will look like

$$S = g_V^2 \sum (f_{7/2}|1|n)(n|1|f_{7/2}) + g_A^2 \sum (f_{7/2}|\sigma \tau_-|n)(n|\sigma \tau_-|f_{7/2}). \quad (\text{B.2})$$

If all states  $|n\rangle$  were accessible, this sum would be  $S = g_V^2 + 3g_A^2$ , the 3 coming from the three spin orientations, and this sum would be the same as for an initial plane-wave state. As noted,  $|n\rangle = |f_{7/2}\rangle$  is blocked, which cuts out the Fermi transitions and the  $\sim \frac{1}{3}$  of the Gamow-Teller transitions which are not spin flip. Thus, the shell-model ratio to be compared with  $f$  of (B.1) is

$$f_{\text{SM}} \approx \frac{2g_A^2}{g_V^2 + 3g_A^2} \frac{6}{26} = 0.13, \quad (\text{B.3})$$

where  $\frac{6}{26}$  is the ratio of  $f_{7/2}$  protons to total number of protons. Thus, the available strength, because the protons are in the highly favorable  $f_{7/2}$  shell, is 2.6 times that in the Fermi gas model used in sect. 3. Of course, this shell will become partially depleted as the electron capture proceeds, but there will be a general development toward heavier nuclei, which adds some protons.

Note that we have not used the Fermi gas model to estimate the excitation energy of the daughter nucleus; it would give incorrect results. The transitions will be as shown in fig. 1, the  $f_{7/2}$  proton going to the  $f_{5/2}$  neutron orbital. Now, the  $f_{7/2}$  orbital is  $\sim 2 \text{ MeV}$  above the average of  $2p_{3/2}$  and  $2p_{1/2}$  orbitals in  $f_{7/2}$  shell nuclei.

Another effect comes into play. In the capture on  $^{56}\text{Fe}$ , a collective Gamow-Teller state,  $[f_{7/2,p}^{-1}, f_{7/2,n}]_{T=1}^{1+}$  will be formed. The particle-hole interaction between neutron particle and proton hole is expected to push this state up in energy. Many learned theoretical works have been written about the nuclear forces in this channel, which is also the channel relevant to pion condensation, but here we prefer to look at experiment. In  $^{48}\text{Ca}$ , where all neutron  $f_{7/2}$  levels are filled, the reaction  $^{48}\text{Ca}(^3\text{He}, ^3\text{H})^{48}\text{Sc}$  populates preferentially the proton  $f_{7/2}$  and  $f_{5/2}$  states. States ascribed to  $[f_{7/2,n}^{-1}f_{7/2,p}]^{1+}$  and  $[f_{7/2,n}^{-1}f_{5/2,p}]^{1+}$  are seen at  $\sim 2$  and  $10 \text{ MeV}$  excitation energy in the final nucleus<sup>36</sup>). The spin-orbit splitting between the  $f_{7/2}$  and  $f_{5/2}$  shells is  $\sim 6 \text{ MeV}$ , so it can be seen that the latter state has been moved up  $\sim 2 \text{ MeV}$  by the particle-hole forces. This

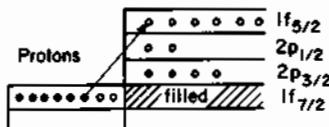


Fig. 1. Shell-model description of the electron capture. In the capture, protons will go to the  $f_{7/2}$  level of the daughter nucleus, which then decays by  $\gamma$ -emission down to the  $2p$  orbitals.

shift involves 8 possible neutron holes times 14 possible proton particles, giving a high degeneracy for the possible particle-hole states. Our initial degeneracy in  $^{56}\text{Fe}$  is 6 proton holes times 6 neutron particles so adding  $\sim 1$  MeV for the collective shift to the  $\sim 2$  MeV for the  $f_{\frac{1}{2}} - 2p$  shell splitting should suffice. In this way we end up at an estimated  $\Delta_n = 3$  MeV.

Had we estimated electron capture rates from nuclear beta decays, and this seems to have been done in most of the literature, we would have obtained much lower rates. Let us consider the  $\beta$ -decay of  $^{49}\text{Ca}$  to  $^{49}\text{Sc}$ , by way of example. The  $J = \frac{3}{2}$  ground state of  $^{49}\text{Ca}$  decays preferably to the  $J = \frac{3}{2}$  excited state of  $^{49}\text{Sc}$ . The shell model would describe both nuclei as  $^{48}\text{Ca}$  plus one  $p_{\frac{1}{2}}$  particle. If the  $\beta$ -decay is then calculated as a single-particle transition of  $p_{\frac{1}{2}}$  neutron to  $p_{\frac{1}{2}}$  proton, the rate is a factor of 30 larger than that found experimentally. Calculation<sup>37)</sup> gives a hindrance factor of 8; this retardation is due to admixture of configurations  $[p_{\frac{3}{2},n}(f_{\frac{3}{2},p}f_{\frac{3}{2},n}^{-1})^{1+}]_{\frac{3}{2}}$  and  $[p_{\frac{3}{2},n}(f_{\frac{3}{2},p}f_{\frac{3}{2},n}^{-1})^{1+}]_{\frac{1}{2}}$ . The admixed configurations are just the components which make up the giant Gamow-Teller collective state. This admixture, through the residual nucleon-nucleon forces, produces a diamagnetism which cuts down strongly matrix elements of  $\sigma$ . Whether one obtains a factor of 8 or of 30 is a delicate matter, since the cancellation is large in either case, and the effective nucleon-nucleon interaction is not known to the necessary accuracy to calculate precisely the small remaining amplitude.

Had we taken the transition  $^{49}\text{Ca} \rightarrow ^{49}\text{Sc}$  to estimate the empirical rate of  $\beta$ -decay and used it in our electron capture calculations, we would have underestimated the rate by a factor of  $\sim 180$  for  $^{56}\text{Fe}$ , a factor of 30 as explained above, and a factor of 6 because all 6  $f_{\frac{1}{2}}$  protons can make the transition to the  $f_{\frac{1}{2}}$  shell-model level in  $^{56}\text{Ca}$ . (The factor of 30 will not be universal, factors of 10 to 100 being typical of the hindrance.)

What we are saying in shell-model language is that ground-state to ground-state or low excited-state transitions are hindered through admixture of the giant Gamow-Teller state, whereas in the electron capture process the transition occurs to the giant Gamow-Teller state itself.

In its role of hindering low-lying  $\beta$ -transitions, the Gamow-Teller state is analogous to the giant dipole state, admixture of which retards low-lying E1 transitions by factors of  $\sim 100$  to 1000. On the other hand, the Gamow-Teller strength has not been found in one relatively concentrated state, as in the case of the giant dipole resonance. This should not matter for the discussion here, since the strength, even if fragmented, should play the same role in hindering low-lying transitions.

In the work of Epstein and Arnett<sup>38)</sup> capture on free protons competed favorably with capture on nuclei, despite the paucity of the former. Whereas the  $X$  relevant for free protons (in this case, just a kinematical factor) is

$$X_{t.p.} = \frac{3}{5} \frac{\mu_e^2}{m_e^2}, \quad (\text{B.4})$$

using (13) and (14) we find the ratio of  $X_{t.p.}$  to the  $X$  for nuclei, eq. (7.2), to be

$$\frac{X_{t.p.}}{X} = 0.99 Y_e^{\frac{3}{2}} \mu_e^{\frac{1}{2}}, \quad (\text{B.5})$$

with  $\mu_e$  in MeV.

For Arnett's <sup>7)</sup> zone 13 ( $\rho = 5.9 \times 10^{11} \text{ g/cm}^3$ ,  $Y_e = 0.356$ ,  $\mu_e = 30.6 \text{ MeV}$ ) in the region just before trapping, this ratio is 1089. From table 10 we see that  $T = 1.51 \text{ MeV}$  for this zone, and from table 1 that  $\hat{\mu} = 19 \text{ MeV}$  for this value of  $Y_e$ ; using (25.2) this gives a ratio

$$n_p/n_n = 2.5 \times 10^{-6}, \quad (\text{B.6})$$

implying  $X_p = 3 \times 10^{-7}$  for  $X_n = 0.1$ . Even with the factor of  $\sim 1000$  retardation found from eq. (B.5), this means that capture on nuclei predominates by a factor of  $3 \times 10^3$  in this zone.

Comparison with table 2 of Mazurek, Truran and Cameron <sup>39)</sup> shows their electron capture rates for nuclei in the region of <sup>56</sup>Fe to be  $\sim 100\text{--}300$  times smaller than ours. Their capture rate on free protons is identical to that given by our eq. (6) (which is slightly larger than that used by Hansen <sup>40)</sup>). Our  $Y_e$ , table 3, are not very different from those of Arnett <sup>7)</sup>, who used the results of Epstein and Arnett <sup>38)</sup> which were based on the rates of Mazurek *et al.* <sup>39)</sup>. This can be understood by the compensation between our large electron capture matrix elements and our large retardation factor arising from the near tracking of  $\mu_e$  by  $\hat{\mu}$ .

Note that our net capture rates and, hence, the  $Y_e$  in table 2 are nearly independent of our input matrix elements. Were we to decrease the numerical factor 2.12 in eq. (8.1) by a factor of 10, then  $\Delta$  as determined by eq. (13.2), would change by a factor of  $(10)^{\frac{1}{2}}$ . This change in  $\Delta$  changes  $X$  in such a way as to compensate for the change in numerical factor, so that the net capture rate does not change much. Such a change in capture rate would, however, change the maximum neutrino energy, which goes as  $\Delta$ .

It should be noted that our low entropy increase in the capture, calculated in sect. 2, does not depend upon the capture proceeding on nuclei, that it would be even lower, were the capture to proceed on free protons. In the latter case, the neutrino is emitted with average energy  $\frac{5}{6} \mu_e$ , a result that follows immediately from phase space arguments and from the fact that the neutron, being heavy, takes up only a negligible part of the electron energy. In terms of our argument, eq. (3) and following, this means that the energy  $\frac{1}{6} \mu_e$  is available for heating the system. On the other hand, as noted in the footnote in sect. 3, free protons used up in the capture have to be replenished from the nuclei, the neutrons formed in electron capture becoming attached to nuclei, and this costs energy  $\hat{\mu} = \mu_n - \mu_p$  per capture, cooling the nuclei. (This is just the familiar process of cooling by evaporation.)

Assuming the neutrinos to escape freely, which they do in the regions of low density where, because of the low temperature, the entropy increase is greatest, this means a net *decrease* in entropy of

$$\left(\frac{\delta S}{k}\right)_{\text{per capture}} = \frac{\frac{1}{6}\mu_e - \hat{\mu}}{kT}. \quad (\text{B.7})$$

(Note also eq. (F.1).) Again for Arnett's zone 13,  $\mu_e = 30.6$  MeV,  $\hat{\mu} = 19$  MeV, and taking a temperature of 1.5 MeV as we did in sect. 2, this gives

$$(\delta S/k)_{\text{per capture}} = -9.3, \quad (\text{B.8})$$

in magnitude more than double the increase in entropy per capture on nuclei, estimated in sect. 2. There is, of course, no reason why the entropy should not decrease, because the system is not closed, with neutrinos escaping. Of course, a decrease in entropy will lower the temperature, shutting the process off since, according to (25.2), the ratio of free protons to free neutrons goes as  $\exp(-\hat{\mu}/kT)$ .

## Appendix C

### DIFFUSION OF NEUTRINOS

The neutrino density  $n_\nu$  is high in the center, and falls off toward the outside. Diffusion may take place, and will be governed by the diffusion equation:

$$\frac{\partial}{\partial t} n_\nu = \frac{1}{r^2} \frac{\partial}{\partial r} \left( r^2 \frac{1}{3} c \lambda_\nu \frac{\partial n_\nu}{\partial r} \right). \quad (\text{C.1})$$

For  $\lambda_\nu$ , we use eq. (16); in this we replace  $\varepsilon_\nu$  by  $\mu_\nu$  because effectively only the neutrinos near the top of the Fermi sea can diffuse, and we write:

$$\lambda_\nu = \lambda_0 \left( \frac{\mu_\nu}{10 \text{ MeV}} \right)^{-2}, \quad (\text{C.2})$$

$$\lambda_0 = 1.0 \times 10^6 \text{ cm} \left( \frac{\rho}{10^{12} \text{ g/cm}^3} \right)^{-1} (0.12 X_h \bar{N} + X_n)^{-1}, \quad (\text{C.3})$$

where  $\bar{N}$  is the average number of neutrons in the heavy nucleus. We introduce

$$\mu' = \frac{\mu_\nu}{10 \text{ MeV}}. \quad (\text{C.4})$$

Since  $n_\nu$  is proportional to  $\mu_\nu^3$ , we replace it by  $\mu'^3$  in (C.1) and have,

$$\frac{\partial}{\partial t} (\mu'^3) = \frac{1}{3r^2} \frac{\partial}{\partial r} \left( r^2 c \lambda_0 \mu'^{-2} \frac{\partial}{\partial r} (\mu'^3) \right). \quad (\text{C.5})$$

We shall consider  $\lambda_0$  as constant, although in reality  $\rho$ , the matter density, varies with  $r$ . Then:

$$\mu'^2 \frac{\partial \mu'}{\partial t} = \frac{1}{2} c \lambda_0 \frac{1}{r^2} \frac{\partial}{\partial r} \left( r^2 \frac{\partial \mu'}{\partial r} \right). \quad (\text{C.6})$$

A particular solution can be obtained by separation of variables; write:

$$\mu' = \mu'_0 \psi(r) \phi(t) \quad (\text{C.7})$$

Then (C.6) becomes

$$\psi^3 \phi^2 \frac{d\phi}{dt} = \frac{c \lambda_0}{3 \mu'_0} \frac{\phi}{r^2} \frac{d}{dr} \left( r^2 \frac{d\psi}{dr} \right) \quad (\text{C.8})$$

We divide by  $\psi^3 \phi$  and set each side of the equation equal to a separation constant which we call  $-\frac{1}{2}a$ , thus:

$$\frac{\phi}{dt} \frac{d\phi}{dt} = -\frac{1}{2}a, \quad (\text{C.9})$$

$$\frac{1}{r^2} \frac{d}{dr} \left( r^2 \frac{d\psi}{dr} \right) = -\frac{3 \mu'_0^2 a}{2 \lambda_0 c} \psi^3. \quad (\text{C.10})$$

(C.9) has the solution,

$$\phi = (1 - at)^{\frac{1}{2}}. \quad (\text{C.11})$$

(C.10) is of the type of the Lane-Emden equation of a polytrope. We define:

$$x = \left( \frac{3 \mu'_0^2 a}{2 \lambda_0 c} \right)^{\frac{1}{3}} r, \quad (\text{C.12})$$

and obtain,

$$\frac{1}{x^2} \frac{d}{dx} \left( x^2 \frac{d\psi}{dx} \right) = -\psi^3. \quad (\text{C.13})$$

This is the Lane-Emden equation<sup>41)</sup> of index 3 whose solution is known. The separation constant is determined by the condition that  $\mu_r$  (and hence  $\psi$ ) must be 0 at some radius  $R$ , i.e. the radius at which neutrino trapping occurs. The zero of the Lane-Emden function of order 3 is at  $x = 6.9$ ; hence from (C.12):

$$a = \frac{2 \lambda_0 c}{3 \mu'_0^2} \left( \frac{6.9}{R} \right)^2. \quad (\text{C.14})$$

For the Lane-Emden equation of order 3, the ratio of the central to the average density is,

$$\frac{n_c}{n_{Av}} = \left( \frac{\mu_0}{\mu_{Av}} \right)^3 = 54.2, \quad (\text{C.15})$$

i.e. there is strong central condensation.

Typical values<sup>7)</sup>, relating to the middle of the region of trapped neutrinos (Arnett's mass point 9) are†:

$$\rho = 5.9 \times 10^{12}, \quad R = 3.2 \times 10^6, \mu_\nu = 35 \text{ MeV}, \quad \bar{N} = 85. \quad (\text{C.16})$$

We take this value of 35 MeV to be  $\mu_{\text{Av}}$ ; then (C.15) gives

$$\mu_0 = 133 \text{ MeV}, \quad \mu'_0 = 13.3. \quad (\text{C.17})$$

This is close to the central  $\mu_\nu$  we calculated in (18.1). Inserting in (C.3), we find  $\lambda_0 = 1.7 \times 10^4 \text{ cm}$ , and

$$a = 9 \text{ sec}^{-1}. \quad (\text{C.18})$$

So it takes about  $\frac{1}{9}$  sec for the neutrinos to get out of the material. This is many times longer than the time scale of the hydrodynamics, cf. eq. (10). This shows that the neutrino diffusion, in the bulk of the core, is negligible during the times relevant to the hydrodynamics, as stated at the end of sect. 3. It may be noted that  $\lambda_0/R^2$  in (C.14) is proportional to  $1/\rho R^2$  which is not strongly dependent on  $\rho$ , thus the time scale (C.18) will not change much if the star is further compressed or expanded.

It is likely that the radial distribution of  $\mu_\nu$  is not as peaked at the center as the Lane-Emden function,  $\psi$ , eq. (C.13). In this case, (C.6) makes it likely that  $\mu'$  decreases fastest with time where  $\mu'$  is small, i.e. near the surface of the region of trapped neutrinos. Therefore we may expect that this surface gradually moves inward, relative to the matter; there is a "neutrino rarefaction wave" moving inward. This result has also been obtained by Lichtenstein *et al.*<sup>13)</sup>.

## Appendix D

### TRAPPING CRITERION

The diffusion theory which we discussed in appendix C gives an expression for the radial current of neutrinos,

$$S = -\frac{1}{3}c\lambda_\nu \frac{\partial n_\nu}{\partial r}. \quad (\text{D.1})$$

We use (C.2) for  $\lambda_\nu$  and set

$$n_\nu = K\mu_\nu^3, \quad K = \frac{1}{6\pi^2(\hbar c)^3}. \quad (\text{D.2})$$

then

$$S = -100c\lambda_0 K \frac{\partial \mu_\nu}{\partial r}. \quad (\text{D.3})$$

† For  $R$ , we have chosen  $2^{\frac{1}{3}}$  times the radius of mass point 9, which would be the radius of the trapped-neutrino sphere if the density were constant at the value in zone 9.

We may define a drift velocity of the neutrinos

$$v_d = \frac{S}{n_\nu}; \quad (D.4)$$

it gives the average velocity of the neutrinos relative to the nuclear material which does the scattering;  $v_d$  is positive which indicates a relative motion outward. Inserting (D.3), (D.2), (C.2),

$$v_d = -100 \frac{c\lambda_1}{\rho_{12}\mu_\nu^3} \frac{\partial\mu_\nu}{\partial r}, \quad (D.5)$$

where  $\lambda_1$  is the value of  $\lambda_0$ , (C.3), for  $\rho = 10^{12}$ , and  $\rho_{12}$  the density in units of  $10^{12} \text{ g/cm}^3$ ; with  $X_b = 1$ ,  $\bar{N} = 72$ , we have

$$\lambda_1 = 1.15 \text{ km.} \quad (D.6)$$

Table 8 gives the relevant quantities in (D.5), using Arnett's Tables I and II. It is possible that the values of  $\mu_\nu$  should actually be somewhat lower than those given by Arnett because Arnett obtains too low values for  $Y_e$  for high densities (as compared to ours), indicating that he assumes too much transformation of electrons into neutrinos and therefore too high a neutrino density. However, Arnett's is the best calculation available, and his numbers should at least indicate the trend.

TABLE 8  
Neutrino drift velocity

Zone* $k$	$-\partial\mu_\nu/\partial r$ (MeV/km)	$\rho_{12}$	$(\mu_\nu/10)^3$ (MeV)	$v_d$ ( $10^8 \text{ cm/sec}$ )	$-u_k$ ( $10^8 \text{ cm/sec}$ )
8	1.15	7.9	61	0.08	5.9
9	1.25	5.9	43	0.17	7.4
10	1.02	3.6	24	0.41	9.8
11	1.17	2.25	11	1.66	12.5
12	0.82	1.35	4.9	4.3	15.2
13	0.35	0.59	1.73	11.9	18.2
14	0.162	0.214	0.50	52	20.2

\* Arnett denotes by  $k$  the zone between his mass points  $k-1$  and  $k$ .  
 $v_d$  is the neutrino drift velocity,  $u_k$  the material velocity.

Table 8 shows that  $\partial\mu_\nu/\partial r$  is nearly constant up to zone 11, then drops fairly rapidly. Both  $\rho_{12}$  and  $\mu_\nu^3 \approx n_\nu$  decrease rapidly toward the outside. Therefore, the drift velocity  $v_d$ , calculated from (D.5), increases very rapidly with zone number  $k$  (for  $k = 14$ , the diffusion theory underlying (D.5) may no longer be valid). The

velocity of the nuclear matter,  $u_k$ , changes only slowly by comparison. Therefore, the net velocity of the neutrinos relative to the star center,

$$\dot{r}_\nu = u + v_d, \quad (16.5)$$

goes through zero with a big slope. The point  $\dot{r}_\nu = 0$  can thus be determined with high accuracy; it lies between zones 13 and 14, at a density

$$\rho_{12} \approx 0.43. \quad (D.7)$$

If Arnett's values for  $\mu_\nu$  are decreased, as proposed above,  $v_d$  will be increased, and hence the point  $\dot{r}_\nu = 0$  will be shifted to somewhat higher density, but we do not believe that this shift can be as much as a full zone. Thus we think it is safe to say that the critical density for trapping,

$$\rho_{\text{trap}} < 10^{12} \text{ g/cm}^3. \quad (D.8)$$

## Appendix E

### DISSOCIATION OF NEUTRON-RICH NUCLEI

In order to see how neutron richness affects the breakup of nuclei, we discuss fictitious  $A_H = 56$  nuclei, with  $Y_e = 0.393$  and  $Y_e = 0.357$ , corresponding to 22 and 20 protons, respectively. We take  $\rho = 5.9 \times 10^{11} \text{ g/cm}^3$  so that we can compare with the  $^{56}\text{Fe}$  breakup, top of table 3.

In order to calculate the  $Q$ -value for the breakup, we must calculate the lowering of the binding energies of our  $A_H = 56$  nuclei with  $Y_e = 0.393$  and  $Y_e = 0.357$  relative to that of  $^{56}\text{Fe}$ . We use Epstein's formula (4.3) obtained from empirical binding energies in this region of  $A$ .

The binding energy of the  $A = 56$  nucleus is then lowered by

$$\delta E = 56 (0.46 - Y_e)^2 \frac{1}{2}(144 \text{ MeV}). \quad (E.1)$$

We have approximated here the  $Y_e = 0.464$  appropriate for  $^{56}\text{Fe}$  by 0.46. On the other hand, there will be a decrease in the number of  $\alpha$ -particles in the breakup and, therefore, a decrease in the binding energy of the dissociated particles of

$$\delta E_\alpha = \frac{56}{2}(0.46 - Y_e) 28.3 \text{ MeV}. \quad (E.2)$$

This leads to an increase in  $Q$ -value for the breakup of

$$\delta Q = \delta E_\alpha - \delta E = \frac{56}{2} \{28.3 - (0.46 - Y_e)144 \text{ MeV}\}(0.46 - Y_e). \quad (E.3)$$

For  $Y_e = 0.393$ ,  $\delta Q = 35 \text{ MeV}$ , so that our  $Q$ -value is

$$Q = 124 + 35 = 159 \text{ MeV}, \quad (E.4)$$

and for  $Y_e = 0.357$ ,

$$Q = 163 \text{ MeV.} \quad (\text{E.5})$$

We next carried out calculations using eq. (23), appropriately changed for 11  $\alpha$ -particles and 12 neutrons for the case of  $Y_e = 0.393$  and for 10  $\alpha$ -particles and 16 neutrons for  $Y_e = 0.357$ . The results are given in the lower part of table 3. These entropies should be increased  $\sim 50\%$  because of neutron evaporation from nuclei

$$Z^A \neq Z^{A-1} + n$$

according to

$$\frac{n_{A-1}n_n}{n_A} = \frac{\Phi_{A-1}}{\Phi_A} 2 \left( \frac{mkT}{2\pi\hbar^2} \right)^{\frac{3}{2}} e^{-Q_n/kT}, \quad (\text{E.6})$$

where  $Q_n$  is the neutron binding energy,  $\geq 5$  MeV in these neutron rich nuclei with  $Y_e = 0.35-0.40$ . There will be many atomic species in the medium, but their abundance will clearly have a maximum at some value of  $Z$  and  $A$ . Near that maximum,  $n_{A-1} = n_A$ ; assuming then also  $\Phi_{A-1} = \Phi_A$ , we can easily calculate  $X_n$  from (E.6).

To this entropy due to nuclear breakup we should add two terms which we calculate for  $\delta = 0.1$  and  $Y_e = 0.393$  and 0.357, hence for temperatures of 1.40 and 1.35 MeV (see table 3). They are: (a) the entropy due to the excitation of the nuclei, using (25.9) which gives 0.35 or 0.34, respectively, and (b) the entropy in the electrons, 0.17 or 0.16; thus, we end up with a total entropy of 1.32 or 1.33, just about right. Our temperatures of 1.40 and 1.35 MeV are substantially lower than Arnett's 1.90 MeV for this density; we shall come back to a discussion of this later.

From our results in table 3 for the entropy in translational motion, we see that for a given  $\delta$ , the entropy is increased as  $Y_e$  is lowered. Thus, for a given entropy,  $\delta$  must be lowered. The entropy increase comes from the larger number of neutrons in the breakup. Already when  $Y_e$  has decreased to 0.393 we see relatively little breakup with  $\alpha$ -particles ( $\delta \approx 0.1$  as compared with  $\delta \approx 0.2$  for  $^{56}\text{Fe}$ ).

Arnett<sup>7</sup>), with  $X_h = 0.57$ ,  $X_n = 0.10$  and  $X_\alpha = 0.33$  and  $kT = 1.90$  MeV for  $\rho = 5.9 \times 10^{11} \text{ g/cm}^3$  has a greater breakup, substantially more  $\alpha$ -particles and a higher temperature. Arnett neglected the entropy from the level density in nuclei; had he included it, less entropy would have been in translational motion and his temperature would have been lower, although still somewhat above ours. Our eqs. (25.5) and (25.6) show that Arnett's effective  $\mu$  was  $\sim 7$  MeV, much too small for his  $Y_e = 0.356$  where 19 MeV would be more appropriate.

From our model calculations in this appendix, we see then how the increased number of neutrons in the breakup of neutron-rich nuclei inhibits, for a given entropy, that breakup. Even if  $Y_e$  remained at 0.464, the value appropriate for  $^{56}\text{Fe}$ ,  $\alpha$ -particles would tend to go back into nuclei with increasing density  $\rho$ . Our considerations here show this tendency to be much more pronounced for smaller values of  $Y_e$ . The role of the  $\alpha$ -particle is very minor.

### Appendix F

#### ELECTRON NUMBERS, ENTROPIES AND TEMPERATURES, AS FUNCTIONS OF DENSITY $\rho$ .

In this appendix we present results of an explicit integration of eq. (11). To determine also the temperature, we use the first law of thermodynamics to follow the entropy change,

$$T \left( \frac{dS}{d \ln \rho} \right)_{\text{per nucleon}} = - \frac{d Y_e}{d \ln \rho} (\mu_e - \mu_n + \mu_p - 1.297 - \bar{\epsilon}_\nu), \quad (\text{F.1})$$

where 1.297 is the neutron-proton mass difference and  $\bar{\epsilon}_\nu$  is the average energy of the neutrino created in electron capture. This equation is valid only when the matter is completely transparent to neutrinos. Instead of using eq. (17.2) for  $X$  in eq. (11), eq. (7) has been integrated without the assumption that  $\varepsilon_e = \mu_e$ , but rather that

$$\varepsilon_e = \varepsilon_\nu + \hat{\mu} + \Delta_n + \Delta_p + 1.297. \quad (\text{F.2})$$

Thus, we use here

$$X = \frac{1}{6} \frac{(\mu_e - Q)^4}{m_e^2 (k_f^2 / 3m) \mu_e} \left( 1 + \frac{2}{5} \frac{Q}{\mu_e} + \frac{1}{10} \frac{Q^2}{\mu_e^2} \right), \quad (\text{F.3})$$

where  $Q = \hat{\mu} + \Delta_n + 1.297$ . To find  $\bar{\epsilon}_\nu$ , we replace  $\varepsilon_\nu^2$  in eq. (7) by  $\varepsilon_\nu^3$  and call the integral  $Y$ :

$$Y = \frac{3}{28} \frac{(\mu_e - Q)^5}{m_e^2 (k_f^2 / 3m) \mu_e} \left( 1 + \frac{1}{3} \frac{Q}{\mu_e} + \frac{1}{15} \frac{Q^2}{\mu_e^2} \right). \quad (\text{F.4})$$

Then,

$$\bar{\epsilon}_\nu = \frac{Y}{X}. \quad (\text{F.5})$$

As  $\hat{\mu}$  follows  $\mu_e$  more and more closely, one then finds that  $\bar{\epsilon}_\nu \approx \frac{2}{3}(\mu_e - Q)$ , but we do not assume this here.

Eqs. (11) and (F.1) have been integrated from the initial conditions  $\rho = 3.7 \times 10^9 \text{ g/cm}^3$ ,  $Y_e = 0.42$  and  $T = 0.69 \text{ MeV}$  up to  $\rho = 10^{12} \text{ g/cm}^3$ , using  $\Delta_n = 3 \text{ MeV}$ . We first consider the case in which dripped neutrons are neglected. The entropy per nucleon  $S$  is the sum of electron (eq.(2.2)), nuclear translational (eq. (2)) and internal nuclear (eq. (25.9)) terms. Results are given in table 9.

We see for conditions near Arnett's zone 12,  $\rho = 1.35 \times 10^{12} \text{ g/cm}^3$  that  $Y_e \approx 0.32$ . This is slightly higher than the  $Y_e^{(0)} = 0.31$  derived in eq. (15.2) and given in table 2 because the righthand side of eq. (F.3) is always less than that of eq. (7.2). Also, the temperature is quite high near  $10^{12} \text{ gm/cm}^3$  because the entropy of dripped neutrons has been neglected. In turn, the entropy generation is too small because of the high temperature.

TABLE 9  
Electron numbers, temperatures and entropies, calculated without inclusion of neutron drip

$\rho_{10}$	$Y_e$	T(MeV)	S/k	$\hat{\mu}$ (MeV)	$\mu_e$ (MeV)	$A$ (MeV)
0.37	0.420	0.69	0.905	4.38	5.96	
0.50	0.420	0.75	0.905	4.42	6.59	
0.67	0.420	0.81	0.905	4.44	7.28	
0.91	0.420	0.87	0.905	4.47	8.04	
1.22	0.420	0.93	0.905	4.49	8.88	0.09
1.65	0.420	1.00	0.905	4.53	9.81	0.99
2.22	0.420	1.07	0.905	4.57	10.83	1.97
2.99	0.420	1.15	0.906	4.65	11.97	3.03
4.03	0.419	1.23	0.909	4.86	13.21	4.06
5.43	0.417	1.33	0.918	5.35	14.56	4.92
7.32	0.412	1.45	0.933	6.23	16.03	5.51
9.87	0.406	1.59	0.954	7.48	17.62	5.84
13.30	0.399	1.74	0.977	9.03	19.35	6.02
17.93	0.391	1.91	1.002	10.79	21.22	6.13
24.16	0.382	2.08	1.026	12.75	23.26	6.21
32.57	0.372	2.26	1.051	14.88	25.47	6.29
43.90	0.362	2.45	1.075	17.21	27.88	6.37
59.17	0.351	2.64	1.098	19.73	30.48	6.45
79.75	0.339	2.83	1.122	22.46	33.29	6.53
107.5	0.327	3.02	1.145	25.40	36.33	6.63
131.2	0.318	3.16	1.160	27.49	38.49	6.70

Entropy is given as entropy per nucleon.

We now wish to investigate the effects of neutron drip on these results. Eqs. (11) and (F.1)–(F.5) are still valid, as is eq. (14.1) for  $\mu_e$ , but eq. (4.4) for  $\hat{\mu}$ , eq. (5) for  $A$  and eq. (5.3) for  $\mu_n$  must be modified by replacing  $Y_e$  in them by  $x = Z/A$ , related to  $Y_e$  by

$$x = Y_e/(1 - X_n). \quad (\text{F.5.1})$$

Since the drip neutrons are essentially free in the density range of interest, one can make use of eq. (25.6) to find  $X_n$ :

$$X_n = \frac{n_n}{\rho} = 79 \frac{T^{\frac{3}{2}}}{\rho_{10}} e^{\mu_n/T}, \quad (\text{F.6})$$

with  $T$  measured in MeV. Alpha particles were included, but they turn out to be negligible throughout these calculations (see eq. (25.5)). Nuclear entropy was included according to eq. (25.9). The entropy of dripped neutrons is included using a formula similar to that for nuclei, eq. (2):

$$\left(\frac{S}{k}\right)_{\text{per dripped nucleon}} = \frac{1}{2} + \ln \left[ \frac{2V}{N_n} \left( \frac{mkT}{2\pi\hbar^2} \right)^{\frac{3}{2}} \right]. \quad (\text{F.6.1})$$

TABLE 10  
Electron numbers, temperatures and entropies, calculated with neutron drip

$\rho_{10}$	$Y_e$	$T(\text{MeV})$	$X_n$	$S/k$	$\hat{\mu}(\text{MeV})$	$\mu_n(\text{MeV})$	$\Delta(\text{MeV})$
0.37	0.420	0.69	1.7(-4)	0.905	4.38	-9.30	
0.50	0.420	0.74	3.6(-4)	0.905	4.39	-9.30	
0.67	0.420	0.80	6.9(-4)	0.905	4.39	-9.31	
0.91	0.420	0.85	1.21(-3)	0.905	4.37	-9.32	
1.22	0.420	0.90	1.81(-3)	0.905	4.35	-9.34	0.23
1.65	0.420	0.96	2.5(-3)	0.905	4.33	-9.36	1.18
2.22	0.420	1.01	3.4(-3)	0.905	4.31	-9.37	2.22
2.99	0.420	1.07	4.4(-3)	0.906	4.35	-9.37	3.32
4.03	0.418	1.12	5.9(-3)	0.910	4.54	-9.30	4.37
5.43	0.415	1.18	8.7(-3)	0.922	5.01	-9.11	5.23
7.32	0.409	1.24	0.0127	0.943	5.84	-8.77	5.86
9.87	0.401	1.30	0.0192	0.973	6.98	-8.31	6.27
13.30	0.391	1.34	0.0283	1.012	8.36	-7.77	6.56
17.93	0.379	1.38	0.0402	1.057	9.92	-7.16	6.80
24.16	0.366	1.42	0.0553	1.109	11.61	-6.52	7.02
32.57	0.351	1.45	0.0742	1.167	13.43	-5.84	7.25
43.90	0.334	1.47	0.0978	1.233	15.35	-5.15	7.50
59.17	0.315	1.51	0.127	1.307	17.33	-4.46	7.79
79.75	0.295	1.54	0.165	1.390	19.34	-3.77	8.13
107.5	0.271	1.60	0.211	1.481	21.33	-3.12	8.52
131.2	0.255	1.65	0.248	1.547	22.61	-2.71	8.82

Entropy is given as entropy per nucleon.

The integrations including neutron drip are displayed in table 10. At  $\rho = 10^{12} \text{ g/cm}^3$ ,  $Y_e^{(t)} \approx 0.28$ , so that the effect of drip neutrons is small, even though  $X_n \approx 0.2$  here. The partial dissociations of nuclei acts as a thermostat to keep the temperature low. Because  $x > Y_e$ ,  $\hat{\mu}$  is less in this case and neutrino energy is higher in table 10 than in table 9. As a result, the entropy generation is more than doubled. Even so, the entropy does not exceed  $\approx 1.5$  per nucleon, in agreement with our earlier discussion.

In the above calculation of entropy production we have neglected contributions from downscattering in energy of the emitted neutrinos. As noted above, the average neutrino energy is  $\bar{\epsilon}_\nu = \frac{3}{5} \Delta$  so that a downscattered neutrino which deposits all of its energy in the system would increase the entropy by

$$\delta S = \bar{\epsilon}_\nu / kT = \frac{3}{5} \Delta / kT. \quad (\text{F.7})$$

Neutrinos elastically scattered from nuclei will deposit only recoil energy

$$E_R \approx \bar{\epsilon}_\nu^2 / 2Am, \quad (\text{F.8})$$

where  $m$  is the nucleon mass, and this is negligibly small. Down-scattering in energy will proceed only through neutrino-electron scattering.

For the ratio of mean free paths for scattering off electrons to coherent scattering off nuclei we find

$$\frac{\lambda_e}{\lambda_{el}} = \frac{2}{Y_e} \left[ \frac{\bar{N}^2}{6A} X_h + X_n \right] \frac{\mu_e}{\varepsilon_\nu}, \quad (\text{F.9})$$

from eq. (14) of Lamb and Pethick<sup>16</sup>, with replacement through our eq. (16.3). The final factor ( $\mu_e/\varepsilon_\nu$ ) results from electron degeneracy. For Arnett's zone 12, this ratio is

$$\frac{\lambda_e}{\lambda_{el}} \approx 640, \quad (\text{F.10})$$

where we have used  $\bar{N} = 88$ ,  $\varepsilon_\nu = \bar{\varepsilon}_\nu = 4.15$  MeV,  $\mu_e = 39.1$  MeV,  $Y_e = 0.31$ . Since, as found earlier,  $\lambda_{el}/R \approx \frac{1}{10}$ , where  $R = 4.27 \times 10^6$  cm is the radius of this zone, the probability of neutrino-electron downscattering at this density looks small.

The possible importance of neutrino downscattering in increasing the entropy was pointed out to us by Don Lamb, and we would like to thank him for extensive discussions.

### References

- 1) T. J. Mazurek, *Astrophys. Space Sci.* **35** (1975) 117; *Astrophys. J. Lett.* **207** (1976) L87
- 2) K. Sato, *Prog. Theor. Phys.* **54** (1975) 1325
- 3) J. R. Wilson, *Astrophys. J.* **163** (1971) 209; *Phys. Rev. Lett.* **32** (1974) 849
- 4) S. Weinberg, *Phys. Rev. Lett.* **19** (1967) 1264
- 5) A. Salam, in *Elementary particle physics*, ed. N. Svartholm (Almqvist and Wiksell, Stockholm, 1968)
- 6) D. Z. Freedman, *Phys. Rev.* **D9** (1974) 1389
- 7) W. D. Arnett, *Astrophys. J.* **218** (1977) 815
- 8) T. A. Weaver, G. B. Zimmerman and S. E. Woosley, Lawrence Livermore Laboratory preprint (1978).
- 9) G. Baym, H. A. Bethe and C. J. Pethick, *Nucl. Phys.* **A175** (1971) 225
- 10) J. M. Lattimer and D. G. Ravenhall, *Astrophys. J.* **223** (1978) 314
- 11) R. I. Epstein, private communication
- 12) D. G. Ravenhall, C. D. Bennett and C. J. Pethick, *Phys. Rev. Lett.* **28** (1972) 978
- 13) I. Lichtenstein, A. Ron, N. Sach, J. J. Wagshal and S. A. Bludman, *Astrophys. J.* **226** (1978) 222
- 14) R. I. Epstein, H. Nørsgaard and J. R. Bond, NORDITA preprint, 1978
- 15) W. D. Arnett, private communication
- 16) D. Q. Lamb and C. J. Pethick, *Astrophys. J.* **209** (1976) L77
- 17) H. A. Bethe, *Rev. Mod. Phys.* **9** (1937) 69
- 18) A. Gilbert and A. G. W. Cameron, *Can. J. Phys.* **43** (1965) 1446
- 19) T. J. Mazurek, J. M. Lattimer and G. E. Brown, *Astrophys. J.*, to appear in 15 April 1979 issue
- 20) A. Bohr and B. R. Mottelson, *Nuclear Structure*, vol. I (Benjamin, New York, 1969)
- 21) G. E. Brown, J. H. Gunn and P. Gould, *Nucl. Phys.* **46** (1963) 598
- 22) G. F. Bertsch and T. T. S. Kuo, *Nucl. Phys.* **A112** (1968) 204;  
I. Hamamoto and P. Siemens, *Nucl. Phys.* **A269** (1976) 199;
- 23) W. Fowler and F. Hoyle, *Astrophys. J. Supp.* **91**, No. 9 (1974) 201
- 24) D. Q. Lamb, J. M. Lattimer, C. J. Pethick and D. G. Ravenhall, *Phys. Rev. Lett.* **41** (1978) 1623

- 25) J. P. Blaizot, D. Gogny and B. Grammaticos, Nucl. Phys. **A265** (1976) 315
- 26) A. B. Migdal, ZhETF (USSR) **61** (1971) 2210; JETP (Sov. Phys.) **34** (1972) 1184;  
R. F. Sawyer, Phys. Rev. Lett. **29** (1972) 382;  
D. J. Scalapino, Phys. Rev. Lett. **29** (1972) 386
- 27) G. E. Brown and W. Weise, Phys. Reports **27C** (1976) 1
- 28) W. F. Durand, Aerodynamic theory (Springer-Verlag, Berlin)
- 29) P. J. Siemens, Nucl. Phys. **A141** (1970) 225
- 30) P. J. Siemens, private communication
- 31) S.-O. Bäckman, private communication
- 32) A. D. Jackson, D.-O. Riska and B. Verwest, Nucl. Phys. **A249** (1975) 397
- 33) W. Myers and W. J. Swiatecki, Nucl. Phys. **81** (1966) 1
- 34) F. D. Mackie and G. Baym, Nucl. Phys. **A285** (1977) 332
- 35) F. Serduke, 1974 Atomic Mass Table, Argonne Natl. Lab
- 36) K. Gaarde, preliminary results
- 37) G. Bertsch and J. Damgaard, Phys. Lett. **23** (1966) 342
- 38) R. I. Epstein and W. D. Arnett, Astrophys. J. **201** (1975) 202
- 39) T. J. Mazurek, A. G. W. Cameron and J. Truran, Astrophys. Space Sci. **27** (1974) 261
- 40) C. J. Hansen, Astrophys. Space Sci. **1** (1968) 499
- 41) D. D. Clayton, Principles of stellar evolution and nucleosynthesis (McGraw-Hill, New York, 1968)

This page is intentionally left blank

## Chapter 2

# How a Supernova Explodes

H.A. Bethe and G.E. Brown

Reprinted with permission from *Scientific American*, May 1985, 252, 60-68  
and Ian Worpole

Copyright (1985) by Scientific American, Inc. All rights reserved.

Figure copyright of Ian Worpole.

### Commentary by G.E. Brown

Little did we know when we wrote "How a Supernova Explodes" for *Scientific American* in 1985 that an explosion in the Large Magellanic Cloud, essentially in our galaxy, would go off only two years later, the first "nearby one" since Kepler's in 1604. On the whole, our theory worked out pretty well.

This page is intentionally left blank

# How a Supernova Explodes

*When a large star runs out of nuclear fuel, the core collapses in milliseconds. The subsequent "bounce" of the core generates a shock wave so intense that it blows off most of the star's mass*

by Hans A. Bethe and Gerald Brown

The death of a large star is a sudden and violent event. The star evolves peacefully for millions of years, passing through various stages of development, but when it runs out of nuclear fuel, it collapses under its own weight in less than a second. The most important events in the collapse are over in milliseconds. What follows is a supernova, a prodigious explosion more powerful than any since the big bang with which the universe began.

A single exploding star can shine brighter than an entire galaxy of several billion stars. In the course of a few months it can give off as much light as the sun emits in a billion years. Furthermore, light and other forms of electromagnetic radiation represent only a small fraction of the total energy of the supernova. The kinetic energy of the exploding matter is 10 times greater. Still more energy—perhaps 100 times more than the electromagnetic emission—is carried away by the massless particles called neutrinos, most of which are emitted in a flash that lasts for about a second. When the explosion is over, most of the star's mass has been scattered into space, and all that remains at the center is a dense, dark cinder. In some cases even that may disappear into a black hole.

Such an outline description of a supernova could have been given almost 30 years ago, and yet the detailed sequence of events within the dying star is still not known with any certainty. The basic question is this: A supernova begins as a collapse, or implosion; how does it come about, then, that a major part of the star's mass is expelled? At some point the inward movement of stellar material must be stopped and then reversed; an implosion must be transformed into an explosion.

Through a combination of computer simulation and theoretical analysis a coherent view of the supernova mechanism is beginning to emerge. It

appears the crucial event in the turnaround is the formation of a shock wave that travels outward at 30,000 kilometers per second or more.

Supernovas are rare events. In our own galaxy just three have been recorded in the past 1,000 years; the brightest of these, noted by Chinese observers in 1054, gave rise to the expanding shell of gas now known as the Crab Nebula. If only such nearby events could be observed, little would be known about supernovas. Because they are so luminous, however, they can be detected even in distant galaxies, and 10 or more per year are now sighted by astronomers.

The first systematic observations of distant supernovas were made in the 1930's by Fritz Zwicky of the California Institute of Technology. About half of the supernovas Zwicky studied fitted a quite consistent pattern: the luminosity increased steadily for about three weeks and then declined gradually over a period of six months or more. He designated the explosions in this group Type I. The remaining supernovas were more varied, and Zwicky divided them into four groups; today, however, they are all grouped together as Type II. In Type I and Type II supernovas the events leading up to the explosion are thought to be quite different. Here we shall be concerned primarily with Type II supernovas.

The basis for the theory of supernova explosions was the work of Fred Hoyle of the University of Cambridge. The theory was then developed in a fundamental paper published in 1957 by E. Margaret Burbidge, Geoffrey R. Burbidge and William A. Fowler, all of Caltech, and Hoyle. They proposed that when a massive star reaches the end of its life, the stellar core collapses under the force of its own gravitation. The energy set free by the collapse expels most of the star's mass, distributing the chemical elements formed in

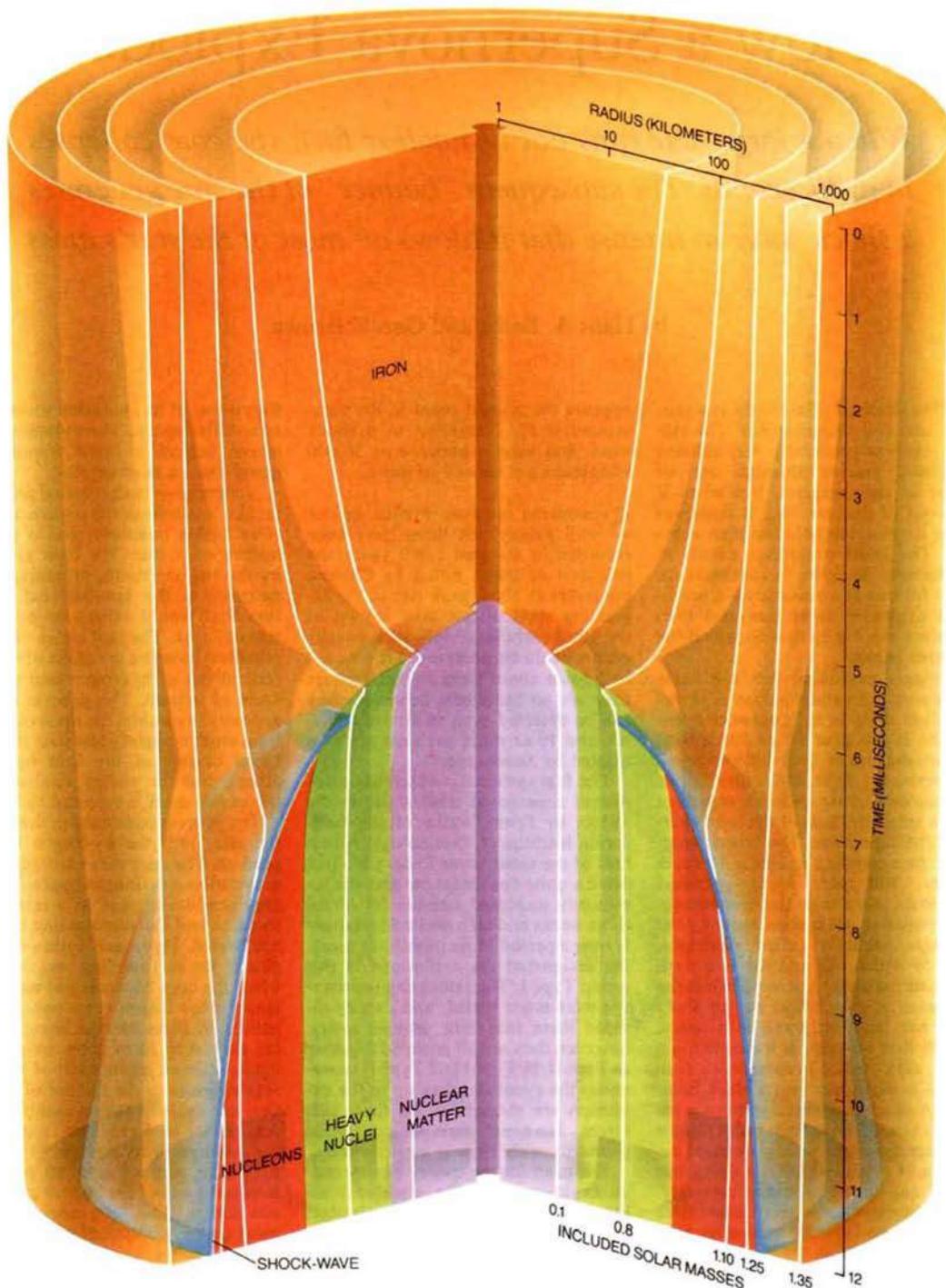
the course of its evolution throughout interstellar space. The collapsed core leaves behind a dense remnant, in many cases a neutron star.

A supernova is an unusual and spectacular outcome of the sequence of nuclear fusion reactions that is the life history of a star. The heat given off by the fusion creates pressure, which counteracts the gravitational attraction that would otherwise make the star collapse. The first series of fusion reactions have the net effect of welding four atoms of hydrogen into a single atom of helium. The process is energetically favorable: the mass of the helium atom is slightly less than the combined masses of the four hydrogen atoms, and the energy equivalent of the excess mass is released as heat.

The process continues in the core of the star until the hydrogen there is used up. The core then contracts, since gravitation is no longer opposed by energy production, and as a result both the core and the surrounding material are heated. Hydrogen fusion then begins in the surrounding layers. Meanwhile the core becomes hot enough to ignite other fusion reactions, burning helium to form carbon, then burning the carbon to form neon, oxygen and finally silicon. Again each of these reactions leads to the release of energy. One last cycle of fusion combines silicon nuclei to form iron, specifically the common iron isotope  $^{56}\text{Fe}$ , made up of 26 protons and 30 neutrons. Iron is the end of the line for spontaneous fusion. The  $^{56}\text{Fe}$  nucleus is the most strongly bound of all nuclei, and further fusion would absorb energy rather than releasing it.

At this stage in the star's existence it has an onionlike structure. A core of iron and related elements is surrounded by a shell of silicon and sulfur, and beyond this are shells of oxygen, carbon and helium. The outer envelope is mostly hydrogen.

Only the largest stars proceed all the



**COLLAPSE AND REBOUND** are the initiating events in a supernova explosion. Here the core of a massive star is shown as it passes through the moment of "maximum scrunch," when the center reaches its highest density. Each contour represents a shell of matter whose radial position is followed through a period of 12 milliseconds. The included mass, or total mass inside the contour, does not change as the shells contract and expand. Initially the core is iron,

but the extreme compression of the collapse converts the innermost few kilometers into nuclear matter, the stuff of the atomic nucleus. Surrounding this region is a shell made up of various heavy nuclei, including iron. At maximum scrunch the contraction stops with a jolt, creating a shock wave (blue line) that travels outward at 30,000 kilometers per second or more. In the wake of the shock nuclei are broken up into individual nucleons (protons and neutrons).

way to the final, iron-core stage of the evolutionary sequence. A star the size of the sun gets no further than helium burning, and the smallest stars stop with hydrogen fusion. A larger star also consumes its stock of fuel much sooner, even though there is more of it to begin with; because the internal pressure and temperature are higher in a large star, the fuel burns faster. Whereas the sun should have a lifetime of 10 billion years, a star 10 times as massive can complete its evolution 1,000 times faster. Regardless of how long it takes, all the usable fuel in the core will eventually be exhausted. At that point heat production in the core ends and the star must contract.

When fusion ends in a small star, the star slowly shrinks, becoming a white dwarf: a burned-out star that emits only a faint glow of radiation. In isolation the white dwarf can remain in this state indefinitely, cooling gradually but otherwise changing little. What stops the star from contracting further? The answer was given more than 50 years ago by Subrahmanyan Chandrasekhar of the University of Chicago.

Loosely speaking, when ordinary matter is compressed, higher density is achieved by squeezing out the empty

space between atoms. In the core of a white dwarf this process has reached its limit: the atomic electrons are pressed tightly together. Under these conditions the electrons offer powerful resistance to further compression.

Chandrasekhar showed there is a limit to how much pressure can be resisted by the electrons' mutual repulsion. As the star contracts, the gravitational energy increases, but so does the energy of the electrons, raising their pressure. If the contraction goes very far, both the gravitational energy and the electron energy are inversely proportional to the star's radius. Whether or not there is some radius at which the two opposing forces are in balance, however, depends on the mass of the star. Equilibrium is possible only if the mass is less than a critical value, now called the Chandrasekhar mass. If the mass is greater than the Chandrasekhar limit, the star must collapse.

The value of the Chandrasekhar mass depends on the relative numbers of electrons and nucleons (protons and neutrons considered collectively): the higher the proportion of electrons, the larger the electron pressure and so the larger the Chandrasekhar mass. In small stars where the chain of fusion reactions stops at carbon the ratio is approximately 1/2 and the Chandra-

sekhar mass is 1.44 solar masses. This is the maximum stable mass for a white dwarf.

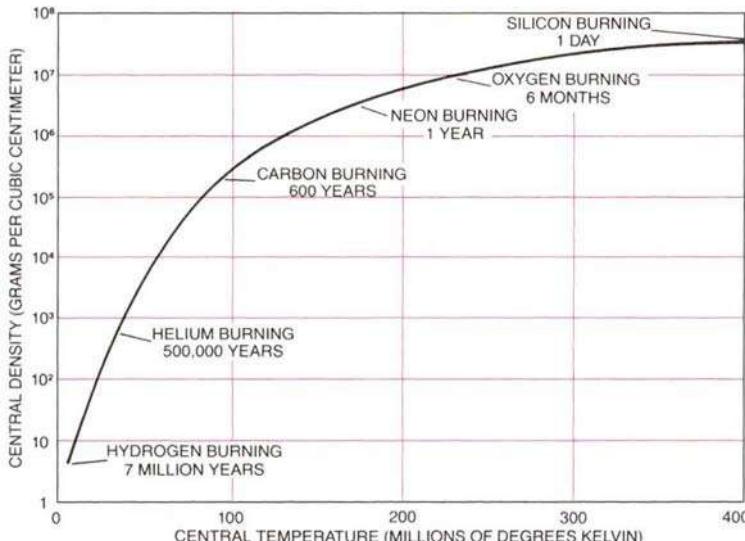
A white dwarf with a mass under the Chandrasekhar limit can remain stable indefinitely; nevertheless, it is just such stars that are thought to give rise to Type I supernovas. How can this be? The key to the explanation is that white dwarfs that explode in supernovas are not solitary stars but rather are members of binary star systems. According to one hypothesis, matter from the binary companion is attracted by the intense gravitational field of the dwarf star and gradually falls onto its surface, increasing the mass of the carbon-and-oxygen core. Eventually the carbon ignites at the center and burns in a wave that travels outward, destroying the star.

The idea that explosive carbon burning triggers Type I supernovas was proposed in 1960 by Hoyle and Fowler. More detailed models have since been devised by many astrophysicists, most notably Icko Iben, Jr., and his colleagues at the University of Illinois at Urbana-Champaign. Recent calculations done by Ken'ichi Nomoto and his colleagues at the University of Tokyo suggest that the burning is actually not explosive. The wave of fusion reactions propagates like the burning of a fuse rather than like the explosion of a firecracker; it is a deflagration rather than a detonation.

Even though the burning is less violent than a detonation, the white dwarf is completely disrupted. The initial binding energy that holds the star together is approximately  $10^{50}$  ergs; the energy released by the burning is 20 times greater ( $2 \times 10^{51}$  ergs), enough to account for the 10,000-kilometer-per-second velocity of supernova remnants. In the course of the deflagration nuclear reactions create about one solar mass of the unstable nickel isotope  $^{56}\text{Ni}$ , which decays into  $^{56}\text{Co}$  and then  $^{56}\text{Fe}$  over a period of months. The rate of energy release from the radioactive decay is just right to account for the gradually declining light emission from Type I supernovas.

The Type II supernovas that are our main concern here arise from much more massive stars. The lower limit is now thought to be about eight solar masses.

In tracing the history of a Type II supernova it is best to begin at the moment when the fusion of silicon nuclei to form iron first becomes possible at the center of the star. At this point the star has already passed through stages of burning hydrogen, helium, neon, carbon and oxygen, and it has the onionlike structure described above.



**EVOLUTION OF A MASSIVE STAR** is a steadily accelerating progress toward higher temperature and density in the core. For most of the star's lifetime the primary energy source is the fusion of hydrogen nuclei to form helium. When the hydrogen in the core is exhausted, the core contracts, which heats it enough to ignite the fusion of helium into carbon. This cycle then repeats, at a steadily increasing pace, through the burning of carbon, neon, oxygen and silicon. The final stage of silicon fusion yields a core of iron, from which no further energy can be extracted by nuclear reactions. Hence the iron core cannot resist gravitational collapse, leading to a supernova explosion. The sequence shown is for a star of 25 solar masses. Data in this illustration and the one on the opposite page are based on calculations done by Thomas A. Weaver of the Lawrence Livermore National Laboratory.

The star has taken several million years to reach this state. Subsequent events are much faster.

When the final fusion reaction begins, a core made up of iron and a few related elements begins to form at the center of the star, within a shell of silicon. Fusion continues at the boundary between the iron core and the silicon shell, steadily adding mass to the core. Within the core, however, there is no longer any production of energy by nuclear reactions; the core is an inert sphere under great pressure. It is thus in the same predicament as a white dwarf: it can resist contraction only by electron pressure, which is subject to the Chandrasekhar limit.

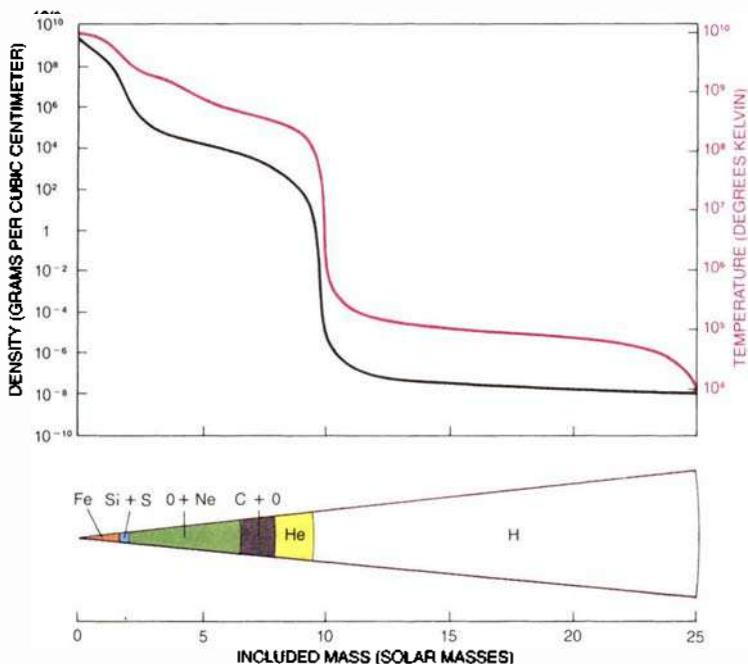
Once the fusion of silicon nuclei begins, it proceeds at an extremely high rate, and the mass of the core reaches the Chandrasekhar limit in about a day. We noted above that for a white dwarf the Chandrasekhar mass is equal to 1.44 solar masses; for the iron core of a large star the value may be somewhat different, but it is probably in the range between 1.2 and 1.5 solar masses.

When the Chandrasekhar mass has been attained, the pace speeds up still more. The core that was built in a day collapses in less than a second. The task of analysis also becomes harder at this point, so that theory relies on the assistance of computer simulation. Computer programs that trace the evolution of a star have been developed by a number of workers, including W. David Arnett of the University of Chicago and a group at the Lawrence Livermore National Laboratory led by Thomas A. Weaver of that laboratory and Stanford Woosley of the University of California at Santa Cruz. They are the "burners" of stars; we and our colleagues in theoretical physics are "users" of their calculations.

The simulations furnish us with a profile of the presupernova core, giving composition, density and temperature as a function of radius. The subsequent analysis relies on applying familiar laws of thermodynamics, the same laws that describe such ordinary terrestrial phenomena as the working of a heat engine or the circulation of the atmosphere.

It is worthwhile tracing in some detail the initial stages in the implosion of the core. One of the first points of note is that compression raises the temperature of the core, which might be expected to raise the pressure and slow the collapse. Actually the heating has just the opposite effect.

Pressure is determined by two factors: the number of particles in a system and their average energy. In the



**ONIONLIKE STRUCTURE** is characteristic of a massive star at the end of its evolution, just before the gravitational collapse. The iron core is embedded in a mantle of silicon, sulfur, oxygen, neon, carbon and helium, surrounded by an attenuated envelope of hydrogen. Temperature and density fall off steadily in the mantle, then drop precipitously at the envelope. Fusion has stopped in the core but continues at boundaries between layers.

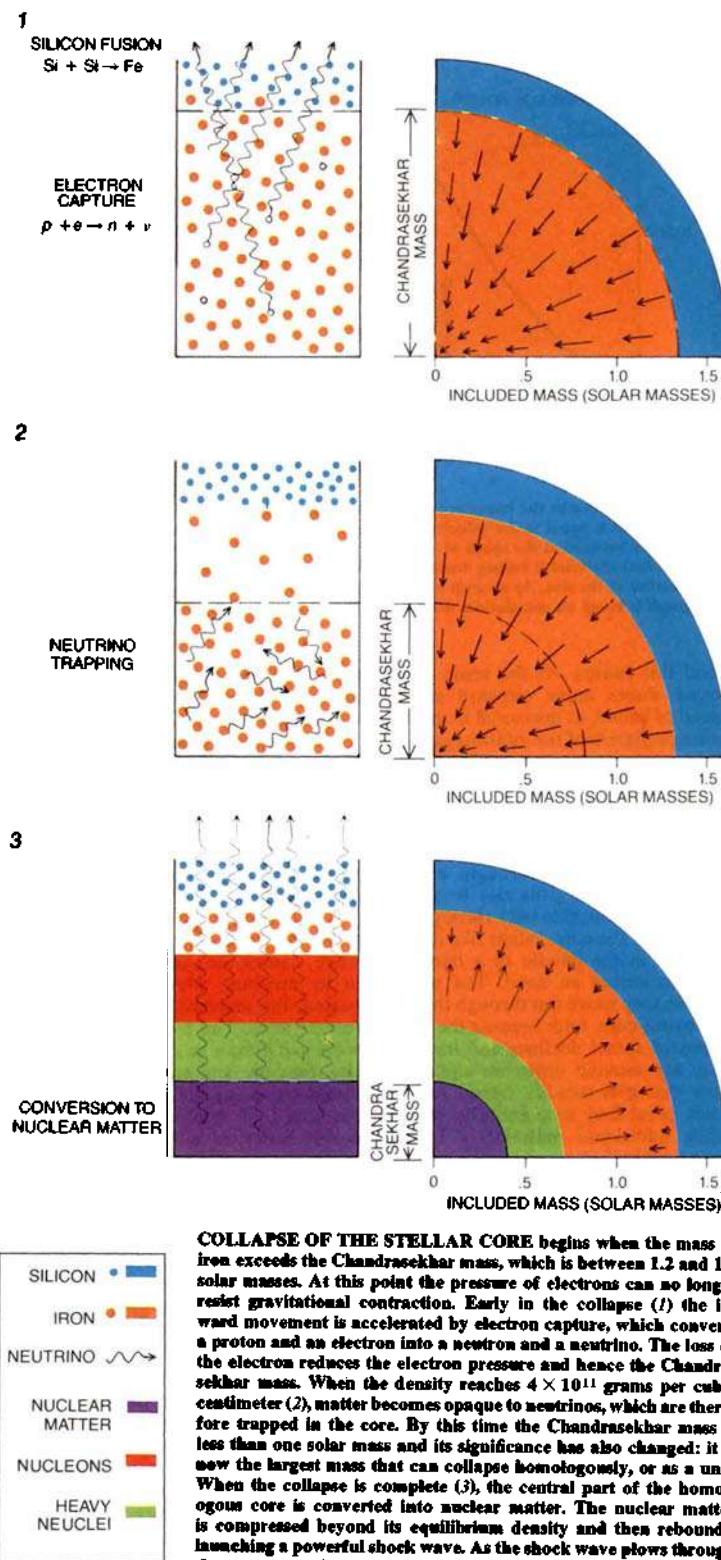
core both nuclei and electrons contribute to the pressure, but the electron component is much larger. When the core is heated, a small fraction of the iron nuclei are broken up into smaller nuclei, increasing the number of nuclear particles and raising the nuclear component of the pressure. At the same time, however, the dissociation of the nuclei absorbs energy; since energy is released when an iron nucleus is formed, the same quantity of energy must be supplied in order to break the nucleus apart. The energy comes from the electrons and decreases their pressure. The loss in electron pressure is more important than the gain in nuclear pressure. The net result is that the collapse accelerates.

It might seem that the implosion of a star would be a chaotic process, but in fact it is quite orderly. Indeed, the entire evolution of the star is toward a condition of greater order, or lower entropy. It is easy to see why. In a hydrogen star each nucleon can move willy-nilly along its own trajectory, but in an iron core groups of 56 nucleons are bound together and must move in lockstep. Initially the entropy per nucleon, expressed in units of Boltzmann's constant, is about 15; in the presupernova core it is less than 1. The

difference in entropy has been carried off during the evolution of the star by electromagnetic radiation and toward the end also by neutrinos.

The low entropy of the core is maintained throughout the collapse. Nuclear reactions continually change the species of nuclei present, which one might think could lead to an increase in entropy; the reactions are so fast, however, that equilibrium is always maintained. The collapse takes only milliseconds, but the time scale of the nuclear reactions is typically from  $10^{-15}$  to  $10^{-22}$  second, so that any departure from equilibrium is immediately corrected.

Another effect was once thought to increase the entropy, but it now seems likely that it actually reduces it somewhat. The high density in the collapsing core favors the reaction known as electron capture. In this process a proton and an electron come together to yield a neutron and a neutrino. The neutrino escapes from the star, carrying off both energy and entropy and cooling the system just as the evaporation of moisture cools the body. There are several complications to this process, so that its effect on the entropy is uncertain. In any case, the loss of the electron diminishes the electron pres-



sure and so allows the implosion to accelerate further.

The first stage in the collapse of a supernova comes to an end when the density of the stellar core reaches a value of about  $4 \times 10^{11}$  grams per cubic centimeter. This is by no means the maximum density, since the core continues to contract, but it marks a crucial change in physical properties: at this density matter becomes opaque to neutrinos. The importance of this development was first pointed out by T. J. Mazurek of the Mission Research Laboratory in Santa Barbara, Calif., and by Katsuhiko Sato of the University of Tokyo.

The neutrino is an aloof particle that seldom interacts with other forms of matter. Most of the neutrinos that strike the earth, for example, pass all the way through it without once colliding with another particle. When the density exceeds 400 billion grams per cubic centimeter, however, the particles of matter are packed so tightly that even a neutrino is likely to run into one. As a result neutrinos emitted in the collapsing core are effectively trapped there. The trapping is not permanent; after a neutrino has been scattered, absorbed and reemitted many times, it must eventually escape, but the process takes longer than the remaining stages of the collapse. The effective trapping of neutrinos means that no energy can get out of the core.

The process of electron capture in the early part of the collapse reduces not only the electron pressure but also the ratio of electrons to nucleons, the quantity that figures in the calculation of the Chandrasekhar mass. In a typical presupernova core the ratio is between .42 and .46; by the time of neutrino trapping it has fallen to .39. This lower ratio yields a Chandrasekhar mass of .88 solar mass, appreciably less than the original value of between 1.2 and 1.5.

At this point the role of the Chandrasekhar mass in the analysis of the supernova also changes. At the outset it was the largest mass that could be supported by electron pressure; it now becomes the largest mass that can collapse as a unit. Areas within this part of the core can communicate with one another by means of sound waves and pressure waves, so that any variations in density are immediately evened out. As a result the inner part of the core collapses homologously, or all in one piece, preserving its shape.

The theory of homologous collapse was worked out by Peter Goldreich and Steven Weber of Caltech and was further developed by Amos Yahil and James M. Lattimer of the State Uni-

**COLLAPSE OF THE STELLAR CORE** begins when the mass of iron exceeds the Chandrasekhar mass, which is between 1.2 and 1.5 solar masses. At this point the pressure of electrons can no longer resist gravitational contraction. Early in the collapse (1) the inward movement is accelerated by electron capture, which converts a proton and an electron into a neutron and a neutrino. The loss of the electron reduces the electron pressure and hence the Chandrasekhar mass. When the density reaches  $4 \times 10^{11}$  grams per cubic centimeter (2), matter becomes opaque to neutrinos, which are therefore trapped in the core. By this time the Chandrasekhar mass is less than one solar mass and its significance has also changed: it is now the largest mass that can collapse homologously, or as a unit. When the collapse is complete (3), the central part of the homologous core is converted into nuclear matter. The nuclear matter is compressed beyond its equilibrium density and then rebounds, launching a powerful shock wave. As the shock wave plows through the outer core, iron nuclei "evaporate" to form a gas of nucleons.

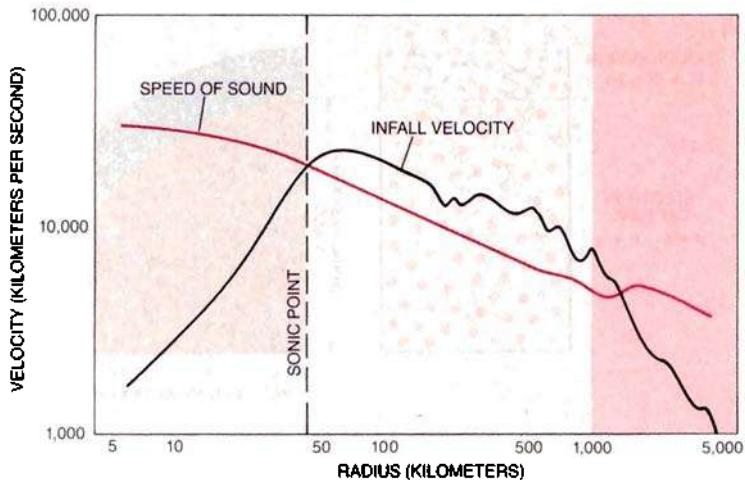
versity of New York at Stony Brook. The shock wave that blows off the outer layers of the star forms at the edge of the homologous core. Before we can give an account of that process, however, we must continue to trace the sequence of events within the core itself.

Chandrasekhar's work showed that electron pressure cannot save the core of a large star from collapse. The only other hope for stopping the contraction is the resistance of nucleons to compression. In the presupernova core nucleon pressure is a negligible fraction of electron pressure. Even at a density of  $4 \times 10^{31}$  grams per cubic centimeter, where neutrino trapping begins, nucleon pressure is insignificant. The reason is the low entropy of the system. At any given temperature, pressure is proportional to the number of particles per unit volume, regardless of the size of the individual particles. An iron nucleus, with 56 nucleons, makes the same contribution to the pressure as an isolated proton does. If the nuclei in the core were broken up, their pressure might be enough to stop the contraction. The fissioning of the nuclei is not possible, however, because the entropy of the core is too low. A supernova core made up of independently moving protons and neutrons would have an entropy per nucleon of between 5 and 8, whereas the actual entropy is less than 1.

The situation does not change, and the collapse is not impeded, until the density in the central part of the core reaches about  $2.7 \times 10^{14}$  grams per cubic centimeter. This is the density of matter inside a large atomic nucleus, and in effect the nucleons in the core merge to form a single gigantic nucleus. A teaspoonful of such matter has about the same mass as all the buildings in Manhattan combined.

Nuclear matter is highly incompressible. Hence once the central part of the core reaches nuclear density there is powerful resistance to further compression. That resistance is the primary source of the shock waves that turn a stellar collapse into a spectacular explosion.

**W**ithin the homologously collapsing part of the core, the velocity of the infalling material is directly proportional to distance from the center. (It is just this property that makes the collapse homologous.) Density, on the other hand, decreases with distance from the center, and as a result so does the speed of sound. The radius at which the speed of sound is equal to the infall velocity is called the sonic point, and it marks the boundary of the homologous core. A disturbance inside the core can have no influence be-



**SONIC POINT** marks the boundary of the homologous core. It is the radius at which the speed of sound is equal to the velocity of the infalling material. A sound wave at the sonic point moves outward at the speed of sound in relation to the material it is passing through, but since that material is falling inward at the same speed, the wave stands still in relation to the center of the star. As a result a disturbance inside the core cannot reach the outside. The graph is based on calculations done by W. David Arnett of the University of Chicago.

yond this radius. At the sonic point sound waves move outward at the speed of sound, as measured in the coordinate system of the infalling matter. This matter is moving inward at the same speed, however, and so the waves are at a standstill in relation to the center of the star.

When the center of the core reaches nuclear density, it is brought to rest with a jolt. This gives rise to sound waves that propagate back through the medium of the core, rather like the vibrations in the handle of a hammer when it strikes an anvil. The waves slow as they move out through the homologous core, both because the local speed of sound declines and because they are moving upstream against a flow that gets steadily faster. At the sonic point they stop entirely. Meanwhile additional material is falling onto the hard sphere of nuclear matter in the center, generating more waves. For a fraction of a millisecond the waves collect at the sonic point, building up pressure there. The bump in pressure slows the material falling through the sonic point, creating a discontinuity in velocity. Such a discontinuous change in velocity constitutes a shock wave.

At the surface of the hard sphere in the heart of the star infalling material stops suddenly but not instantaneously. The compressibility of nuclear matter is low but not zero, and so momentum carries the collapse beyond the point of equilibrium, compressing the central core to a density even higher than that of an atomic nucleus. We call

this point the instant of "maximum scrunch." Most computer simulations suggest the highest density attained is some 50 percent greater than the equilibrium density of a nucleus. After the maximum scrunch the sphere of nuclear matter bounces back, like a rubber ball that has been compressed. The bounce sets off still more sound waves, which join the growing shock wave at the sonic point.

A shock wave differs from a sound wave in two respects. First, a sound wave causes no permanent change in its medium; when the wave has passed, the material is restored to its former state. The passage of a shock wave can induce large changes in density, pressure and entropy. Second, a sound wave—by definition—moves at the speed of sound. A shock wave moves faster, at a speed determined by the energy of the wave. Hence once the pressure discontinuity at the sonic point has built up into a shock wave, it is no longer pinned in place by the infalling matter. The wave can continue outward, into the overlying strata of the star. According to computer simulations, it does so with great speed, between 30,000 and 50,000 kilometers per second.

**U**p to this point in the progress of the supernova essentially all calculations are in agreement. What happens next, however, is not yet firmly established. In the simplest scenario, which we have favored, the shock wave rushes outward, reaching the surface of the iron core in a fraction of a

second and then continuing through the successive onionlike layers of the star. After some days it works its way to the surface and erupts as a violent explosion. Beyond a certain radius—the bifurcation point—all the material of the star is blown off. What is left inside the bifurcation radius condenses into a neutron star.

Alas! Using presupernova cores simulated in 1974 by Weaver and Woosley, calculations of the fate of the shock wave are not so accommodating. The shock travels outward to a distance of between 100 and 200 kilometers from the center of the star,

but then it becomes stalled, staying at roughly the same position as matter continues to fall through it. The main reason for the stalling is that the shock breaks up nuclei into individual nucleons. Although this process increases the number of particles, which might be expected to raise the pressure, it also consumes a great deal of energy; the net result is that both temperature and pressure are sharply reduced.

The fragmentation of the nuclei contributes to energy dissipation in another way as well: it releases free protons, which readily capture electrons. The neutrinos emitted in this process can

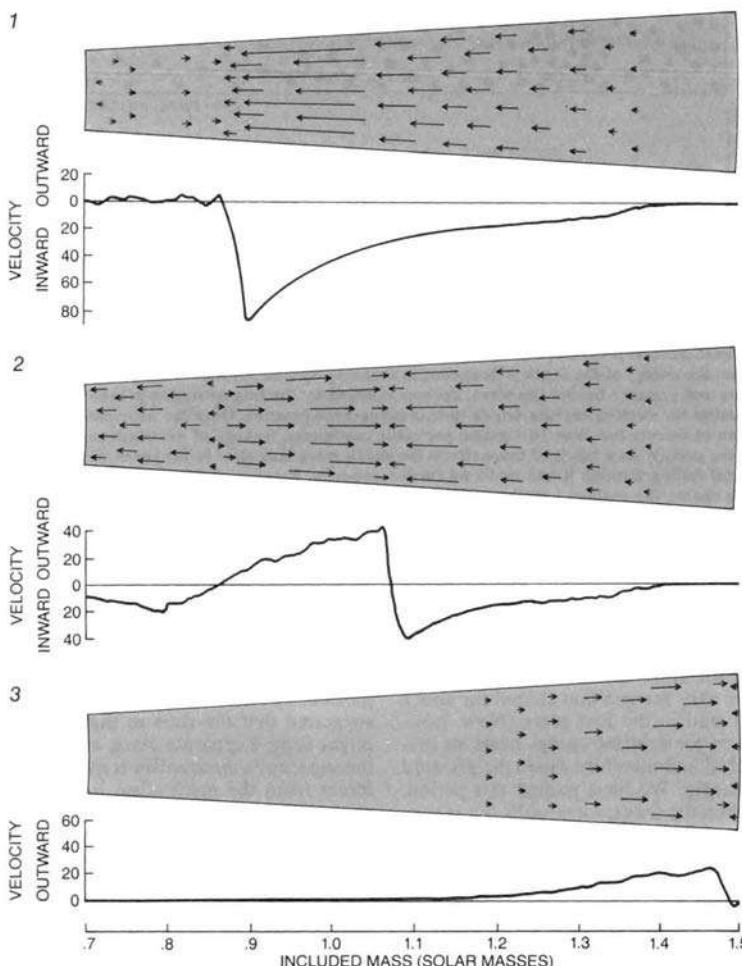
escape, removing their energy from the star. The escape is possible because the shock has entered material whose density is below the critical value for neutrino trapping. The neutrinos that had been trapped behind the shock also stream out, carrying away still more energy. Because of the many hazards to the shock wave in the region between 100 and 200 kilometers, we have named this region of the star the "minefield."

It would be satisfying to report that we have found a single mechanism capable of explaining for all Type II supernovas how the shock wave makes its way through the minefield. We cannot do so. What we have to offer instead is a set of possible explanations, each of which seems to work for stars in a particular range of masses.

The place to begin is with stars of between 12 and about 18 solar masses. Weaver and Woosley's most recent models of presupernova cores for such stars differ somewhat from those they calculated a decade ago; the most important difference is that the iron core is smaller than earlier estimates indicated—about 1.35 solar masses. The homologous core, at whose surface the shock wave forms, includes .8 solar mass of this material, leaving .55 solar mass of iron outside the sonic point. Since the breaking up of iron nuclei has the highest energy cost, reducing the quantity of iron makes it easier for the shock to break out of the core.

Jerry Cooperstein and Edward A. Baron of Stony Brook have been able to simulate successful supernova explosions in computer calculations that begin with Weaver and Woosley's model cores. The main requirement, first surmised by Sidney H. Kahana of the Brookhaven National Laboratory, is that the homologous core be very strongly compressed, so that it can rebound vigorously and create an intense shock. Two factors cooperate to achieve this result in the simulations. The first factor is the use of general relativity rather than the force field of Newtonian gravitation. The second is the assumption that nuclear matter is much more compressible than had been thought.

Baron's first result showed that a star of 12 solar masses would explode if the compressibility of nuclear matter is 1.5 times the standard value. This seemed rather arbitrary, but then one of us (Brown) examined the problem with a sophisticated method of nuclear-matter theory. It turned out that the most consistent interpretation of the experimental findings yields a compressibility of 2.5 times the standard value! We then found that in 1982 An-



**SHOCK WAVE** can move faster than sound and so it can carry the energy and momentum of the rebound past the sonic point. Just before the bounce (1) the inner core has reached the density of nuclear matter and has stopped contracting, but overlying matter is about to fall onto the core at speeds of up to 90,000 kilometers per second. Two milliseconds later (2) the core is being driven further inward, but at the same time much of the infalling matter has rebounded to form a shock wave. After 20 milliseconds (3) the shock has reached the edge of the core. This mechanism of supernova explosion, in which the shock succeeds directly in bursting through the core, seems to work for stars of between 12 and 18 solar masses. Velocity profiles shown were calculated by Jerry Cooperstein of the State University of New York at Stony Brook. Velocities are given in thousands of kilometers per second.

drew D. Jackson, E. Krotscheck, D. E. Meltzer and R. A. Smith had reached the same conclusion by another method, but no one had recognized the relevance of their work to the supernova problem. We consider the higher estimate of nuclear compressibility quite reliable.

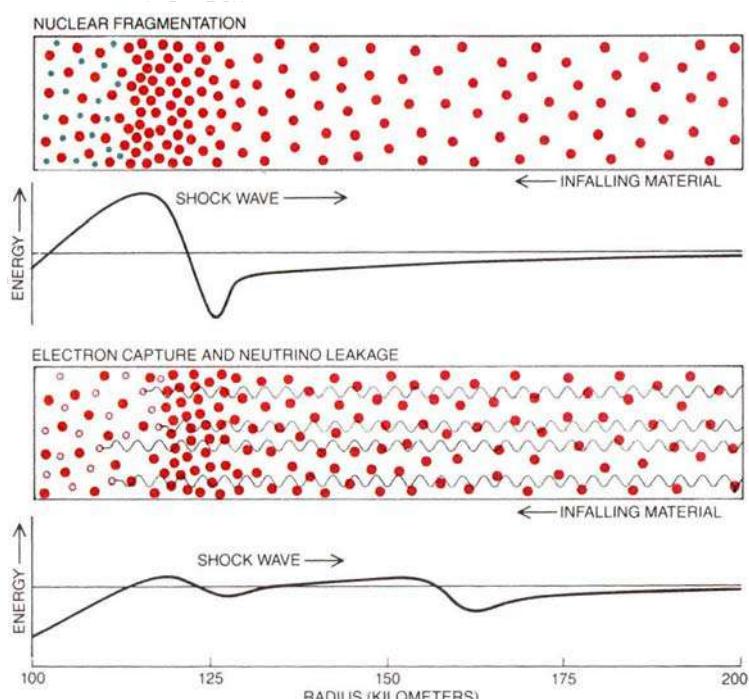
The mechanism described by Baron, Cooperstein and Kahana seems to work for stars of up to about 18 solar masses. With still larger stars, however, even the powerful shock wave created in their simulations becomes stalled in the minefield. A star of 25 solar masses has about two solar masses of iron in its core, and so the shock wave must penetrate 1.2 solar masses of iron rather than .55 solar mass. The shock does not have enough energy to dissociate this much iron.

A plausible explanation of what might happen in these massive stars has recently emerged from the work of James R. Wilson of Lawrence Livermore, who has done extensive numerical simulations of supernova explosions. For some time it had seemed that when the shock wave failed, all the mass of the star might fall back into the core, which would evolve into a black hole. That fate is still a possible one, but Wilson noted a new phenomenon when he continued some of his simulations for a longer period.

In the collapsing stellar core it takes only 10 milliseconds or so for the shock wave to reach the minefield and stall. A simulation of the same events, even with the fastest computers, takes at least an hour. Wilson allowed his calculations to run roughly 100 times longer, to simulate a full second of time in the supernova. In almost all cases he found that the shock wave eventually revived.

The revival is due to heating by neutrinos. The inner core is a copious emitter of neutrinos because of continuing electron capture as the matter is compressed to nuclear density. Adam S. Burrows and Lattimer of Stony Brook and Mazurek have shown that half of the electrons in the homologous core are captured within about half a second, and the emitted neutrinos carry off about half of the gravitational energy set free by the collapse, some  $10^{48}$  ergs. Deep within the core the neutrinos make frequent collisions with other particles; indeed, we noted above that they are trapped, in the sense that they cannot escape within the time needed for the homologous collapse. Eventually, though, the neutrinos do percolate upward and reach strata of lower density, where they can move freely.

At the radius where the shock wave



**SHOCK WAVE SEEKS TO STALL** in stars whose mass is greater than about 18 solar masses. Several processes sap the wave's energy. The most important is nuclear fragmentation: the energy of the shock is dissipated in breaking up iron nuclei, lowering the temperature and pressure behind the wave. Protons released by the fragmentation provide opportunities for electron capture, which further reduces the pressure. Once the wave enters a region of density less than  $10^{11}$  grams per cubic centimeter, leakage of neutrinos carries off more energy. As a result of these effects the shock wave may slow to the speed of the material falling through it and make no further progress. Because of the various hazards to the shock, the authors call the region between 100 and 200 kilometers the "minefield."

stalls only one neutrino out of every 1,000 is likely to collide with a particle of matter, but these collisions nonetheless impart a significant amount of energy. Most of the energy goes into the dissociation of nuclei into nucleons, the very process that caused the shock to stall in the first place. Now, however, the neutrino energy heats the material and therefore raises the pressure sharply. We have named this period, when the shock wave stalls but is then revived by neutrino heating, "the pause that refreshes."

**N**eutrino heating is most effective at a radius of about 150 kilometers, where the probability of neutrino absorption is not too low and yet the temperature is not so high that the matter there is itself a significant emitter of neutrinos. The pressure increase at this radius is great enough, after about half a second, to stop the fall of the overlying matter and begin pushing it outward. Hence 150 kilometers becomes the bifurcation radius. All the matter within this boundary ultimately falls

into the core; the matter outside, 20 solar masses or more, is expelled.

The one group of stars left to be considered are those of from eight to 11 solar masses, the smallest stars capable of supporting a Type II supernova explosion. In 1980 Weaver and Woosley suggested that the stars in this group might form a separate class, in which the supernova mechanism is quite different from the mechanism in heavier stars.

According to calculations done by Nomoto and by Weaver and Woosley, in the presupernova evolution of these lighter stars the core does not reach the temperature needed to form iron; instead fusion ends with a mixture of elements between oxygen and silicon. Energy production then stops, and since the mass of the core is greater than the Chandrasekhar limit, the core collapses. The shock wave generated by the collapse may be helped to propagate by two circumstances. First, breaking up oxygen or silicon nuclei robs the shock of less energy than the dissociation of iron nuclei would. Second, far-

ther out in the star the density falls off abruptly (by a factor of roughly 10 billion) at the boundary between the carbon and the helium shells. The shock wave has a much easier time pushing through the lower-density material.

For a star of nine solar masses Nomoto finds that the presupernova core consists of oxygen, neon and magnesium and has a mass of 1.35 solar masses. Nomoto and Wolfgang Hillebrandt of the Max-Planck Institute for Physics and Astrophysics in Munich have gone on to investigate the further development of this core. They find that the explosion proceeds easily through the core, aided by the burning of oxygen nuclei, and that a rather large amount of energy is released.

Two recent attempts to reproduce the Nomoto-Hillebrandt results have been unsuccessful, and so the status of their model remains unclear. We think the greater compressibility of nuclear matter assumed in the Baron-Cooperstein-Kahana program should be helpful here. Of course it is possible that stars this small do not give rise to supernovas; on the other hand, there are suggestive arguments (based on mea-

urements of the abundance of various nuclear species) that the Crab Nebula was created by the explosion of a star of about nine solar masses.

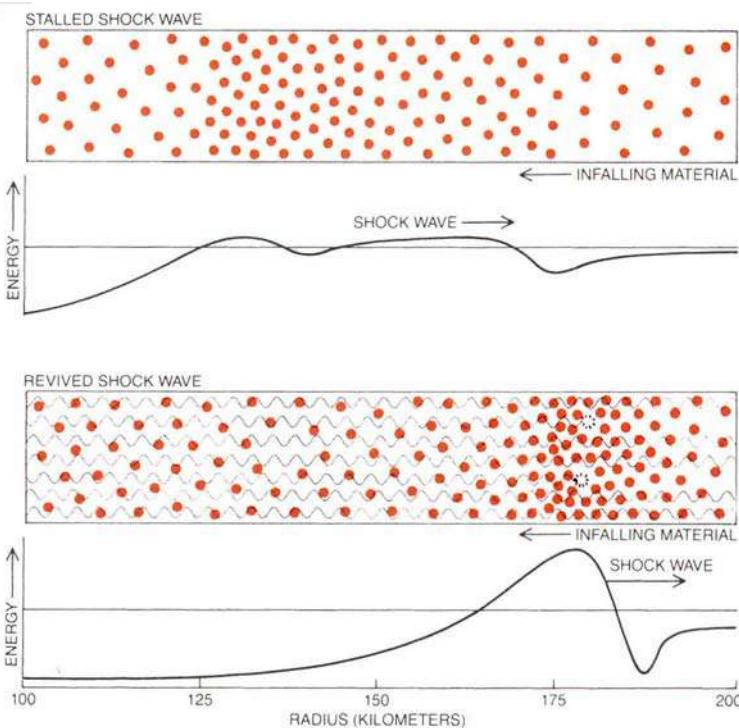
After the outer layers of a star have been blown off, the fate of the core remains to be decided. Just as gravitation overwhelms electron pressure if the mass exceeds the Chandrasekhar limit, so even nuclear matter cannot resist compression if the gravitational field is strong enough. For a cold neutron star—one that has no source of supporting pressure other than the repulsion of nucleons—the limiting mass is thought to be about 1.8 solar masses. The compact remnant formed by the explosion of lighter stars is well below this limit, and so those supernovas presumably leave behind a stable neutron star. For the larger stars the question is in doubt. In Wilson's calculations any star of more than about 20 solar masses leaves a compact remnant of more than two solar masses. It would appear that the remnant will become a black hole, a region of space where matter has been crushed to infinite density.

Even if the compact remnant ultimately degenerates into a black hole, it begins as a hot neutron star. The central temperature immediately after the explosion is roughly 100 billion degrees Kelvin, which generates enough thermal pressure to support the star even if it is larger than 1.8 solar masses. The hot nuclear matter cools by the emission of neutrinos. The energy they carry off is more than 100 times the energy emitted in the explosion itself: some  $3 \times 10^{58}$  ergs. It is the energy equivalent of 10 percent of the mass of the neutron star.

The detection of neutrinos from a supernova explosion and from the subsequent cooling of the neutron star is one possible way we might get a better grasp of what goes on in these spectacular events. The neutrinos originate in the core of the star and pass almost unhindered through the outer layers, and so they carry evidence of conditions deep inside. Electromagnetic radiation, on the other hand, diffuses slowly through the shells of matter and reveals only what is happening at the surface. Neutrino detectors have recently been set up in mines and tunnels, where they are screened from the background of cosmic rays.

Another observational check on the validity of supernova models is the relative abundances of the chemical elements in the universe. Supernovas are probably the main source of all the elements heavier than carbon, and so the spectrum of elements released in simulated explosions ought to match the observed abundance ratios. Many attempts to reproduce the abundance ratios have failed, but earlier this year Weaver and Woosley completed calculations whose agreement with observation is surprisingly good. They began with Wilson's model for the explosion of a star of 25 solar masses. For almost all the elements and isotopes between carbon and iron their abundance ratios closely match the measured ones.

In recent years the study of supernovas has benefited from a close interaction between analytic theory and computer simulation. The first speculations about supernova mechanisms were put forward decades ago, but they could not be worked out in detail until the computers needed for numerical simulation became available. The results of the computations, on the other hand, cannot be understood except in the context of an analytic model. By continuing this collaboration we should be able to progress from a general grasp of principles and mechanisms to the detailed prediction of astronomical observations.



**REVIVAL OF THE STALLED SHOCK WAVE** in heavy stars may be due to heating by neutrinos. Their source is the collapsed core, which radiates the energy equivalent of 10 percent of its mass in the form of neutrinos. Only a small fraction of them are absorbed, but the flux is so intense that many iron nuclei are dissociated. Earlier in the evolution of the supernova the breakup of iron nuclei took energy from the shock wave, but since the process is now powered by external neutrinos, the dissociation no longer decreases shock energy.

This page is intentionally left blank

---

### Chapter 3

## Accretion onto and Radiation from the Compact Objects Formed in SN 1987A

G.E. Brown and J. Weingartner

Reprinted with permission from *Astrophysical Journal*, 436 (1994) 843–847  
Copyright (1994) by American Astronomical Society

### Commentary by G.E. Brown

After “Babble” Hans and I spent every January together on the west coast either at Santa Barbara, Santa Cruz, or at Caltech, chiefly at the latter. In 1987, after working a week or two on supernova theory, we decided we could get no further without more observations, and we began working on the theory of relativistic heavy ion collisions. I remember shaking Hans’ hand as I parted with him at the end of January, and saying, “All right, it’s time for a nearby supernova explosion.” February 23, 1987 one went off. People said, “You are very lucky.” Actually, we were, since the star had exploded  $\sim 150,000$  years earlier, the neutrinos and radiation taking that long to reach us. I replied, “Luck comes to those who work.”

In any case we were on a merry ride, full of excitement, following the news of the arrival of neutrinos, etc. Hans kept writing papers on supernova theory, some of which are published in his “Selected Works”.

I was waiting expectantly for the neutron star, which should emerge from the implosion, to be observed. Every January when we were in Caltech I asked the observers how long it would take the material which shrouded the neutron star to clear out so that we could observe it. The answer was always, “about another year”. I decided that if one year was always the remaining time, they really didn’t know what they were talking about. With Joe Weingartner, a senior at Stony Brook, I decided to work out the time at which we should see light from the neutron star from the matter falling onto it. Roger Chevalier had already shown numerically that we should see light at the Eddington limit, about two orders of magnitude greater than we actually saw (which could be well explained by radioactivity from the expanding envelope) in a time slightly less than one year. Using classical accretion theory, Joe and I could reproduce Chevalier’s result analytically.

There have been recurrent reports of seeing evidence of a pulsar in 1987A, the most recent one by Middleditch *et al.* (2000). If, indeed, a pulsar is present, we would expect to see a plerion on the clouds surrounding the neutron star.

Perhaps the most fascinating alternative explanation was given in a paper by Stirling Colgate and Chris Fryer, which I believe never got published. The cut in matter expelled comes in Fe (we know that about  $0.07M_{\odot}$  of Fe was given off). Fe lines produce a tremendously high opacity  $\kappa$ . The Eddington limit is proportional to  $\kappa^{-1}$ . With a  $\kappa$  of two orders of magnitude greater than that used by Brown and Weingartner, the Eddington limit could be brought down to the upper limit given by observation.

I return to the black hole concept in the next Commentary.

#### Reference

- (1) Middleditch, J., Kristian, J.A., Kunkel, W.E., Hill, K.M., Watson, R.D., Lucinio, R., Imamura, T.N., Steiman-Cameron, T.Y., Sheaver, A., Butler, R., Redfens, M., and Danks, A.C. (2000), *New Astronomy* **5**, 243.

## ACCRETION ONTO AND RADIATION FROM THE COMPACT OBJECT FORMED IN SN 1987A<sup>1</sup>

G. E. BROWN AND J. C. WEINGARTNER

Department of Physics, State University of New York at Stony Brook, Stony Brook, NY 11794

Received 1994 February 22; accepted 1994 June 6

### ABSTRACT

We follow the work of Chevalier, Houck, Blondin, and Park on hypercritical spherical accretion onto compact objects, applying their work to the case of the compact object and remnant formed in SN 1987A.

We begin by motivating the Bondi spherical accretion theory, obtaining an expression for the mass accretion rate. We take SN 1987A parameters from Woosley and Bethe to evaluate the expression for this particular case and find hypercritical accretion on the order of  $10^4$  times the Eddington rate. The Eddington rate can be (greatly) exceeded because neutrinos carry off the energy.

In this situation photons within a certain distance from the compact object are trapped; we derive an expression for this trapping radius, which decreases in time. For the case that the compact object is a neutron star, even though photons are trapped, neutrinos can still escape and carry off accretion energy, allowing for self-consistent solutions for hypercritical accretion. We use the work of Dicus on neutrino cooling to derive an expression for the shock radius, that is, the distance of the accretion shock front from the neutron star. The shock radius increases with time, so that at some critical time the shock radius equals the trapping radius. We find this critical time to be about 0.6 yr. After this time the luminosity in photons should increase to the Eddington limit,  $3.8 \times 10^{38}$  ergs s<sup>-1</sup>.

For the case that the compact object is a black hole, only the internal energy produced by the  $p\Delta V$  work on the infalling matter outside of the trapping radius can be radiated. This would result in a luminosity  $\sim 10^{34}$ – $10^{35}$  ergs s<sup>-1</sup>.

The observed light curve of SN 1987A is explained by radioactive decays with a current luminosity of a few times  $10^{36}$  ergs s<sup>-1</sup>. The expected contribution from spherical accretion onto a neutron star is clearly not present, while the expected contribution for a black hole would be too small to detect. Our considerations thus support the hypothesis that the compact object formed in SN 1987A is a black hole rather than a neutron star.

**Subject headings:** accretion, accretion disks — black hole physics — stars: neutron — supernovae: individual (SN 1987A)

### 1. INTRODUCTION

The compact object formed in the explosion SN 1987A has yet to be seen (Chevalier 1992). Conventional wisdom is that neutron stars are formed with strong magnetic fields. Even if the beaming of the pulsar radiation is in the wrong direction to be observed on Earth, contributions to the light curve should have been observed or the surrounding synchrotron nebula from pulsar accelerated particles (Helfand & Becker 1984) should be observable, had a neutron star been formed. The compact object could conceivably be a neutron star with weak magnetic field, or it could be a black hole (Brown, Bruenn, & Wheeler 1992).

The light curve of SN 1987A can be explained by radioactive decays (Suntzeff et al. 1992) down to a level of luminosity of a few times  $10^{36}$  ergs s<sup>-1</sup> (see Fig. 1). It has flattened off, and the present luminosity can be explained by the decay of  $^{44}\text{Ti}$  with a half-life of 47 yr.

We consider the level at which accretion can contribute to the light curve. Whereas a strong magnetic field could prevent accretion onto a neutron star, throwing off particles in the neighborhood of the Alfvén radius by a propeller effect (Illarionov & Sunyaev 1985), there is nothing to prevent accretion of ambient matter onto a neutron star with weak magnetic field.

In § 2 we consider the rate of mass accretion onto the compact object, which should be the same for both a neutron star and a black hole. We use the Bondi (1952) spherical accretion theory, and we will find that the accretion is hypercritical, that is, exceeds the Eddington accretion rate by a factor  $\sim 10^4$ .

In § 3 we consider the consequence of the hypercritical nature of the accretion for the luminosity in photons. We follow the work of Blondin (1986), Chevalier (1989), and Houck & Chevalier (1991) for the case of the neutron star. For the black hole case, we briefly note the results of Park (1990).

Conclusions and discussion will be given in § 4.

### 2. BONDI ACCRETION ONTO THE COMPACT CORE OF SN 1987A

Scenarios for the supernova explosion involve an outgoing shock wave that, initially, travels with speed  $\approx 10,000$  km s<sup>-1</sup>. At  $\approx 4$  s after the explosion, there is a bifurcation in the matter behind the shock (Bethe 1993). The shock takes some time to form; matter within a radius  $R \sim 20,000$  km falls onto the compact object, while matter outside this radius is ejected. This scenario would seem to leave a vacuum in the region of the compact object.

However, calculations by Woosley (1988) show that a deceleration of matter is formed when the initial outgoing shock enters the hydrogen envelope. This deceleration sharpens into a reverse shock, and this shock has been used to deposit  $\sim 0.1$

<sup>1</sup> Supported by the U.S. Dept. of Energy grant DE-FG02-88ER 40388.

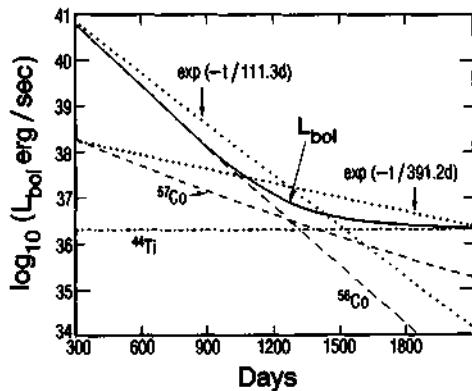


FIG. 1.—Bolometric light curve (UV to IR) powered by the decays of  $^{56}\text{Co}$  ( $0.076 M_{\odot}$ ),  $^{57}\text{Co}$  ( $0.0043 M_{\odot}$ ), and  $^{44}\text{Ti}$  ( $1.15 \times 10^{-4} M_{\odot}$ ), from Nomoto et al. (1993).

$M_{\odot}$  back onto the compact object (Chevalier 1989). Bethe (1990) gives a simple explanation for the mechanism leading to the reverse shock (see § 5.J of his review). The reverse shock thus fills in the hole that the initial shock leaves.

Meanwhile, the compact object has been imparted a kick velocity, typically  $\sim 100$ – $400 \text{ km s}^{-1}$ , because of asymmetry in the supernova explosion (Dewey & Cordes 1987; Fomalont et al. 1992; Harrison, Lyne, & Anderson 1993). When material sent radially inward by the reverse shock reaches the supernova center, the compact object is no longer there, and the material proceeds to move radially outward. Eventually, all the material will be moving radially outward, though with different speeds depending on distance from the center. Thus, Chevalier (1989) suggests that the compact object will end up comoving with matter traveling at the same speed as its kick velocity. Therefore, we may, to good approximation, consider the compact object to be at rest with respect to its ambient matter and use the Bondi (1952) spherical accretion theory in determining  $\dot{M}$ , the rate of mass accretion onto the compact object. If the accretion were not spherical one would expect effects from density gradients to result in an accretion disk. Efficiency of accretion from a disk is rather large,  $\sim 0.1$ , and luminosity from this would be seen.

The derivation of the formula for  $\dot{M}$  can be found in § 14.3 of Shapiro & Teukolsky's (1983) book; we will sketch the argument. We consider an approximately adiabatic hydrodynamical flow, described by a continuity equation and Euler equation and an adiabatic index  $\Gamma$ . The continuity equation yields

$$\dot{M} = 4\pi r^2 \rho v, \quad (2.1)$$

where  $v$  is the inward radial velocity and the value of  $\dot{M}$  is the same at any distance  $r$  from the compact object. The Euler equation is integrated to yield the Bernoulli equation,

$$\frac{1}{2} v^2 + \frac{1}{\Gamma - 1} c^2 - \frac{GM}{r} = \frac{1}{\Gamma - 1} c_{\infty}^2, \quad (2.2)$$

linking the motion at  $r$  to the conditions at infinity.  $M$  is the mass of the compact object,  $c$  is the local sound speed, and the subscript  $\infty$  refers to the value at infinity. In order to have a smooth, monotonic increase in  $v$  from 0 at infinity to free-fall values near the compact object and to avoid singularities in the

flow, the solution must pass through some critical radius  $r_c$  at which the flow speed equals the local speed of sound. The continuity and Euler equations yield

$$v_c^2 = c_c^2 = \frac{1}{2} \frac{GM}{r_c} \quad (2.3)$$

for the flow speed at the critical, or transonic, radius. Combining equations (2.2) and (2.3) yields

$$v_c^2 = c_c^2 = \left( \frac{2}{5 - 3\Gamma} \right) c_{\infty}^2; \quad r_c = \left( \frac{5 - 3\Gamma}{4} \right) \frac{GM}{c_{\infty}^2}. \quad (2.4)$$

Since, for an ideal Maxwell-Boltzmann gas,

$$\rho = \rho_{\infty} \left( \frac{c}{c_{\infty}} \right)^{2(\Gamma - 1)}, \quad (2.5)$$

we can now determine  $\dot{M}$  in terms of  $\rho_{\infty}$  and  $c_{\infty}$  by using equation (2.1) at  $r_c$ :

$$\dot{M} = 4\pi \lambda (GM)^2 \rho_{\infty} c_{\infty}^{-3}, \quad (2.6)$$

with  $\lambda = 0.25$  for  $\Gamma = 5/3$ , which we take to be the case. We now use that

$$c^2 = \frac{\Gamma k T}{\mu m_H} \quad (2.7)$$

for an ideal Maxwell-Boltzmann gas to replace the  $c_{\infty}$  in equation (2.6) with an expression in  $T_{\infty}$ , the temperature at infinity. Here  $k$  is Boltzmann's constant,  $\mu$  is the mass number of the accreting particles, and  $m_H$  is the mass of a hydrogen atom. We now find that, for  $\rho_{\infty}$  in  $\text{g cm}^{-3}$ ,

$$\dot{M} = 3.46 \times 10^{34} \mu^{3/2} \left( \frac{M}{M_{\odot}} \right)^2 \rho_{\infty} \left( \frac{T}{10^4 \text{ K}} \right)^{-3/2} \text{ g s}^{-1}. \quad (2.8)$$

Now we estimate the ambient gas density. Woosley (1988) finds the final velocity (after being slowed down by the reverse shock) of the CO core to be

$$v_f \approx 6 \times 10^7 \text{ cm s}^{-1}. \quad (2.9)$$

Woosley's calculation fits the light curve of SN 1987A well, including Doppler shifts of the relevant lines, so this  $v_f$  can be taken as an empirical quantity. Also from Woosley (1988), the mass of the CO core (after about  $1.5 M_{\odot}$  is removed in forming the compact object) is about  $2.5 M_{\odot}$ . The compact object will likely find itself in material from the CO core, since the  $v_f$  of the CO core is greater than the typical kick velocities imparted in supernovae. Thus, we estimate the density of the gas around the compact object by taking a  $2.5 M_{\odot}$  CO core to be evenly distributed over a sphere of radius  $R = v_f t$ , where  $t$  is the time after the supernova explosion. We estimate

$$\rho_{\infty} = \frac{2.5 M_{\odot}}{(4/3)\pi(v_f t)^3} = 1.78 \times 10^{-13} \left( \frac{t}{\text{yr}} \right)^{-3} \text{ g cm}^{-3}. \quad (2.10)$$

We next estimate  $T_{\infty}$ . Bethe (1994) finds a temperature of  $T = 70 \text{ keV}$  soon after the supernova explosion for a shock radius  $4 \times 10^9 \text{ cm}$ . Since the shock is radiation dominated, we expect  $T$  to scale as  $R^{-3/4}$  (Bethe & Pizzochero 1990), where  $R$  is the distance of the CO core radius from the supernova center:

$$R = v_f t = 1.88 \times 10^{15} t/\text{yr}. \quad (2.11)$$

Thus, we estimate  $T_\infty$  as

$$\begin{aligned} T_\infty &= 70 \text{ keV} \left( \frac{4 \times 10^9}{1.88 \times 10^{15} t/\text{yr}} \right)^{3/4} \\ &= 3.90 \text{ eV} \left( \frac{t}{\text{yr}} \right)^{3/4} \\ &= 4.52 \times 10^4 \text{ K} \left( \frac{t}{\text{yr}} \right)^{-3/4}. \end{aligned} \quad (2.12)$$

Now we need just determine  $\mu$ . From the Timmes, Woosley, & Weaver (1994) presupernova  $18 M_\odot$  star, we find the composition of matter at the bifurcation radius to be

$$X_O = 0.38; \quad X_{N_2} = 0.48, \quad (2.13)$$

with smaller amounts of other elements. For convenience, we will consider the gas to be oxygen. At temperatures as given in equation (2.12), the oxygen will be triply ionized for several years. Thus,  $\mu = 16/4 = 4$ .

Combining this with equations (2.8), (2.10), and (2.12) and taking  $M = 1.5 M_\odot$ , we find

$$\begin{aligned} \dot{M} &= 1.15 \times 10^{22} (t/\text{yr})^{-15/8} \text{ g s}^{-1} \\ &= 1.81 \times 10^{-4} (t/\text{yr})^{-15/8} M_\odot \text{ yr}^{-1}. \end{aligned} \quad (2.14)$$

### 3. HYPERCRITICAL ACCRETION ONTO A NEUTRON STAR OR BLACK HOLE

The Eddington accretion rate is

$$\dot{M}_{\text{Edd}} = \frac{4\pi c R}{\kappa_\infty} = 5.92 \times 10^{-8} M_\odot \text{ yr}^{-1}, \quad (3.1)$$

where  $R \approx 10^6$  cm is the neutron star radius and  $\kappa_\infty$  is the opacity, which we take to be  $\kappa_\infty \approx 0.1 \text{ cm}^2 \text{ g}^{-1}$ . This is an estimate that applies over a range of temperature and density similar to that present here (Chevalier 1981). The Eddington luminosity  $L_{\text{Edd}}$ , the luminosity for which the pressure of outward traveling photons balances the inward gravity force on the material, is obtained from  $L_{\text{Edd}} = \dot{M}_{\text{Edd}} c^2$ . If  $\dot{M}$  exceeds  $\dot{M}_{\text{Edd}}$ , then some of the accretion energy must be removed by means other than photons. In the present case,

$$\dot{m} \equiv \frac{\dot{M}}{\dot{M}_{\text{Edd}}} = 0.31 \times 10^4 \left( \frac{t}{\text{yr}} \right)^{-15/8}. \quad (3.2)$$

When  $\dot{M}$  exceeds  $\dot{M}_{\text{Edd}}$  by so large a factor, the accretion is called hypercritical, and was considered by Blondin (1986).

Blondin finds a trapping radius  $r_t$  such that photons within  $r_t$  are advected inward with accreting matter faster than they can diffuse outward. We follow his derivation in slightly modified form. We start with the same type of equation as we used in deriving the Bondi  $\dot{M}$ :

$$\dot{M} = 4\pi r^2 \rho v. \quad (3.3)$$

Here, however, we consider any radius  $r$  and take  $v$  to be the free fall velocity, that is,  $v = (2GM/r)^{1/2}$ . We divide equation (3.3) by equation (3.1) to find

$$\rho \kappa_\infty = R r_i^{1/2} r^{-3/2} \dot{m}, \quad (3.4)$$

where  $r_i = 2GM/c^2$  is the Schwarzschild radius;  $r_i$  is about 4.4 km for a  $1.5 M_\odot$  compact object. We now find the optical

depth for electron scattering to be

$$\tau_{ee} \equiv \int_r^\infty \rho \kappa_{ee} dr = 2R\dot{m}(rr_i)^{-1/2}, \quad (3.5)$$

where  $r$  is the distance from the compact object where the photon begins its journey.

From random walk, the time for the photon to diffuse over a distance  $d$  is

$$\tau_{\text{diff}} = \frac{d}{c} \frac{d}{\lambda_{ee}} \cong \frac{d}{c} \tau_{ee}, \quad (3.6)$$

where  $\lambda_{ee}$  is the photon mean free path. The approximation in the last step follows if we assume  $d$  is large enough that most of the scatterings that the photon undergoes on its trip to infinity have already occurred by the time it has traveled distance  $d$ . In order to obtain the photon trapping radius, we set this equal to the dynamical time, which is the time it would take for a piece of matter at position  $r$  to travel distance  $d$  at its instantaneous speed at  $r$ , and approximates the time it takes for a photon to be advected inwards distance  $d$ :

$$t_{\text{dyn}} = \frac{d}{v(r)} = d \left( \frac{r}{2GM} \right)^{1/2}, \quad (3.7)$$

where  $v(r)$  is the free-fall velocity at  $r$ . Using equation (3.5) for  $\tau_{ee}$  and substituting  $r_t$ , the trapping radius, for  $r$  we find

$$r_t = 2R\dot{m}. \quad (3.8)$$

Since solution of the diffusion equation for radial diffusion in three dimensions decreases the diffusion time by a factor of  $\pi^2/3$ , we must likewise decrease  $r_t$  by this factor. We finally obtain

$$r_t = 0.6R\dot{m} = 1.86 \times 10^9 (t/\text{yr})^{-15/8} \text{ cm}. \quad (3.9)$$

Any photon flux emitted much below  $r_t$  is unable to diffuse upstream and thus can have no effect on the luminosity reaching the observer at infinity.

Chevalier and collaborators took into account that neutrinos can carry away accretion energy and developed self-consistent solutions for hypercritical accretion (Chevalier 1989, 1990; Houck & Chevalier 1991). In particular, they find an expression for the radius of the accretion shock in terms of  $\dot{M}$ , for a neutron star. We follow the derivation in Chevalier (1989, p. 854) with small modification. We first derive an expression for  $p_{\text{sh}}$ , the pressure at the surface of the neutron star, in terms of  $\dot{M}$  and  $r_{\text{sh}}$ , the shock radius. We then examine neutrino cooling near the surface of the neutron star, producing an equation in terms of  $p_{\text{sh}}$ . Insertion of our expression for  $p_{\text{sh}}$  gives  $r_{\text{sh}}$  in terms of  $\dot{M}$ .

Since the pressure is radiation dominated, the accretion envelope forms an  $n = 3$  ( $\Gamma = 4/3$ ) polytrope. Thus, inside the shock

$$\rho = \rho_{\text{sh}} \left( \frac{r}{r_{\text{sh}}} \right)^{-3}; \quad p = p_{\text{sh}} \left( \frac{r}{r_{\text{sh}}} \right)^{-4}; \quad v = v_{\text{sh}} \left( \frac{r}{r_{\text{sh}}} \right), \quad (3.10)$$

where the subscript sh refers to the value at the shock front. Because of the adiabatic compression by factor  $(\Gamma + 1)/(\Gamma - 1)$ ,

$$\rho_{\text{sh}} = 7\rho_0, \quad (3.11)$$

where  $\rho_0$  is the density just outside the shock front. We neglected the (small) decrease in  $\Gamma$  because of increased ionization of the material going through the shock. As in Blondin

(1986),  $\rho_0$  is calculated as follows:

$$\rho_0 = \frac{\dot{M}}{4\pi r_{sh}^2 v_{in}}, \quad (3.12)$$

where  $v_{in}$  is the free-fall velocity at the shock radius. From conservation of mass flow across the shock front,

$$v_{sh} = -\frac{1}{2}v_{in}. \quad (3.13)$$

Thus, the kinetic energy of the accreting matter is diminished by a factor of 49; that is, it is almost entirely converted into thermal energy, so we can estimate the thermal energy density as

$$\epsilon_{sh} \approx \frac{7}{2}\rho_0 v_{in}^2, \quad (3.14)$$

the factor of 7 entering because of the compression (eq. [3.11]). In this derivation we neglected  $v_{sh}$  as compared with  $v_{in}$ . The small correction that a complete consideration would produce would tend to cancel the small correction that would be generated had we taken account of the decrease in  $\Gamma$  due to ionization across the shock in obtaining equation (3.11). Since the pressure is radiation dominated,

$$p_{sh} = \frac{1}{3}\epsilon_{sh} \approx \frac{7}{6}\rho_0 v_{in}^2. \quad (3.15)$$

Inserting equation (3.12) and  $(2GM/r_{sh})^{1/2}$  for  $v_{in}$ , and using the second of equations (3.10), we find the pressure at the surface of the neutron star

$$p_{sh} = 1.86 \times 10^{-12} \dot{M} r_{sh}^{3/2} \text{ dyn cm}^{-2}, \quad (3.16)$$

where  $\dot{M}$  is expressed in  $\text{g s}^{-1}$  and  $r_{sh}$  in cm. We took the radius of the neutron star to be  $\approx 10$  km.

The energy loss by neutrino pair production per unit volume is (Dicus 1972)

$$\dot{e}_n = 1.06 \times 10^{25} T^9 C\left(\frac{\mu_e}{T}\right) \text{ ergs cm}^{-3} \text{ s}^{-1}, \quad (3.17)$$

where  $\mu_e$  is the electron chemical potential,  $T$  is in MeV, and  $C(x)$  is a slowly varying function of  $x$  which can be computed from the paper of Dicus. For  $x = 0$ ,  $C = 0.92$ ; we shall use this value, because the electrons are not very degenerate. In the region where  $\dot{e}_n$  is operative,  $T \sim 1$  MeV, so that  $e^+, e^-$  pairs, as well as photons, contribute to the radiation pressure. With  $T$  in MeV, the photon blackbody energy density is

$$w = 1.37 \times 10^{26} T^4 \text{ ergs cm}^{-3}. \quad (3.18)$$

Inclusion of the  $e^+, e^-$  pairs multiplies this by a factor of 11/4, and we divide by three to obtain the pressure

$$p = 1.26 \times 10^{26} T^4 \text{ ergs cm}^{-3}. \quad (3.19)$$

Combining equations (3.17) and (3.19) yields

$$\dot{e}_n = 1.83 \times 10^{-34} p^{2.25}, \quad (3.20)$$

where  $\dot{e}_n$  is in  $\text{ergs cm}^{-3} \text{ s}^{-1}$  when  $p$  is in  $\text{ergs cm}^{-3}$ . The neutrino cooling is taken to occur within a pressure scale height  $r_{ns}/4$  of the neutron star, or in a volume of  $\approx \pi r_{ns}^3$ . Energy conservation gives

$$\pi r_{ns}^3 \times 1.83 \times 10^{-34} p^{2.25} = \frac{G M \dot{M}}{r_{ns}}, \quad (3.21)$$

with, again, everything in cgs units. Inserting equation (3.16) into equation (3.21), we solve for  $r_{sh}$ :

$$r_{sh} \cong 6.4 \times 10^8 \left( \frac{\dot{M}}{M_\odot \text{ yr}^{-1}} \right)^{-10/27} \text{ cm}. \quad (3.22)$$

The power  $-10/27 = -0.370$  is the same as that obtained by Houck & Chevalier (1991) using the accurate neutrino cooling function, and not that of Chevalier (1989).

The detailed calculation of Houck & Chevalier (1991) finds the only substantial correction to our schematic calculation to arise from general relativity, which can be taken into account by multiplying the expression for  $r_{sh}$  by 0.4. Thus,

$$r_{sh} \cong 2.6 \times 10^8 \left( \frac{\dot{M}}{M_\odot \text{ yr}^{-1}} \right)^{-0.370} \text{ cm} = 6.3 \times 10^9 \left( \frac{\dot{M}}{\text{yr}} \right)^{0.694} \text{ cm}, \quad (3.23)$$

where we have used equation (2.14) for  $\dot{M}$ .

From equation (3.9),  $r_{tr}$  decreases with time, while from equation (3.23),  $r_{sh}$  increases with time. Thus, while  $r_{sh}$  starts off smaller than  $r_{tr}$ , at some time  $t_{cr}$  the two radii will be equal, and thereafter  $r_{sh} > r_{tr}$ . We find

$$t_{cr} = 0.62 \text{ yr}. \quad (3.24)$$

Our  $t_{cr}$  is somewhat smaller than the  $t_{cr} = 1$  yr of Chevalier (1990).

As the shock radius increases beyond  $r_{tr}$  for  $t > t_{cr}$ , the luminosity will increase to  $L_{Edd} = 3.8 \times 10^{38} \text{ ergs s}^{-1}$  (Chevalier 1990). The gas is rich in metals, as in Type Ia supernovae. Using the Type Ia opacity, the  $0.1 \text{ cm}^2 \text{ g}^{-1}$  we used here, Chevalier (1989) estimates the optical depth is reached at a radius  $r_{ph} \sim 2.7 \times 10^{13} \text{ cm}$  with an effective temperature at  $r_{ph}$  of  $\sim 4000$  K. Although the luminosity is produced by neutron star accretion, the temperature of the escaping radiation is low, a bit below solar.

Although the radiative effects of a luminosity near  $L_{Edd}$  will reduce the infall velocity, and thus, the mass infall rate, the Eddington limiting luminosity should persist for some time, down to  $\dot{M} \cong 2 \times 10^{-5} M_\odot \text{ yr}^{-1}$ , where steady state solutions are no longer possible because of strong heating of the infalling gas (Houck & Chevalier 1991). From equation (2.14) we estimate that the time to come down to this  $\dot{M}$  is  $\sim 3$  yr.

At about this time the neutrino dominated steady accretion phase ends. A mass of  $10^{-4}$  to  $10^{-3} M_\odot$  is bound to the neutron star. One might expect the accretion of this mass onto the neutron star to produce a luminosity of order  $L_{Edd}$  for many years if it is not ejected by some dynamical instability. In the latter case the ejection should have observational consequences.

Luminosity from accretion at the level of  $L_{Edd}$  is excluded by observation of the light curve (Fig. 1). Sources additional to radioactive decay are excluded at a level of  $\sim 6 \times 10^{36} \text{ ergs s}^{-1}$  (Suntzeff et al. 1992).

For the case of spherical accretion onto a black hole, there is a comprehensive review by Park (1990). For the case of  $\dot{m} \sim 10^4$  relevant to us, Park finds that the radiation is trapped inside  $r_{tr}$ , as in the case of the neutron star. Only the internal energy produced by the  $p dV$  work done on the infalling matter outside of the trapping radius can be radiated. Park's curves, both from his calculations and those of Blondin (1986), go up to  $\dot{m} = 10^4$ ; continuing them to our values of  $\dot{m}$  gives

$$L \sim 10^{34} - 10^{35} \text{ ergs s}^{-1}. \quad (3.25)$$

Because the large distance matter is isothermal, the effective temperature of the radiation will be  $\sim 10^4$  K. This would clearly not be observable at present, as it would be overwhelmed by the luminosity from radioactivity, in particular, from  $^{44}\text{Ti}$ .

#### 4. CONCLUSIONS

Formation of an accretion disk about the compact object formed in SN 1987A is ruled out, since it would contribute to the visible light curve at the level of the Eddington luminosity, and this is not observed in the light curve. A weakly magnetized neutron star accreting material ejected during the supernova also should contribute to the visible light curve at about the Eddington luminosity. A black hole, however, should only contribute to the visible light curve at a level  $\sim 10^{34}\text{--}10^{35}$  ergs  $\text{s}^{-1}$ , which would be rendered unobservable because of the much greater luminosity generated by the decay of  $^{44}\text{Ti}$ .

Our considerations support the suggestion (Brown et al. 1992) that the compact object formed in SN 1987A was a black hole. Recently, Brown & Bethe (1994) showed that there is a good possibility that stars of main-sequence masses of 18–30  $M_\odot$  first explode, returning matter to the galaxy, and then later go into black holes. In their argument, SN 1987A, with progenitor mass  $\sim 18 M_\odot$ , is near the lower limit, and no strong case could be made from the calculations as to whether it went into a neutron star or a black hole.

Calculations by Thielemann, Nomoto, & Hashimoto (1992) estimate the mass of the compact object formed in SN 1987A to be within the range of gravitational masses 1.30–1.50  $M_\odot$ , for a main-sequence mass of 18  $M_\odot$ . The higher values follow from accretion during the formation of the delayed shock. The

delayed scenario is now believed to be the correct one. Therefore, we believe gravitational masses near 1.5  $M_\odot$  to pertain to SN 1987A. This is supported by the work of Timmes, Woosley, & Weaver (1994), who find a gravitational mass of 1.49  $M_\odot$  for the compact core of an 18  $M_\odot$  main-sequence star. Bethe (1990, § 8.B) gives a general argument from Aufderheide, Woosley, & Weaver (1990) that the iron core for an 18  $M_\odot$  star will be of mass 1.32 or 1.33  $M_\odot$  upon collapse. Adding the  $(\Delta M)_{\text{acc}} \approx 0.17 M_\odot$  of Thielemann et al. (1992) brings the core mass up to  $\sim 1.5 M_\odot$ . Note that it is the considerable amount of accretion which brings it up to this value.

Our considerations here support the conclusion that the maximum neutron star mass is  $M_{\text{max}} \sim 1.5 M_\odot$ , with a possible uncertainty of  $\pm 0.1 M_\odot$  or  $-0.05 M_\odot$ , the lower limit being given by the 1.44  $M_\odot$  measured mass of the larger “neutron” star in the binary pulsar PSR 1913+16 (Taylor & Weisberg 1989). Although we should not draw conclusions from astronomical observations about processes that can be measured in the laboratory (W. Fowler 1993, private communication), nucleon stars are our only laboratory for investigation of the equation of state of dense matter. Most relevant for this equation of state is  $M_{\text{max}}$ .

We would like to thank Hans Bethe for many discussions, Stu Shapiro for advice and help, and, especially, Roger Chevalier for his patient tuition on the accretion mechanism. Both he and Jim Lattimer gave helpful criticism on our manuscript. This work was begun in discussion with Lars Bildsten, and we gratefully acknowledge his help. We thank Stan Woosley for the 18  $M_\odot$  progenitor before publication (Timmes et al. 1994) and for helpful criticism.

#### REFERENCES

- Aufderheide, M. B., Woosley, S. E., & Weaver, T. A. 1990, preprint
- Bethe, H. A. 1990, Rev. Mod. Phys., 62, 801
- . 1993, ApJ, 419, 197
- Bethe, H. A., & Pizzochero, P. 1990, ApJ, 350, L33
- Blondin, J. M. 1986, ApJ, 308, 755
- Bondi, H. 1952, MNRAS, 112, 195
- Brown, G. E., & Bethe, H. A. 1994, ApJ, 423, 659
- Brown, G. E., Bruenn, S. W., & Wheeler, J. C. 1992, Comm. Astrophys., 16, 153
- Chevalier, R. A. 1981, ApJ, 246, 267
- . 1989, ApJ, 346, 847
- . 1990, in SN 1987A and Other Supernovae, ed. J. J. Danziger (Elba: ESO)
- . 1992, Nature, 360, 628
- Dewey, R. J., & Cordes, J. M. 1987, ApJ, 321, 780
- Dicus, D. A. 1972, Phys. Rev. D, 6, 941
- Fomalont, E. B., Goss, W. M., Lyne, A. G., Manchester, R. N., & Jusitanon, K. 1992, MNRAS, 258, 497
- Harrison, P. A., Lyne, A. G., & Anderson, B. 1993, MNRAS, 261, 113
- Helfand, D. J., & Becker, R. H. 1984, Nature, 307, 215
- Houck, J. C., & Chevalier, R. A. 1991, ApJ, 376, 234
- Illarionov, A. F., & Sunyaev, R. A. 1985, A&A, 39, 185
- Nomoto, K., Shigeyama, T., Kumagai, S., Itoh, M., Nishimura, J., Hashimoto, M., Saio, M., & Kato, M. 1993, in Physics of Neutron Stars and Black Holes, ed. Y. Tanaka (Universal Academy)
- Park, M.-G. 1990, ApJ, 354, 64
- Shapiro, S. L., & Teukolsky, S. A. 1983, Black Holes, White Dwarfs, and Neutron Stars (New York: John Wiley)
- Suntzeff, N. B., Phillips, M. M., Elias, J. H., DePoy, D. L., & Walker, A. R. 1992, ApJ, 384, L33
- Taylor, J. H., & Weisberg, J. M. 1989, ApJ, 345, 434
- Thielemann, F.-K., Nomoto, K., & Hashimoto, M. 1992, in Origins and Evolutions of the Elements, in Honor of H. Reeves' 60th Birthday, ed. N. Prantzos & E. Vangioni-Flam (Cambridge: Cambridge Univ. Press)
- Timmes, F. X., Woosley, S. E., & Weaver, T. A. 1994, ApJS, submitted
- Woosley, S. E. 1988, ApJ, 330, 218

This page is intentionally left blank

---

## Chapter 4

# A Scenario for a Large Number of Low-Mass Black Holes in the Galaxy

G.E. Brown and H.A. Bethe

Reprinted with permission from *Astrophysical Journal*, 423 (1994) 659–664  
Copyright (1994) by American Astronomical Society

### Commentary by G.E. Brown

In this paper we introduced the concept of low-mass black hole, one in which the helium envelope has first been blown off by an explosion and a delayed collapse of the compact object into a black hole follows. The black hole mass is then roughly the mass of the Fe core of the progenitor.

In the autumn of 1993 I phoned Hans Bethe one evening — we often talk on the phone — and told him that SN 1987A had ended with the compact object going into a low-mass black hole. This was a somewhat revolutionary idea, because both the Woosley and Nomoto groups had calculated the Fe core mass of the  $18M_{\odot}$  progenitor to be  $\sim 1.45\text{--}1.55M_{\odot}$ , even with some allowance for fallback. From Paper 3 with Joe Weingartner I was sure that no neutron star was left. Furthermore, I had graduate student Vesteinn Thorsson who, with Prakash and Lattimer (Thorsson *et al.*, 1994), had worked out the maximum neutron star mass that could be supported in case of kaon condensation in neutron stars, which I describe below. Kaon condensation softens the equation of state a lot. Thorsson *et al.* had worked out the strangeness condensation for three sets of parameters, the most important parameter being the strangeness content of the nucleon. These developments are described in detail in our Appendix B, “Kaon Condensation in Dense Stellar Matter”, by C.-H. Lee. I believed, on the basis of results from lattice calculations, the intermediate value, that the strangeness content of the nucleon is about 1/3 that of the up or down quark content. Using this and a compression modulus  $K \sim 210$  MeV for nuclear matter, I could deduce from the Thorsson *et al.* work that the maximum neutron star mass should be  $\sim 1.5M_{\odot}$ , substantially lower than most people believed.

I remember that we were walking up a trail on a hike near Santa Barbara in 1993, when I explained the idea of kaon condensation to Hans. He immediately understood the idea, saying, “You mean that you’re squeezing electrons into  $K^-$ -mesons?” This was amusing, in

that structure in electrons, at distances of many orders of magnitude less than the measured  $K^-$  radius of  $\sim 0.6$  fermi, had still not been discovered. Electrons are tiny. In fact the squeezing of electrons into  $K^-$ -mesons does not go directly, but in a roundabout way with the participation of nucleons. Still, there is equilibration, so that when all is said and done

$$\mu_n - \mu_p = \mu_{K^-},$$

where the  $\mu$ 's are chemical potentials.

One evening I phoned Hans to tell him that SN 1987A had gone into a black hole. Hans listened carefully to my arguments over the phone, but objected to my conclusion, citing work of the Nomoto group that a pulsar could still be accommodated, the amount of radiation not being too large to exceed the observed light curve. I was somewhat dismayed and went back to my work. An hour later, my wife, Betty, called me to the phone. Hans had called. All he said, but he said it emphatically, was "You're right!" and he hung up. (Our telephone conversations have never been unnecessarily long.)

The main idea of strangeness condensation is as follows: Whereas the  $K^-$  mesons has a mass of 495 MeV in free space, its mass comes down in a high density medium, roughly by 100 MeV for each increase of the density by nuclear matter density  $\rho_0$  (the density in the center of heavy nuclei). Thus, at a density of  $3\rho_0$ ,  $m_{K^-}^* \sim 200$  MeV, somewhat lower than the electron chemical potential  $\mu_e$  in a neutron star at that density. At the point where  $m_{K^-}^*(\rho) = \mu_e$  it becomes energetically favorable for electrons to change into  $K^-$ -mesons. The latter are bosons, so they can go into a Bose condensate of zero momentum. This substantially softens the equation of state, decreasing the pressure, and, therefore, the maximum neutron star mass. This is explained in great detail in the Appendix.

A short summary of our results for maximum neutron star mass is given in Appendix B. The experimental confirmation obtained in heavy ion collisions that create matter up to densities of  $3\rho_0$  shows a substantial enhancement in  $K^-$  production because of its reduced mass is included here. The heavy ion work is a large activity, producing very exciting results, which we shall describe in the Appendix.

When we proposed the maximum neutron star mass of  $1.5M_\odot$ , the observed masses of radio pulsars were all over the place, many with large error bars. Since then Thorsett & Chakrabarty (1999) have summarized measurements of the masses of neutron stars in binaries with degenerate (neutron star or white dwarf) companions. We show their results in the figure.

The measurements of the neutron star mass in the high-mass X-ray binary Vela X-1 give (Barziv *et al.* 2000) a higher mass,

$$M_{NS} = 1.87^{+0.23}_{-0.17} M_\odot.$$

The situation in which the neutron star mass is measured here is much more complicated than that where the companion is degenerate; i.e., small and relatively structureless. The massive companion star is floppy, the center of gravity of the system is inside this star, and the center of light is substantially different from the center of the star because of the

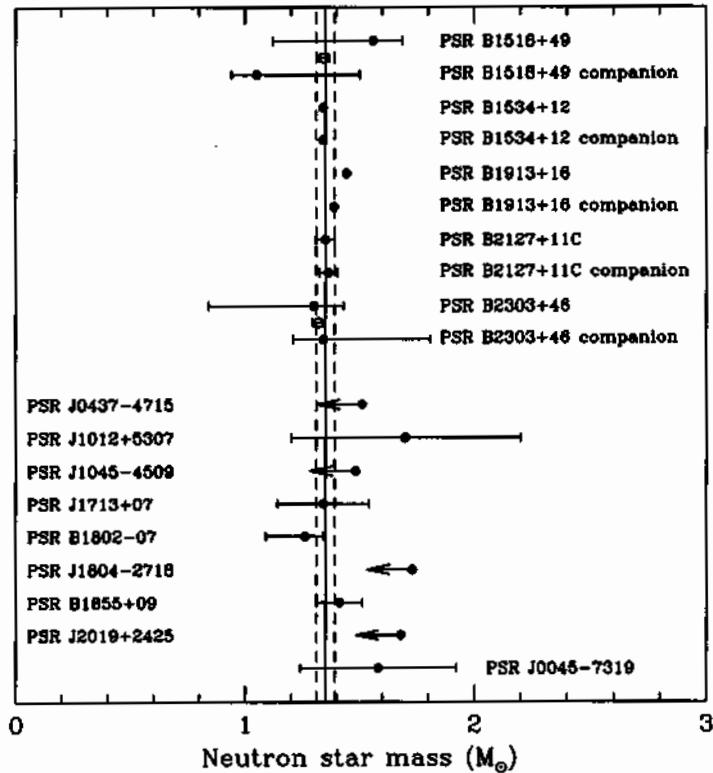


Figure. Neutron star masses from observations of radio pulsar system (Thorsett & Chakrabarty 1999). All error bars indicate central 68% confidence limits, except upper limits are one-sided 95% confidence limits. Five double neutron star systems are shown at the top of the diagram. In two cases, the average neutron star mass in a system is known with much better accuracy than the individual masses; these average masses are indicated with open circles. Eight neutron star-white dwarf binaries are shown in the center of the diagram, and one neutron star-main sequence star binary is shown at the bottom. Vertical lines are drawn at  $m = 1.35 \pm 0.04 M_{\odot}$ . It should be noted that PSR B2303+46 has since been shown to have a white dwarf companion. Furthermore, in J2019+2425 the upper limit on the neutron star mass has been brought down to  $1.51 M_{\odot}$  (Nice, Splaver & Stairs 2001). As noted in this paper, neutron star-white dwarf systems go through an extended period of mass transfer during which the secondary loses several tenths of a solar mass of matter, at least some of which is accreted onto the neutron star. Thus one might expect the neutron stars to be more massive than those in neutron star-neutron star binaries.

strong gravitational attraction by the neutron star. These difficulties are well known to the observers involved, who should be working through the necessary corrections. These have been laid out in detail by Zuiderwijk *et al.* (1977). In present observations the radial velocity curve has wide variations from night to night, which are included in the statistical analyses of the data. However, until the origin of these variations is understood dynamically, we are not comfortable with the mass determination of the neutron star.

Whereas the formation of the low-mass black hole is described well in Paper 4, we later made more precise how high-mass black holes are formed, in Papers 14, 15 and especially in Papers 20 and 21. In the paper commented on here we concluded (Eq. (3)) that return

of matter to the Galaxy must be cut off at a mass of

$$M_{\text{cutoff}} = 25 \pm 5 M_{\odot}.$$

Paper 21 shows that  $M_{\text{cutoff}} = 19\text{--}20 M_{\odot}$ , rather precisely, and also that there is only a narrow mass range of  $\sim 18\text{--}20 M_{\odot}$  in which low-mass black holes can be made. All masses here are zero age main sequence (ZAMS) masses.

As far as production of elements such as oxygen, etc. are concerned, this production is cut off at the  $19\text{--}20 M_{\odot}$ , but, because of high wind loss, such production takes place again in a region of ZAMS masses  $\sim 40 M_{\odot}$ . Thus, there is a bimodal production, from  $\sim 10\text{--}19 M_{\odot}$  and then, again, a region around  $40 M_{\odot}$ . The idea of a bimodal production of elements is correctly outlined in Paper 13 by Bethe & Brown. This paper still utilized He-star wind losses which were  $\sim 3$  times too large. Better wind losses were employed in Paper 21, without changing qualitatively any of the Paper 13 scenario. Quantitative calculations using our bimodal element production remain to be carried out.

It is amusing that Woosley & Weaver (1988) arrived at a value for the  $^{12}\text{C}(\alpha, \gamma)^{16}\text{O}$  rate, which is very important for determining element abundances, by matching these latter from the explosion of a ZAMS  $25 M_{\odot}$  star to the observed ones. The reaction rate they arrived at was

$$^{12}\text{C}(\alpha, \gamma)^{16}\text{O} \text{ rate} = 175 \text{ keV barns},$$

quoted for a reaction energy of 300 keV. However, we established in Paper 13, and more definitely in Paper 20, that a ZAMS  $25 M_{\odot}$  star evolves into a high-mass black hole. None the less, Paper 20 will show that the  $^{12}\text{C}(\alpha, \gamma)^{16}\text{O}$  rate of Woosley & Weaver is correct, to an accuracy of a few keV barns. We believe this determination (of 175 keV barns) from evolutionary arguments to be at least as accurate as that from experiments. The latter determine the E1 transition well, but the E2 contribution, of about the same magnitude, is known only roughly.

Recent developments on kaon condensation are discussed in Appendix D and its commentary.

## References

- (1) Barziv *et al.* (2000), as reported by van Kerkwijk (2000); Van Kerkwijk, M.H. (2000), in *Proc. ESO Workshop on Black Holes in Binaries and Galactic Nuclei*, ed. L. Kaper, E.P.J. van den Heuvel, and P.A. Woudt (Berlin, Springer), in press (astro-ph/0001077).
- (2) Nice, D.J., Splaver, E.M., and Stairs, I.H. (2001), *Astrophysical Journal* **549**, 516.
- (3) Thorsett, S.E., and Chakrabarty D. (1999), *Astrophysical Journal* **512**, 288.
- (4) Thorsson, V., Prakash, M., and Lattimer, L. (1994), *Nuclear Physics A* **572**, 693.
- (5) Woosley, S.E., and Weaver, T.A. (1988), *Physics Reports* **163**, 79.

- (6) Zuiderwijk, E.J., Hammerschlag-Hensberge, G., van Paradijs, J., Sterken, C., and Hensberge, H. (1977), *Astronomy & Astrophysics* **54**, 167; Van Paradijs, J., Takens, R., and Zuiderwijk, E. (1977), *Astronomy & Astrophysics* **57**, 221; Van Paradijs, J., Zuiderwijk, E.J., Takens, R.J., Hammerschlag-Hensberge, G., van den Heuvel, E.P.J., and de Loore, C. (1977), *Astronomy & Astrophysics Supplement* **30**, 195.

This page is intentionally left blank

## A SCENARIO FOR A LARGE NUMBER OF LOW-MASS BLACK HOLES IN THE GALAXY

G. E. BROWN<sup>1</sup> AND H. A. BETHE<sup>2</sup>

California Institute of Technology, W. K. Kellogg Radiation Laboratory, 106-38, Pasadena, CA 91125

Received 1993 March 24; accepted 1993 July 27

### ABSTRACT

We point out that stars within a fairly large range of masses, roughly  $18\text{--}30 M_{\odot}$ , can both explode as supernovae (and give off neutrinos as observed in SN 1987A) and then go into black holes. The masses of the black holes so formed are only slightly above  $1.5 M_{\odot}$ , the maximum mass for compact cores, according to our arguments. These masses are substantially smaller than those of the best candidates for black holes observed thus far. We discuss the possibility that SN 1987A produced one of these small black holes.

We review the requirements that observed nucleosynthesis puts on theoretical nucleosynthesis. Especially the empirical value of  $\Delta Y/\Delta Z = 4 \pm 1.3$  gives a sensitive determination of the mass at which nucleosynthesis must be cut off, by heavier stars collapsing into black holes without returning matter to the Galaxy, because otherwise heavy stars would produce chiefly metals and unduly lower the calculated value of  $\Delta Y/\Delta Z$ . We arrive at a value of  $M_{\text{cutoff}} = 25 \pm 5 M_{\odot}$  for the mass at which nucleosynthesis must be cut off.

We show that the introduction of kaon condensation sufficiently softens the equation of state of dense matter, so that compact cores of mass greater than  $M_{\text{max}} \approx 1.5 M_{\odot}$  will not be stable. For the first  $\gtrsim 12$  s, however, compact cores up to  $\sim 1.84 M_{\odot}$  are stabilized; this gives the nucleosynthesis mass cutoff as  $\sim 30 M_{\odot}$ , at the upper end of our limit from nucleosynthetic demands. Our soft equation of state extends black hole production to stars of lower mass than previously estimated, and, therefore, increases the estimated number of black holes by an order of magnitude or more, to  $\sim 10^9$  in the Galaxy.

**Subject headings:** black hole physics — Galaxy: stellar content — nuclear reactions, nucleosynthesis, abundances — stars: evolution

### 1. INTRODUCTION

It was suggested (Brown 1988) that the maximum neutron star mass could be  $M_{\text{max}} \sim 1.5 M_{\odot}$ , not much larger than the  $1.442 \pm 0.03 M_{\odot}$  of the heavier pulsar (Taylor & Weisberg 1989) in PSR 1913+16. This suggestion was based on arguments that the high-density equation of state of dense matter is “softer” than generally thought. Recently (Brown 1992), with the implementation of the  $K^-$  condensation in the equation of state, this suggestion has become more compelling, as we shall discuss later. Rather than “neutron star” in what follows we shall use the term nucleon star, because with  $K^-$  condensation, nearly as many protons as neutrons are present.

Brown, Bruenn, & Wheeler (1992b) asked, “Is There a Black Hole in Supernova 1987A?” They suggested that the late time accretion or additional matter deposited by the reverse shock might send the core of the neutron star in SN 1987A over the limit of  $\sim 1.50 M_{\odot}$  into a black hole after the nucleosynthesis and 12 s neutrino emission was completed. An alternative discussion of the late time accretion has since been given by Bethe (1993), which makes this scenario more definite.

For some time Twarog & Wheeler (1982, 1987) and Wheeler, Sneden, & Truran (1989) have argued that heavy stars at the end of their life must disappear from the Galaxy without ejecting their matter in a supernova explosion, to avoid over-production of oxygen. They call this the “death function” and propose that masses above about  $24\text{--}28 M_{\odot}$  disappear and presumably turn into black holes. These arguments were made more definite by Maeder (1992), who showed that stellar yields integrated over the mass spectrum are able to reproduce the

observed  $\Delta Y/\Delta Z$  (the relative helium to metal enrichment) for low metallicities, only if there are black holes formed above  $20\text{--}25 M_{\odot}$ . Here  $\Delta Y$  is the increase of  $Y$ , the weight fraction of He, from the primordial value after the big bang, to its actual value,  $\Delta Z$ , the same for “metals” where the primordial value is zero.

In this note we wish to make the scenario of black hole formation more definite, pointing out that stars within a fairly large range of masses can first explode as supernovae and give off neutrons, as observed in SN 1987A, and then go into black holes. The masses of the black holes so formed will not greatly exceed the  $M_{\text{max}} \sim 1.5 M_{\odot}$  for nucleon stars, and, therefore, will be substantially smaller than those of the best candidates for black holes observed thus far.

Our paper will be organized in the following way. In § 2, we shall discuss the possibility that SN 1987A, as well as progenitors of similar mass, have gone into black holes. In § 3, we shall review the nucleosynthetic demands for a “death function” that sets in at relatively small stellar masses. In § 4, we discuss the soft equation of state of dense matter that results once a kaon condensate is introduced at several times the nuclear matter density. In § 5 we discuss the postshock mechanism for the formation of black holes. In § 6, we discuss the implications of our new scenario for the estimated number of black holes in the Galaxy.

### 2. NEUTRON STARS OR BLACK HOLES?

Helfand & Becker (1984), from radio, optical, X-ray, and  $\gamma$ -ray surveys of supernova remnants, came to the startling conclusion that nearly half of the supernovae in the Galaxy leave no observable compact remnant. In terms of historical supernovae in our galaxy, this was not necessarily surprising, since about half of them have been Type Ia, which are not expected to leave compact remnants. However, as discussed in

<sup>1</sup> Permanent address: Physics Department, State University of New York, Stony Brook, NY 11794.

<sup>2</sup> Permanent address: Laboratory of Nuclear Studies, Cornell University, Ithaca, NY 14853.

detail by van den Bergh, McClure, & Evans (1987), the observed sample of supernovae in our Galaxy has been strongly biased against Type II supernovae (SN II's) because most of them are at low galactic latitude so their light is heavily absorbed on its way to us. These authors estimate from extragalactic studies (the supernovae rate in Shapley-Ames galaxies) that SN II's outnumber SN I's by a factor of about 6. (Type Ib supernovae, presumed to have heavy progenitors, are included with SN II in this estimate.) We will return to a discussion of Type Ib supernovae at the end of § 5.) In any case, we can say that the SN II's outnumber the SN I's by a factor of several. Thus the Helfand-Becker conclusion can be understood only if almost half of the explosions from SN II's do not form compact remnants. We assume here the only viable alternative to be black holes. The conclusion—that roughly as many black holes as compact remnants are formed—is indeed startling, and we will spend the rest of this note to substantiate this possibility.

There is circumstantial evidence that black hole formation is associated with the supernova explosion Cas A. There is no sign of a nucleon star in Cas A as a thermal source of X-rays, a pulsar, or an X-ray synchrotron nebula. The fast-moving knots in Cas A show the Mg/CNO ratio to be down by a factor of 2 from solar while Si, S, Ar, and Ca are enhanced compared to CNO by a factor of a few (Chevalier & Kirshner 1979). This is roughly consistent with the ejecta expected from a  $20 M_{\odot}$  star (Woosley & Weaver 1986). Thus Cas A is a good candidate for a star of about the same mass as the progenitor of SN 1987A that left a black hole remnant after exploding and producing nucleosynthesis. It should be noted that El Eid & Prantzos (1988) have suggested another scenario for Cas A. In it, the supernova results from disruption of a very massive star due to pair creation instability.

The light curve in SN 1987A is still being powered by radioactive decay and an additional source of luminosity greater than  $8 \times 10^{36} \text{ ergs s}^{-1}$  is now ruled out by the bolometric light curve at day 1500 (Suntzeff et al. 1992). This lies far below early modeling of an interior pulsar which was assumed to deposit energy in the innermost layers at a constant rate of  $8.5 \times 10^{41} \text{ ergs s}^{-1}$  (Shigeyama et al. 1987), and means that if a pulsar is present, its magnetic field is very weak and the star is rotating slowly. A more precise determination of the nature of the compact remnant must await the clearing of debris so that soft X-rays from the central star can be detected, but it seems reasonable to suggest that it is a black hole (Brown et al. 1992b).

Estimates of the compact core mass formed in the delayed explosion in SN 1987A (Mayle & Wilson 1988; Thielemann, Nomoto, & Hashimoto 1992) place it close to the  $M_{\max} \sim 1.5 M_{\odot}$ , so that the late time accretion, to be discussed in the § 5, could have pushed it over the limit.

All measured nucleon star masses, with the exception of Vela X-1 (Lamb 1991; Thorsett et al. 1993), are consistent with  $M_{\max} \sim 1.5 M_{\odot}$ , and the rather large 2  $\sigma$  error bar on Vela X-1 stretches below  $1.6 M_{\odot}$ . Thus, we believe the  $M_{\max} \sim 1.5 M_{\odot}$  to be reasonable from the standpoint of nucleon star masses, although Vela X-1 is still above our limit. It should be noted that most equations of state in the literature have  $M_{\max}$  between 2 and  $3 M_{\odot}$ , easily stabilizing Vela X-1.

### 3. NUCLEOSYNTHETIC DEMANDS ON THE STELLAR DEATH FUNCTION

As noted in § 1, Twarog & Wheeler (1982, 1987) have raised the problem that oxygen is overproduced unless there is a

cutoff at about  $\sim 24-28 M_{\odot}$  on the main-sequence mass, above which the star does not contribute its content to the galaxy at the end of its life. This cutoff is called the "death function" (Wheeler et al. 1989). Recent work by Maeder (1992) has sharpened the nucleosynthetic requirements. Maeder focused on the ratio  $\Delta Y/\Delta Z$  of the (relative) helium-to-metal enrichment. This ratio has been measured in extragalactic low-metallicity H II regions by Pagel et al. (1992), who give  $\Delta Y/\Delta Z$  in the range of 3-6, with a preferred value between 4 and 5.

The quantity  $\Delta Y/\Delta Z$  is sensitive to the maximum mass of a star that will give the products of the nucleosynthesis back to the galaxy. Stars of mass  $M \lesssim 8 M_{\odot}$  shed their mass in the giant stage, mostly in the form of helium, before becoming white dwarfs. (They also shed hydrogen, but this is irrelevant for  $\Delta Y/\Delta Z$ .) Some helium, but mostly metals, are produced in heavier stars. The amount of produced oxygen, which comprises the main part of the metals, increases roughly linearly with the mass of the exploding star. If all the material in heavy stars, at the end of their lives, were to be given back to the Galaxy by supernova explosions, there would be too much metals, and  $\Delta Y/\Delta Z$  would be between 1 and 2 (Maeder 1992). Therefore, many of the heavy stars must simply disappear; i.e., become black holes; we assume that all above a certain cutoff mass do so. The measured  $\Delta Y/\Delta Z$  determines rather precisely where the cutoff should be made.

For low-metallicity regions, in which the influence of stellar winds is small, this ratio provides a stringent requirement on the cutoff. Maeder finds that stellar yields integrated over the mass spectrum are able to reproduce the observed  $\Delta Y/\Delta Z$  if black holes are formed above  $20-25 M_{\odot}$ . (Note that this estimate refers to the lower limit on collapses which eject photosynthesized material.)

The initial work of Maeder suffered from normalization problems, resulting from his choice of a minimum mass  $m_1 = 1 M_{\odot}$  in his equations. Most of the mass is, however, in stars with masses less than solar. Taking this into account (Maeder 1993) changes the absolute yields considerably. However, the relative enrichment  $\Delta Y/\Delta Z$  is not affected by the normalization choice, so the main conclusion of Maeder's earlier paper—that the observed  $\Delta Y/\Delta Z$  ratios in low-metallicity regions can be accounted for only with a mass cutoff of  $M_{\text{cutoff}} \sim 20 M_{\odot}$ —is unchanged.

As we will point out, over a region of about  $10 M_{\odot}$ , collapsing stars eject photosynthesized material and then go into light-mass black holes. Thus our limit for black hole formation will lie somewhat below Maeder's and may encompass the  $\sim 18 M_{\odot}$  estimated for the progenitor of SN 1987A.

A major uncertainty in Maeder's estimates is in the initial mass function (IMF). This is taken as a power law

$$\frac{dN}{dM} = M^{-(1+x)}, \quad (1)$$

i.e., the classical Salpeter mass function. In order to estimate uncertainties, calculations were also performed for  $x = 1.70$ . Maeder's results for  $\Delta Y/\Delta Z$  show rather large differences between the two values,  $x = 1.35$  and  $1.70$ , for the slope of the IMF. Of course, the cutoff is at a substantially higher mass in the case of the 1.70 slope, since the larger slope already decreases the frequency of heavy-mass stars so their contribution to nucleosynthesis does not have to be cutoff so early.

In a series of recent papers, Massey et al. (1989a); Massey, Parker, & Garmany (1989b); Massey & Thompson (1991), and

Parker et al. (1992) have measured the IMFs of stars up to  $80 M_{\odot}$  in OB associations in the LMC, SMC, and in our Galaxy. These authors find that the IMF can be fitted, in the Salpeter form with slopes ranging from  $x = 1.0 \pm 0.1$  to  $1.8 \pm 0.1$  (Bildsten 1993). We find the total mass not ejected to be

$$\langle M \rangle = \int_{M_1}^{M_{\max}} M \left( \frac{dN}{dM} \right) dM = \frac{C}{x-1} \left[ 1 - \left( \frac{M_{\max}}{M_1} \right)^{1-x} \right], \quad (2)$$

where  $M_{\max}$  is the maximum mass of a stable star,  $M_1$  the cutoff mass, and  $C = M_1^x (dN/dM)|_{M_1}$  is a normalization constant. Offhand, small  $x$  seem to play a disproportionately large role in equation (2). But equation (2) is easy to evaluate. For  $M_1 = 25$  and  $M_{\max} = 130$ , we have

$$x = 1.1 \quad 1.35 \quad 1.8$$

$$\langle M \rangle / C = 1.50 \quad 1.13 \quad 0.91.$$

Thus, the dependence on  $x$  is actually moderate, and Salpeter's  $x = 1.35$  seems a good average. We do not believe that larger  $x$ 's need to be considered.

From an entire grid of evolved stars of  $0.1$ – $100 M_{\odot}$ , Timmes, Woosley, & Weaver (1993) have calculated Galactic chemical evolution. They have extended their calculation (private communication) to compute  $\Delta Y/\Delta Z$ , confirming Maeder's results, but with a mass cut slightly higher by  $\sim 5 M_{\odot}$  in order to fit the Pagel values. We tend to confirm their higher mass cut (see § 5).

We conclude from the above arguments that return of elements to the Galaxy from stellar collapse must be cut off at a mass

$$M_{\text{cutoff}} = 25 \pm 5 M_{\odot}. \quad (3)$$

It should be noted that Timmes et al. (1993) do not find the mass of their compact core to increase monotonically with stellar mass. Thus, for zero metallicity stars of mass 15, 20, 25, and  $30 M_{\odot}$ , the baryon number masses of the core are 1.62, 1.42, 1.90, and  $1.87 M_{\odot}$ , respectively. For solar metallicity, stars of mass, 15, 18, 20, 25, and  $30 M_{\odot}$  have baryonic core masses of 1.57, 1.64, 1.54, 1.67, and  $1.99 M_{\odot}$ . Although the metallicity dependence of these masses is extremely interesting, we will neglect it in our present discussion. In order to convert these masses into gravitational masses, we use the binding energy given by Lattimer & Yahil (1989)

$$E = 0.075 M_{\odot} \left( \frac{M}{M_{\odot}} \right)^2, \quad (4)$$

which is a  $0.17 M_{\odot}$  correction for  $M = 1.5 M_{\odot}$ ,  $0.3 M_{\odot}$  for  $M = 2 M_{\odot}$ . In our scenario, then the stars with baryon masses greater than  $1.67 M_{\odot}$  will go ultimately into black holes. However, in § 5 we shall show that all of the stars of masses listed above will return matter to the galaxy, so that those that do go into black holes will go into light mass ones,  $M_{\text{BH}} \gtrsim 1.5 M_{\odot}$ .

Note that the  $1.64 M_{\odot}$  for the  $\sim 18 M_{\odot}$  progenitor of SN 1987A lies just below the limit of  $1.67 M_{\odot}$  for stars to go into a black hole, although the late time accretion (see § 5) would send it into one. Since the mass of the progenitor was known only to within  $\sim \pm 2 M_{\odot}$ , it is obviously a delicate matter whether SN 1987A went into a black hole or not, but the possibility certainly exists.

#### 4. $K^-$ CONDENSATION AND THE EQUATION OF STATE OF DENSE MATTER

Central to our scenario is that the equation of state of dense matter, for  $\rho \gtrsim 3\rho_0$ , where  $\rho_0$  is nuclear matter density, is very different from the conventional one. In the latter, most of the protons change into neutrons at the higher densities, because it is too expensive in terms of energy to keep many electrons. The ratio of electrons to nucleons at high densities is typically  $Y_e \sim 0.1$  and the same number of protons as electrons will be present. Therefore, one talks about the compact remnants formed in stellar collapse as "neutron stars." In our new scenario, based on the developments of Brown et al. (1992a), it is energetically favorable for protons to remain at the higher densities  $\rho \neq 3\rho_0$ . Their charge is chiefly balanced by  $K^-$  mesons, which form a Bose condensate, although a few electrons may still be present. The energy of this Bose condensate is lowered because of strongly attractive interactions between nucleons and the  $K^-$  mesons. The net result of kaon condensation is that at the higher densities at which the stability of the nucleon stars is determined, nuclear matter, rather than neutron matter, is present. In other words, the compact objects are nuclear matter stars (or nucleon), rather than neutron stars. We will now briefly sketch the history of kaon condensation.

Some years ago, Kaplan & Nelson (1986) showed, starting from chiral Lagrangians, that kaons obtain an effective mass  $m_K^*$  which decreases in dense matter, going to zero at several times nuclear matter density. Explicit solution of the dense matter problem was given by Politzer & Wise (1991). Quantitative application to stellar collapse has been worked out by Brown et al. (1992a) and detailed calculations of the equation of state, with application to compact cores of collapsing stars, by Thorsson (1992) and Thorsson, Prakash, & Lattimer (1993). The calculations are somewhat involved, so we refer the reader to these papers in detail.

The chief result of the drop in  $m_K^*$  for us is that  $K^-$  mesons can replace electrons in carrying negative charge in the dense matter of collapsing stars. Thus, "kaon condensation" sets in when the *in-medium* kaon energy  $\omega_{K^-}$  becomes equal to the electron chemical potential

$$\omega_{K^-} = \mu_- = \mu_e - \mu_p, \quad (5)$$

where the final equality is set by beta equilibrium. In addition to the kaon mass  $m_K^*$  being lowered, the  $K^-$  meson experiences an attractive vector mean field potential from the nucleon. This mean field originates from the same exchange of  $\rho$ - and  $\omega$ -mesons that produce the  $\Lambda(1405)$  as a bound state of  $K$  and nucleon (Siegel & Weise 1988; Müller-Groeling, Holinde, & Speth 1990). In somewhat conservative calculations, Thorsson (1992) and Brown et al. (1992a) find that kaon condensation sets in at a density  $\rho_c \approx 3\rho_0$ , where  $\rho_0 = 0.16 \text{ fm}^{-3}$  is nuclear matter density.

Of chief concern to us here is the fact that the effect of kaon condensation is to soften the equation of state of dense matter. Reasons for a low-compression modulus of nuclear matter were summarized in Brown (1988). The kaon condensate comes in at higher densities in order to keep the equation of state soft. This softness, as compared with neutron matter, chiefly results from the lower symmetry energy resulting from the nearly equal numbers of protons and neutrons in kaon-condensed matter.

Thorsson (1992) and Thorsson et al. (1993) have investigated the composition, structure, and evolution of nucleon stars with

kaon condensates. With equations of state that satisfy the empirical properties of nuclear matter, and have a compression modulus of  $K_0 = 180$  MeV, at the lower end of the empirically quoted  $210 \pm 30$  MeV (Blaizot, Gogny, & Grammaticos 1976), and with the kaon condensate, the authors find compact core masses of

$$1.42 M_\odot < M_{\max} < 1.45 M_\odot, \quad (6)$$

the precise value depending upon the assumed density dependence of the nuclear symmetry energy. These values are barely able to stabilize the heavier pulsar of mass  $1.442 \pm 0.03$  in PSR 1913+16 (Taylor & Weisberg 1989). An increase of  $K_0$  up to  $\sim 190$  MeV would bring  $M_{\max}$  up to our estimated  $1.5 M_\odot$ . Clearly, our knowledge of the equation of state is not accurate enough to distinguish between this value and those of equation (6).

Brown (1992) has shown that with the presence of the kaon condensate, hadronic matter has such a low energy density that it does not merge to quark matter until very high densities,  $\sim 10\rho_0$  or greater.

We see, therefore, that the presence of a kaon condensate substantially softens the equation of state of dense matter and brings the maximum nucleon star mass  $M_{\max}$  down substantially from the range  $M_{\max} \lesssim 2 M_\odot$ , generally given by equations of state without kaon condensate present.

Extensive calculations of higher order corrections to the kaon condensate in heavy fermion chiral perturbation theory (Brown et al. 1994) have confirmed the Politzer & Wise (1991) results that the kaon condensate is extremely robust. Thus, we believe in the result that the maximum nucleon star mass is substantially smaller,  $M_{\max} \sim 1.5 M_\odot$ , than obtained previously with conventional equations of state.

##### 5. POSTSHOCK MECHANISM FOR FORMATION OF BLACK HOLES

Once the mass of the compact remnant exceeds  $M_{\max}$ , collapse into a black hole goes in a fraction of a millisecond. However, if this collapse happened in SN 1987A, it must have been more than 12 s after the first collapse into the compact object, because at 12 s a neutrino was still observed at Kamiokande II. The question is, then how the core can be stable for such a long period of nucleosynthesis and neutrino emission.

It has been suggested for some time that stars heavier than  $\sim 25 M_\odot$  may leave black holes (e.g., Wilson et al. 1986), also that such stars first explode, exhibiting light curves of Type II supernovae and returning matter to the galaxy, and then collapse into black holes (Woosley & Weaver 1986). The compact core is, for a certain range of masses, stabilized by thermal pressure during the period of Kelvin-Helmholtz contraction long enough to carry out nucleosynthesis, going into a black hole after cooling and deleptonization. Calculations by Bombaci, Prakash, & Brown (1993) show that with their equation of state and additional  $\sim 0.1 M_\odot$  over and above the  $M_{\max}$ , where  $M_{\max}$  is the maximum mass of the cold deleptonized core, can be stabilized by thermal pressure sufficiently long to accomplish nucleosynthesis (and to emit neutrinos for the 12 s period in which they were observed in Kamiokande during SN 1987A). The additional  $\sim 0.1 M_\odot$  is a "window" in the core mass, in which nucleosynthesis can take place, with the core later going into black hole.

In the "conventional scenario" (i.e., without  $K^-$  condensation) the window is only that produced by thermal pressure. One might think that the considerable pressure pro-

duced by trapped neutrinos in the explosion would stabilize some additional mass. However, just after the explosion  $Y_i$  is about 0.4, and therefore the nuclear symmetry energy is much smaller than for neutron matter,  $Y_i \sim 0$ . In fact, the lower symmetry energy for the higher  $Y_i$  decreases the pressure more than the leptons build it up. In the calculations of Bombaci et al. (1993) the "window" created by thermal pressure is decreased by  $\sim 0.05 M_\odot$ . This "window" is, therefore, relatively small in the conventional theory, although equations of state other than that used by Bombaci et al. might give a somewhat larger one.

With kaon condensation the situation is very different. The high-density matter remains nuclear matter, as we discussed in the last section, so that our  $M_{\max} \sim 1.5 M_\odot$  applies to this nuclear matter star. The proton fraction  $Y_p \sim 0.4$ , which exists after explosion, remains the same as we go to nuclear matter with  $K^-$  condensation, so that no symmetry energy needs to be paid. Thus, as the leptons are emitted, the nuclear interaction energy remains roughly the same; we simply lose kinetic energy of the leptons. Consequently, the leptons can stabilize a substantially larger core than  $M_{\max}$  for short times. This was recently pointed out to us by Lattimer & Prakash (1993) and we summarize their work here. They find that with the nuclear compression modulus of 190 MeV, which gives the cold,  $Y_i \sim 0$   $M_{\max} \sim 1.5 M_\odot$ , a mass of  $1.7 M_\odot$  is stable for short times, because of the lepton pressure. Inclusion of the thermal pressure mentioned above will bring this mass up to  $\sim 1.8 M_\odot$ .

As the trapped neutrinos are emitted, bringing the lepton pressure down over a period of greater than 12 s, the lepton pressure decreases, but the nucleons in the core are heated because of entropization of the matter by the neutrino flux, analogous to Joule heating in a wire (Burrows & Latimer 1986) and the thermal pressure is substantially increased, sufficiently so that the compact core continues to be stabilized until it cools.

The postshock accretion, discussed by Brown et al. (1992b) plays a lesser role. This accretion has recently been discussed by Bethe (1993), who had argued that in the supernova shock there is vigorous convection, but that at some time (2 s in SN 1987A) convection stops because heat is no longer supplied by neutrinos. When this happens, a substantial fraction of the previously convecting material falls into the neutron star at the center. He estimates this fraction to be about 10% of the mass in the shock wave, or about  $0.04 M_\odot$ . Most of this mass will accrete after the  $\sim 12$  s it took for the neutrons to be emitted in SN 1987A.

The combination of stabilization of an additional  $0.2 M_\odot$  by leptons for  $t > 12$  s,  $0.1 M_\odot$  by thermal pressure and the postshock accretion, means that compact cores which have late-time cold masses of up to  $1.84 M_\odot$  will be stable during this  $\sim 12$  s and possibly longer, even if the maximum stable cold mass is only  $M_{\max} \sim 1.5 M_\odot$ .

From these arguments and the baryon number masses of Timmes et al. (1993) which we discussed in § 3, we see that stars of mass up to  $\sim 30 M_\odot$  will be able to become supernovae and thus return elements to the galaxy, but most of those in the range of  $18-30 M_\odot$  will later go into low-mass black holes. Most stars heavier than  $30 M_\odot$  will go directly into high-mass black holes.

We return to a discussion of Type Ib supernovae. Wheeler & Levereault (1985) came to the conclusion that the physics of Ib explosions was similar to that of Type II supernovae. Enzman & Woosley (1988) found, through detailed light curve calcu-

lations, that the Type Ib progenitors were  $4-7 M_{\odot}$  He stars, corresponding to  $15-25 M_{\odot}$  mainsequence stars. These authors suggested that the Type Ib progenitors are Wolf-Rayet stars. Swartz & Wheeler (1991) suggest that it may be better to discern the mass of exploding Wolf-Rayet stars from their late-time light curve, rather than near maximum, as studied by Enzman & Woosley. Assuming the  $V$  light curve to be parallel to the bolometric light curve, Swartz & Wheeler conclude that the mass of the ejecta of the Type Ib SN 1984L must have been in excess of  $10 M_{\odot}$ , and all of this in He or heavier elements. A He star of  $10 M_{\odot}$  corresponds to a main-sequence star of mass  $\sim 30 M_{\odot}$ . A similar argument can be made for SN 1985F. From our estimates, this should be possible, with the core SN 1984L having gone, following nucleosynthesis, into a black hole. On the other hand, stars somewhat heavier than  $\sim 30 M_{\odot}$  will go directly into black holes, as suggested by Wheeler & Shields (1976) for the progenitor of the black hole of mass  $\gtrsim 9 M_{\odot}$  in Cygnus X-1.

We are led to conclude that the cutoff for the return of elements to the Galaxy should be at  $\sim 30 M_{\odot}$ , at the upper end of the estimate, equation (3). This is in agreement with the estimate of S. E. Woosley (private communication). As noted in § 3, the compact core masses do not, however, increase monotonically with stellar mass. Thus, some stars of mass greater than  $30 M_{\odot}$  may explode, returning matter to the Galaxy, whereas others of mass greater than  $30 M_{\odot}$  may go directly into black holes.

#### 6. IMPLICATIONS OF OUR SCENARIO FOR BLACK HOLE FORMATION

Properties of the best candidates for black holes; e.g., Cyg X-1, LMC X-3, A0620-00, and GS 2023+338, have been reviewed recently by van den Heuvel (1992). These are relatively heavy, the most likely masses falling in the range of  $9-16 M_{\odot}$ . In our scenario, such heavy mass black holes would be formed by the nearly instantaneous collapse, without nucleosynthesis, of large stars  $M \gtrsim 30 M_{\odot}$ .

Lighter black holes with mass  $M_{\text{BH}} \simeq 1.5 M_{\odot}$  should result from the collapse of most of the stars in mass range  $18 M_{\odot} \lesssim M \lesssim 30 M_{\odot}$ , according to our scenario. These events are

accompanied by nucleosynthesis. If they occur on binaries, then the binary may be disrupted in the explosion.

Taking the present supernova rate in late-type spiral galaxies like our own to be  $\sim 0.03 \text{ yr}^{-1}$  and correcting for a higher rate for the first few billion years after the formation of the Galaxy, van den Heuvel (1992) estimates that there are at least some  $10^9$  neutron stars in the Galaxy. Whereas van den Heuvel estimates the ratio  $f$  of neutron stars to black holes in the Galaxy to be

$$10 < f < 45, \quad (7)$$

we find roughly equal numbers of neutron stars and black holes; i.e.,  $f \sim 1$ . Thus, we expect about  $10^9$  black holes in the Galaxy.

*Note added in manuscript.*—After completion of our manuscript, we received a paper by N. Prantzos to be published in A&A in which he shows that observations of the carbon/oxygen ratio in halo stars require all stars up to  $\sim 90-100 M_{\odot}$  to contribute to galactic nucleosynthesis. Uncertainties in the  $^{12}\text{C}(\alpha, \gamma)^{16}\text{O}$  rate do, however, affect this argument. N. Prantzos (private communication) stressed also that some oxygen from SN II's, locked up in grains, may have escaped the shallow potential of dwarf galaxies where  $\Delta Y/\Delta Z$  was measured, changing the value of this quantity. Whereas Pagel et al. (1992) did correct for this loss, some uncertainty remains. In view of Prantzos' arguments, we believe that the arguments of Maeder, although extremely interesting, may not be as firm as we indicated in the text.

We would like to thank Lars Bildsten, S. R. Kulkarni, J. C. Wheeler, B. E. J. Pagel, and E. S. Phinney for useful discussion. We are extremely grateful to J. M. Lattimer and M. Prakash for discussion of their work, before publication, on the short-time stability of compact cores. We are grateful to Steve Koonin and the W. K. Kellogg Radiation Laboratory for their hospitality. The work of G. E. Brown was partially supported by the Department of Energy under grant DE-FG02-88ER40388.

#### REFERENCES

- Bethe, H. A. 1993, ApJ, submitted
- Bildsten, L. 1993, private communication
- Blazquez, J. P., Gorgay, D., & Grammaticos, B. 1976, Nucl. Phys. A, 265, 315
- Bombaci, I., Prakash, M., & Brown, G. E. 1993, to be published
- Brown, G. E. 1988, Nature, 336, 519
- . 1992, in Proc. 1st Symp. on Nucl. Phys. of the Universe, Oak Ridge, TN, September 24-26, ed. M. Strayer (Singapore: World Scientific), in press
- Brown, G. E., Bruenn, S. W., & Wheeler, J. C. 1992b, Comm. Ap., 16, 153
- Brown, G. E., Kubodera, K., Kho, M., & Thorsson, V. 1992a, Phys. Lett. B, 291, 355
- Brown, G. E., Lee, C. H., Rho, M., & Throsson, V. 1994, Nucl. Phys. A, in press
- Burrows, A., & Lattimer, J. M. 1986, ApJ, 307, 178
- Chevalier, R. A., & Kirshner, R. P. 1979, ApJ, 233, 154
- Dewey, R. J., & Cordes, J. M. 1987, ApJ, 321, 780
- El Eid, M. F., & Prantzos, N. 1988, in Origins and Distribution of the Elements, ed. G. J. Mathews (Singapore: World Scientific), 485
- Enzman, L. M., & Woosley, S. E. 1988, ApJ, 333, 754
- Helfand, D. J., & Becker, R. H. 1984, Nature, 307, 215
- Kaplan, D. B., & Nelson, A. E. 1986, Phys. Lett. B, 175, 57
- Lamb, F. K. 1991, in Frontiers of Stellar Evolution, ed. D. L. Lambert (ASP Conf. Ser., 20), 299
- Lattimer, J. M., & Prakash, M. 1993, to be published
- Lattimer, J. M., & Yahil, A. 1989, ApJ, 340, 420
- Maeder, A. 1992, A&A, 264, 105
- . 1993, A&A, 268, 833
- Massey, P., Garmany, C. D., Silkey, M., & Degioria-Eastwood, K. 1989a, AJ, 97, 107
- Massey, P., Parker, J. W., & Garmany, C. D. 1989b, AJ, 98, 1305
- Massey, P., & Thompson, A. J. 1991, AJ, 101, 1405
- Mayle, R. W., & Wilson, J. R. 1988, ApJ, 334, 909
- Müller-Groeling, A., Holinde, K., & Speth, J. 1990, Nucl. Phys. A, 513, 557
- Pagel, B. E. J., Simonson, E. A., Terlevich, R. J., & Edmunds, M. G. 1992, MNRAS, 255, 325
- Parker, J. W., Garmany, C. D., Massey, P., & Walborn, N. R. 1992, AJ, 103, 1205
- Potter, H. D., & Wise, M. B. 1991, Phys. Lett. B, 273, 156
- Prantzos, N. 1993, A&A, in press
- Shigeyama, T., Nomoto, K., Hashimoto, M., & Sugimoto, D. 1987, Nature, 328, 1
- Siegel, P. B., & Weiss, W. 1988, Phys. Rev. C, 38, 221
- Sontz, N. B., Phillips, M. M., Elias, J. H., DePoy, D. L., & Walker, A. R. 1992, ApJ, 384, L33
- Swartz, D. A., & Wheeler, J. C. 1991, ApJ, 379, L13
- Taylor, J. H., & Weisberg, J. M. 1989, ApJ, 345, 434
- Thielemann, F.-K., Nomoto, K., & Hashimoto, M. 1992, in Origins and Evolutions of the Elements, in honor of H. Reeves 60th Birthday, ed. N. Prantzos & E. Vangioni-Flam (Cambridge: Cambridge Univ. Press)
- Thorsett, S. E., Arzoumanian, Z., McKinnon, M. M., & Taylor, J. H. 1993, ApJ, 405, L29
- Thorsson, V. 1992, thesis, State University of New York at Stony Brook
- Thorsson, V., Prakash, M., & Lattimer, J. M. 1993, to be published
- Timmer, F. X., Woosley, S. E., & Weaver, T. A. 1993, preprint
- Twarog, B. A., & Wheeler, J. C. 1982, ApJ, 261, 636
- . 1987, ApJ, 316, 153

- van den Bergh, S., McClure, R. D., & Evans, R. 1987, *ApJ*, 323, 44  
van den Heuvel, E. P. J. 1992, in *Proc. Int. Space Year Conf.*, ed. T. D. Guyenne & J. J. Hunt (Satellite symp. No. 3; ESA ISY-3) (Noordwijk: ESA Publ. Div., ESTEC), 29  
Wheeler, J. C., & Levereault, R. 1985, *ApJ*, 294, L17  
Wheeler, J. C., & Shields, G. A. 1976, *Nature*, 259, 642  
Wheeler, J. C., Sneden, C., & Truran, J. W. 1989, *ARA&A*, 27, 279  
Wilson, J. R., Mayle, R., Woosley, S. E., & Weaver, T. A. 1986, in *Proc. Texas Rel. Astrophys. Symp.* II, ed. N. Livio & G. Shaviv, Jerusalem, Israel (New York Acad. Sci. 470), 267  
Woosley, S. E., & Weaver, T. A. 1986, *ARA&A*, 24, 205

---

## Chapter 5

# Neutron Star Accretion and Binary Pulsar Formation

G.E. Brown

Reprinted with permission from *Astrophysical Journal*, 440 (1995) 270-279  
Copyright (1995) by American Astronomical Society

### Commentary by G.E. Brown

This paper led us ultimately into binaries and binary evolution. It also led into the evolution of black holes, since they were formed naturally in our binary evolution.

I had gone to the Institute for Theoretical Physics in Santa Barbara to participate in a meeting on Hot QCD. A parallel program in Dense Stellar Systems was going on at the same time. I found it to be very exciting. I'd read one of the first papers on common envelope evolution in binaries and happened to discuss it with Fred Rasio, whose office was just across the hall from mine. Fred told me of much more recent work. Astronomers discussed common envelope evolution as something mysterious, but I thought I could calculate it analytically, except for the initial immersion of the neutron star in the envelope of the massive convective companion. Ironically, this initial immersion was calculated very elegantly later by Fred Rasio and Mario Livio (1996). I like to think that I stimulated Fred to do so.

Roger Chevalier (1989) had already estimated that when a neutron star went into common envelope evolution with a convective massive star, the densities would be so great that the accretion rate would exceed  $10^4$  times the Eddington rate which was sufficient to trap the photons emitted to carry off the binding energy of the accreted matter and carry them back onto the neutron star with adiabatic inflow. Accretion at the rate of  $> 10^4 \dot{M}_{\text{Edd}}$  is known as hypercritical accretion in the literature. The infalling matter is so dense that it literally traps the photons, carrying them back onto the neutron star. Zel'dovich *et al.* (1972) had carried out the first calculation of hypercritical accretion before common envelope evolution was introduced. What the accreted matter does is somewhat complicated. We return to this in the Commentary of Paper 18.

This side activity of hypercritical accretion continued with a slow stream of papers parallel to the general stream, which limited accretion to the Eddington limit.

The Eddington rate of accretion is  $\dot{M}_{\text{Edd}} \simeq 1.5 \times 10^{-8} M_{\odot} \text{ yr}^{-1}$ . Since the dynamical time

for common envelope evolution (Paper 13) is  $\tau_{\text{dyn}} \sim 1$  yr, this means that accretion at the Eddington rate can be neglected. However, the actual accretion is highly hypercritical, more like  $1M_{\odot} \text{ yr}^{-1}$ . I was used to the concept of photon trapping because neutrino trapping, in a very different density regime, takes place in the collapse of large stars.

The continued resistance of much of the astrophysical community to our work comes chiefly because they do not believe in hypercritical accretion. I carried Chevalier's work somewhat further, using techniques developed in Paper 3, and estimated that the neutron star would accrete  $\sim 1M_{\odot}$  of material, enough to send it into a low-mass black hole.

The only way in which a neutron star binary could survive would be if the neutron star could avoid going through common envelope evolution. This could only happen if the original progenitors were so close in mass (within  $\sim 5\%$ ) that they burned helium at the same time, in which case common envelope evolution took place while they were helium stars. The resulting two neutron stars in the binary would then be very close to each other in mass.

In Paper 12 Hans Bethe and I completed the explicit calculation of common envelope evolution.

When I submitted this paper to the *Astrophysical Journal*, I wondered how the referee would react to my first attempt at binary evolution. I quote some of the report of the referee, Dipankar Bhattacharya, who had written the Physics Report "Formation and Evolution of Binary and Millisecond Radio Pulsars" with Ed van den Heuvel (1991).

"The author attempts to understand the consequences of heavy accretion inside a common envelope and confirms the starting conclusion of Chevalier (1993) that a neutron star is unlikely to survive the accretion process. If this is true, then contrary to the widely held view, double neutron stars cannot be the products of common envelope evolution....

"The result is a very important one, and is likely to have a major impact on the evolutionary models of X-ray binaries and recycled pulsars. The publication of the paper is therefore highly recommended."

## References

- (1) Bhattacharya, D., and van den Heuvel, E. (1991), *Physics Reports* **203**, 1.
- (2) Rasio, F., and Livio, M. (1996), *Astrophysical Journal* **471**, 366.
- (3) Chevalier, R. (1989), *Astrophysical Journal Letters* **346**, 847.
- (4) Zel'dovich, Ya. B., Ivanova, L.N., and Nadezhin, D.K. (1972), *Soviet Astronomy* **16**, 209.

## NEUTRON STAR ACCRETION AND BINARY PULSAR FORMATION

G. E. BROWN<sup>1</sup>

Institute for Theoretical Physics, University of California, Santa Barbara, Santa Barbara, CA 93106-4030

Received 1994 June 3; accepted 1994 August 15

### ABSTRACT

In the standard scenario for binary pulsar formation, the neutron star from the explosion of the primary supergiant moves through the envelope of the companion star. The envelope is then expelled through the hydrodynamic coupling of the dynamical friction. With loss of energy, the orbit of the neutron star tightens.

In moving through the companion star envelope, the neutron star accretes a substantial amount of matter. Chevalier has estimated that this will generally send the neutron star into a black hole. We are able to confirm his estimates showing that the amount of accretion can be simply related to the energy loss through dynamical friction. This connection shows that if the latter is sufficient to expel the envelope, then the neutron star will accrete  $\geq 1 M_{\odot}$ , which is sufficient to convert it into a black hole.

A new scenario, in which binary pulsars generally result from the explosion of helium star binaries, is suggested. Not only can black hole formation be avoided in this scenario, but it would result in nearly equal masses for the two neutron stars in the binary. There is evidence for this.

Statistics of the (rare) binary pulsar formation are discussed.

**Subject headings:** accretion, accretion disks — binaries: close — black hole physics — pulsars: general — stars: neutron

### 1. INTRODUCTION

Studies of possible neutron star formation in SN 1987A have raised the issue of the amount of accretion onto a neutron star soon after formation. The initial rate of accretion, following the supernova explosion, can be estimated to be  $\dot{M} \cong 1 M_{\odot} t^{-1}$   $s^{-1}$  with  $t$  measured in seconds. This is 15–16 orders of magnitude greater than Eddington. Indeed, it is this accreting matter which is directly responsible for the neutrinos emitted during this time, which are responsible for about half of the total energy in the detected neutrinos (Bethe 1993). Following this accretion through until later times, Chevalier (1989) finds that  $\sim 0.15 M_{\odot}$  of material has accreted onto a possible neutron star during the first year of its existence.

Colgate (1971) showed that neutrino emission could carry off energy, allowing accretion at a high rate, and Zel'dovich, Ivanova, & Nadezhin (1972) considered the optically thick accretion of  $\sim 10^{-5} M_{\odot}$  of matter over a time of 36 s. The problem of hypercritical accretion, with neutrinos carrying off the accreted energy, was taken up again by Chevalier (1989). There was, however, not much recognition in the literature that accretion could proceed much faster than the Eddington limit,  $\sim 10^{-8} M_{\odot} \text{yr}^{-1}$  for neutron stars.

As long as photons carry off the accreted energy, the Eddington limit will generally apply, at least in spherically symmetric situations. (SS 433 is often cited as a case where the energy can be carried off in jets, once spherical symmetry is given up.)

Accretion onto a neutron star possibly formed in SN 1987A has been worked out in detail by Houck & Chevalier (1991), who obtain self-consistent solutions of the fluid dynamical equations down to an accretion rate of  $\sim 2 \times 10^{-5} M_{\odot} \text{yr}^{-1}$ , beginning from the high rates established above. Between this rate and the Eddington rate of  $\sim 3 \times 10^{-8} M_{\odot} \text{yr}^{-1}$  (for  $N = Z$  matter), accretion may be strongly opposed by the radi-

ation pressure of the accreting material; i.e., there may be a gap in possible rates of accretion. We shall schematize their results in our § 2, in terms of an analytical model. The calculation is straightforward because the initial very high accretion rates overwhelm the pressure from photons, carrying them inward, and temperatures near the proto-neutron star are high enough so that neutrinos are easily produced. The situation is particularly simple because, at most, only a weak magnetic field can be present. Otherwise, the plerion (hot synchrotron nebula) formed by the ionized particles emitted along magnetic field lines would have been seen. It is well known that a strong rotating magnetic field can hinder the accretion of matter through the so-called propeller effect (Illarionov & Sunyaev 1985). We shall return to this later.

Chevalier (1993) has extended his considerations to the case of a neutron star in the dense environment of a stellar envelope. Such a situation occurs in the scenario for making binary pulsars, in which the spiral-in of the neutron star through the companion supergiant expels the envelope, which is hydrodynamically coupled to the drag from the neutron star, leaving a helium star as companion (Van den Heuvel 1976). Hypercritical accretion rates are encountered in this spiral-in, and Chevalier suggests that except in very extended stars, radiation pressure is unable to limit the accretion to roughly the Eddington rate. In other cases, the neutron star accretes sufficient material to carry it into a black hole.

The spiral-in situation is more complex than that of spherical accretion in the case of the compact object formed in SN 1987A. First, the neutron star moves through the matter, so that Bondi-Hoyle-Lyttleton, rather than Bondi, accretion must be considered. (The motion is mildly supersonic.) Carrying over the considerations from spherical accretion must, therefore, be justified. Second, the neutron star initially moves into low-density matter in which radiation pressure will certainly limit accretion to the Eddington limit, and it must be shown dynamically how the “smothering” overcomes this. Third, the neutron star may have a strong magnetic field. In § 3, we shall discuss these points. We shall refer to results of numerical

<sup>1</sup> Permanent address: Physics Department, State University of New York, Stony Brook, NY 11794; supported in part by the Department of Energy under grant DE-FG02-88ER40388.

calculations in order to show that the extension of the results from spherical accretion to the situation of mildly supersonic flow is justified. We then demonstrate that in typical situations, the ram pressure of the flow of matter from the envelope of the large star, in the system in which the neutron star is at rest, is sufficient to break down the radiative envelope which limits the accretion to the Eddington rate. We shall discuss effects from magnetic fields.

In § 4 we find a simple relation between the total energy produced by dynamical friction and the total mass accreted onto the neutron star, in terms of the average squared velocity of the neutron star through the envelope of the large star. This relation holds only to the extent that radiation pressure cannot prevent hypercritical accretion and that magnetic fields do not inhibit it substantially. Section 3 has justified these assumptions.

In § 5 we will discuss unusual circumstances which allow for binary pulsar formation. We point out that proceeding from a helium star binary, after explosion of one helium star the resulting neutron star does not have to go through a hydrogen envelope. We review work on the evolution of W-R + W-R binaries. We point out that the neutron stars produced by a helium star binary are likely to be very close to each other in mass, as appears to be the observed outcome of binary pulsar formation.

In § 6 we discuss the probability of forming neutron star binary pulsars and compare with the three observed ones.

We give conclusions in § 7.

## 2. SPHERICAL HYPERCRITICAL ACCRETION AND HYPERCRITICAL BONDI-HOYLE ACCRETION

We begin by following the scenario of Chevalier (1989) for hypercritical spherical accretion. In the case of a supernova explosion such as SN 1987A, the ejecta are not at rest, but after a few doubling times of the radius they are expanding uniformly. The compact object may be formed with substantial kick velocity in the explosion, but, in time, will tend to comove with the ejecta. Using standard Bondi spherical accretion theory (eq. [14.3.20] of Shapiro & Teukolsky 1983), assuming the accreting gas to consist mainly of oxygen and taking initial conditions from the supernova kinematics of Woosley (1988), Brown & Weingartner (1994) find an accretion rate of

$$\dot{M} = \frac{1.82 \times 10^{-4}}{t_{15/8}} M_{\odot} \text{ yr}^{-1} \quad (2.1)$$

where the time  $t$  is measured in years. This is appreciably larger than the Eddington accretion rate

$$\dot{M}_{\text{Edd}} = \frac{4\pi c R}{\kappa_{\text{ee}}} = 1.5 \times 10^{-8} M_{\odot} \text{ yr}^{-1}, \quad (2.2)$$

where the neutron star radius  $R$  has been taken to be  $10^6$  cm and the opacity, for ionized hydrogen,  $\kappa = 0.4 \text{ cm}^2 \text{ g}^{-1}$ . The ratio of the accretion to Eddington, at a time of 1 yr is

$$\dot{m} = \frac{\dot{M}}{\dot{M}_{\text{Edd}}} = 1.21 \times 10^4. \quad (2.3)$$

This number is close to that of Chevalier (1989), who used different kinematics for his input. We find his scenario to be insensitive to the precise initial conditions. When  $\dot{M}$  exceeds  $\dot{M}_{\text{Edd}}$  by such a large amount, the accretion is called hypercritical, and was considered by Blondin (1986). Since this is essen-

tial for our later development, we sketch the argument of Blondin.

From

$$\dot{M} = -4\pi r^2 \rho v \quad (2.4)$$

and the  $\dot{M}_{\text{Edd}}$  of equation (2.2), one finds the optical depth to be

$$\tau_{\text{ee}} = \int \rho \kappa_{\text{ee}} dr = \frac{2R\dot{m}}{\sqrt{r_s}}, \quad (2.5)$$

where the Schwarzschild radius  $r_s = 2GM/c^2$  has been introduced. The free-fall velocity

$$v = \sqrt{2GM/r} \quad (2.6)$$

has been used. From random walk, the photon diffusion time is

$$t_{\text{diff}} = \frac{r}{c} \frac{r}{\lambda_{\text{ee}}} \approx \frac{r}{c} \tau_{\text{ee}}, \quad (2.7)$$

where  $\lambda_{\text{ee}}$  is the mean free path for electron scattering. In order to obtain the photon trapping radius, we set this equal to the dynamical time

$$t_{\text{dyn}} = \frac{r_{\text{tr}}}{v_{\text{tr}}} \cong r_{\text{tr}} \sqrt{\frac{r_{\text{tr}}}{2GM}} \approx \frac{r_{\text{tr}}}{c} \tau_{\text{ee}}, \quad (2.8)$$

where we have labeled the trapping radius by  $r_{\text{tr}}$ . Setting  $t_{\text{diff}} = t_{\text{dyn}}$  and using equation (2.5) for  $\tau_{\text{ee}}$ , with the help of equation (2.2), we find

$$r_{\text{tr}} = 2R\dot{m}. \quad (2.9)$$

The accuracy of equation (2.9) can be increased by solving the diffusion equation in three dimensions, assuming spherical symmetry, and keeping only the lowest eigenvalue. This gives a factor of  $3/\pi^2$  multiplying the right-hand side of equation (2.9); the photons escape more rapidly when the three-dimensionality of the problem is taken into account. Our final expression is

$$r_{\text{tr}} \cong 0.6R\dot{m}. \quad (2.10)$$

We next discuss how the accretion envelope forms around the neutron star, following Chevalier (1989). An accretion shock develops at a radius  $r_{\text{sh}}$ , for which we shall solve. As material flows toward the neutron star inside of the trapping radius, adiabatic compression of the radiation-pressure-dominated gas raises the temperature sufficiently high that neutrino pairs are produced, in a thin shell about the neutron star. These immediately escape, carrying away the accretion energy.

Since the pressure is radiation-dominated, the accretion forms an  $n = 3$  ( $\gamma = 4/3$ ) polytrope. Thus, inside of the shock

$$\rho = \rho_{\text{sh}} \left( \frac{r}{r_{\text{sh}}} \right)^{-3}, \quad P = P_{\text{sh}} \left( \frac{r}{r_{\text{sh}}} \right)^{-4}, \quad v = v_{\text{sh}} \left( \frac{r}{r_{\text{sh}}} \right), \quad (2.11)$$

where the subscript "sh" refers to the value just inside of the shock front. Because of the adiabatic compression by the factor  $(\gamma + 1)/(\gamma - 1)$  of the matter flowing through the shock

$$\rho_{\text{sh}} = 7\rho_{\text{in}}, \quad (2.12)$$

where  $\rho_{\text{in}}$  is the density of infalling matter, and, from equations (2.4) and (2.9),

$$\rho_{\text{in}} = \frac{\dot{M}}{4\pi r_{\text{sh}}^2 v_{\text{in}}}, \quad v_{\text{in}} = \left( \frac{2GM}{r_{\text{sh}}} \right)^{1/2}. \quad (2.13)$$

From conservation of matter flow through the shock

$$v_{sh} = -\frac{1}{2}v_{in}, \quad (2.14)$$

so that the kinetic energy per unit volume,  $\frac{1}{2}\rho v_{in}^2$ , is almost completely ( $\sim 98\%$ ) changed into thermal energy,

$$\epsilon_{sh} \cong \frac{7}{2}\rho_{in}v_{in}^2, \quad (2.15)$$

the factor of 7 coming from the compression, equation (2.12). Here  $\epsilon_{sh}$  is the energy density just behind the shock. We neglect  $v_{sh}$  as compared with  $v_{in}$ . Since the energy and pressure are radiation-dominated,

$$p_{sh} = \frac{1}{3}\epsilon_{sh} \cong \frac{7}{6}\rho_{in}v_{in}^2. \quad (2.16)$$

The energy loss by neutrino pairs per unit volume is (Dicus 1972)

$$\dot{\epsilon} = 1.06 \times 10^{25} T^9 C(\mu_e/T) \text{ ergs cm}^{-3} \text{ s}^{-1} \quad (2.17)$$

with  $T$  on MeV and  $C(x)$  a slowly varying function of  $x$  which can be computed from the paper of Dicus. For  $x = 0$ ,  $C = 0.92$ ; we shall use this value because the electrons are not very degenerate. In the region of the neutron star where  $\dot{\epsilon}_n$  is operative, the temperature is  $T \approx 1 \text{ MeV}$ , so that  $e^+, e^-$  as well as photons contribute to the pressure. With  $T$  in MeV, the photon blackbody energy density is

$$W = 1.37 \times 10^{26} T^4 \text{ ergs cm}^{-3}. \quad (2.18)$$

We multiply by 11/4 to include effects from  $e^+$  and  $e^-$  and divide by 3 to obtain the pressure:

$$p = 1.26 \times 10^{26} T^4 \text{ ergs cm}^{-3}. \quad (2.19)$$

We can then write

$$\dot{\epsilon}_n = cp^{2.25}, \quad (2.20)$$

with  $c = 1.833 \times 10^{-34}$  in cgs units.

The neutrino cooling is taken to occur within a pressure scale height  $r_m/4$  of the neutron star, or a volume of  $\pi r_m^3$ . Energy conservation gives

$$\pi r_m^3 c p_m^{2.25} = \frac{GMM}{r_m}. \quad (2.21)$$

Using equations (2.11)–(2.17),  $p_m$  can be written in terms of  $r_{sh}$  and  $r_{cr}$  ( $\sim 10 \text{ km}$ ). Equation (2.19) can then be solved for  $r_{sh}$ , giving

$$\begin{aligned} r_{sh} &= \dot{M}^{-10/27} (GM)^{-1/27} \left( \frac{12\sqrt{2}\pi}{7} \right)^{2/3} \frac{1}{(c\pi)^{8/27}} r_m^{40/27} \\ &= 6.4 \times 10^8 \text{ cm} \left( \frac{\dot{M}}{M_\odot \text{ yr}^{-1}} \right)^{-10/27}. \end{aligned} \quad (2.22)$$

This calculation of the power  $-10/27 \approx -0.37$  reproduces that of the detailed calculation of Houck & Chevalier (1991) using the accurate neutrino cooling function. The only appreciable correction to our schematic model found in the Houck & Chevalier calculation arises from general relativity, which changes the region near the neutron star. This can be taken into account by multiplying the expression for  $r_{sh}$  by 0.4. Thus

$$r_{sh} \cong 2.6 \times 10^8 \text{ cm} \left( \frac{\dot{M}}{M_\odot \text{ yr}^{-1}} \right)^{-0.37}. \quad (2.23)$$

As long as  $r_{sh} < r_{cr}$ , the photons will be carried in by the adiabatic flow, with little radiation pressure at distances  $r > r_{cr}$  to oppose the inflow from larger distances. The critical value of  $\dot{M}$  is then given by

$$r_{cr} = 0.6R\dot{M} = r_{sh} = r_{cr}, \quad (2.24)$$

with  $r_{sh}$  given by equation (2.23). We include here the correction factor  $3/\pi^2$  discussed following equation (2.10). Solution of equation (2.24) gives

$$\dot{M}_{cr} = 1.62 \times 10^{-4} M_\odot \text{ yr}^{-1} (\dot{M}_{cr} = 1.09 \times 10^4). \quad (2.25)$$

It can be seen from equation (2.1) that for SN 1987A this critical rate of accretion is reached at a time  $t = 1.1 \text{ yr}$ , in rough agreement with Chevalier (1989) and Houck & Chevalier (1991). For our work in the remaining part of this paper, the important result is that at accretion rates of  $\sim 1.09 \times 10^4$  times Eddington and higher, photons will be trapped and carried adiabatically inward with the flow. At higher accretion rates, photon pressure cannot stop the inflow of material. The trapping radius corresponding to  $\dot{M}_{cr}$  can be obtained from equation (2.24):

$$r_{cr} = 6.5 \times 10^9 \text{ cm}. \quad (2.26)$$

### 3. SPIRAL-IN OF A NEUTRON STAR IN A BINARY

In the standard scenario for making binary pulsars such as PSR 1913+16 (Flannery & Van den Heuvel 1975), the neutron star formed in the explosion of the primary supergiant enters into the envelope of the companion. (The hydrogen envelope of the primary, which evolves first, has already been transferred to the companion.) The neutron star spirals in, losing orbital energy through dynamical friction (drag). The latter will expel the envelope of the star through hydrodynamical coupling to the envelope. Because of the energy loss, the orbit of the neutron star will be tightened, down to a radius of several  $R_\odot$ .

In discussions of this scenario, possible accretion of matter onto the neutron star has generally been limited, in order of magnitude, to the Eddington limit (although Van den Heuvel & DeLoore 1973 considered the possibility that the neutron star went into a black hole). Chevalier (1993) has, however, recently pointed out that sufficiently high densities will be encountered by the neutron star that photons will be trapped in the inflow of the matter. The amount of accretion will generally be large enough to send the neutron star into a black hole. Only in special circumstances can this be avoided.

In the spiral-in process, the relative velocity of the neutron star to the envelope material  $V$  is usually supersonic, with Mach number  $V/c_s = \mathcal{M} > 1$ , where  $c_s$  is the velocity of sound in the matter. The case of accretion where  $V > c_s$  was considered by Hoyle & Lyttleton (1939) and Bondi & Hoyle (1944). Bondi (1952) suggested for the intermediate case of  $\mathcal{M} \sim 1$ , the approximation

$$\dot{M} = \frac{4\pi\lambda(GM)^2\rho\infty}{(V^2 + c_s^2)^{3/2}}, \quad (3.1)$$

with  $\lambda = 1/2$ . Detailed numerical calculations (Hunt 1971, 1979; Shima et al. 1985) show  $\lambda$  should be close to unity, for the supersonic situation, and we shall use  $\lambda = 1$ .

The detailed picture of Bondi & Hoyle (1944) was not reproduced by the numerical calculations (Hunt 1971; Shima et al. 1985). In particular, no accretion column developed behind the

neutron star, although accretion was larger at back angles than at forward ones. Instead, bow shocks tended to form in front of the accreting object. In spite of the flow being much more complicated than radially symmetrical free fall, even for sizable Mach number  $M$ , the density profiles were nearly spherically symmetrical, as were the streamlines, at small radii  $r < 0.1R_A$ , where  $R_A$  is the accretion radius

$$R_A = \frac{2GM}{V^2}. \quad (3.2)$$

The near radial symmetry of the flow at these short distances can be understood by the fact that the velocities developed in the infall are much greater than the critical velocities, so that by the time the matter reaches the compact object it has little memory of its initial velocity. Chevalier (1993) estimates accretion radii of  $1 \times 10^{11}$  cm and  $1 \times 10^{12}$  cm for the two stellar envelopes he considers, so that  $0.1R_A$  exceeds  $r_{cr}$  of equation (2.26). Consequently, the density and the flow is reasonably spherically symmetric in the photon-trapping region. Although the matter velocity is only  $\sim$  half of free fall, Hunt (1971) and Shima et al. (1985) find the total accretion to be given by equation (3.1) with  $\lambda \sim 1$ , as noted.

As the neutron star spirals in, it first encounters low densities, at the outer edge of the envelope of the large star. Something like a nascent Thorne-Zytkow star (Thorne & Zytkow 1977), i.e., a neutron star with envelope supported by radiation pressure, starts to form. We now argue that the envelope of the forming Thorne-Zytkow object will be overwhelmed by material ram pressure, once the neutron star spirals in to higher densities.

The envelope of the nascent Thorne-Zytkow star is initially supported by radiation pressure. We can simply estimate this pressure as follows. In the equation of hydrostatic equilibrium

$$\frac{1}{\rho} \frac{dp}{dr} = - \frac{GM}{r^2} \quad (3.3)$$

we replace, for order of magnitude estimate,  $dp/dr$  by  $p/r$ . We then estimate the pressure to be

$$p \approx \frac{GM}{r} \rho. \quad (3.4)$$

The accretion for some region around the neutron star, in particular out to the region  $r \sim r_{cr} = 6.5 \times 10^9$  cm which is crucial for photon trapping, will be at the Eddington limit,  $1.5 \times 10^{-6} M_\odot \text{ yr}^{-1} = 0.95 \times 10^{16} \text{ g s}^{-1}$ . Assuming a free-fall velocity for the material, we can obtain the density  $\rho$  from

$$\rho_{Edd} = \frac{\dot{M}_{Edd}}{4\pi r_{cr}^2 v_{ff}} = \frac{3.9 \times 10^{-6}}{r_{cr}^{5/2}} \text{ g cm}^{-3}, \quad (3.5)$$

giving

$$\rho_{Edd} = \rho_{Edd} \frac{GM}{r} = \frac{7.28 \times 10^{14} \text{ dyn cm}^{-2}}{r_{cr}^{5/2}}, \quad (3.6)$$

where  $r_{cr}$  is the radius in units of  $10^6$  cm. Of course, pressures will increase from this initial value as density piles up.

Detailed calculations for a  $14.45 M_\odot$  Thorne-Zytkow star by Cannon et al. (1992) show the accretion rate to indeed be constant from near the neutron star surface out to  $r > 10^{12}$  cm, with a steady state pressure of

$$p \approx 3 \times 10^{10} \text{ dyn cm}^{-2} \quad (3.7)$$

at  $r_{cr} = 6 \times 10^9$  cm. (The density is  $\sim 10^{-6} \text{ g cm}^{-3}$  at this radius.) In the nascent Thorne-Zytkow star which is just forming, the  $\rho_{Edd}$  of equation (3.5), which is only  $\sim 10^5 \text{ dyn cm}^{-2}$ , is more appropriate, because matter has not had time to pile up. In any case, it is safe to say that the pressure will be somewhere between  $10^5$  and  $10^{10} \text{ ergs cm}^{-3}$ , closer to the lower value as the nascent Thorne-Zytkow star begins to form, before pileup of the density.

Velocities  $V$  of the neutron star relative to envelope material are  $\sim 2 \times 10^7 \text{ cm s}^{-1}$ . In the frame in which the neutron star is at rest, the envelope gas will give a ram pressure

$$p_{ram} = \rho V^2. \quad (3.8)$$

Initially, as the neutron star enters the stellar envelope, this pressure will be small, because of the low-density  $\rho$ , and will not sensibly disturb the beginning formation of the Thorne-Zytkow star. However, later, as the neutron star spirals in, higher densities will be encountered.

We will use the results of Bodenheimer & Taam (1984) who made a detailed two-dimensional calculation of the spiral-in of a neutron star through the envelope of a  $16 M_\odot$  supergiant. Bodenheimer & Taam considered two cases, the first in which the  $16 M_\odot$  companion was brought to the beginning of helium core burning. The main-sequence evolution lasted  $1.03 \times 10^7$  yr and was followed by an overall contraction lasting  $1.56 \times 10^6$  yr. The helium burning at the center of a helium core of  $3.1 M_\odot$  set in  $3.28 \times 10^4$  yr later. The stellar radius at that time was  $4.3 \times 10^{12}$  cm. Some of these characteristics will be important for us later.

For our present discussion, their case 2, where the star is evolved an additional  $1.68 \times 10^5$  yr into the red giant phase, with  $R = 3.7 \times 10^{13}$  cm, a factor of  $\sim 10$  larger than in case 1, is more relevant. At least, with the lower envelope densities, it will give much less accretion onto the neutron star than case 1. We carry out our estimates for the red giant, case 2, as the best case to avoid excessive accretion and shall simply quote results for case 1.

Although the common envelope time is given as  $\sim 10^3$  yr in the literature, soon after the neutron star is inserted into the red giant envelope at a distance of  $3.9 \times 10^{11}$  cm from the center of the companion a luminosity of  $\sim 5 \times 10^{41} \text{ ergs s}^{-1}$  (of the order of luminosities in early supernova light curves) is generated by the dynamical friction. The hydrodynamical coupling of the latter to the stellar envelope is good, with an efficiency of  $\sim 50\%$ . In less than a year enough frictional energy is produced to expel the envelope, which is initially bound by  $1.8 \times 10^{48}$  ergs, although it takes  $\sim 10$  yr for the envelope to be expelled. The dynamical time for the main process is thus several years.

The envelope density at point of insertion star is  $6 \times 10^{-4} \text{ g cm}^{-3}$ . With an assumed neutron star velocity of  $\sim 10^7 \text{ cm s}^{-1}$  this gives a ram pressure of

$$p_{ram} = \rho V^2 \sim 10^{11} \text{ dyn cm}^{-2}. \quad (3.9)$$

This is somewhat greater than the pressure, equation (3.7), in the  $14.45 M_\odot$  Thorne-Zytkow star at  $r_{cr} = 6 \times 10^9$  cm and about 6 orders of magnitude larger than the initial  $\rho_{Edd} \simeq 10^5 \text{ dyn cm}^{-2}$  at that radius, from equation (3.6). Thus, already at the point of insertion, the ram pressure will overwhelm that of the nascent Thorne-Zytkow star. We emphasize that we have used a very diffuse red giant envelope, with low densities, and taken the density at an early point in the spiral-in; higher densities will be encountered later. Already here, the ram pres-

sure is about as large as that in a  $14.45 M_{\odot}$  Thorne-Zytkow star at the relevant radius  $r_{cr}$ , and in this star the pressure has been greatly increased by pileup, as compared with the nascent Thorne-Zytkow star.

At  $r_{cr} = 6 \times 10^9$  cm,  $\rho_{\text{BD}}$  of equation (3.5) is only  $\sim 10^{-11}$  g cm $^{-3}$ , a factor of  $\sim 10^9$  below the envelope density at the point Bodenheimer & Taam (1984) inserted the neutron star. In the frame in which the neutron star is at rest, the matter with initial density  $\rho = 6 \times 10^{-4}$  g cm $^{-3}$  brought in at  $r_{cr}$  will fall onto the neutron star with an  $\dot{m} \sim 10^8$ , much higher than the  $\dot{m}_{cr} = 1.09 \times 10^4$  of equation (2.25). Thus, the neutron star will be blanketed, the photons being swept in with the flow.

It is generally believed that neutron stars are formed with strong magnetic fields  $B \sim 10^{12}$  G, in rapid rotation. Ilyarionov & Sunyaev (1985) have used this strong, rapidly rotating magnetic field to keep material from accreting onto the neutron star through the so-called "propeller effect." At the low densities first encountered by the neutron star as it enters the envelope of the giant, this effect may be operative, but as the pressure of the envelope material becomes larger than the magnetic pressure  $B^2/8\pi$ , effects from the magnetic field will no longer be important. The magnetic pressure is tiny compared with that from the hot radiation-dominated accretion envelope in the neighborhood of the neutron star, where the temperature reaches  $T \sim 1$  MeV.

Chevalier dealt with the usually difficult problem of angular momentum (1993), finding that taking it into account does not sensibly change the accretion scenario. Because of the density gradient in stellar envelopes, angular momentum is clearly encountered in the accretion. Chevalier shows that the necessary accretion disk will be formed not far outside the neutron star, and that the matter will drop onto the neutron star. We would point out that the accretion disk would form well within the radius of the accretion shock,  $r_{sh}$ . Angular momentum will be transferred to the accretion envelope, mostly at quite small radii, but this will not change our general picture.

#### 4. RELATION OF ACCRETED MASS TO DYNAMICAL FRICTION ENERGY

Hypocritical mass accretion and dynamical friction arise from the same basic physical process, the deceleration of the ambient gas by the gravitational field of the neutron star. In the case of mass accretion, the matter inside of the accretion radius  $R_A$  of equation (3.2) ultimately comes to rest on the neutron star. The deceleration of matter outside of  $R_A$  contributes to the dynamical friction, but not to the accreted mass.

More specifically, the accretion rate can be written, for  $V \gg c_s$ , as

$$\dot{M} \cong \pi R_A^2 \rho_{\infty} V, \quad (4.1)$$

whereas the drag force  $F$  for supersonic flow is (Dokuchev 1964; Ruderman & Spiegel 1971; Rephaeli & Salpeter 1980)

$$F = \pi R_A^2 \rho_{\infty} V^2 \ln(b_{\max}/b_{\min}), \quad (4.2)$$

where  $b_{\max}$  and  $b_{\min}$  are the gravitational upper and lower cutoff distances, respectively. The drag coefficient is, then,

$$C_d = 2 \ln(b_{\max}/b_{\min}). \quad (4.3)$$

With this definition,

$$F = \frac{1}{2} C_d V^2 \dot{M}. \quad (4.4)$$

The neutron star, therefore, expends energy in dynamical friction at a rate

$$\dot{E} = FV = \frac{1}{2} C_d V^2 \dot{M}. \quad (4.5)$$

As noted earlier, the main expenditure of energy takes place during a time of  $\sim 1$  yr. Integrating over time,

$$E \cong \frac{1}{2} C_d \langle V^2 \rangle \Delta M, \quad (4.6)$$

where  $\langle V^2 \rangle$  is the mean square velocity of the neutron star during the relevant time, and  $\Delta M$  is the total accreted mass. Direct numerical calculation (Shima et al. 1985) shows  $C_d$  to have a value between 6 and 7, for Mach number  $M \geq 2$ , not very dependent on  $\gamma$ , and a value of 4-5 is found for Mach number unity, the 4 for  $\gamma = 5/3$ , 5 for  $\gamma = 4/3$ . We choose  $C_d = 6$  for our present discussion, giving

$$E \cong 3 \langle V^2 \rangle \Delta M. \quad (4.7)$$

As noted, the dynamical friction should contribute the energy to remove the envelope of the supergiant. The envelope is bound with energy  $E_b = 1.8 \times 10^{48}$  ergs. Given coupling efficiency of  $\sim 50\%$  discussed above,  $E$  equals twice the binding energy. We have

$$\Delta M \approx \frac{1}{2} E_b / \langle V^2 \rangle = 6 M_{\odot} / \langle V^2 \rangle, \quad (4.8)$$

where  $V^2$  is the value with  $V$  measured in  $10^7$  cm. For values of  $\langle V^2 \rangle$  of several, this is  $\geq 1 M_{\odot}$ .

Although  $C_d$  drops to  $\sim 0.4$  for the subsonic  $V/C_s \equiv M = 0.6$  calculated by Shima et al., we should remember that we assumed  $M \gg 1$  in deriving the above formulae. Using equation (3.1) with  $\lambda = 1$  as an interpolation formula in order to go to subsonic velocities, we find that the relations (4.5) and (4.6) are valid provided that  $C_d$  is replaced by

$$(C_d)_{\text{eff}} = C_d \left( \frac{M^2 + 1}{M^2} \right)^{3/2}. \quad (4.9)$$

Since  $[(M^2 + 1)/M^2]^{3/2} = 7.34$  for  $M = 0.6$ ,  $(C_d)_{\text{eff}}$  drops much less rapidly than  $C_d$  with decreasing  $M$ . Consequently, our above results are not greatly changed by the motion going slightly subsonic.

In the first case of Bodenheimer & Taam of the more compact star, densities are much higher, the envelope binding energy is  $2.25 \times 10^{49}$  ergs, and the  $6 M_{\odot}$  in equation (4.8) would be replaced by  $75 M_{\odot}$ ! Since Brown & Bethe (1994) argue that the maximum mass for neutron stars is  $M_{\max} \simeq 1.5 M_{\odot}$ , and since they are born with masses  $1.25 M_{\odot} < M < 1.45 M_{\odot}$  (Burrows & Woosley 1986) accretion of  $\sim 1 M_{\odot}$  will almost certainly send the neutron star into a black hole.

Chevalier (1993) estimated accretion rates of  $0.2 M_{\odot} \text{ yr}^{-1}$  for the red giant extended star and  $5 M_{\odot} \text{ yr}^{-1}$  for the compact star. By connecting our accretion rate with the total amount of dynamical friction, we have made the case more definite. Special conditions are clearly needed if the neutron star is to survive. More likely, we believe, is that the neutron star goes into a black hole, sucking up the companion star. In this way, a heavy black hole would be the end result. This is one of the two possibilities noted by Van den Heuvel & De Loore (1973). Indeed, in their seminal paper in 1975 Flannery & Van den Heuvel (1975) said "The absence of any known binary pulsar among the 160 pulsars within 3 kpc distance therefore implies either that the probability for disruption is large for these systems or that the remnants of most helium stars are black holes instead of neutron stars."

### 5. SPECIAL CIRCUMSTANCES THAT ALLOW FOR BINARY PULSTAR EVOLUTION

As noted by Chevalier (1993), special conditions may be needed to prevent black hole formation. In calculations to date, neutron stars are born with masses of  $\sim 1.25$  to  $1.45 M_{\odot}$  (Burrows & Woosley 1986). Brown & Bethe (1994) suggested that the maximum neutron star mass is  $1.5 M_{\odot}$ . At the time, only the Vela X-1 star did not fall within this limit. Recent measurement and analysis by van Kerkwijk et al. (1994) brings the lower limit on the error bar down to  $1.37 M_{\odot}$ , so even this pulsar does not violate the suggested limit. There is, therefore, but little room for accretion to increase the neutron star mass without the neutron star going into a black hole. Our estimates in the last section suggest that even with a diffuse red giant envelope, the neutron star entering this envelope may go into a black hole in the spiral-in.

As noted earlier, the standard mechanism for tightening the binary orbit, by a large factor, is by giving up orbital energy through dynamical friction. Assuming a  $1.4 M_{\odot}$  neutron star and a  $4 M_{\odot}$  helium core, the binding energy equal to half the energy in dynamical friction (taking the efficiency of hydrodynamical coupling to the envelope to be 0.5) is

$$\frac{GM_1 M_2}{2a} = 3.6 \times 10^{48} \text{ ergs}, \quad (5.1)$$

giving  $a \sim 3 R_{\odot}$  as the final separation. We have neglected here the small binding energy of the initial, much wider, binary. Coincidentally, this  $a$  is near the  $a = 2.8 R_{\odot}$  of PSR 1913+16.

Recent work by Terman, Taam, & Hernquist (1994) considering three-dimensional effects in the merger of a red giant with a dwarf companion indicates that the spiral-in is not well described by the dynamical friction picture. The non-axisymmetric gravitational torque is more effective in shrinking the orbit and expelling the common envelope. At least for the case of the red giant envelope where density gradients are large, our picture outlined above is likely to overestimate the accretion by a large factor (Taam 1994). We have, however, made a simple connection between the energy that would have to be furnished by dynamical friction in order to expel the envelope and the accretion. Minimally, our development shows that the standard scenario, which uses the dynamical friction to expel the envelope, does not work.

How the spiral-in of the neutron star in the common envelope goes in detail is not known, and more extensive calculations are needed to pin this down. It is clear from observations that the outcome of almost none of these spiral-ins is a neutron star plus helium star binary. The high-mass X-ray binary phase lasts  $2-5 \times 10^4$  yr (Van den Heuvel 1976), whereas the neutron star, helium star binary lasts for the average helium burning time of  $\sim 5 \times 10^5$  yr. Therefore, according to the standard scenario, there should be 10–25 more of the latter than the former, or  $\sim 300$ –750 helium star plus neutron star binaries. In fact, only one has been observed, Cygnus X-3 (in which the compact object may be a black hole). This discrepancy between anticipated numbers and the one (possible) observed such system was noted already by Van den Heuvel (1976). We do not know whether the neutron star in spiral-in tidally disrupts the core, or whether it merges with the core, before or after turning into a black hole, etc. However, it is clear that essentially no neutron stars survive the spiral-in.

Our suggestion, given the above arguments, is that the neutron star should not transverse a hydrogen envelope, if it is

to survive. If the two supergiants can eject their hydrogen envelopes and burn helium at the same time, then the above problem of too much accretion might be avoided. In short, we suggest that binary pulsars may be products of helium star binaries.

The circumstances forming a helium star binary are, indeed, special. The helium-burning time for a  $16 M_{\odot}$  main-sequence star (with helium core of  $\sim 4 M_{\odot}$ ) is only  $1.25 \times 10^6$  yr (Paczynski 1971; Habets 1986). The Taam et al.  $16 M_{\odot}$  supergiant spent  $1.03 \times 10^7$  yr in main-sequence evolution. Consequently, the helium-burning time is only  $\sim 12\%$  of the stellar lifetime. If the primary and secondary are to burn helium at the same time, their lifetimes can differ by, at most, 12%. Since lifetimes scale roughly as the inverse cube of the mass, this means that the primary can be no more than 4% heavier than the secondary. We shall later review work evolving WR + WR binaries that bears out our simple estimate.

The closeness in mass of primary and secondary means that the resulting neutron stars in the binary, if they survived the explosion, would be nearly equal in mass. This is certainly true in the two relativistic binary pulsars and is likely true in PSR 2303+46. In PSR 1913+16,  $16 M_{\odot}$  stars are good progenitors (Burrows & Woosley 1986). In fact, the neutron star masses in the binary are rather close,  $1.442 \pm 0.001 M_{\odot}$  and  $1.386 \pm 0.001 M_{\odot}$  (Thorsett et al. 1993). The larger neutron star would have resulted from the first explosion; presumably it was spun up by accretion from the unexploded star (Smarr & Blandford 1976). Helium stars lose mass at a rate of  $\sim 10^{-6}$  to  $10^{-5} M_{\odot} \text{ yr}^{-1}$  by wind (Langer 1989) and this is sufficient to give accretion at roughly Eddington limit onto the neutron star. In our scenario, the second helium star would be less massive than the first to explode, so one would expect the second neutron star to be somewhat less massive, although the mass of the resulting neutron star changes rather slowly (and, possibly, nonmonotonically) with change in helium star mass. Furthermore, the first neutron star would be estimated to have accreted  $\sim 0.015 M_{\odot}$  from the helium star wind in the average helium star burning time of  $5 \times 10^5$  yr, assuming accretion of ionized  $N - Z$  matter at the Eddington limit. We will return to a discussion of mass loss in helium stars later. A hydrogen envelope does not have to be around for the accretion.

The masses of the relativistic binary PSR 1534+12 are  $1.337 \pm 0.003 M_{\odot}$  and  $1.341 \pm 0.003 M_{\odot}$  (Wolszczan 1991; Arzoumanian, Taylor & Wolszczan 1994). It is likely that the 12 day PSR 2303+46 binary is composed of two neutron stars, of similar masses, although only the average mass  $m_{av} = 1.27 M_{\odot}$  is well constrained (Thorsett et al. 1993). The large distance ( $\sim 500$  pc) of the pulsar from the galactic plane constrains the probable main-sequence mass of the secondary to be  $\leq 10 M_{\odot}$  (Stokes, Taylor, & Dewey 1985). The scenario in which such small mass stars produce neutron stars is not well understood, because an iron core in hydrostatic equilibrium never develops in their evolution. However, it is believed that neutron stars do result. Burrows & Woosley (1986) find a gravitational mass range of  $1.27$ – $1.32 M_{\odot}$  for the neutron star resulting from an  $11 M_{\odot}$  main-sequence star, and it is likely that the range is not much different for a  $10 M_{\odot}$  star.

In detailed calculations Woosley & Weaver (1994) find that an  $11 M_{\odot}$  main-sequence star with solar metallicity develops a  $1.32 M_{\odot}$  baryon number core. This translates, with the binding energy correction of  $E = 0.075 M_{\odot} (M/M_{\odot})^2$  of Lattimer & Yahil (1989) into a gravitational mass of  $1.22 M_{\odot}$ . This could be increased somewhat by fallback in the explosion.

A helium star progenitor mass of  $\sim 2.1 M_{\odot}$  is found for PSR 1534+12 (Wolszczan 1991), which is already on the small side, and that for a  $1.27 M_{\odot}$  neutron star would probably be smaller. Thus, the  $1.27 M_{\odot}$  is probably near the lower limit for a neutron star mass, and if one of the neutron stars in the binary were to have a substantially larger mass, the other would have to fall below this limit. Both of the stars thus should be close in mass to  $m_{\text{ev}}$ . Consequently, there would appear to be merit in a scenario which makes the two neutron stars close in mass in a binary.

It should be noted that calculations (Weaver & Woosley 1994) find remnant core masses in supernova explosions to fluctuate somewhat, as a function of stellar main-sequence mass. Thus, for solar metallicity, stars of mass 15, 18, 20, 25, and  $30 M_{\odot}$  have baryonic Fe core masses of 1.32, 1.46, 1.74, 1.78, and  $1.83 M_{\odot}$ , respectively. (Gravitational masses are  $\sim 10\%$  lower.) Such variations are expected to result from shell burning, which changes irregularly with change in mass. There may be other causes. Given such variations, the difference between neutron star masses found in PSR 1913+16 may be more typical than the near equality in PSR 1534+12.

We believe our helium stars to be relatively low mass Wolf-Rayet stars. The general scenario for the possible evolution of W-R + W-R binaries has recently been outlined by Pols et al. (1991). The (larger) primary evolves first, transferring mass to the secondary as it goes through the red giant phase. The secondary then evolves, and a reverse mass transfer to the helium star takes place. The consequences are uncertain, but it is expected that the helium star will expand upon accretion and that a common envelope will form (Iben & Tutukov 1985). A double helium star in a close orbit may be the result of such evolution. There is, in fact, substantial literature on the formation of W-R + W-R binaries.

Doom & DeGreve (1982) estimated that a fraction of 0.5%–3% of all massive O-type binaries would become W-R + W-R binaries. Yungelson & Tutukov (1991) arrived at essentially the same results using different models. The relative rarity of W-R + W-R binaries results, of course, from proximity in mass required for the two stars in the progenitor binary. Vrancken et al. (1991) show that it would be difficult to see W-R + W-R binaries and estimate that less than 1% can be observed. Nonetheless, a good candidate, HD 5980, has been observed (Koenigsberger et al. 1994). This binary has a 19 day period.

An interesting possibility is that Cygnus X-3 resulted from the explosion of one of the stars in a W-R + W-R binary. Van Kerkwijk (1993) finds a mass loss through wind of  $\dot{M} \simeq 4 \times 10^{-5} M_{\odot} \text{ yr}^{-1}$ . In later work (van Kerkwijk et al. 1994) the observed mass loss was found to be a factor  $\sim 3$  larger, but a discussion of possible effects of inhomogeneities suggests that the estimate may be too large.

In a comprehensive study of mass-loss rates of Wolf-Rayet stars, Langer (1989) finds

$$\dot{M} = (0.6 - 1.0)10^{-7} \left( \frac{M_{\text{WR}}}{M} \right)^{2.5} M_{\odot} \text{ yr}^{-1}. \quad (5.2)$$

Taking 0.8, the central value for the coefficient, we find  $M_{\text{WR}} = 12 M_{\odot}$  in order to reproduce  $\dot{M} = 4 \times 10^{-5} M_{\odot} \text{ yr}^{-1}$ . (With the coefficient unity, an  $11 M_{\odot}$  W-R would be needed.) If the present mass is  $\sim 12 M_{\odot}$ , it is likely that the initial W-R mass was heavier, since a substantial amount could have blown away with the above rate of mass loss. This means

that the initial mass of the present helium star in Cyg X-3 resulted from a main-sequence mass, before mass loss, of  $M_{\text{ZAMS}} > 35 M_{\odot}$ . If Cyg X-3 evolved from a W-R + W-R binary, the mass of the primary, which exploded into the neutron star, would have been similar. Brown, Weingartner, & Wijers (1995) show that stars in this range will generally go into low-mass (rather than high-mass) black holes provided that their hydrogen envelope is lifted off before significant helium is burned. Indeed, given the adopted mass-loss rate, accretion on to a low-mass black hole, via an accretion disk, is easily calculated to be somewhat super-Eddington. The presence of jets in Cyg X-3, as we remark later, indicates such an accretion rate. Thus a high-mass black hole is probably ruled out. (I am indebted to Marten van Kerkwijk for this observation.)

Cygnus X-3 does not pulse, so there is no need for a neutron star. However, with the ambient density from  $\dot{M} = 4 \times 10^{-5} M_{\odot} \text{ yr}^{-1}$  and the  $R < 4.8 \times 10^{11} \text{ cm}$  found by Kitamoto et al. (1992) from the diplike structure in the X-ray light curve, we find an optical depth

$$\tau < 2.26, \quad (5.3)$$

and  $\tau \sim 2$  may be sufficient to preclude detection of pulsations. Also, our following argument about magnetic field decay because of accretion would bring  $B$  down substantially, the precise value depending upon the time since the primary exploded.

Although the lack of pulsation is not a strong argument, the large W-R mass presently necessary to produce the observed wind does preclude a neutron star if Cyg X-3 resulted from a W-R + W-R binary.

Of course, the helium star + neutron binary Cyg X-3 is not only a rare, but also a unique event, which might be the only presently observed survivor of common envelope phase. Our earlier arguments indicate that the neutron star would have accreted enough matter to become a black hole, in this case also. We note that Cherepashchuk & Moffat (1994) have given observational arguments that the compact object in Cyg X-3 is a black hole. The ultimate fate of Cyg X-3 will be, we believe, a black hole + black hole binary.

Although the average lifetime of a helium star,  $\sim 5 \times 10^5 \text{ yr}$ , is long compared with the  $\sim 2\text{--}5 \times 10^4 \text{ yr}$  existence of a high-mass X-ray binary, it is short compared with typical evolutionary times in the standard scenario for relativistic binary pulsars. In the  $\log B(G)$  versus  $\log P(s)$  diagram (Fig. 2 of Bhattacharya & van den Heuvel 1991) the pulsar will have to spin down, experience field decay and be spun up again by accretion to its present frequency. In the  $\sim 5 \times 10^5 \text{ yr}$  long enough for all of this to take place?

Taan & van den Heuvel (1986) have shown that empirically field decay is inversely correlated with mass accretion, although there is up to now no fundamental basis for this correlation. Let us follow this argument in order to see whether the helium star time is long enough to fit this correlation. Winds are sufficient from a helium star that accretion should take place at roughly the Eddington limit  $\dot{M}_{\text{Edd}}$ .

The (empirical) relationship between field decay and accretion has been modeled by Shibazaki et al. (1989) by

$$B = \frac{N_0}{1 + \dot{M} t_{\text{acc}}/m_B}, \quad (5.4)$$

where  $t_{\text{acc}}$  is the time accretion proceeds and  $m_B$  is a constant. The 1 in the denominator is unimportant for our argument and

we take  $\dot{M} = \dot{M}_{\text{Edd}}$  also constant, so we can set

$$B \cong CB_0/t_{\text{acc}}. \quad (5.5)$$

The  $C$  can now be determined by noting that the millisecond pulsars such as PSR 1937+21 and PSR 1957+20, which are found to be spun up to their short periods in  $t_{\text{acc}} \approx 10^7$  yr (Cook, Shapiro, & Teukolsky 1993), have magnetic fields  $B \sim 5 \times 10^6$  G. For  $B_0 = (1-5) \times 10^{12}$  G,

$$C = (1-5) \times 10^3, \quad (5.6)$$

the lower value of  $C$  pertaining to the higher value of  $B_0$ .

Using  $t_{\text{acc}} \approx 5 \times 10^5$  yr, the average helium star lifetime, gives us a magnetic field of  $\sim 10^{10}$  G (for either  $B_0 = 10^{12}$  G or  $5 \times 10^{12}$  G, once the appropriate value of  $C$  is used), just about right for the relativistic binary pulsars. Shibasaki et al. (1989) showed that the binary pulsar periods (and those of the millisecond pulsars), determined from the spin up by accretion are consistent with their magnetic fields. Consequently, the above  $\sim 5 \times 10^5$  yr is presumably also time to spin up the binary pulsars to their observed periods.

As noted above, helium stars of masses  $2-4 M_\odot$  would be required to make the observed binary pulsars. In the literature, workers have favored higher masses for Wolf-Rayet stars; e.g., Vrancken et al. (1991) find that taking a minimum mass  $(M_{\text{WR}})_{\text{min}} = 5 M_\odot$  results in an overproduction of W-R stars in the Galaxy by a factor of 2 or 3, contrary to the case with  $(M_{\text{W-R}})_{\text{min}} = 10 M_\odot$ . However, recent hydrodynamical calculations (Enzman & Woosley 1988; Shigeyama et al. 1990) show that lower mass helium stars formed in binaries reproduce the light curves of Type Ib supernova explosions. Podsiadlowski, Joss, & Hsu (1992) find that  $\sim 15\%-30\%$  of all massive stars (with initial masses between 8 and  $20 M_\odot$ ) end their lives as helium stars with masses between 2 and  $6 M_\odot$ . They note that this frequency is consistent with the frequency of Type Ib/Ic supernovae.

In summary, the special circumstances that we have adduced under which black hole formation can be avoided resulted in the masses of the two neutron stars in a binary to be nearly equal. This seems to be observed in the binary pulsars observed to date.

## 6. GENERAL DISCUSSION

In this paper we have presented arguments which substantially revise the scenario of common-envelope evolution, so one may have a number of worries. Our thesis is that only under exceptional conditions can a neutron star survive the common envelope with a large star and our example has been with O stars overflowing their Roche lobes. The paucity of neutron star-helium star binaries—Cyg X-3 being the only known example—supports our contention that there are not many common-envelope phases in which the neutron star survives. Even in Cyg X-3 the compact object is probably a low-mass ( $\sim 1.5 M_\odot$ ) black hole.

Our argument depended upon hypercritical accretion, where  $\dot{M} \geq 10^4 \dot{M}_{\text{Edd}}$ . It has been believed that accretion can proceed somewhat faster than Eddington in consideration of binary systems, although our large hypercritical accretion has not generally been considered. In the system SS 433 the massive donor seems to be overflowing the Roche lobe at a super-Eddington rate, with no sign of the system evolving toward a common-envelope phase. Matter is carried off in jets. Unpublished calculations by the author, referred to in the last section, indicate that accretion on to the compact object in Cyg X-3 is

somewhat super-Eddington. Again, jets occur in Cyg X-3. (The maximum observed X-ray flux in Cyg X-3 occurs at the times of maximum radio flares; i.e., at the time of the jets, and is roughly the Eddington flux for this source.) It thus seems that for accretion which exceeds the Eddington rate up to a factor of several, jets can carry off the energy.

Following supernova explosions, approximately spherically symmetric accretion occurs at many orders of magnitude greater than Eddington. We believe, and this belief is supported by the quoted calculations of Chevalier and collaborators, that this hypercritical accretion can be carried down to  $\sim 10^4 \dot{M}_{\text{Edd}}$ , the temperatures at the neutron star being high enough so that neutrinos can carry off the energy. How accretion proceeds in the region between  $\dot{M}_{\text{Edd}}$  and  $\sim 10^4 \dot{M}_{\text{Edd}}$  is probably complicated in detail, and only possible in certain circumstances. From the example of SS 433 we suspect that this cannot be handled in spherical symmetry.

We have not considered progenitors of neutron star plus heavy white dwarf, although we referred to the work of Cook et al. (1993) which shows that accretion from a white dwarf onto neutron star via an accretion disk will, in general, not send the neutron star into a black hole, because the maximum mass of the neutron star grows with the angular momentum delivered. Our arguments may suggest that in the standard evolutionary model for, e.g., PSR 0655+64 (van den Heuvel & Taam 1984) the neutron star would not survive a common envelope with the companion star, as the latter evolved.

Indeed, PSR 0655+64 and Her X-1 are often adduced as old systems in which the magnetic field of the neutron star has not decayed appreciably. The scenario for the Her X-1 evolution (Sutanto 1975) does not have the neutron star going through a common envelope.

The  $\sim 8.6 \times 10^{10}$  G magnetic field of 0655+64 (Van den Heuvel & Taam 1984) is just where it should be on the diagram of Shibasaki et al. (1989) for the 0.2 s period, if the neutron star were speeded up by accretion from its evolved companion. Bhattacharya & Van den Heuvel (1991), p. 65, note that 0655+64 has the characteristics of a “recycled” pulsar. The neutron star could not, therefore, have experienced substantial accretion during a common-envelope phase, or its magnetic field would have been lowered. (Remember that the accretion is not limited to Eddington.) There consequently seem to be reasons for evolving 0655+64 so that the neutron star does not have to go through a common-envelope phase.

Other than the few remarks above, we have not considered the evolution of low-mass X-ray binaries, or even high-mass X-ray binaries involving Be companions, and these undoubtedly bring in many new features. What we have learned during this investigation is that each star has its own peculiarities, and we offer our general formalism with full knowledge that many modifications will have to be introduced for each particular binary considered.

## 7. FREQUENCY OF BINARY PULSTAR FORMATION; CONSEQUENCES

The occurrence of binary pulsars is rare. Within 3 kpc, only two binary pulsars, the relativistic PSR 1534+12 at  $\sim 0.5$  kpc and PSR 2303+46 at 2.3 kpc, have been discovered. PSR 1913+16 is  $\sim 5$  kpc away.

On the other hand, at least two high-mass X-ray binaries, Vela X-1 and 4U 1700-37, lie within 3 kpc. The high-mass X-ray binaries have an expected lifetime of only  $2-5 \times 10^4$  yr, whereas that of the binary pulsars is  $\sim 5 \times 10^8$  yr. (That of

PSR 1913+16 can be estimated from the formulae of Verbunt, Wijers, & Burns 1990 to be  $3 \times 10^8$  yr, but gravitational radiation is particularly strong in this relativistic binary pulsar, so we believe our above estimate to be more correct, on the average.) Thus, there should be  $\geq 1000$  times more binary pulsars than high-mass X-ray binaries, to the extent that the former remain intact in the explosion. Van den Heuvel (1976) estimates that for the Galaxy there should be  $\sim 3000/fb$  binary pulsars, where  $f$  is the survival fraction for the binaries in the final explosion of the Wolf-Rayet, and  $b$  is the probability that the pulsar is beamed so that it can be observed. He used a lifetime of only  $\sim 5 \times 10^6$  yr, for the binary pulsars, so his number should be increased to  $3 \times 10^5/fb$ .

The factor  $b$  may not be much smaller than unity. Beginning with Helfand & Becker (1986) it has been understood that the central emission region due to the synchrotron nebula powered by a pulsar, will be observable even if the pulsar is pointed in the wrong direction to be observed.

Pulsars are observed to have high kick velocities, resulting from nonspherical explosions. Harrison, Lyne, & Anderson (1993) find rms kick velocities of  $\sim 300$  km s $^{-1}$  and Fomalont et al. (1992) found a mean value of 355 km s $^{-1}$ . In both cases there was a long tail, up to more than 1000 km s $^{-1}$ . These velocities are substantially higher than those previously quoted (Dewey & Cordes 1987). The minimum kick velocity needed for PSR 1913+16 is estimated by Burrows & Woosley to be  $V_{min} = 170-250$  km s $^{-1}$ , derived from the equations of Flannery & Van den Heuvel (1975). The actual kick velocity may be  $\sim$  twice these values, depending on the preexplosion period  $P_0$ .

Even with kick velocities  $\sim 400$  km s $^{-1}$ , an appreciable fraction of binary pulsars are calculated to survive the explosion. With disruption occurring in  $\sim 90\%$  of the explosions, we would still be left with  $\sim 100$  times more binary pulsars in the standard evolution scenario than are observed.

In any case, it is clear that relativistic binary pulsars, composed of neutron stars, are rare objects. Our suggestion that neutron star binary pulsars come from helium star binaries cuts down the number expected from the standard scenario. Given a high-mass X-ray binary in which one of the stars is a neutron star, the main-sequence mass from which the companion helium star evolved was estimated in the last section to be within  $\leq 4\%$  of the mass of the main-sequence progenitor of the neutron star. Brown & Bethe (1994) estimated the main-sequence range from which neutron stars are produced to be 8 to  $18 M_\odot$ , most higher masses giving black holes. Thus, 4% in mass covers  $\sim 0.5 M_\odot$  in this range, or about 5% of these stars. Given a Salpeter IMF, about half of the main-sequence stars of mass  $M > 8 M_\odot$  lie in the range 8– $18 M_\odot$ . Given that the progenitor of the neutron star is known to lie in this range, then in an uncorrelated distribution of the helium star progenitor, only  $\sim 2\%$  will satisfy our criterion of proximity in mass. Thus the number estimated to follow from our scenario is diminished by  $\sim$  two orders of magnitude from that estimated for the standard scenario.

This estimate assumes no correlation in masses of primary and secondary in binaries. There are indications that a large fraction of observed close binaries contain two stars of nearly equal masses (Abt & Levy 1978; Tutukov & Yungelson 1980).

Therefore, the actual fraction of stars close enough in mass may exceed 2%.

Mustering the courage to make statistics for small numbers (unity), we can say that the one known double Wolf-Rayet binary HD 5980 (Koenigsberger et al. 1994), which is estimated to live the average helium burning time of  $\sim 5 \times 10^5$  yr, is the progenitor of  $\sim 1000/fb$  binary pulsars, since the latter live  $\sim 10^3$  times longer. Clearly, a larger number of W-R+W-R binaries would be an embarrassment for our new scenario.

### 8. CONCLUSIONS

Our calculations and estimates suggest that the standard scenario for forming binary pulsars, in which the neutron star from the first explosion spirals in through the hydrogen envelope of the secondary, does not, in general, work. In almost all cases the neutron star will accrete sufficient mass to turn into a black hole. We do not know whether the black hole causes tidal disruption of the helium core, merges with it, or has some other fate, but the compact core, helium star binary seldom survives. Cygnus X-3 is the only observed case, and even here the compact star may be a black hole.

We suggest that neutron star binaries can evolve from helium star binaries. The latter can form only when the primary and secondary undergo helium burning at the same time. This restricts their masses to be within  $\sim 4\%$  of each other. The outcome of the explosions of the stars in the helium star binary would be two neutron stars with nearly equal mass. This is a property which observed neutron star binary pulsars seem to possess, giving some support to our suggestion. (The heavier W-R+W-R binaries should evolve into black hole, black hole binaries.

Even with our suggested strong selection criterion of helium star binary formation prior to binary pulsar formation, most of the binaries may be disrupted by the high-kick velocities in the explosion. Our estimates suggest, however, that our scenario gives, if anything, too high a probability of neutron star binary pulsar formation, which is a very rare event.

We would like to thank Lars Bildsten for help in the early stages of this work. Fred Rasio furnished a great deal of helpful advice and criticism over the period of this work. Marten van Kerkwijk was a great source of useful information. Roger Chevalier not only developed the accretion scenario on which this work is based, but gave extremely helpful advice. Tony Moffat provided us with useful information about HD 5980. Ed Van den Heuvel gave useful advice and Hans Bethe suggested useful modifications in this paper. Ralph Wijers added greatly to the believability of this paper. Dipankar Bhattacharya gave detailed criticism and useful suggestions, which greatly improved the manuscript. We would like to thank Joe Weingartner for checking the calculations and for useful suggestions on styling. Finally, the author would like to thank the director, Jim Langer, and the Institute for Theoretical Physics in Santa Barbara, for excellent working conditions under which this work was carried out. At the Institute for Theoretical Physics, this work was supported in part by the National Science Foundation under grant PHY89-04035.

### REFERENCES

- Abt, H. A., & Levy, S. G. 1978, ApJS, 36, 241
- Arzoumanian, Z., Taylor, J. H., & Wolszczan, A. 1994, in preparation
- Bethe, H. A. 1993, ApJ, 419, 197
- Bhattacharya, D., & Van den Heuvel, E. P. J. 1991, Phys. Rep., 203, 1
- Blondin, J. M. 1986, ApJ, 308, 755
- Bodenheimer, P., & Taam, R. E. 1984, ApJ, 280, 771

- Bondi, H. 1952, MNRAS, 112, 195
- Bondi, H., & Hoyle, F. 1944, MNRAS, 104, 273
- Brown, G. E., & Bethe, H. A. 1994, ApJ, 423, 659
- Brown, G. E., & Weisgarber, J. C. 1994, ApJ, 436, 843
- Brown, G. E., Weisgarber, J. C., & Wijers, R. A. M. J. 1995, in preparation
- Burrows, A., & Woosley, S. E. 1986, ApJ, 308, 680
- Cannat, R. C., Eggleton, P. P., Zytikow, A. N., & Podsiadlowski, P. 1992, ApJ, 386, 206
- Cherepashchuk, A. M., & Moffat, A. F. J. 1944, ApJ, 424, L53
- Chevalier, R. A. 1989, ApJ, 346, 847
- . 1993, ApJ, 411, L33
- Colgate, S. A. 1971, ApJ, 163, 221
- Cook, G. B., Shapiro, S. L., & Teukolsky, S. A. 1993, preprint
- Dewey, R. J., & Cordes, J. M. 1987, ApJ, 321, 780
- Dicus, D. A. 1972, Phys. Rev. D, 6, 941
- Dokuchaev, V. P. 1964, Soviet Astron., 8, 23
- Doom, C., & DeGreve, J. P. 1982, Ap&SS, 87, 357
- Enman, L. M., & Woosley, S. E. 1988, ApJ, 333, 754
- Flannery, B. P., & Van den Heuvel, E. P. J. 1975, A&A, 39, 61
- Fomalont, E. B., Goss, W. M., Lyne, A. G., Manchester, R. N., & Justtanont, K. 1992, MNRAS, 258, 497
- Habets, G. M. J. H. 1986, A&A, 167, 91
- Harrison, P. A., Lyne, A. G., & Anderson, B. 1993, MNRAS, 261, 113
- Holland, D. J., & Becker, R. H. 1984, Nature, 307, 215
- Houck, J. C., & Chevalier, R. A. 1991, ApJ, 376, 234
- Hoyle, F., & Lyttleton, R. A. 1939, Proc. Cambridge Phil. Soc., 35, 405
- Hunt, R. 1971, MNRAS, 154, 141
- . 1979, MNRAS, 188, 83
- Iben, I., & Tutukov, A. V. 1985, ApJS, 58, 661
- Illarionov, A. F., & Sunyaev, R. A. 1985, A&A, 39, 185
- Kitamoto, S., Mizobuchi, S., Yamashita, K., & Nakamura, H. 1992, ApJ, 384, 263
- Koenigberger, G., Moffat, A. F. J., St-Louis, N., Auer, L. H., Drissen, L., & Seggewiss, W. 1994, ApJ, 436, 301
- Langer, N. 1989, A&A, 220, 135
- Lattimer, J. M., & Yahil, A. 1989, ApJ, 340, 420
- Paczynski, B. 1971, Acta Astron., 21, 1
- Podsiadlowski, Ph., Joss, P. C., & Hsu, J. J. L. 1992, ApJ, 391, 246
- Pols, O. R., Cott, J., Waters, L. B. F. M., & Heise, J. 1991, A&A, 241, 419
- Raphael, Y., & Salpeter, E. E. 1980, ApJ, 240, 20
- Ruderman, M. A., & Spiegel, E. A. 1971, ApJ, 165, 1
- Shapiro, S. L., & Teukolsky, S. A. 1983, Black Holes, White Dwarfs and Neutron Stars (New York: Wiley)
- Shibasaki, N., Murakami, T., Shaham, J., & Nomoto, K. 1989, Nature, 342, 636
- Shigayama, T., Nomoto, K., Tsujimoto, T., & Hashimoto, M. 1990, ApJ, 361, L23
- Shima, E., Matsuda, T., Takeda, H., & Sawada, K. 1985, MNRAS, 217, 367
- Smaarr, L. L., & Blandford, R. 1976, ApJ, 207, 154
- Stokes, G. H., Taylor, J. H., & Dewey, R. J. 1985, ApJ, 294, L21
- Sutanto, W. 1975, A&A, 41, 7
- Taan, R. E. 1994, private communication
- Taan, R. E., & Van den Heuvel, E. P. J. 1986, ApJ, 305, 235
- Terman, J. L., Taam, R. E., & Hernquist, L. 1994, ApJ, 422, 729
- Thorne, K. S., & Zytikow, A. N. 1977, ApJ, 212, 832
- Thorsett, S. E., Arzoumanian, Z., McKinnon, M. M., & Taylor, J. H. 1993, ApJ, 405, L29
- Tutukov, A. V., & Yungelson, L. R. 1980, in Close Binary Stars: Observations and Interpretations, ed. J. M. Plavec, D. M. Popper, & R. K. Ulrich (Dordrecht: Reidel), 15
- Van den Heuvel, E. P. J. 1976, in IAU Symp. 73, Structure and Evolution of Close Binary Systems, ed. P. Eggleton, S. Mitton, & J. Whelan (Dordrecht: Kluwer), 35
- Van den Heuvel, E. P. J., & De Loore, C. 1973, A&A, 25, 387
- Van den Heuvel, E. P. J., & Taam, R. E. 1984, Nature, 309, 235
- van Kerkwijk, M. H. 1993, A&A, 276, L9
- . 1994, in preparation
- van Kerkwijk, M. H., Geballe, T. R., King, D. L., Van der Klis, M., & Van Paradijs, J. 1994, A&A, submitted
- Verbiest, F., Wijers, R. A. M. J., & Burns, H. M. G. 1990, A&A, 234, 195
- Vrancken, J., DeGreve, J. P., Yungelson, L., & Tutukov, A. 1991, A&A, 249, 411
- Weaver, T. A., & Woosley, S. E. 1994, in preparation
- Wijers, R. A. M. J., Van Paradijs, J., & Van den Heuvel, E. P. J. 1992, A&A, 261, 145
- Wolazzan, A. 1991, Nature, 350, 688
- Woosley, S. E. 1988, ApJ, 330, 218
- Yungelson, L., & Tutukov, A. 1991, Wolf-Rayet Stars and Relations with Other Massive Stars in Galaxies, ed. der Hucht & B. Hidayat (Dordrecht: Kluwer), 459
- Zeldovich, Ya. B., Ivanova, L. N., & Nadezhin, D. K. 1972, Soviet Astron., 16, 209

## Chapter 6

# How Collapsing Stars Might Hide Their Tracks in Black Holes

Reprinted with permission from *Science*, 261 (1993) 831-832  
Copyright (1993) by American Association for the Advancement of Science

This page is intentionally left blank

---

## How Collapsing Stars Might Hide Their Tracks in Black Holes

By Gary Taubes

Where have all the neutron stars gone? As many as 150 shells of gas and dust left by the recent collapse and explosion of massive stars litter our galaxy, and the standard wisdom predicts that each of them should hold a dense cinder, or neutron star. But fewer than two dozen of these supernova remnants have revealed neutron stars. The most notorious absence involves the great supernova of 1987. Under all standard stellar evolution scenarios, a pulsar — a radio-emitting neutron star — should have shown itself by now in the debris of this supernova, the brightest in 400 years. Yet none has been found.

Now Hans Bethe of Cornell University — the renowned physicist, still productive at 87 — and Gerald Brown of the State University of New York at Stony Brook have proposed a new way of accounting for the missing neutron stars: gone to black holes every one. Astrophysicists have long thought that only the very largest stars could collapse all the way down to black holes. But in a paper to be published in the *Astrophysical Journal*, Bethe and Brown argue that even run-of-the-mill supernovas, in stars having only 18 to 30 times the mass of the sun, could create black holes.

"If their theory is really right," says Shri Kulkarni, an astronomer at the California Institute of Technology who has read the paper, "it's pretty profound." It would imply that the galaxy harbors as many as a billion "small" black holes, measuring a few kilometers in diameter. The theory also has implications for the cinders left when stars of less than 18 solar masses collapse. It suggests that they leave behind not neutron stars but composites of neutrons and protons (nucleons) that should properly be called nucleon stars — though they would appear the same as neutron stars to Earthbound observers.

Traditionally, astrophysicists have envisioned two ways for a star weighing more than 8 solar masses to end its life. In both cases, the star's fate is sealed when nuclear burning in its core converts the last of its fuel into iron. When the iron core grows to between 1.5 and 2 solar masses, it can no longer sustain itself against gravity, and it collapses. If the core weighs less than 1.8 solar masses, as it will in

a star of less than 30 solar masses, the collapse ends abruptly when the core has formed a neutron star — in essence, a giant ball of neutrons. The jolt apparently generates a shock wave, which blows off the remainder of the star. The result is a Type II supernova, with a neutron star left over. (Type I supernovas occur in white dwarf stars and leave no remnant.)

In a progenitor star that has a mass greater than 30 times that of the sun, the core grows to 1.8 solar masses or more before collapsing, and the collapse doesn't stop with the formation of a neutron star. In a matter of milliseconds, the massive star vanishes into a black hole. As Brown puts it, "The whole mass of the star just goes poof, straight into a black hole." Thus, as far as most astrophysicists were concerned, a massive star could end its life either as a supernova and a neutron star, or as a massive black hole — without a supernova.

But the surplus of supernova remnants over neutron stars left some researchers wondering whether there might not also be a third way for a star to terminate its life: by lingering at the neutron star stage for long enough to trigger a supernova, then vanishing into a black hole. In 1986, for example, Stan Woosley of the University of California, Santa Cruz, and Jim Wilson and Tom Weaver of the Lawrence Livermore National Laboratory suggested two scenarios for such "hit-and-run" supernovas. One applied to large stars (more than 30 solar masses): The stellar core, though massive enough to form a black hole, would temporarily stop collapsing when it had formed a neutron star because the thermal energy of the collapsed material would generate too much pressure for the collapse to continue. After the neutron star had cooled, it would collapse to a black hole. In the other scenario, the collapse would yield a neutron star at first, but star material blown out by the explosion would later fall back onto the neutron star, pushing it over the threshold for further collapse.

Having suggested these notions, however, Woosley, Wilson, and Weaver never pushed them with any vigor. When Brown and Bethe learned of them this spring after developing their own scenario, they agreed with the end point, supernovas and black holes from the same star, but offered a different mechanism. It is built on a new equation of state for the matter inside a collapsed star — a mathematical description of how it responds to pressure — that implies "softer" material than theorists had pictured.

Because the nuclear matter provides less outward pressure for a given density, the new equation of state allows a massive star smaller than 30 solar masses, with a core as small as 1.5 solar masses, to collapse all the way into a small black hole. The collapse doesn't go off without a hitch, though: It pauses at the collapsed star stage for as long as 20 seconds, long enough to launch a supernova on its way. That would allow astrophysicists to have their supernovas without worrying about the lack of neutron stars. "Without our equation of state," Bethe said, "we would have said that whenever you get a [Type II] supernova the massive object left behind should be a neutron star. This we now no longer believe."

---

The new equation of state reflects Brown and Bethe's picture of how subatomic particles behave during the collapse at the end of a massive star's life. Astrophysicists have assumed that when a stellar core collapses, all the electrons in the core are captured by protons, which emit neutrinos and become neutrons — hence the term neutron star. But last year, Brown and several colleagues suggested that the electrons in a collapsing core might meet a different fate. As the core reaches a density three times that of nuclear matter, they proposed, it becomes energetically favorable for electrons to turn into negatively charged  $\bar{k}$  mesons, or kaons (particles consisting of a strange quark paired with a down antiquark) and emit neutrinos in the process.

Because the protons don't absorb the kaons as they would electrons, the protons remain protons rather than turning into neutrons. The result is a mixture of protons and neutrons steeped in what's called a kaon condensate, a state of matter described first in 1986 by David Kaplan and Ann Nelson, then junior fellows at Harvard, and in more detail in 1991 by David Politzer and Mark Wise of Caltech. And because the attraction between protons and neutrons is stronger than between neutrons and neutrons, the collapsed material is more compressible than theorists had previously assumed — and easier to squeeze down into a black hole.

The collapse into a black hole doesn't happen immediately, though, because the neutrinos trapped in this nuclear matter heat it, temporarily stabilizing it from further collapse. "Only when the neutrinos all leave," says Brown, "does the star cool down" — and, if the core weighs more than 1.5 solar masses, collapse further into a black hole. If the core weighs less than 1.5 solar masses, it will remain what Bethe and Brown call a nucleon star, made of both protons and neutrons, rather than a neutron star.

How Brown and Bethe's revisionist view of supernovas will be accepted remains to be seen. Woosley thinks "we could still live in a universe with a maximum mass of a neutron star at 1.8 or even 2 solar masses, and not be in contradiction with anything we know." Adds Weaver, "There's a fair amount of uncertainty in the physics of the core collapse process. [The] idea is a good one, but it's premature to say it solves the problem."

Still, there may be a way to test the theory, says Kulkarni of Caltech. The places to look, he says, are globular clusters, the tight knots of old stars that form a halo around the galaxy and dot its disk. Over their long life-spans, globular clusters have experienced a lot of supernovas, and if Brown and Bethe are right, they should be rich in small black holes. In the crowded environment of a globular cluster, the black holes would tend to pair up with ordinary stars, forming binary systems that might be revealed by rare, powerful x-ray outbursts, caused when material from the companion star falls into the black hole.

Watching for these outbursts isn't a perfect test, though. One catch is that their violent birth process might have left many of the black holes traveling fast enough

to fly out of the globular clusters before they paired up. Another is that the x-rays from a black hole binary may look a lot like those from a neutron star binary. Concludes Kulkarni, "It will take a while to say anything."

Bethe isn't worried. For now, he says, the scenario will stand or fall not on observational tests but on the plausibility of its picture of collapsing nuclear matter. When asked whether other astrophysicists will come around, he said, "Yes. Maybe not all, but many."

## Chapter 7

# Mystery of the Missing Star

Reprinted with permission from *Discover*, December 1996, 111–115  
and Adam Frank  
Copyright (1996) by Adam Frank and Discover Magazine

This page is intentionally left blank

## Mystery of the Missing Star

Once, astronomers thought the spectacular supernova of 1987 would confirm their most cherished theories of star death. Instead it left a hole they've tried to plug for nine years.

By Adam Frank

On February 23, 1987, an unprepossessing star called Sanduleak -69°202, which had sat without incident in the southern skies for eons, suddenly blew itself to bits in the most spectacular supernova to be seen in nearly three centuries. SN1987A, as Sanduleak was now designated, was the closest supernova to our solar system ever to be observed with modern telescopes, and it offered astronomers an unparalleled view of the death of a massive star. It also gave them dramatic, unequivocal confirmation of what by then had become the "classic" theory of supernova explosions. There was only one problem. Something fundamental was missing from SN1987A. Nine years later, it's still missing.

According to the classic theory, there should be a neutron star where Sanduleak used to be. Neutron stars are "dead" stars, stellar cinders made of neutrons squeezed through the bars of their atomic cages and thus able to achieve extraordinary densities — a neutron star just 10 miles across contains as much matter as our sun. Physicists don't know very much about how these odd beasts behave, but they know that the stars usually appear in the sky as pulsars, rotating stars that project beams of intense radio waves into space, like cosmic lighthouses. In the case of SN1987A, however, no pulsars have been detected.

The only alternative suggested by classic theory is that Sanduleak, rather than forming a neutron star, collapsed into a black hole, but astronomers don't put much stock in this option. For one thing, conventional wisdom says that Sanduleak was too small to turn into a black hole. And besides, black holes tend to swallow everything

in their vicinity, supernovas included. If there was a black hole where Sanduleak used to be, we would never have seen the supernova that produced it in the first place. Yet we did, and its hollow remnant haunts us still.

When conventional wisdom falls so far short of explaining what astronomers see in the sky, it often means that it's time for some radical new idea. Such ideas usually come from the younger, hungrier researchers in astrophysics, the ones who are struggling for job, tenure, and reputation and who, by virtue of the youthful combination of inexperience and exuberance, are apt to see old problems in a new light. SN1987A is no exception, at least not in its demand for novel explanations. What is surprising, however, is that those explanations are coming from two of astronomy's most respected — and most senior — scientists.

Between them, Hans Bethe and Gerry Brown have 120 years of experience in physics research, not to mention a Nobel Prize — Bethe won it in 1967 for a theory articulating the physics of nuclear reactions in stars. At 90, he is now a professor of physics at Cornell in Ithaca, New York. "We had a conference a few years back," says Brown — a mere tot at 70 and a professor of physics at the State University of New York at Stony Brook — "and someone pointed out that I published my first paper before many of the speakers were born, and Hans had published his first paper before *I* was born." Drawing on their vast experience, and with the help of some fresh insight, Brown and Bethe have offered an ingenious solution to the puzzle of SN1987A. If they are correct, they will have rewritten completely the physics of neutron stars.

Brown and Bethe believe that the failure to find a neutron star in SN1987A is not atypical at all. They cite a study in which about half of all known supernova remnants were shown to lack conclusive evidence of neutron stars. "If all supernovas produce neutron stars," says Brown, "then why do so many supernova remnants [the giant smoke rings left over from the blast] lack evidence for neutron stars at their centers?" The answer, he and Bethe believe, is that there is indeed a black hole at the center of SN1987A — a small one, formed in a fundamentally different way than classical theory suggests.

To all outward appearances, Sanduleak's behavior was precisely what you'd expect of any ordinary, medium-size dying star — that is, one with only about 20 times the mass of our sun. Having run out of fuel for the nuclear fusion reactions in its core, it succumbed to its own gravity and collapsed. In half a second the core, which consisted of silicon, carbon, oxygen, and other heavy elements, was squeezed so tightly that the individual nuclei, the clusters of protons and neutrons at the heart of each atom, merged. Protons and electrons began to turn into neutrons, and the core became one big nucleus — a proto-neutron star.

The core continued to collapse, compressing itself to a state as much as ten times denser than an atomic nucleus, or 10,000 trillion times more compact than water. ("I don't know how he figured this out," says Brown, "but Hans once calculated

---

that a teaspoon of neutron star material weighs as much as all the buildings in Manhattan.”) Once the core crossed this threshold of “nuclear density,” the strong nuclear force — the same force that binds protons and neutrons to the atomic nucleus — began to repel with a vengeance, bringing the collapse to an abrupt halt. “Once you pass nuclear densities,” says Brown, “the core acts like a dense rubber ball. The harder you squeeze it, the harder it pushes back.” The hydrogen and helium gases that made up the bulk of the star’s outer layers weren’t massive enough to overcome the repulsive strong force. Instead of falling into the core, they bounced off and went hurtling into space at great speed. This was the supernova explosion.

After that there is some disagreement as to what in fact happened to the star formerly known as Sanduleak. The core should have been left behind as a neutron star, but it is, of course, nowhere to be seen. Instead of turning into a cosmic rubber ball, the core could have continued to collapse, becoming a black hole, except that classic theory requires a tremendously heavy torrent of infalling matter from the outer layers to overcome the strong nuclear force — far more than puny Sanduleak could have mustered.

Brown and Bethe offer a third option: they think that the “rubber” of the cosmic hard rubber ball is much “softer” than most astrophysicists think. The term that astrophysicists use to refer to the hardness or softness of the core is its “equation of state.” “A softer equation of state means that something doesn’t push back as hard when you squeeze it,” says Stan Woosley, a supernova expert at the University of California at Santa Cruz. In other words, a softer equation of state would mean that when collapsing, even a moderately massive star, such as Sanduleak, can overcome the strong nuclear force and turn itself into a black hole.

Brown and Bethe, to be sure, are not observers but theorists, and so it is somewhat misleading to couch their insights in terms of observational data. Their ideas derive from an intuition about the way nature behaves on its most fundamental level, the kind of “feeling,” or hunch — almost a personal aesthetic — that is every bit as important for the good theorist as the ability to solve equations. To turn this hunch into a theory, however, Brown and Bethe needed to come up with a physical mechanism that would soften the equation of state. And they needed to explain another annoying observation that belied their assertion that a black hole formed at SN1987A: during the supernova explosion, scientists detected a burst of neutrinos lasting about ten seconds emanating from Sanduleak. Neutrinos are ghostly subatomic particles that weigh almost nothing. They are produced in prodigious numbers when protons and electrons merge into neutrons during the core’s collapse. Since neutrinos have little if any mass, they can easily escape the gravitational clutches of a neutron star, but — and here’s the rub — not a black hole’s. “A black hole would swallow the neutrinos,” says Woosley. “So if one was going to form in SN1987A, it would have to wait at least ten seconds to let the neutrinos we saw escape.” To make their idea work, Brown and Bethe needed to find some trick

hidden deep in the unknown structure of matter that could soften up a neutron star — but one that would wait precisely ten seconds before doing so.

Every January, Brown and Bethe meet in Pasadena, at Caltech. There the two friends spend a month working together on problems involving supernovas and nuclear physics. "The way we usually work," says Brown, "I cook dinner for Hans and then give him the problems I want him to solve. Then he puts off thinking about them until he can take a bath."

"I do my best thinking in the bath," confirms Bethe. "We get together the next morning and then we discuss the problem."

In this manner the two physicists began in 1993 to work on the problem of soft neutron stars. While they were walking in the hills above the campus one afternoon, Brown mentioned a remarkable idea, which later became a cornerstone of their work on SN1987A. "There was this theory of kaon condensates two kids from Harvard had come up with," recalls Brown. "I told Hans about it, and he quickly pictured how it could work in a collapsing star."

These two "kids" were David Kaplan and Ann Nelson, a husband-and-wife team of physicists now at the University of Washington, and their theory described how a particular chain of subatomic transformations might occur deep within a neutron star. The idea rested on the quantum mechanical notion that particles do not have stable identities and that they can, with enough energy, spontaneously transform themselves from one type of particle to another. "What we suggested," says Kaplan, "was that when ordinary matter is squeezed to the densities you find in a neutron star, electrons in the star can be transformed into particles called kaons."

Kaons are exotic, very heavy particles with many unusual qualities, and one that is particularly relevant to the soft equation of state that Brown and Bethe were seeking: two or more kaons can occupy the same energy level at the same time. This is to say, in physicist's jargon, that kaons do not obey a law of quantum mechanics known as the exclusion principle. Electrons, on the other hand, do: the more electrons you cram together, the higher their energy levels go. Thus in a collapsing stellar core, the more electrons you had occupying the same space, the higher the energy and the greater the resistance that would be generated against infalling gas — making the core "harder." Kaons, by contrast, could coexist peacefully in the core, all occupying the lowest energy level without exerting any resistance at all against infalling matter. They would, in fact, form a Bose-Einstein condensate, a form of matter long predicted but created in a laboratory for the first time just last year. And the appearance of a kaon condensate would have an immediate and dramatic effect. "Kaon condensates don't bounce around," says Kaplan. "The weird thing about them is they have essentially zero temperature. They are ice cold and don't help support the neutron star at all." The core, in a word, would be soft.

It's a lot to ask of an electron, a very light particle, to make the leap to the very heavy kaon. To physicists, the transformation Kaplan and Nelson proposed was as

---

startling as having your dog suddenly transform itself into your mailman. Turning something so light into something so heavy would require tremendous energy. How would it take place? Kaplan and Nelson found the answer in Einstein's famous legacy — that mass and energy are interchangeable.

If you think of mass and energy as money in a bank account, the electron is very poor, and the heavy kaon is very dear indeed. But that's the valuation only here, on Earth. In the core of a neutron star, the situation is much different. There the potential exists for very light (cheap) kaons to pop into being. The difference arises because in dense nuclear matter a kaon would exert an attractive force on the nearby neutrons, and attraction is a form of energy. Thus the energy needed to create the kaon in the first place is lowered, and the event becomes a possibility. And in quantum mechanics, possibility equals actuality — if there's potential for an electron to turn into a light kaon, it will happen. "Deep inside the nuclear matter," says Kaplan, "the kaons weigh less, so the electrons can more easily make the leap."

Although kaons made on Earth in particle accelerators exist for only 10 billionths of a second, kaons in a neutron star would be held stable by their attraction to the neutrons. "The kaon is such an evanescent and unstable particle on Earth," says Kaplan. "For this rare, exotic, and barely understood form of matter to play such a major role in a neutron star was quite a surprise. For a physicist it's really delightful."

Brown and Bethe quickly recognized the kaon condensate theory to be the mechanism that just might soften the neutron star's equation of state. Along with other collaborators, the two physicists hammered out the details of how the condensates might trigger the collapse of the neutron star into a black hole. The theory was both powerful and elegant. But to explain SN1987A there was still the problem of timing. What mechanism could keep the kaon condensate from forming long enough to let the ten-second burst of neutrinos escape?

As it turns out, the neutrinos themselves provide the answer through the exclusion principle. Neutrinos are an inevitable by-product of the compressing core of a dying star. As the core is squeezed, protons and electrons undergo the quantum mechanical quick change and merge into neutrons, each time creating a neutrino as well. Electrons that turn into kaons also release neutrinos, but by the time the core is dense enough to allow kaons to form, the neutron star is already facing a neutrino overpopulation crisis. Since neutrinos obey the exclusion principle, there's a limit to how many you can cram together into the core. Once that limit is reached, the production of new neutrinos comes to a halt. Since you can't make a kaon without releasing a neutrino, the production of kaons comes to a stop as well.

"Because of the exclusion principle," says Brown, "you can only make new kaons at the rate that old neutrinos leave. This means you have to open up holes for the new neutrinos or new kaons won't be allowed to form."

"Once the old neutrinos leave en masse," says Kaplan, "the electrons will turn into kaons, the kaons will turn into a condensate, and within a millisecond the young neutron star goes screaming into oblivion." What's needed, though, is some way to get rid of the neutrinos.

In an ordinary star, neutrinos fly right through matter, but in the superdense core they must bull their way out, wiggling slowly from the center toward the edge. "Adam Burrows [a supernova expert at the University of Arizona in Tucson] did a fine paper some years back that calculated the time it takes for neutrinos to diffuse out of the young neutron star," says Brown. "We used his result in our calculations." The time worked out to be, surprisingly enough, about ten seconds. With the ten-second timing delay out of the way, Brown and Bethe could now explain why there should be a black hole at the center of SN1987A. They also estimated that there should be 50 times more black holes in the galaxy than previously estimated.

Despite the neat fit, not everybody is convinced. "The absence of evidence for a neutron star in SN1987A is certainly not evidence of absence," says Stan Woosley. "My prejudice is that there is an as-yet-undetected neutron star in SN1987A."

And Brown, for one, concedes that the theory he and Bethe have conceived is still very speculative. "Nobody understands what matter is like at nuclear densities, and here we are, cheerfully talking about densities five or ten times greater. Nobody believes us at those densities, and we really don't believe ourselves. But in our scenario, the formation of a kaon condensate really has a qualitative effect. It's a new idea, so we are allowed some poetic license."

Brown, of course, is being modest. His and Bethe's theory about the physics of neutron stars may have flaws, but it is one of the most complete and plausible theories around. For all the uncertainty, their work may be only the first fruit of a larger symbiosis between nuclear physics and astrophysics that is just now getting under way. For instance, new machines called heavy ion colliders can now smash whole nuclei rather than just individual particles, and they are giving physicists their first experimental glimpse of matter at nuclear densities. "We are entering a new era," says Brown. "Soon we may get to learn directly about nuclear matter at three, four times nuclear density." This information will in turn help astrophysicists understand the extreme states of stellar evolution. All this activity just adds fuel to Brown and Bethe's enthusiasm.

"Dr. Brown got me into supernovas 20 years ago," says Bethe. "Now here I am, 90 years old, and I am still working on them. I am just fascinated." Brown seconds the sentiment. "I am more excited than I ever have been," he says. "After all these years, I still just want to understand the universe." It looks like the retirement parties will have to wait.

## Chapter 8

# Observational Constraints on the Maximum Neutron Star Mass

H.A. Bethe and G.E. Brown

Reprinted with permission from *Astrophysical Journal*, 445 (1995) L129–L132  
Copyright (1995) by American Astronomical Society

### Commentary

In this *ApJ* letter we tried to show that knowing where the mass cut was in SN 1987A, from the  $\sim 0.075M_{\odot}$  of Ni that was ejected, we could deduce the mass of the compact object to be  $< 1.56M_{\odot}$ . Of course we were trying to show that our maximum neutron star mass  $M_{\max} \simeq 1.5M_{\odot}$  of Paper 4 was supported by observation, since we believe that SN 1987A went into a low-mass black hole.

Whereas some nice simple connections between important quantities are made, especially in the use of the virial theorem, and we show that the separation distance inside of which matter falls back onto the neutron star is  $\sim 3900$  km, we no longer believe that our upper limit is as quantitatively well determined in this paper as we wrote at the time.

Note that an up-to-date figure of observed neutron star masses is given in our Commentary to Paper 4, and should replace Fig. 1 of this paper.

One of the more important results not brought out in this paper is that the fallback of material following the shock (see also Table 3 of Paper 9) is roughly equal in magnitude to the binding energy of the compact object, given by Eq. (15) of this paper. Thus, the Fe core mass is a good estimate of the mass of the compact object.

This page is intentionally left blank

## OBSERVATIONAL CONSTRAINTS ON THE MAXIMUM NEUTRON STAR MASS

H. A. BETHE<sup>1</sup> AND G. E. BROWN<sup>2</sup>

California Institute of Technology, W. K. Kellogg Radiation Laboratory, 106-38, Pasadena, CA 91125

Received 1994 October 3; accepted 1995 March 17

### ABSTRACT

We review estimates of the mass of the compact core in SN 1987A and conclude that the most accurate determination can be obtained from the known value of  $\sim 0.075 M_{\odot}$  of Ni production in the explosion. With binding energy correction, this gives an upper limit of gravitational mass of  $\sim 1.56 M_{\odot}$ , slightly larger than Brown & Bethe's previous estimate of  $\sim 1.5 M_{\odot}$ . Observation by OSSE of the ratio of  $\gamma$ -rays from  $^{57}\text{Co}$  and  $^{56}\text{Co}$  indicates that neutron-rich material from the inner regions does not reach the mass cut by convection or Rayleigh-Taylor instability.

Arguments that the core of SN 1987A went into a black hole are reviewed. If one accepts this to be true, then the maximum compact core mass gives an upper limit on neutron star masses of

$$(M_{\text{NS}})_{\text{max}} \cong 1.56 M_{\odot}$$

(gravitational), in rough agreement with the previous result of Brown & Bethe.

*Subject headings:* stars: neutron — supernova remnants

### 1. INTRODUCTION

The delayed shock mechanism, in which energy is brought from outside the neutrinosphere up to repower the shock wave, has now been shown to provide a viable mechanism for supernova explosions (Herant et al. 1995; see also Janka & Müller 1993, 1994). Results sufficiently quantitative to describe the details of the supernova explosion SN 1987A, however, have not yet been obtained. Since we are particularly interested in the gravitational mass of the compact object in SN 1987A, which we believe to have gone into a black hole, we have searched for observational constraints on this mass. With the assumption that the core did go into a black hole, these constraints then indicate a limiting maximum mass for neutron stars.

We will outline in this Letter how the known value of  $\sim 0.075 M_{\odot}$  of  $^{56}\text{Ni}$ , which decayed radioactively to Fe, gives us quite accurately the range of gravitational masses

$$1.44 M_{\odot} < (M_{\text{p}}) < 1.56 M_{\odot}, \quad (1)$$

for the core in SN 1987A. The necessary arguments involve the explosion energy, which can be obtained from the supernova light curve but is independent of the precise explosion mechanism and depends rather insensitively on the presupernova model. If our arguments that the compact object in SN 1987A went into a black hole are accepted, then the range (1) gives a value for maximum neutron star mass. The lower limit of  $1.44 M_{\odot}$ , the mass of the Hulse-Taylor pulsar, is firm. Our discussion will be to establish the upper limit of  $1.56 M_{\odot}$ .

### 2. DEVELOPMENT

We consider the outside edge of the  $^{56}\text{Ni}$  formation, assuming spherical symmetry. The minimum temperature for synthesis of  $^{56}\text{Ni}$  is

$$4 \times 10^9 \text{ K} = 0.35 \text{ MeV}. \quad (2)$$

<sup>1</sup> Permanent address: Laboratory of Nuclear Studies, Cornell University, Ithaca, NY 14853.

<sup>2</sup> Permanent address: Physics Department, State University of New York, Stony Brook, NY 11794.

The radiation pressure, including electron-positron pairs, at this temperature is

$$P_s = 1.87 \times 10^{24} \text{ ergs cm}^{-3}. \quad (3)$$

The material pressure is small in comparison; the electron pressure is already included in  $P_s$ . The nuclear pressure is very small because the mass per nucleon is large (56) and the density is low ( $\sim 3 \times 10^6 \text{ g cm}^{-3}$ ).

We wish to determine the volume in the star which reaches the pressure given in equation (3). This can be obtained from the shock energy. The most powerful way to do this is by means of the virial theorem (Cooperstein, Bethe, & Brown 1984), which is

$$E' - K - D = 4\pi P_s R_s^3 - 4\pi P_g R_g^3. \quad (4)$$

Here  $E'$  is the total energy in the shock at the time the Ni is formed,  $K$  is the kinetic energy, and  $D$  is the nuclear dissociation energy. The shock is assumed to be spherical; the subscript  $s$  denotes quantities at the shock front, and the subscript  $g$  denotes quantities at the base of the shock. This case is at the gain radius  $R_g$ , the point where energy gain from neutrino capture equals loss by capture of ambient electrons (+ or -) by nucleons,

$$e^- + p \rightarrow n + \nu, \quad (5)$$

with the neutrino escaping.  $R_g$  is estimated to be 120–140 km (Bethe 1993, 1995), while  $R_s$  will turn out to be about 4000 km. Therefore, it is likely that

$$P_g R_g^3 \ll P_s R_s^3, \quad (6)$$

and  $P_g R_g^3$  will be neglected.

$D$  is the energy absorbed when ambient nuclei (say O or Si) dissociate into  $\alpha$ -particles and nucleons. Dissociation into  $\alpha$ -particles is actually unimportant (Bethe 1995, § 12). When the shock is as far out as 4000 km, only a very small part of its volume and mass is dissociated into nucleons; hence, this part of  $D$  is also negligible.

The kinetic energy can be estimated from the shock pressure (eq. [3]) and the density of matter outside the shock (Bethe 1995, eq. [4.1]). Combination of these two quantities gives the

material velocity directly behind the shock,  $u_s$ , and it is then assumed that the material velocity at smaller  $r$  is proportional to  $r$ . On this basis, we calculate

$$K = 0.4 \text{ foe} \quad (1 \text{ foe} = 10^{51} \text{ ergs}). \quad (7)$$

The energy  $E'$  includes the gravitational potential (this, in fact, is the strength of the virial theorem). The shock energy  $E_s$  is almost the observed energy  $E_f$ , except that the shock loses some energy  $B$  on the way out by removing the bound envelope of the star; this has been estimated in Bethe (1995, § 10) to be

$$B = 0.4 \text{ foe}, \quad (8)$$

$$E' = E_f + B. \quad (9)$$

This  $B$  just cancels  $K$  in equation (4); therefore, neglecting all small terms,

$$E_f = 4\pi P_s R_s^3. \quad (10)$$

The observed energy of SN 1987A (Bethe & Pizzochero 1990) is

$$E = 1.4 \pm 0.4 \text{ foe}. \quad (11)$$

The central value in equation (11) corresponds to  $M_{\text{env}} \approx 10 M_\odot$ , consistent with Woosley's most successful models (Woosley 1988), and the upper limit, 1.8 foe, corresponds to  $M_{\text{env}} \approx 14 M_\odot$ , about the maximum acceptable envelope mass for SN 1987A. The pressure at the shock is almost exactly the radiation pressure; the pressure in the nuclear component is small because of low density and because of the high atomic mass of Ni (56). Thus, using the pressure (eq. [3]) with equation (11), we find

$$R_s = 3900 \pm 400 \text{ km}. \quad (12)$$

We now use the computation by Wilson & Mayle (1990) to find the mass which is inside  $R_s$  at the time when the shock reaches this distance, about 0.5 s after bounce according to present SN theory. This turns out to be

$$M_b = 1.815 \pm 0.024 M_\odot. \quad (13)$$

This is the mass enclosed by the outer edge of the  $^{56}\text{Ni}$  region. We have to subtract the mass of Ni in the ejecta, known to be  $0.075 M_\odot$ . With Ni go other elements in the neighborhood, such as  $^{57}\text{Co}$  and  $^{54}\text{Fe}$ . Regarding Co, the measurements by OSSE, which we discuss below, show that the ratio  $^{57}\text{Co}/^{56}\text{Co} = 1.5$  times the solar ratio of their radioactive daughters,  $^{57}\text{Fe}/^{56}\text{Fe}$ . The latter ratio is 0.024; thus,  $^{57}\text{Co}$  is 3.6% of  $^{56}\text{Ni}$ . For  $^{51}\text{Fe}/^{56}\text{Fe}$ , the solar ratio is 6.4%. Therefore it is reasonable to estimate the amount of  $^{57}\text{Co} + ^{54}\text{Fe}$  in the ejecta to be  $\sim 10\%$  of  $^{56}\text{Ni}$ . Subtracting  $1.1 \times 0.075 M_\odot$  from equation (13), we find for the baryon number mass of the massive object left behind by SN 1987A,

$$M_b = 1.733 \pm 0.024 M_\odot. \quad (14)$$

The subscript  $b$  indicates that this is baryonic mass.

It should be noted that the data from the computation of Wilson & Mayle are essentially pre-SN data. At the time that we use these data, 0.5 s past bounce, their shock has only reached about 1000 km; thus, the place we use these data is unaffected by their shock; it is simply a hydrodynamical extrapolation of the pre-SN conditions. The time of 0.5 s, on the other hand, is taken from the modern theory of SN due to Herant et al. (1995), but our results are not sensitive to this

time, as can be seen from the relatively small error in the determination given in equation (13), even though the radius of the mass cut is uncertain to  $\geq 10\%$  (eq. [12]). The reason is the low density of material in the region of the mass cut.

In order to convert the baryonic into gravitational mass, we use the formula for binding energy due to Lattimer & Yahil (1989),

$$E = 0.084 M_\odot (M_b/M_\odot)^2, \quad (15)$$

where the gravitational mass  $M_b$  is used. From the range in equation (14), we obtain the range of gravitational masses

$$M_g = 1.535 \pm 0.02 M_\odot. \quad (16)$$

However, from the mass of the Hulse-Taylor pulsar, we know that the lower limit must be at least  $1.44 M_\odot$ .

It can be checked, from the calculations of Keil & Janka (1995), that for cores of gravitational masses in the range of  $\sim 1.5 M_\odot$ , the Lattimer-Yahil binding energy correction is accurate to  $\sim 10\%$ , giving a final gravitational mass to within  $\sim 1\%$ . Unpublished Stony Brook calculations for  $\sim 1.5 M_\odot$  cores with the kaon-condensed equation of state (EOS) show a similar accuracy.

Our result for the compact core mass of SN 1987A is, consequently,

$$1.44 M_\odot < M_g < 1.56 M_\odot. \quad (17)$$

No signals have been seen from the compact core of SN 1987A. Consequently, Brown & Bethe (1994) argued that this compact core collapsed into a low-mass black hole. Assuming this to be so, the compact core mass would have to be above the maximum mass of a neutron star (more accurately called, according to Brown & Bethe, a nucleon star). The upper limit in equation (17) is, then, our estimate of the maximum neutron star mass.

Thielemann, Nomoto, & Hashimoto (1994) found a lower mass cut than we did,  $M_b = 1.7 M_\odot$ , compared to our  $M_b = 1.815 M_\odot$ . The difference is mainly due to their use of Nomoto's pre-SN structure, which generally assigns lower masses to any feature in the star than does Woosley's, which we use. Similarly, Kumagai et al. (1993) and Thielemann et al. (1994), also using Nomoto's pre-SN numbers and an actual SN computation, find that only  $M_b = 1.63-1.67 M_\odot$  satisfies the OSSE ratio for  $^{57}\text{Co}/^{56}\text{Co}$ . We wish to emphasize that our procedure of using Woosley's calculation gives the most conservative upper limit for the mass of the compact object in SN 1987A.

The assumption of Brown & Bethe (1994) that the compact core of SN 1987A went into a black hole has been strengthened by detailed calculations of Brown & Weingartner (1994), based on the Chevalier (1989) work, which show that hypercritical accretion could hide a possible neutron star for only  $\sim 1$  yr, but after this time SN 1987A should be observed with a luminosity  $\sim L_{\text{Edd}} = 4 \times 10^{38} \text{ ergs s}^{-1}$ , 2 orders of magnitude above the present light curve of  $L \sim 4 \times 10^{36} \text{ ergs s}^{-1}$ . On the other hand, spherical accretion onto a low-mass black hole would contribute only at the level of  $10^{34}-10^{35} \text{ ergs s}^{-1}$  (Brown & Weingartner 1994) and, therefore, would not be seen. It is clear that our claimed maximum mass of  $1.56 M_\odot$  depends on the core having gone into a black hole.

### 3. GAMMA-RAY OBSERVATIONS

Information on the material in the Ni region is provided by cobalt  $\gamma$ -rays from SN 1987A. These have been measured in the

OSSE experiment on the *Gamma Ray Observatory* (GRO) (Kurfess et al. 1992; Clayton et al. 1992). It is found, at days 1617–27 and 1767–81 after SN 1987A, that

$$\langle {}^{57}\text{Co}/{}^{56}\text{Co} \rangle = 1.5 \pm 0.3 \pm 0.2, \quad (18)$$

relative to the Sun's ratio of the stable isotopes  ${}^{57}\text{Fe}/{}^{56}\text{Fe}$ , which are the results of the cobalt decays.

The result of equation (18) enables us to draw conclusions regarding the original chemistry of the material which became Ni and Co in the explosion. The critical quantity is

$$\eta = \frac{N - Z}{A} = 1 - 2Y_e. \quad (19)$$

With the measured ratio of  ${}^{57}\text{Co}/{}^{56}\text{Co}$ , the value of  $\eta$  in the Ni region is very small. This is compatible with the assumption that the material was originally Si (as in Nomoto's pre-SN model); "normal" Si on Earth has 4.7%  ${}^{29}\text{Si}$  and 3.1%  ${}^{30}\text{Si}$ , hence, an average  $\eta = 3.7 \times 10^{-4}$ . In Woosley's pre-SN model, the region which later becomes Ni is half-oxygen and half-silicon, giving about half that value of  $\eta$ .

However, the material in the original prompt shock was strongly neutronized, having  $\eta$  as high as 0.4 or more. The OSSE results indicate that very little of this material can have made its way, by convection, into the Ni region. This gives a severe constraint on convection: We need it to get the shock started quickly, but it must not be strong enough to get appreciable prompt-shock material into the Ni region.

This conclusion may be somewhat mitigated by the fact that the OSSE data stand in conflict with photometric arguments. Recent photometric data of Suntzeff et al. (1992) and Dwek et al. (1992) have led them to conclude that the  $({}^{57}\text{Co}/{}^{56}\text{Co})$  ratio is larger than 1.5, i.e., closer to 5. Kurfess et al. (1992) and Clayton et al. (1992) discuss possibilities to resolve the conflict, one of them being that the  ${}^{57}\text{Co}$  is more deeply bound in the core than in the theoretical models used. These more opaque models require larger production ratios, up to  $2.6 \pm 0.6$ , to account for the luminosity and the  $\gamma$ -rays.

#### 4. DISCUSSION

In the past, it has generally been thought that the reason accurately measured neutron star masses lie at  $1.44 M_\odot$  and below is evolutionary in nature. Large stars collapse when the iron core exceeds the Chandrasekhar limit

$$M_{\text{CH}} = 5.76 Y_e^2, \quad (20)$$

where  $Y_e$  is the electron fraction, which is  $Y_e \approx 0.43$  at collapse. Thermal pressures increase the maximum stable mass by 10%–15%.

Accretion has generally been assumed to proceed only up to the Eddington limit

$$\dot{M}_{\text{Edd}} = 1.5 \times 10^{-8} M_\odot \text{ yr}^{-1}, \quad (21)$$

so that even in the very old millisecond pulsars only  $\sim 0.1$ – $0.2 M_\odot$  is estimated to have been accreted. However, Chevalier (1993) and Brown (1995) have shown that in the common envelope phase of binary pulsar evolution, accretion can proceed at hypercritical rates

$$\dot{M} \geq 10^4 \dot{M}_{\text{Edd}}. \quad (22)$$

Therefore, if evolutionary history determines their mass, neutron stars of much higher mass than  $1.5 M_\odot$  (gravitational)

should exist, but none have been found. Therefore, there should be an intrinsic limit on the mass of a neutron star, as in equation (16).

It is surprising that this limit is so close to the Chandrasekhar limit (eq. [20]) for  $Y_e = 0.5$ . This must be accidental because the Chandrasekhar limit is due to the Fermi energy of electrons, while the neutron star limit is due to the interaction between nucleons.

At the time of the Brown & Bethe (1994) suggestion that  $(M_{\text{NS}})_{\text{max}} \cong 1.5 M_\odot$ , the compact object in the high-mass binary Vela X-1 lay outside the limit given by equation (1). New observations by Van Kerkwijk et al. (1995) find that observed velocities in Vela X-1 deviate substantially from the smooth radius-velocity curve expected for pure Keplerian motion. The deviations seem to be correlated with each other within one night, but not from one night to another. The excursions suggest something like pulsational coupling to the radial motion and make it difficult to obtain an accurate mass measurement. The lower limit for the mass of the compact object in Vela X-1 is now found to be  $1.43 M_\odot$  at a 95% confidence level, or  $1.37 M_\odot$  at this confidence interval around the most probable value. Consequently, Vela X-1 no longer violates the value of equation (16).

It is striking that well-measured neutron star masses lie below  $1.5 M_\odot$  (see Fig. 1). However, the central value in the compact object in 4U 1700–37 lies at  $1.8 M_\odot$ , although the error bars encompass  $1.5 M_\odot$ . Brown, Weingartner, & Wijers (1995) give arguments that this compact object could well be a low-mass black hole.

#### 5. CONCLUSIONS

From the Ni production, we can set an upper limit of  $\sim 1.56 M_\odot$  gravitational mass for the compact object in

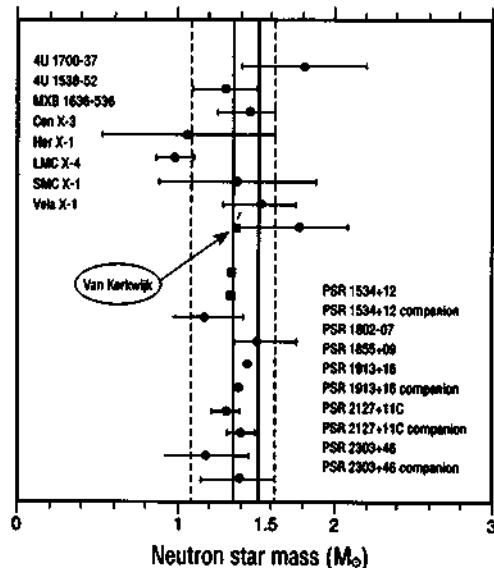


FIG. 1.—Measured mass of 17 neutron stars from Arzoumanian (1995), with the lower limit from Van Kerkwijk et al. (1994). Objects in high-mass X-ray binaries are at the top, and radio pulsars and their companions are at the bottom.

SN 1987A. We reviewed arguments that this compact object went into a black hole. Based on these, we believe  $\sim 1.56 M_{\odot}$  to give an upper limit on gravitational neutron star masses, only slightly larger than the  $1.5 M_{\odot}$  suggested by Brown & Bethe (1994).

We are grateful to K. Nomoto and F.-K. Thielemann for helpful discussions and for furnishing us with the unpublished

results on which the Kumagai et al. (1993) paper was based, as well as a preprint of Thielemann et al. (1994). We would like to thank M. P. Ulmer of the OSSE Collaboration for an informative discussion. Critical remarks of Ralph Wijers were useful. The work of G. E. B. was partially supported by the Department of Energy under grant DE-FG02-88ER40388 and the National Science Foundation under grant PHY 94-20470.

## REFERENCES

- Arzoumanian, Z. 1995, Ph.D. thesis, Princeton Univ.  
 Bethe, H. A. 1993, ApJ, 412, 192  
 ————, 1995, ApJ, submitted  
 Bethe, H. A., & Pizzochero, P. M. 1990, ApJ, 350, L33  
 Brown, G. E. 1995, ApJ, 440, 270  
 Brown, G. E., & Bethe, H. A. 1994, ApJ, 423, 659  
 Brown, G. E., & Weinberg, J. C. 1994, ApJ, 436, 843  
 Brown, G. E., Weinberg, J. C., & Wijers, R. A. M. J. 1995, in preparation  
 Chevalier, R. A. 1989, ApJ, 346, 847  
 ————, 1993, 411, L33  
 Clayton, D. B., Leising, M. D., The, L.-S., Johnson, W. N., & Kurfeß, J. D. 1992, ApJ, 399, L141  
 Cooperstein, J., Bethe, H. A., & Brown, G. E. 1984, Nucl Phys. A, 429, 527  
 Dwek, E., Mossley, S. H., Giacconi, W., Graham, J. R., Loewenstein, R. F., Silverberg, R. F., & Smith, R. K. 1992, ApJ, 389, L21  
 Herant, M., Benz, W., Hix, W. R., Fryer, C. L., & Colgate, S. A. 1995, ApJ, in press  
 Janka, H.-T., & Müller, E. 1993, in *Frontiers of Neutrino Astrophysics*, ed. Y. Suzuki & K. Nakamura (Tokyo: Universal Academy Press), 203  
 ————, 1994, in *The Physics of Supernovae*, Proc. of the Symposium at the Franklin Inst., Philadelphia, April 7-8  
 Keil, W., & Janka, H.-T. 1995, A&A, in press  
 Kumagai, S., Nomoto, K., Shigeyama, T., Hashimoto, M., & Itoh, M. 1993, A&A, 273, 153  
 Kurfeß, J. D., et al. 1992, ApJ, 399, L137  
 Lattimer, J. M., & Yahil, A. 1989, ApJ, 340, 426  
 Soutzlos, N. B., Phillips, M. M., Elata, J. H., DePay, D. L., & Walker, A. R. 1992, ApJ, 384, L33  
 Thielemann, F.-K., Nomoto, K., & Hashimoto, M. 1994, ApJ, submitted  
 Van Kerwijk, M. H., Van Paradijs, J., Zuidewijk, E. J., Hammeringh-Hensberge, G., Kaper, L., & Sterken, C. 1995, A&A, submitted  
 Wilson, J. R., & Mayle, R. 1990, private communication  
 Woosley, S. E. 1988, ApJ, 330, 218

---

## Chapter 9

# On the Formation of Low-Mass Black Holes in Massive Binary Stars

G.E. Brown, J.C. Weingartner and R.A.M.J. Wijers

Reprinted with permission from *Astrophysical Journal*, 463 (1996) 297–304

Copyright (1996) by American Astronomical Society

### **Commentary**

The main point of this article was to show that helium stars clothed with their hydrogen envelope end up as much more massive Fe cores than helium stars which are not so clothed; i.e., “naked” He stars. This contradicted conventional wisdom in the literature, which maintained that how the He core burned was independent of whether it was clothed by hydrogen or not. However, Timmes, Woosley & Weaver (1996) had evolved compact cores resulting from Type II and Type Ia supernova explosions, finding the former to peak  $\sim 1.7\text{--}1.8M_{\odot}$ , and the latter,  $\sim 1.3M_{\odot}$ . In Type II supernova explosions the He core is, of course, clothed by hydrogen during the helium core burning, whereas in Type Ia, which show no hydrogen in their spectra, but none the less come from massive stars, the hydrogen envelope is missing so that the He burns in a naked star.

This main point of this paper can be immediately seen from Fig. 1, which plots the Fe core mass as a function of ZAMS mass. The curve for the Fe core mass from the clothed He star rises rapidly whereas that from the naked He core basically saturates at an Fe core mass of  $\sim 1.5M_{\odot}$ . In the Commentary for Paper 4 we suggest that the Fe core mass is essentially that of the final compact core, the lowering of the core mass by gravitational binding energy being compensated for by fallback. In Table 3 of the paper commented on here we give this fallback for separation distances of the order of those arrived at in Paper 8 and the binding energy of the compact core is given in Eq. (15) of Paper 8.

In Paper 4 we suggested that an Fe core mass of  $\gtrsim 1.7\text{--}1.8M_{\odot}$  would go into a high-mass black hole, and one of  $1.5\text{--}1.7$  to  $1.8M_{\odot}$ , into a low-mass black hole. These limits will be discussed in more detail in the Commentary for Paper 21. In Paper 21, a much improved evolution of Fe cores resulting from both clothed and naked He stars is reported. In particular, wind losses during He core burning are reduced there to  $1/2\text{--}1/6$  of those used here, it being realized in the time between papers that He wind losses must be corrected

downwards for “clumpiness”, as is discussed in Paper 14. Although jagged, because of the small number of stars evolved, and crude, the curves of Fig. 1 are almost exactly what we end up with in Paper 21 after much improved and more complete calculations.

In the usual scenario for the evolution of high-mass X-ray binaries (the high mass belonging to the companion), the more massive of the two primaries first expands in the red giant stage and transfers mass to the secondary, which ends up as the massive star pouring matter onto the neutron star or black hole, the remnant of the more massive primary. This mass transfer from the red giant stage of the progenitor is called Roche Lobe Overflow (RLOF) or, equivalently, Case B mass transfer. A naked He core is left, with massive companion, which may at this stage be more massive than the originally more massive primary, depending on how much mass it accepts. Figure 1 of Paper 9 shows that this He star will go into either a neutron star, or possibly a low-mass black hole, but certainly not into a high-mass black hole. We will show in Papers 20 and 21 how high-mass black holes can be made.

Of particular interest in Paper 9 are the two high-mass X-ray binaries 4U 1700-37 and 4U 1223-62. The compact object in the former has a somewhat inaccurately measured mass of  $1.8 \pm 0.4 M_{\odot}$ , and falls nicely into the category of low-mass black hole. In Paper 9 we held back from claiming it was a low-mass black hole, but since that time Ralph Wijers (1997) has studied the accretion histories of HMXBs which are known to contain a neutron star, and from similarities of these histories compared with that of 1700-37, the pulsing of the rotating magnetic field of 1700-37 should be seen, if the compact object were a neutron star. We therefore feel more confident now in claiming it to be a low-mass black hole. In Paper 9, we find the most plausible range of the black hole progenitor to ZAMS  $24 - 48 M_{\odot}$ . From Fig. 1 of Paper 9 this is seen to be possible, but only if the maximum neutron star mass is not much higher than  $1.5 M_{\odot}$ . In general terms, Wellstein & Langer (1999) confirm this evolutionary scenario.

The X-ray pulsar 4U 1223-62 has the highest-mass optical companion known in a HMXB. The mass of the companion Wray 977, a luminous blue variable, is found to be at least  $48 M_{\odot}$ . In Paper 9 the binary evolved by RLOF with mass loss, and a mass for the progenitor of the neutron star is found to be at least  $36 M_{\odot}$ . Wellstein and Langer point out, however, that a more satisfactory evolution is for mass exchange between the two massive progenitors in the binary while they are on the main sequence (Case A mass transfer). This mass exchange can be conservative, and in the case they work out the progenitor of the neutron star can be of ZAMS mass  $\sim 26 M_{\odot}$ .

## References

- (1) Timmes, F.X., Woosley, S.E., and Weaver, T.A. (1996), *Astrophysical Journal* **457**, 843.
- (2) Wellstein, S., and Langer, N. (1999), *Astronomy & Astrophysics* **350**, 148.
- (3) Wijers, R.A.M.J. (1997), *Monthly Notices of the Royal Astronomical Society* **287**, 607.

## ON THE FORMATION OF LOW-MASS BLACK HOLES IN MASSIVE BINARY STARS

G. E. BROWN<sup>1</sup> AND J. C. WEINGARTNER<sup>2</sup>

Department of Physics, State University of New York at Stony Brook, Stony Brook, NY 11794

AND

RALPH A. M. J. WIJERS<sup>3</sup>

Princeton University Observatory, Peyton Hall, Princeton, NJ 08544-1001, and Institute of Astronomy, Madingley Road, Cambridge CB3 0HA, UK

Received 1995 April 7; accepted 1995 December 6

### ABSTRACT

Recently, Brown & Bethe suggested that most stars with main-sequence mass in the range of  $\sim 18\text{--}30 M_{\odot}$  explode, returning matter to the Galaxy, and then go into low-mass ( $\geq 1.5 M_{\odot}$ ) black holes. Even more massive main-sequence stars would chiefly go into high-mass ( $\sim 10 M_{\odot}$ ) black holes. The Brown-Bethe estimates gave  $\sim 5 \times 10^8$  low-mass black holes in the Galaxy. We here address why none of these have been seen, with the possible exception of the compact objects in SN 1987A and 4U 1700–37.

Our main point is that the primary star in a binary loses its hydrogen envelope by transfer of matter to the secondary and loss into space, and the resulting “naked” helium star evolves differently than a helium core, which is at least initially covered by the hydrogen envelope in a massive main-sequence star. We show that primary stars in binaries can end up as neutron stars even if their initial mass substantially exceeds the mass limit for neutron star formation from single stars ( $\sim 18 M_{\odot}$ ). An example is 4U 1223–62, in which we suggest that the initial primary mass exceeded  $35 M_{\odot}$ , yet X-ray pulsations show a neutron star to be present.

We also discuss some individual systems and argue that 4U 1700–37, the only example of a well-studied high-mass X-ray binary that does not pulse, could well contain a low-mass black hole. The statistical composition of the X-ray binary population is consistent with our scenario, but due to the paucity of systems it is consistent with more traditional models as well.

*Subject headings:* binaries: close — black hole physics — stars: evolution — stars: neutron — stars: Wolf-Rayet

### 1. INTRODUCTION

The formation of compact objects from massive stars is a difficult topic, because it depends on a number of physical processes of which our theoretical understanding is not yet satisfactory, the sample of available objects to test the theory on is small, and it is very difficult to get accurate observational determinations of the fundamental data required. We therefore use a mixed approach here in trying to constrain the outcomes of massive stellar evolution, using a combination of observational and theoretical arguments, whichever appears more reliable in each situation.

Chief among the theoretical uncertainties are (1) convective and semiconvective mixing in massive stars, which influences the sizes of the helium to iron cores that a star of given mass gets, (2) a quantitative calculation of supernova explosions and fallback, which should tell us what fraction of a core will eventually end up in the compact remnant, (3) the influence of binarity on the evolution of a stellar core, which determines how single stars will differ from ones in close binaries, (4) the precise dynamics of mass transfer in close binaries, and (5) the equation of state of matter at and above nuclear density, which determines the maximum mass of neutron stars and therefore helps determine which stellar masses can be progenitors of which types of compact object. The Brown & Bethe (1994) scenario for formation of a large number of low-mass black holes in the Galaxy used

a rather soft equation of state. We review the evidence for that here and in addition take account of new developments regarding the influence of binarity, using the best available treatments of uncertainties 1, 3, and 5. Until recently, the view was that loss of the hydrogen envelope in a close binary had very little influence on the eventual outcome of the evolution of a star, but recent work on mass loss of helium stars (Woosley, Langer & Weaver 1993, 1995) has prompted us to review this. We conclude that hardly any of the many low-mass black holes predicted by Brown & Bethe will be found in X-ray binaries (§ 2).

The observational uncertainties are mostly (1) that very few neutron stars have reliable mass determinations, and none of the ones in X-ray binaries have small errors (§ 3) and (2) that the optical companions of massive X-ray binaries seldom have well-known masses (§ 3.1).

The Brown & Bethe (1994) scenario was painted with broad strokes. Whereas the chief points may be correct, individual events have special features, such as the fluctuation of Fe cores with main-sequence mass. Even though quantitative calculations of the entire supernova have not been carried out to date, and it may take some time until accurate ones are completed, it is interesting to try to correlate observations with their general picture.

### 2. CORE EVOLUTION AND REMNANTS OF MASSIVE STARS

As the core of a massive star moves to more and more advanced burning stages it becomes hotter and denser, and the star becomes more centrally condensed. Density gradients above the core are so strong that the evolution of the

<sup>1</sup> E-mail: popenec@nuclear.physics.sunysb.edu.

<sup>2</sup> E-mail: josephw@phoenix.princeton.edu.

<sup>3</sup> E-mail: ramjw@ast.cam.ac.uk.

central density and temperature become independent of what the envelope does. As a result, the more advanced burning stages of massive stars become more and more independent of the mass of the star, most of which is locked up in the envelope. This phenomenon, called convergence, has long led researchers in binary evolution to make a crucial simplifying assumption in their models: that a star may lose its envelope to a companion when it expands to become a giant without this event causing any alteration of the further evolution of its naked helium core.

The main new stellar evolution ingredient in this paper is the fact that strong mass loss of such naked helium cores implies that this is not quite true (Woosley et al. 1993, 1995). To form an X-ray binary, the star whose core later becomes the accreting compact object nearly always loses its envelope and evolves as a naked He star. This implies that the compact objects in X-ray binaries have formed in an essentially different way than compact objects from single stars. Therefore, all the black holes and many of the neutron stars that we know form a population whose characteristics are not derivable from single stellar evolution. In the next few sections, we detail and attempt to quantify this effect.

### 2.1. Formation of High-Mass X-Ray Binaries

High-mass X-ray binaries consist of a compact object accompanied by an O- or B-type massive star, on or close to the main sequence. We briefly consider the standard model for their formation, to estimate the initial mass of the progenitor of the compact object given the mass of the current optical companion. We follow the work of van den Heuvel & Habets (1984) with small modifications and use the evolution tracks of Maeder (1990), since he used the same mass-loss prescription as in the work of Woosley and collaborators discussed later. A high-mass X-ray binary as observed now starts out as a close binary with primary mass  $M_p$  and mass ratio  $q$ . When the primary reaches the end of the main sequence, it expands and transfers its hydrogen envelope to its companion (so-called case B mass transfer); a fraction  $f$  of the transferred mass is lost from the binary. Then the now naked helium star primary evolves rapidly to a supernova and explodes. We neglect the short time this takes. The now more massive secondary evolves to the end of its main-sequence life. We accounted for rejuvenation by the added mass (van den Heuvel 1969; but see Braun & Langer 1995) when computing the time from mass transfer to core hydrogen exhaustion in the secondary. We assume that core hydrogen exhaustion in the secondary marks the start of the X-ray binary phase. This is reasonable because the observed high-mass X-ray binaries are in fairly close binaries, and the expansion of a star from the end of the main sequence across the Herzsprung gap is fast, so not much time will pass after the end of the main

sequence until substantial accretion starts. (It is even possible that the optical companions are still burning hydrogen in their cores and thus are technically on the main sequence, but have a giant-like structure due to severe mass loss; see Ziolkowski 1979.) The value of  $q$  is unknown, of course, and the value of  $f$  is rather uncertain: while mass transfer between roughly equal-mass stars is often thought to be conservative, there are indications that it may not be in practice, especially if the donor is a giant. We will vary these unknown parameters to estimate their importance.

As an example, consider a binary with initial masses 45 and  $36 M_\odot$  ( $q = 0.8$ ). When the primary reaches terminal age main sequence (TAMS), wind losses have reduced the masses to 40 and  $33 M_\odot$ . The  $22 M_\odot$  envelope of the primary is now transferred, during which 20% (say) is lost from the system. Now the stars are 18 and  $51 M_\odot$ , and soon thereafter the  $18 M_\odot$  helium star explodes, leaving a  $1.5 M_\odot$  compact object. When the rejuvenated secondary reaches TAMS, wind losses have reduced it to  $48 M_\odot$ , implying that in this case an X-ray binary has formed with an optical companion of mass  $48 M_\odot$ , precisely the minimum mass for  $1223 - 62$  (§ 3.1.5).

The lowest possible value for the initial primary mass (given a target value for the eventual optical companion mass) is obtained by maximizing the initial total mass  $M_p(1+q)$  and minimizing mass loss, which means setting  $q = 1$  and  $f = 0$ . In Table 1, we give the masses of the initial primary required to get a given optical companion mass in an HMXB using the calculation outlined above for different values of  $q$  and  $f$ .

### 2.2. Formation and Collapse of the Iron Core

Stars of main-sequence mass  $M > 12 M_\odot$  collapse when the Fe core reaches the Chandrasekhar limit  $M_{CS} = 5.76 Y_e^2 M_\odot$ , where  $Y_e$  is the ratio of electrons to nucleons. The maximum stable mass can be increased by thermal pressure by the factor  $1 + (\pi k T / \mu_e)^2$ , where  $\mu_e$  is the electron chemical potential, typically an  $\sim 15\%$  enhancement. With  $Y_{e,\text{final}} = 0.43 - 0.44$  this gives a thermally modified Chandrasekhar gravitational mass of  $\tilde{M}_{CS} \approx 1.25 M_\odot$ , scarcely dependent on the mass of the star. The  $Y_{e,\text{final}}$  is set by the strong  $\beta$ -decay of  $^{63}\text{Co}$ , which opposes the electron capture proceeding to lower  $Y_e$  (Aufderheide et al. 1990; Timmes, Woosley & Weaver 1996, § 8.C of Bethe 1990).

Brown & Bethe (1994) developed the scenario, based on the kaon condensation equation of state of dense matter (Thorsson, Prakash, & Lattimer 1994), that in many cases in the collapse of massive stars the compact core is stable for a sufficient time for explosion and the return of matter to the Galaxy, and then goes into a black hole. This was estimated to happen for stars of ZAMS masses of  $\sim 18 - 30 M_\odot$ . This possibility, that a star first explodes and subsequently

TABLE 1  
MASS OF THE PROGENITOR OF THE COMPACT OBJECT IN A HIGH-MASS X-RAY BINARY WITH CURRENT OPTICAL COMPANION MASS  $M_{opt}$ <sup>a</sup>

$M_{opt}$	$q = 1.0$			$q = 0.9$			$q = 0.8$			$q = 0.7$		
	$f = 0.0$	$f = 0.2$	$f = 0.5$	$f = 0.0$	$f = 0.2$	$f = 0.5$	$f = 0.0$	$f = 0.2$	$f = 0.5$	$f = 0.0$	$f = 0.2$	$f = 0.5$
20.....	11.7	12.9	15.1	12.5	13.9	16.4	13.4	15.0	18.0	14.5	16.3	19.9
30.....	18.8	20.7	24.0	20.2	22.3	26.2	21.8	24.3	28.9	23.8	26.7	32.3
40.....	26.9	29.4	33.9	29.1	32.0	37.3	31.7	35.1	41.4	35.0	39.1	46.7
50.....	36.2	39.4	45.0	39.4	43.2	49.7	43.5	48.0	55.8	48.8	54.3	63.8

<sup>a</sup> The initial mass ratio is  $q$ , and  $f$  is the fraction of the envelope mass lost in the first mass transfer phase.

drops into a black hole, had been suggested by Wilson et al. (1986) and Woosley & Weaver (1986). They had in mind the conventional scenario of a neutron star in which thermal pressure and neutrino pressure stabilize the compact object until it cools and collapses into a black hole. Prakash et al. (1995) show that this is possible for a small interval of  $\Delta M \sim 0.05-0.1 M_{\odot}$  in compact core masses. In addition to this "window" from thermal pressure and late time fallback, Brown & Bethe (1994) find approximately an additional  $0.2 M_{\odot}$  from the properties of the kaon condensed EOS, which implies a total window of  $\Delta M = 0.25-0.3 M_{\odot}$ . Chiefly this results because at high densities the matter ends up as nuclear matter, not neutron matter. The former is much "softer" than the latter and sends the core into a black hole.

The Brown & Bethe (1994) scenario indicated that most single stars with main-sequence masses between 18 and  $30 M_{\odot}$  explode, returning matter to the Galaxy, and leave low-mass black holes. In Table 2 we list the baryon number masses of remnants computed by Woosley & Weaver (1995), together with the gravitational masses obtained from the Lattimer & Yahil (1989) binding energy correction  $E = 0.084(M/M_{\odot})^2 M_{\odot}$ , where  $M$  is the gravitational mass of the compact core. Since supernova explosions giving quantitative results have not yet been carried out, they chose mass cuts outside the neutronized iron core and at the location of an abrupt entropy jump if one were nearby. The resulting behavior of remnant mass on initial stellar mass is complicated and not very certain.

The Fe core mass will not give the entire mass of the compact object, as there will be fallback from out to the bifurcation radius. Following Thielemann, Hashimoto, & Nomoto (1990) and Bethe (1990), we can estimate this radius from the fact that a small amount,  $\sim 0.075 M_{\odot}$ , of Fe came off from SN 1987A. This means that bifurcation had to come at a radius slightly inside of that up to which oxygen and silicon were burned to  $^{56}\text{Ni}$ , which later went

into Fe through weak decays. To form  $^{56}\text{Ni}$  from  $^{28}\text{Si}$  by successive addition of  $\alpha$  particles, the temperature must be above  $T = 350 \text{ keV} = 4 \times 10^9 \text{ K}$ . Given that the energy is mostly in radiation and electron pairs,  $T = 350 \text{ keV}$  corresponds to a blackbody energy density of  $w = 3.5 \times 10^{24} \text{ ergs cm}^{-3}$ . The shocked system is approximately isothermal, so the energy density is also  $w = 3E/(4\pi R^3)$ , where  $R$  is the shock radius. It is then straightforward to find that

$$R = (4100 \text{ km})E_{51}^{1/3}, \quad (1)$$

where  $E_{51}$  is the total energy in foes. Estimates for SN 1987A give  $E_{51}$  in the range of 1–1.5; therefore  $4100 \text{ km} < R < 4700 \text{ km}$ . Detailed calculations of Bethe & Brown (1995) give  $R = 3900 \pm 400 \text{ km}$ , not very different from equation (1). Hence, it may be reasonable to choose the enclosed mass somewhere in this range as the mass that will end up in the compact core. The  $M_{3500}$  and  $M_{4500}$  for Wolf-Rayet cores in Table 3 were kindly furnished us by Stan Woosley (1994, private communication).

Woosley et al. (1995) find that for 10 explosions of Wolf-Rayet stars of various masses in the range of  $4-20 M_{\odot}$ , the mass of ejected  $^{56}\text{Ni}$  is small, lying in the narrow range  $0.07-0.15 M_{\odot}$ . Thus, our procedure of obtaining the bifurcation radius near the edge of the iron core, as was done in SN 1987A, finds support. Woosley et al. (1993) also find that the helium core mass and further outcome is influenced by the uncertain  $^{12}\text{C}(\alpha, \gamma)^{16}\text{O}$  rate, but the variation is only a few percent.

The compact core mass for the  $40 M_{\odot}$  star evolved by Woosley et al. (1993) is similar to that for the  $60 M_{\odot}$  one, so there is presumably little difference in the region of masses  $40-60 M_{\odot}$ . In fact, Woosley et al. note that all of their models, with the exception of the  $85 M_{\odot}$  model, have strikingly similar iron cores, cores that are also similar in mass to lighter presupernova stars arising in the  $12-35 M_{\odot}$  range. And they note "Thus it seems likely that whatever mechanism functions to explode the common Type II supernova will also operate for at least some of these stars. There is no apparent mass limit above which one can say that a black hole mass remnant is very probable." According to Brown & Bethe (1994), the dividing line for high-mass black hole formation lies at a gravitational mass of  $1.84 M_{\odot}$ , i.e., baryon number mass  $2.09 M_{\odot}$ . In Woosley et al. (1993) only their highest mass star ( $85 M_{\odot}$ ) satisfies this, but the later results of Woosley & Weaver (1995) in Figure 1 do indicate that this limit can be exceeded at a mass as low as  $30 M_{\odot}$  in single stars. Four good candidates for high-mass black holes are listed by van den Heuvel (1992), and many more have been found recently (see review by Wijers 1996). While our considerations may make their formation somewhat more difficult, there are a number of known very massive

TABLE 2  
COMPACT CORE MASSES FOR SOLAR METALLICITY,  
FROM WOOSLEY & WEAVER 1995<sup>a</sup>

Main-Sequence Mass ( $M_{\odot}$ )	Baryon Number Mass ( $M_{\odot}$ )	Gravitational Mass ( $M_{\odot}$ )
15	1.43	1.30
18	1.76	1.56
20	2.06	1.78
25	2.07	1.79
30	4.24	4.24?
35	7.38	7.38?

<sup>a</sup> Question marks indicate that the conversion to gravitational mass is uncertain for objects that immediately collapse to black holes.

TABLE 3  
IRON CORES IN THE EVOLUTION OF WOLF-RAYET STARS, WITH MASS LOSS,  
FROM WOOSLEY, LANGER, & WEAVER 1995<sup>a</sup>

Initial He Star Mass/ $M_{\odot}$	Final He Star Mass/ $M_{\odot}$	Fe Core Mass/ $M_{\odot}$	$M_{3500}/M_{\odot}$	$M_{4500}/M_{\odot}$
5	2.82	1.38	1.55 (1.39)	1.59 (1.42)
7	3.20	1.42	1.67 (1.485)	1.71 (1.52)
10	3.51	1.49	1.69 (1.50)	1.73 (1.53)
20	3.55	1.49	1.70 (1.51)	1.77 (1.56)

<sup>a</sup> When several cases for a star are given, we have taken only case A. Metallicity 0.02 was considered. Numbers in parentheses are gravitational masses. Masses are baryon number masses. The  $M_{3500}$  and  $M_{4500}$  are the enclosed masses at 3500 and 4500 km, respectively. They were kindly furnished us privately by Stan Woosley.

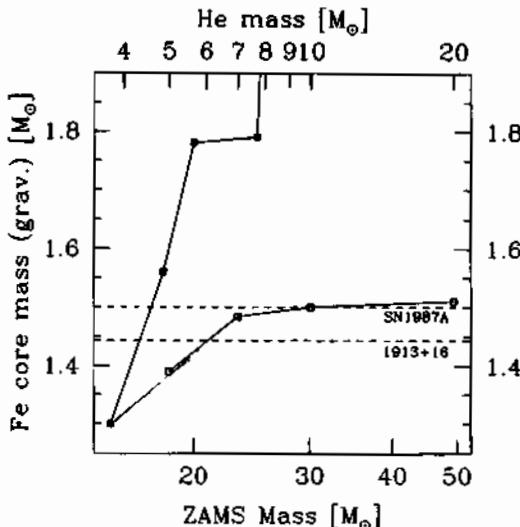


FIG. 1.—A comparison of the compact core masses resulting from the evolution of single stars (filled symbols) and naked helium stars with masses equal to the corresponding case B core masses of those same stars. Data are from Tables 2 and 3. The horizontal dashed lines indicate the mass of the heaviest known well-measured pulsar (PSR B1913+16; see Fig. 2) and the probable core mass of SN 1987A (see text).

W-R stars (Cherepashchuk 1991, and references therein) of which it is hard to imagine that they would not form massive black holes.

Given the very many uncertainties in the evolution of heavy stars with mass loss, it might appear unreasonable to consider the fact that most of the  $M_{3500}$  and  $M_{4500}$  masses exceed the Brown-Bethe  $1.50 M_{\odot}$  limit for neutron star masses. However, given the Bethe & Brown (1995) determination, as an exercise, we will do just that. But first, we must consider the effect of mass loss in helium stars, since all primaries in X-ray binaries spend some time prior to their explosion as naked helium stars (§ 2.1).

### 2.3. Wind Mass Loss: The Difference an Envelope Makes

When the primary in a close binary transfers its envelope to its companion it becomes a pure helium star. This helium star does not, however, evolve like the helium core of the original main-sequence star with hydrogen envelope. The core evolution and nucleosynthesis are altered if substantial mass loss continues, as it usually does, after the helium core is uncovered. According to Woosley et al. (1993) mass loss can lead to final helium star masses as small as  $4 M_{\odot}$  for a wide range of initial masses, such as the  $35\text{--}85 M_{\odot}$  range studied. This occurs because the mass-loss rate is mass-dependent. Simply integrating their mass-loss formula,

$$\dot{M}_{\text{WR}} = 5 \times 10^{-8} \left( \frac{M_{\text{WR}}}{M_{\odot}} \right)^{2.6} M_{\odot} \text{ yr}^{-1}, \quad (2)$$

over  $10^6$  yr we find that  $20$ ,  $10$ , and  $4 M_{\odot}$  helium stars end up at  $4.6$ ,  $4.1$ , and  $2.8 M_{\odot}$ , respectively. These numbers are not far from those arrived at by the full evolution calculation, so it is clear that the final masses are almost completely determined by the mass-loss rate  $\dot{M}_{\text{WR}}$ . This result cannot yet be considered very well established, because measurements of masses and mass-loss rates are usually quite uncertain. The available data sometimes yield a much

shallower dependence of mass-loss rate on mass, in which case the strong mass convergence noted here does not occur (see, e.g., Langer 1989; Schmutz, Hamann, & Wessolowski 1989; Smith & Maeder 1989). We shall nevertheless stick to this mass-loss prescription, since detailed calculations are available for it.

The chief result of Woosley et al. (1993) is that a presupernova star is not uniquely specified by its initial helium core mass. In order to show why a naked helium star ends up with a smaller Fe core mass than an initially "covered" helium core of the same mass, which resulted by loss of mass by wind from a massive main-sequence star, they evolve a  $4.25 M_{\odot}$  naked helium core and a  $4.25 M_{\odot}$  helium core that resulted after mass loss by wind from a  $60 M_{\odot}$  main-sequence star. Their chief point is that the latter core retains a chemical memory (although not a thermal memory) of its earlier history when it was covered up. The convective core size at the end of helium core burning is similar ( $M_{\text{CC}} \approx 2 M_{\odot}$ ) in the two cases. However, the chemical composition just outside this core is very different. In the case of the initially covered core, most of the matter just above the convective core has been burned to carbon and oxygen (presumably the "wraps" have kept the region hotter) so that there is very little helium left. In the case of the naked helium core, the helium concentration rises to 100% immediately beyond the convective zone. In the case of the initially covered helium core, the helium burning shell that develops at core helium exhaustion moves rapidly outward, through the small helium concentration, but for the naked star with  $Y \approx 1$ , it remains almost fixed in mass at the edge of the convective core. Consequently, the carbon-oxygen core masses of the presupernova models are very different in the two cases,  $3.03 M_{\odot}$  for the initially covered case and  $2.12 M_{\odot}$  for the always naked case. This leads to a smaller Fe core for the naked case and a better chance that it will end up as a neutron star.

The iron cores of Woosley et al. (1995) are given in Table 3. Note that these iron core masses are substantially less than those given in Table 2, where the cores were evolved with hydrogen envelope present. For example, the value of  $1.67 M_{\odot}$  for  $M_{3500}$  of a  $7 M_{\odot}$  helium star should be compared with the  $2.07 M_{\odot}$  remnant mass for a  $25 M_{\odot}$  main-sequence star in Table 2. In Figure 1 we show a comparison between the core resulting from covered and naked helium cores. It is evident that the difference increases rapidly with core mass above  $\sim 8 M_{\odot}$ , where mass loss according to equation (2) becomes important. Consequently, we see that "naked" primaries in binaries are much more likely to end up as neutron stars than single stars of the same initial main-sequence mass.

### 2.4. Neutron Star Mass Accretion and Survival

Neutron star masses can be accurately measured only when the neutron star occurs in a binary, and there are many situations in which it can accrete mass from the companion. This was generally not thought to increase the neutron star mass appreciably, because the accretion was assumed to be less than the Eddington limit,  $\dot{M} \approx 1.5 \times 10^{-8} M_{\odot} \text{ yr}^{-1}$ . However, it has been known for some time that if neutrinos can carry off the bulk of the energy, accretion can proceed at a much greater rate (Colgate 1971; Zeldovich, Ivanova, & Nadezhin 1972; Bisnovatyi-Kogan & Lamzin 1984). Chevalier (1993) pointed out that during the common envelope phase of binary evolution, photons

would be trapped and accretion could occur at much higher rates, typically  $10^{-4}$ – $10^{-3} M_{\odot}$  yr $^{-1}$ , and that neutron stars that have to go through this phase generally will go into black holes. Since the standard scenario for binary pulsar evolution has neutron stars going through a common envelope phase,<sup>4</sup> in the usual situation the neutron star may have the opportunity to accrete up to  $1 M_{\odot}$  of matter. Terman et al. (1994) and Taam, Bodenheimer, & Rózyczka (1994) have found, in a three-dimensional treatment of the neutron star in the common envelope, that the neutron star may survive spiral-in. This possibility arises when the neutron star ends up in a low-density region just outside a hydrogen burning shell with the massive companion in its red giant phase. Consequently, the standard scenario of binary pulsar evolution is expected to involve various gradations in the amount of matter accreted onto the neutron star. Thus, the fact that neutron stars of mass greater than the larger one of  $1.44 M_{\odot}$  in PSR 1913+16 have not been observed might be interpreted as evidence that neutron stars heavier than this do not exist. The best test of this issue will be the measurement of the mass of a millisecond pulsar with a white dwarf companion, since in the standard scenario such a pulsar will easily have accreted a few tenths of a solar mass of material (Phinney & Kulkarni 1994).

### 3. LIMITS TO LOW-MASS BLACK HOLES

Tests of the above scenario and the mass limits suggested by theory for the formation of neutron stars, low-mass black holes, and massive black holes can be obtained both from considering individual systems with special properties (§ 3.1) and from statistical analysis of the population of X-ray binaries as a whole (§ 3.2).

#### 3.1. Individual Systems

The current mass of the optical companion in an X-ray binary constrains the mass of the progenitor of the compact object (§ 2.1). Other systems with special properties may likewise provide tests of the mass cuts derived above (§ 2). Here we discuss five such cases in turn.

##### 3.1.1. GRO J0422+32

This object is one of the newly discovered X-ray novae with *GRO*. Most of these turn out to have rather high lower limits on the mass of the compact object in them. It is therefore generally assumed that these compact objects are black holes; a typical mass of these black holes is  $\sim 6 M_{\odot}$  (Wijers 1996). It might therefore be tempting to take the low-mass function of GRO J0422+32,  $f(M) = 1.21 \pm 0.06 M_{\odot}$  (Filippenko, Matheson, & Ho 1995), as evidence that it contains a low-mass black hole. However, the mass function is also proportional to  $\sin^3 i$ , where  $i$  is the angle between the orbital plane and the plane of the sky. If one takes the whole set of six X-ray novae with measured mass functions, then it would be surprising not to find a mass function as low as  $1.2 M_{\odot}$  among them if they were all  $6 M_{\odot}$  in reality.

##### 3.1.2. Supernova 1987A

It has been suggested that if SN 1987A had left a neutron star, it should have showed up due to a very large accretion luminosity within a year (Chevalier 1989; Brown & Weingartner 1994). Based on its nondetection, Brown, Bruenn, & Wheeler (1992) suggested that a black hole was formed instead. Bethe & Brown (1995) obtained an upper limit on the mass of the compact core in SN 1987A of  $1.56$

$M_{\odot}$  from the  $0.075 M_{\odot}$  of Fe production (see § 2.2) using the presupernova core evolved by Woosley. The presupernova core of Thielemann, Nomoto, & Hashimoto (1995) would have given  $\sim 1.443 M_{\odot}$ , just above the mass of the Hulse-Taylor pulsar. This latter mass may be somewhat too small because of the use of Schwarzschild, rather than Ledoux, convection in the calculations of the Nomoto group. On the other hand, using evolutionary calculations of Woosley et al. (1993) together with the evolution of the Hulse-Taylor pulsar by Burrows & Woosley (1986), the pulsar mass comes out as  $1.50 M_{\odot}$ , somewhat larger than the observed  $1.44 M_{\odot}$ . Therefore, we believe that the upper limit on the compact core mass in SN 1987A is somewhat too high, and we shall adopt the estimate of  $1.50 M_{\odot}$  of Brown & Bethe (1994), keeping in mind that it may be wrong, in either direction, by a few percent. Based on the reasoning that the compact object in SN 1987A went into a black hole but only barely so, we adopt the core mass in 1987A as the maximum possible neutron star mass:

$$M_{NS,max} = 1.50 M_{\odot}. \quad (3)$$

In Figure 2 and Table 4 we show the known masses of compact objects. The masses of radio pulsars, at the bottom of Figure 2, all fit nicely in with our estimate of  $M_{NS,max}$ . While the errors are large, it is interesting to discuss 4U 1700–37 and Vela X-1, as the central values of their masses exceed our  $M_{NS,max}$  the most.

##### 3.1.3. Vela X-1

Because of the pulses in the X-ray spectrum, the compact object in Vela X-1 is known to be a neutron star. Its high nominal mass (Fig. 2 and Table 4) may therefore be a worry in view of our low maximum neutron star mass. Van Kerkwijk et al. (1995b) find that observed velocities in Vela X-1 deviate substantially from the smooth radial-velocity curve expected for pure Keplerian motion. The deviations seem to be correlated with each other within one night, but not from one night to the other. The excursions suggest something like pulsational coupling to the radial motion and make it difficult to obtain an accurate mass measurement. The lower limit for the mass of the compact object in Vela X-1 is now found to be  $1.43 M_{\odot}$  at 95% confidence lower limit or

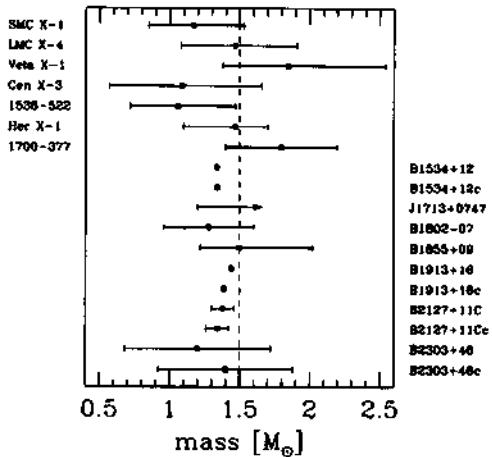


FIG. 2.—Measured masses of 18 compact objects. X-ray binaries are at the top, radio pulsars and their companions at the bottom. The vertical dashed line indicates our preferred value of  $M_{NS,max} = 1.50 M_{\odot}$ .

<sup>4</sup> An alternative scenario is given by Brown (1995).

TABLE 4  
MEASURED MASSES OF 18 COMPACT OBJECTS

Source	Type <sup>a</sup>	Mass ( $M_{\odot}$ ) <sup>b</sup>	Reference
SMC X-1 .....	HMXB	1.17 <sup>+0.36</sup> <sub>-0.32</sub>	1
LMC X-4 .....	HMXB	1.47 <sup>+0.44</sup> <sub>-0.39</sub>	1
Vela X-1 .....	HMXB	1.85 <sup>+0.69</sup> <sub>-0.47</sub>	1
Cen X-3 .....	HMXB	1.09 <sup>+0.57</sup> <sub>-0.32</sub>	1
1538-522 .....	HMXB	1.06 <sup>+0.41</sup> <sub>-0.34</sub>	1
Her X-1 .....	IMXB	1.47 <sup>+0.23</sup> <sub>-0.37</sub>	1
1700-37 .....	HMXB	1.8 (4) <sup>c</sup>	2
B1534+12 .....	BPSR	1.3378 (34)	3
B1534+12c .....	BPSR	1.3405 (34)	3
J1713+0747 .....	BPSR	> 1.2	4
B1802-07 .....	BPSR	1.28 (32)	3
B1855+09 <sup>d</sup> .....	BPSR	1.50 <sup>+0.32</sup> <sub>-0.34</sub>	5
B1913+16 .....	BPSR	1.442 (6)	6
B1913+16c .....	BPSR	1.386 (6)	6
B2127+11C .....	BPSR	1.38 (8)	7
B2127+11Cc .....	BPSR	1.34 (8)	7
B2303+46 .....	BPSR	1.20 (52)	3
B2303+46c .....	BPSR	1.40 (48)	3

<sup>a</sup> The abbreviations mean high mass X-ray binary, intermediate-mass X-ray binary, and binary pulsar, respectively. A lowercase c appended to a pulsar name denotes the unseen companion, which is also thought to be a neutron star.

<sup>b</sup> All errors or limits refer to the 95% confidence region. Numbers in parentheses are errors in the last digits. If a 1  $\sigma$  error was specified in the quoted reference, it was simply doubled. In case of pulsar B1855+09 this is somewhat dubious, because it is the only such case with asymmetric errors. Nonetheless, the confidence contours in the reference show that the limits we quote are roughly correct.

<sup>c</sup> This mass is rather less rigorous and reliable than the others, but it is included because it features in our discussion.

REFERENCES.—(1) van Kerkwijk et al. 1995a; (2) Heap & Corcoran 1992; (3) Arzoumanian 1995; (4) Camilo 1995; (5) Kaspi et al. 1994; (6) Taylor & Weisberg 1989; (7) Deich 1996.

1.37  $M_{\odot}$  at 95% confidence interval around the most probable value (van Kerkwijk et al. 1995b). Hence the data do not yet contradict equation (3) but may do so if the observations on this system improve.

### 3.1.4. 4U 1700-37

The other object in Figure 2 that lies beyond our 1.5  $M_{\odot}$  limit is 4U 1700-37. Its mass determination (Heap & Corcoran 1992) is the poorest in the whole set because it is based on the spectral type and wind properties of the optical companion. Yet with a nominal mass of 1.8  $\pm$  0.4  $M_{\odot}$  it is interesting to note the peculiarity of this source. Contrary to all other high-mass systems with such a low-mass compact star, it is not pulsing in X-rays (Bhattacharya & van den Heuvel 1991), so it does not appear like a rotating magnetized neutron star. Also, its spectrum extends to very high energies, significantly beyond 60 keV (Rubin et al. 1993)<sup>5</sup> and may well be harder than that of other high-mass X-ray binaries. Such a hard spectrum is often associated with black holes, even though some proven neutron stars seem capable of it (Tanaka & Lewin 1995).

The companion star HD 153919 in 4U 1700-37 is an O6I star with a very uncertain mass. Conti (1978) chose a value of 27  $M_{\odot}$  for this star, although he noted that the value was uncertain. Heap & Corcoran (1992) derive a

much higher value, 52  $\pm$  2  $M_{\odot}$ . The situation is unsatisfactory, and we settle on an estimate of 40  $\pm$  10  $M_{\odot}$  with error bars large enough to encompass both of the above estimates. Indeed, Heap & Corcoran say that HD 153919 is much like  $\lambda$  Cep. Herrero (1995) finds a mass for  $\lambda$  Cep in the central part of this range. This then leads to a minimum progenitor mass of the compact object of 19–36  $M_{\odot}$ , and a more plausible range of 24–48  $M_{\odot}$ . If it is a black hole, then its likely lower progenitor mass than 1223–62 (§ 3.1.5) indicates a possible nonmonotonic behavior of the remnant mass with initial mass of the star or the influence of other parameters (such as how far the star evolved to the giant stage before it lost its envelope).

In any case, it is worth considering whether the absence of pulsation cannot simply be the result of a low field of a neutron star, despite the fact that in all other high-mass X-ray binaries the neutron star does manage to pulse. Taam & van den Heuvel (1986) have shown that empirically field decay is inversely correlated with mass accretion (although there is up to now no fundamental theoretical basis for this correlation). Applying a direct proportionality between field and accreted mass given that millisecond pulsars are thought to have accreted 0.01–0.1  $M_{\odot}$  of material and thereby decreased their magnetic field by 4 orders of magnitude, one finds  $B_0/B \sim \Delta M/(10^{-5.5} M_{\odot})$ . 1700-37 may have been accreting material for up to  $5 \times 10^4$  yr, at a rate of perhaps 10% of the Eddington rate, implying  $\Delta M < 10^{-4} M_{\odot}$ . This means that its field could have decayed to a few percent of its initial value, possibly putting it at  $\sim 3 \times 10^{10}$  G now. Other pulsars that descended from massive binaries, e.g., PSR 1913+16 (the Hulse-Taylor binary pulsar), have fields even lower than that, so a low field is quite possible in 1700-37. However, the X-ray spectrum should then be softer than that of X-ray pulsars, rather than harder as observed. With some reserve, we therefore advocate the view that the compact object in 1700-37 is indeed a low-mass black hole.

### 3.1.5. 4U 1223-62

This X-ray pulsar (White et al. 1976) has the highest-mass optical companion known. Sato et al. (1986) determined the mass of the companion, Wray 977, to be  $M_{\text{opt}} \approx 38 M_{\odot}$ , accounting for the limit on the inclination angle due to the absence of X-ray eclipses. More recently, Kaper et al. (1995) revised the spectral classification of Wray 977, claiming it is a hypergiant and thus further away from us. This more than doubles the star's radius and thus forces a smaller inclination ( $i \leq 62^\circ$ ) in order to avoid eclipses. Consequently, the minimum mass is 48  $M_{\odot}$ . From Table 1 and § 2.1, we infer that the mass of the progenitor of the neutron star must have exceeded 36  $M_{\odot}$ . A more plausible progenitor mass, adopting  $q = 0.8$  and  $f = 0.2$ , would have been 45  $M_{\odot}$ . Such massive stars leave helium stars of 13 and 18  $M_{\odot}$ , which according to Table 3 leave compact objects just above our 1.5  $M_{\odot}$  limit for neutron stars. Given the uncertainties, we may still say that the presence of a neutron star is consistent with our understanding of the evolution of this binary.

Incidentally, a 36  $M_{\odot}$  star is already close to the range where binarity ceases to matter much to a star's evolution. Above  $\sim 45 M_{\odot}$ , stars become luminous blue variables and lose their envelopes without the help of a binary companion to become Wolf-Rayet stars (Chiosi & Maeder 1986). Woosley et al. (1993) find that rapid mass loss in the luminous blue variable phase determines the stellar mass at the

<sup>5</sup> We thank D. Chakrabarty for pointing this out to us.

beginning of helium burning. The hydrogen-rich envelope is completely gone, and their situation is similar to that of the naked helium stars formed after mass transfer in a binary.

### 3.2. Statistical Considerations

Let us now gather up a plausible set of mass limits for stars to form neutron stars and black holes, both for the single and close binary case, to estimate how much difference it will make in the numbers we expect to see. In both cases, we shall deem compact objects to form from the mass range  $8-100 M_{\odot}$ . In single stars, we take the mass range for low-mass black hole formation to be  $18-30 M_{\odot}$ , after Brown & Bethe (1994), in fair agreement with the Woosley et al. results (Fig. 1). Neutron stars are formed below this range and high-mass black holes above it. In binary stars, we face the problem that the theoretical curve of core mass versus stellar mass intersects the critical mass of  $1.5 M_{\odot}$  at a very shallow angle, so we do not get a good value for the dividing mass. The uncertainty is even greater because, as we noted earlier, the Woosley et al. (1995) masses give an  $\sim 0.06 M_{\odot}$  too high a mass for PSR 1913+16. If we lower the curve for naked helium stars by that amount, we would get no low-mass black holes at all in the mass range shown. Instead we set the cut at  $36 M_{\odot}$ , the lowest value for the progenitor mass of the neutron star in 4U 1223-62 (§ 3.1.5.). As the upper limit for low-mass black hole formation we choose  $50 M_{\odot}$ , somewhat arbitrarily.

We now assume that single stars and primaries in binaries have a Salpeter initial mass distribution, i.e.,  $N(>M) \propto M^{-1.35}$ , to compute the fractions of stars that will yield each type of compact object. For single stars, we find that 69% form neutron stars, 17% low-mass black holes, and 14% high-mass black holes. For binaries, the results are starkly different: 90% neutron stars, 5% low-mass black holes, and 5% high-mass black holes.

Of course, compact objects are only seen in binaries (except radio pulsars), so this is the only place where we can test the numbers. The largest number of high-mass X-ray binaries are Be/X-ray binaries (15–20; see Tables 3, 5, and 8 of Bhattacharya & van den Heuvel 1991), where the companion is a rapidly rotating B star, i.e., lighter than  $20 M_{\odot}$ . This implies that the initial primary was almost certainly below  $18 M_{\odot}$  (see Table 1), and thus the compact object should be a neutron star, as is observed, whether we use the numbers for binary or single stars. In high-mass X-ray binaries the situation is less clear. There are six with X-ray pulsars, possibly one with a low-mass black hole and two with massive black holes. This implies somewhat higher fractions of black holes than we just derived, but we should account for the fact that by looking only at high-mass X-ray binaries, which have very massive optical companions, we have implicitly limited ourselves to a smaller range of initial masses, probably starting at  $20 M_{\odot}$  or so rather than  $8 M_{\odot}$  (Bhattacharya & van den Heuvel 1991). For the single star mass cuts, this means virtually no neutron stars would be present in them. So the fact that we do see mainly neutron stars in high-mass X-ray binaries is indirect evidence for the increased mass limit for neutron star formation. (This is only true within the context of a soft EOS of neutron stars: stiffer neutron stars can form from stars more massive than  $20 M_{\odot}$ .) The binary scenario mass cuts now would yield 60% neutron stars, 20% low-mass black holes, and 20% massive black holes. This is in reasonable agreement with the observed number ratio of 2:1:1 for nearby systems. It is

difficult to compare the predicted formation rates with the observed numbers because the lifetimes may be different for the different types, as may the selection effects. The lifetimes for the neutron star and low-mass black hole systems will be similar, because the mass ratios are essentially the same, and thus mass transfer is unstable on the same timescale and they are both expected to live as X-ray sources for a thermal timescale of the envelope, i.e., a few times  $10^4$  yr. But the massive black holes are closer in mass to their optical companions; hence the mass transfer in systems like Cyg X-1 could be less unstable and their lifetimes longer, leading to some overrepresentation in the observed sample. There is also some bias in the observed sample toward high-mass black holes because the higher the mass of a black hole is, the more it can accrete: the accretion luminosity from Bondi-Hoyle type wind accretion increases with the square of the mass of the accreting object up to the Eddington luminosity. Above that, the luminosity is limited to the Eddington rate, which increases linearly with mass. High-mass black holes with  $M \sim 10 M_{\odot}$  can thus give  $\sim 2$  orders of magnitude more accretion than the low-mass ones with  $M \sim 1.5 M_{\odot}$ .

The number ratio of massive stars that form low-mass black holes to those that form neutron stars is thus low, 5%–20% depending on what fraction of stars are in binaries. This implies that it is not at all unlikely to find no low-mass black holes among the 10 well-studied X-ray binaries in globular clusters. Hence the absence of known low-mass black holes in the population of globular-cluster X-ray sources is quite consistent with the Brown-Bethe scenario for low-mass black hole formation, and statements to the contrary by Kulkarni, Hut, & McMillan (1993) are incorrect.

### 4. CONCLUSIONS

We have discussed the scenario in which the primary star in a binary, as long as its mass is less than  $\sim 40 M_{\odot}$ , will evolve quite differently from an isolated star of the same mass. This is due to transfer of its hydrogen envelope to the secondary, and subsequent large mass-loss rates of the helium core (which had hitherto not been taken into account). In this way, primaries corresponding to main-sequence masses as massive as  $35 M_{\odot}$  can evolve into neutron stars, whereas single stars in the mass range  $18-30 M_{\odot}$  would go into low-mass black holes. In addition, luminous blue variables in the ZAMS mass range  $50-60 M_{\odot}$  may end up as neutron stars, especially in close binaries. This leaves only a narrow region of masses around  $35-50 M_{\odot}$  for possible evolution into low-mass black holes, in addition, possibly, to some very massive stars above  $60 M_{\odot}$ . Although several aspects of our discussion are uncertain, it does seem clear that few stars in binaries would be expected to go into low-mass black holes. The different behavior of "naked" helium cores may explain why only one possible low-mass black hole has been observed in high-mass X-ray binaries while the Brown-Bethe scenario, in which stars of main-sequence masses  $18-30 M_{\odot}$  explode and then go into low-mass black holes, may still be roughly correct.

We showed that 4U 1700-37, the only example of a well-studied X-ray binary that does not pulse, is a fair candidate for containing a low-mass black hole.

We would like to thank Ed van den Heuvel for suggesting that 1700-37 contained a black hole. We are extremely

grateful to Stan Woosley for sending us nearly all of the evolutionary material on which our arguments are based and for the many private communications, some of which are quoted in the paper. We are grateful to Hans Bethe for

helpful criticism. G. B. and J. W. are supported by the U.S. Department of Energy Grant DE-FG02-88ER 40 388; R. W. is supported by a Compton Fellowship (grant GRO/PFP-91-26) and by a PPARC fellowship.

## REFERENCES

- Arzoumanian, Z. 1995, Ph.D. thesis, Princeton Univ.  
 Auferheide, M. B., Brown, G. E., Kuo, T. T. S., Stout, D. B., & Vogel, P. 1990, ApJ, 362, 241  
 Bethe, H. A. 1990, Rev. Mod. Phys., 62, 801  
 Bethe, H. A., & Brown, G. E. 1995, ApJ, 445, L129  
 Bhattacharya, D., & van den Heuvel, E. P. J. 1991, Phys. Rep., 203, 1  
 Bisnovatyi-Kogan, G. S., & Lamzin, S. A. 1984, Soviet Astron., 28, 187  
 Braun, H., & Langer, N. 1995, A&A, 297, 483  
 Brown, G. E. 1995, ApJ, 440, 270  
 Brown, G. E., & Bethe, H. A. 1994, ApJ, 423, 659  
 Brown, G. E., Bruenn, S. W., & Wheeler, J. C. 1992, Comments Astrophys., 16, 153  
 Brown, G. E., & Weingartner, J. C. 1994, ApJ, 436, 843  
 Burrows, A., & Woosley, S. E. 1986, ApJ, 308, 680  
 Camilo, F. 1995, Ph.D. thesis, Princeton Univ.  
 Cherepashchuk, A. M. 1991, in *Wolf-Rayet Stars and Interrelations with Other Massive Stars in Galaxies*, ed. K. A. van der Hucht & B. Hidayat (Dordrecht: Kluwer), 187  
 Chevalier, R. A. 1989, ApJ, 346, 847  
 ———. 1993, ApJ, 411, L33  
 Chiosi, C., & Maeder, A. 1986, ARA&A, 24, 329  
 Colgate, S. A. 1971, ApJ, 163, 221  
 Conti, P. S. 1978, A&A, 63, 225  
 Deich, W. 1996, in preparation  
 Filippenko, A. V., Matheson, T., & Ho, L. C. 1995, ApJ, 455, 614  
 Heap, S. R., & Corcoran, M. F. 1992, ApJ, 387, 340  
 Herrero, A. 1995, private communication  
 Kaper, L., Lamers, H. J. G. L. M., Ruymakers, E., van den Heuvel, E. P. J., & Zuiderwijk, E. J. 1995, A&A, 300, 446  
 Kulkarni, S. R., Hut, P., & McMillan, S. 1993, Nature, 364, 421  
 Langer, N. 1989, A&A, 220, 135  
 Lattimer, J. M., & Yahil, A. 1989, ApJ, 340, 420  
 Maeder, A. 1990, A&AS, 84, 139  
 Pfeiffer, E. S., & Kulkarni, S. R. 1994, ARA&A, 32, 591  
 Prakash, M., Bombaci, I. M., Ellis, P. J., Lattimer, J. M., & Knorren, R. 1995, Phys. Rept. in press  
 Rubin, B. C., Harmon, B. A., Fishman, G. J., Meegan, C. A., Wilson, R. B., Briggs, M. S., Paciesas, W. S., & Finger, M. H. 1993, in *AIP Conf. Proc.*, Vol. 280, *Compton Gamma-Ray Observatory*, ed. M. Friedlander, N. Gehrels, & D. J. Macomb (New York: AIP), 381  
 Sato, N., Nagase, F., Kawai, N., Kelley, R. L., Rappaport, S., & White, N. E. 1986, ApJ, 304, 241  
 Schmutz, W., Hamann, W., & Wessolowski, U. 1989, A&A, 210, 236  
 Smith, L. F., & Maeder, A. 1989, A&A, 211, 71  
 Taam, R. E., Bodenheimer, P., & Różyczka, M. 1994, ApJ, 431, 247  
 Taam, R. E., & van den Heuvel, E. P. J. 1986, ApJ, 305, 235  
 Tanaka, Y., & Lewin, W. H. G. 1995, in *Cambridge Astrophysics Ser.*, Vol. 26, *X-Ray Binaries*, ed. W. H. G. Lewin, J. van Paradijs, & E. P. J. van den Heuvel (Cambridge: Cambridge Univ. Press), 126  
 Taylor, J. R., & Weinberg, J. M. 1989, ApJ, 345, 434  
 Terman, J. L., Taam, R. E., & Hernquist, L. 1994, ApJ, 422, 729  
 Thielemann, F., Hashimoto, M., & Nomoto, K. 1990, ApJ, 349, 222  
 Thielemann, F., Nomoto, K., & Hashimoto, M. 1996, ApJ, 460, in press  
 Thorsson, V., Prakash, M., & Lattimer, J. M. 1994, Nucl. Phys. A, 572, 693  
 Timmes, F. X., Woosley, S. E., & Weaver, T. A. 1996, ApJ, 457, 834  
 van den Heuvel, E. P. J. 1969, AJ, 74, 1095  
 ———. 1992, in *Proc. Internat. Space Year Conf. Satellite Symp. 3, Space Sciences with Particular Emphasis on High-Energy Astrophysics*, ed. T. D. Guyenne & J. J. Hunt (Noordwijk: ESTEC), 29  
 van den Heuvel, E. P. J., & Habets, G. M. H. J. 1984, Nature, 309, 598  
 van Kerkwijk, M. H., van Paradijs, J., & Zuiderwijk, E. J. 1995a, A&A, 303, 497  
 van Kerkwijk, M. H., van Paradijs, J., Zuiderwijk, E. J., Hammerschlag-Hensberge, G., Kaper, L., & Sterken, C. 1995b, A&A, 303, 483  
 White, N. E., Mason, K. O., Huckle, H. E., Charles, P. A., & Sanford, P. W. 1976, ApJ, 209, L119  
 Wijers, R. A. M. J. 1996, in *Evolutionary Processes in Binary Stars*, ed. R. A. M. J. Wijers, M. B. Davies, & C. A. Tout NATO ASI Ser. C (Dordrecht: Kluwer)  
 Wilson, J. R., Mayle, R., Woosley, S. E., & Weaver, T. A. 1986, in *Proc. Texas Symp. on Relativistic Astrophysics*, ed. M. Livio & G. Shaviv (New York: New York Acad. Sci.), 267  
 Woosley, S. E., Langer, N., & Weaver, T. A. 1993, ApJ, 411, 823  
 ———. 1995, ApJ, 448, 315  
 Woosley, S. E., & Weaver, T. A. 1986, ARA&A, 24, 205  
 ———. 1995, ApJS, 101, 181  
 Zeldovich, Y. B., Ivanova, L. N., & Nadezhin, D. K. 1972, Soviet Astron., 16, 209  
 Ziolkowski, J. 1979, in *IAU Symp. 83, Mass Loss and Evolution of O-Type Stars*, ed. P. Conti & C. W. H. de Loore (Dordrecht: Reidel), 385

---

## Chapter 10

# The Evolution of Relativistic Binary Pulsars

T. Wettig and G.E. Brown

Reprinted with permission from *New Astronomy*, 1 (1996) 17–34

Copyright (1996) by Elsevier Science B.V.

### **Commentary by G.E. Brown**

Tilo Wettig had finished his Ph.D. thesis on Bose Condensation under the supervision of Andy Jackson in the summer of 1994. He was supposed to go on to a postdoctoral position with Hans Weidenmüller in the autumn, but Hans was unable to get financial support for him until the New Year. Thus, I proposed to Tilo, “Why don’t you work with me on the evolution of binary pulsars?”

At that time we knew of only four binary pulsars, PSR 1534+12, PSR 1913+16, PSR 2303+46 and PSR 2127+11C. The final one was in a globular cluster and it was thought that the two neutron stars had not been formed in the binary, but had got together in an exchange. It turned out later that the companion in the reasonably wide PSR 2303+46 was a white dwarf (see Paper 19).

In any case, our question was, “Why are there two relativistic binary pulsars which are so narrow, the separation  $a \sim 3R_{\odot}$ , instead of a whole distribution out over larger distances?” Our tentative answer came in terms of the “Observability Premium”. It is well known (Bhattacharya & van den Heuvel, 1991) that rotating neutron stars slow down by emitting magnetic dipole radiation. Once their period is less than a certain critical period, they no longer generate the voltage to drive the electron-positron cascade which produces the pulsar radiation. The total time that they can be observed is roughly

$$\tau_{\text{obs}} = 5 \times 10^6 B_{12}^{-1} \text{ yrs},$$

where  $B_{12}$  is the magnetic field in units of  $10^{12}$  gauss.

Fresh neutron stars such as the Crab pulsar have high magnetic fields, typically  $10^{12}$ G. (And they can have much higher ones, such as the magnetars with fields up to  $10^{14}$ – $10^{15}$ G. It is clear that these have very short observation times  $\tau_{\text{obs}}$ .) The relativistic binary pulsars have fields  $\sim 10^{10}$ G, and can therefore be seen for  $\sim 100$  times longer than a typical “fresh” neutron star with  $B \sim 10^{12}$ G.

An important ingredient in the argument of Paper 10 is that magnetic fields of pulsars seem to decay in strength only when they are in binaries. An empirical relation, which seems to work quite well, is that the amount of decay is proportional to the amount of accreted matter, as explained in the paper.

The new point, introduced in Paper 5, is that matter can be accreted onto the pulsar by the He star wind of the companion in the He-star, pulsar binary progenitor of the binary pulsar. In order to bring the pulsar magnetic field down in the  $\sim 5 \times 10^5$  yrs that the He-star wind can be accreted onto it, the distance between the two stars (He star and pulsar) in the binary must be small. It then follows that the (generally eccentric) orbit of the binary neutron star must be nearly as tight, since the latter orbit has the point of explosion of the He-star in common with the orbit of the He-star, pulsar binary.

Although not explicitly introduced here in this form, the important point of this paper is to motivate the introduction of an observability premium

$$\Pi = 10^{12} G/B,$$

where  $B$  is the magnetic field of the pulsar; the weighting by  $\Pi$  should be included in population syntheses because the longer observability of low-field pulsars means that we are more likely to see them. In particular, it is clear that if tight orbits are necessary for the magnetic field  $B$  to be brought down by mass accretion, then we will be much more likely to see the tight binaries. This work has been made more quantitative by calculations of J. Fernaschelli, R.A.M.J. Wijers and G.E. Brown in a paper to appear in *ApJ*.

#### Reference

- (1) Bhattacharya, D., and Van den Heuvel, E.P.J. (1991), *Physics Reports* **203**, 1.



ELSEVIER

---



---

**NEW  
ASTRONOMY**


---



---

New Astronomy 1 (1996) 17-34

## The evolution of relativistic binary pulsars\*

Tilo Wettig<sup>1</sup>, Gerald E. Brown

*Department of Physics, State University of New York, Stony Brook, NY 11794-3800, USA*

Received 15 December 1995; accepted 2 February 1996

Communicated by Edward P.J. van den Heuvel

---

### Abstract

We address the question of why two of the four neutron star binaries, PSR 1913+16 and PSR 1534+12, have narrow orbits, with binary separation  $\sim 3 R_{\odot}$ , and relatively low magnetic fields  $B \sim 10^{10}$  G. Beginning from a scenario of double helium-star progenitors for binary pulsars, we show that the orbits will be tightened to  $\sim 3-5 R_{\odot}$  during a common envelope evolution during which the two helium stars expel the common hydrogen envelope. There is then, in the evolution, an "observability premium" if the helium-star, neutron-star intermediate stage is such that the binary is narrow in this stage. If so, accretion onto the neutron star from the helium star wind can lower its magnetic field, lengthening the pulsar spin-down time by doing so. The lower magnetic field  $B$  then has a longer spin-down time and, hence, a longer time of observability.

Evolution of PSR 1913+16 from a binary of rather massive, main-sequence mass  $\sim 24 M_{\odot}$ , O-stars proves to be straightforward, and the probability of such evolution in our model looks reasonable. We are unable to carry the lighter binary PSR 1534+12 through the second common envelope stage, reverse case BB mass transfer, which occurs when light-mass helium stars expand in a red supergiant phase, engulfing the neutron star. We suggest that, in some cases, the binary does not survive in this second common envelope, the neutron star either spiraling in to the helium star, or going into a black hole.

**PACS:** 97.80.-d; 97.10.Gz; 97.60.Bw; 97.60.Gb; 97.60.Jd

**Keywords:** Binaries; Accretion, accretion disks; Supernovae; Pulsars; Stars: neutron

---

### 1. Introduction

In this article we address the question of why two of the four neutron star binaries, PSR 1913+16 and

PSR 1534+12, have narrow orbits, with binary separation  $\sim 3 R_{\odot}$ . The third binary, PSR 2303+46, is wider, with a 13 day period, and so is the newly discovered PSR J1518+4904 (Nice et al., 1996) with an 8.5 day period. Although common envelope evolution of the progenitor stars in the binary can bring them close together, we shall see that in the evolution of the neutron star binary there is an "observability premium" if the helium-star, neutron-star intermediate stage is such that the binary is

\*Supported by the U.S. Department of Energy under grant No. DE-FG02-88ER40388

<sup>1</sup>Present address: Max-Planck-Institut für Kernphysik, Postfach 103980, D-69029 Heidelberg, Germany.

E-mail: tilo@tck.mpi-hd.mpg.de

narrow in this stage. If so, accretion onto the neutron star from the helium star wind can substantially lower its magnetic field, greatly lengthening its spin-down time and, therefore, time of observability. We do not consider the neutron star binary PSR 2127+11C, which is naturally explained as resulting from an exchange reaction between a neutron star and a binary which took place  $< 10^8$  yrs ago in the cluster core of M15 (Phinney & Sigurdsson, 1991).

There is a commonly accepted scenario (Van den Heuvel & Van Paradijs, 1993) for evolving relativistic binary pulsars. It involves high mass X-ray binaries (or Be/X-ray binaries) as an intermediate link between the original O- or B-star binary and the final relativistic binary pulsar, such as PSR 1913+16. In the high mass X-ray binary, following the X-ray phase, the neutron star enters into the envelope of the large companion star. As the neutron star spirals in, it creates dynamical friction in the envelope. The hydrodynamical coupling leads to the expulsion of the envelope as the neutron star tightens its orbit. The neutron star and helium star are left in a relatively narrow orbit.

In this scenario, the helium-star, neutron-star binary is the link between the high mass X-ray binaries and the relativistic binary pulsars. The problem is that, at most, only one of these, Cyg X-3, is seen. Even in this case, it is argued that the compact object is likely a black hole (Brown, 1995a).

The "problem of the disappearing neutron star" was noted already in 1976 by Van den Heuvel (1976). In intervening years, this disappearance was often described by the neutron star surrounding itself by an accretion envelope from the large star and becoming a Thorne-Żytkow star, which would appear as a giant. Brown (1995a) showed, however, that the ram pressure experienced by the nascent Thorne-Żytkow star, as the neutron star entered into the common envelope, would crush it. Support for Brown's result is found by Fryer & Benz (1994), although "neutrino heating ... can play an important role in creating instabilities in some formation schemes, leading to an expulsion of matter rather than rapid accretion."

Chevalier (1993) estimated that a neutron star in a common envelope accretes sufficient matter to evolve into a black hole. The black hole then presumably sucks up the envelope and helium core of the giant. Chevalier's estimates were confirmed by Brown (1995a). Special circumstances are necessary to avoid black hole formation.

Terman et al. (1994) and Taam et al. (1996) find that very special circumstances are necessary to avoid spiral-in to the core. First of all, in a three-dimensional non-axisymmetric calculation, Terman et al. find that most of the energy from the tightening of the orbit of the smaller star, an  $0.94 M_{\odot}$  dwarf in their case, goes into torquing up the envelope of the red giant, rather than into dynamical friction. This presents a new scenario. If the large star is in red giant phase, with sufficiently steep density gradients, the density is low enough so that matter can be expelled from the vicinity of the dwarf by centrifugal force. In Fig. 5 of Taam et al. (1996), the low density region, extending from  $\sim 3.5-10.5 R_{\odot}$ , is visible. The spiral-in of the dwarf ends in the lower part of this region, finding little material to be expelled, so that further tightening of the orbit is avoided. The "tidal catastrophe" goes back to Darwin (1908). If the mass of the lighter star is too small, the orbit of the companion spirals further and further inwards, being unable to furnish enough angular momentum to spin up the red giant envelope, as it tries to bring it into corotation with the orbit. Without special circumstances, the companion will merge with the helium core of the giant. Taam and collaborators have studied these special circumstances. As an example of a neutron star which might survive the common envelope, Taam (1996) considers a  $1.4 M_{\odot}$  neutron star in a wide orbit about a  $16 M_{\odot}$  giant, with initial period of  $\sim 1.3$  yrs.

Although, as noted by Terman et al. (1994), the energy dissipated in dynamical friction decreases rapidly as the velocity of the neutron star is brought subsonic with energy transfer to the envelope, the accretion rate moves from Bondi-Hoyle-Lyttleton wind accretion towards the spherically symmetrical Bondi accretion. Even though the energy dissipation into dynamical friction decreases greatly, with de-

creasing velocity relative to the envelope, the accretion decreases much less rapidly, as described by Eq. (4.9) of Brown (1995a). Thus, the neutron star may well go into a black hole, even with the special circumstances of Taam and collaborators.

Van den Heuvel (1994), in considerations similar to those of Taam and collaborators, reviewed possible progenitor systems and finds that of known systems only the Be/X-ray binary X Per (4U 0352+30) has the requisite  $P_{\text{orb}} > 1 \text{ yr}$  ( $P = 580 \text{ days}$ ) to survive common envelope evolution. So it is clear that only a small fraction of binaries survive.

Terman & Taam (1996a), in extending non-axi-symmetric common envelope calculations to later times, find that the neutron star can eject the entire common envelope and survive spiral-in if the companion is in the red supergiant phase and if the orbital period is similar to those of the long period Be X-ray binaries, say  $\geq 0.8-2 \text{ yrs}$  (for companions of ZAMS mass  $12-24 M_{\odot}$ ).

Van den Heuvel (1994) estimates that even with the survival of only one of the known systems, the "theoretical rate" for evolution of close neutron star binaries is still 15 to 30 times the "observed" birthrate (of the two known systems PSR 1913+16 and 1534+12). He suggests that  $\sim 99\%$  (rather than  $\sim 90\%$ ) of the Be/X-ray binaries may spiral in to complete coalescence, or that  $\sim 90\%$  of the post-spiral-in systems spiral in very deeply in a second (case BB) phase of mass transfer, during helium shell burning. We shall return to this latter suggestion later.

Brown (1995a) suggested an alternative scenario for binary pulsar formation which avoids the neutron star going through the envelope of the O- or B-star secondary, since he found that the neutron star usually went into a black hole in such a common envelope. Brown's scenario has a double helium star binary as progenitor system. In order to burn helium at the same time, the two O- or B-stars in the original binary must be close in mass, to within  $\sim 4\%$ . As we show below, this probably has the consequence that the resulting two neutron stars in a given binary are very close in mass. The required proximity in mass selects  $\sim 2\%$  of the initial binaries

(Brown, 1995b). In order of magnitude, this is similar to the selection, in the standard scenario of binary pulsar formation, of the X-ray binaries that can avoid spiral-in, and we view the two scenarios as competitive, one with the other. We shall employ the double helium star as progenitor here, since the standard scenario has been discussed in many works, and, for definiteness, our numerical work in evolving the relativistic binary pulsars will be carried out for the double helium star scenario. Since the intermediate neutron-star, helium-star binary is common to both scenarios, and since this is the most important stage in our evolution, our numerical work has direct implications, also, for the standard scenario. Thus, if it turns out that it is to be preferred over our double helium star scenario, our conclusions from the numerical work will still be valid.

As noted by Brown (1995a,b), the double helium star scenario gives a natural explanation for why the neutron star masses within a given neutron star binary are close, but differ substantially more between binaries. In the case of PSR 1913+16, the heavier helium star explodes first, so this scenario also explains why the pulsar, which gains a little more mass from the accretion we describe below, is the heavier star in the binary.

Brown (1995b) has given a partial explanation for the narrow orbits of the relativistic binaries and, more importantly, for their low magnetic fields. He points out that if the helium-star, neutron-star orbit is narrow, then the neutron star will accrete appreciably from the helium-star wind. Following the model of Taam & Van den Heuvel (1986), the accretion lowers the magnetic field  $B$  of the pulsar. The spin-down time of the pulsar is greatly lengthened by this, so that it is observable for 30–50 times longer than an isolated pulsar would be. Taking into account the effects of accretion brings the magnetic field  $B$  down and gives us a weighting towards narrow helium-star, neutron-star orbits, which we introduce into our calculations.

In this paper we use  $a_0$  to label the initial binary separation,  $a_i$  to label the separation at the helium-star, neutron-star stage, and  $a_f$  to label the separation of the components in the final neutron-star binary.

Neutron stars are known to be formed with large kick velocities of  $\sim 350 \text{ km s}^{-1}$  in the explosion (Fomalont et al., 1992; Harrison et al., 1993). Even larger kick velocities have been found by Lyne & Lorimer (1994). Narrow final binaries will also be favored by this, since wide binaries have small binding energies and will be less likely to remain bound following the explosion. We take an assumed distribution  $P(v_k)$  of kick velocity magnitudes and choose equal probability for any direction of the kick velocity, relative to the orbital velocity, in the explosion of the helium star.

Before beginning the explosion calculations, we discuss, in the next section, the possible evolution of a helium star binary.

## 2. Double helium star evolution

Very little calculational work has been carried out in the literature for mass exchange, etc., in the evolution of systems of very nearly equal masses (mass ratio  $q \sim 1$ ). The most complete summary is given by Pols (1993) who began with nearly equal mass progenitors in some cases. We believe the general picture to be reasonably clear and will outline it, hoping that this will stimulate more quantitative work.

Quite a lot of work has been carried out with W-R stars, which are heavier than our helium stars. Doom & DeGreve (1982) estimated that a fraction of 0.5–3% of all massive O-type binaries would become W-R + W-R binaries. Yungelson & Tutukov (1991) arrived at essentially the same results, using different models. The relative rarity of W-R + W-R binaries results, of course, from the requirement that both O-stars should shed their H-envelope at about the same time, which means that their initial masses must be nearly equal. Vrancken et al. (1991) show that it would be difficult to see W-R + W-R binaries, and estimate that fewer than 1% can be observed. Nonetheless, the W-R + W-R binary HD 5980 has been identified (Koenigsberger et al., 1994). It is a very interesting binary, which has changed spectral subtype from WNE to WNL between 1978 and 1990,

and the rate of change seems to be accelerating (Barba & Niemela, 1994). It has brightened considerably and completely changed its spectrum, which is now similar to  $\eta$  Car-type variables, with strong P-Cyg lines of HeI and H Balmer, as well as FeIII and NII lines. Considerable hydrogen, which is difficult to observe at the high temperatures involved, may still be around HD 5980 (Moffat, private communication). This W-R + W-R binary is probably too massive for our purposes, also too wide, with a 19 day orbit. W-R's lose mass through wind at such a rate that it would be difficult for the binary to form a common envelope. If one of the binary components is an LBV, as the above may indicate, the hydrogen envelope loss would be even more rapid. (There is a third star in the neighborhood, and this may be the LBV.)

There are many more light stars than heavy stars, so that there should exist double helium-star binaries. They would be very difficult to see, because – since the helium star radii are small – double-lined eclipsing would require a very special orientation. In the case of binaries with single helium stars, lighter-mass helium stars are now widely considered to be progenitors of type Ib supernovae, as reviewed by Brown (1995a). Given the estimate (Brown, 1995b) that in  $\sim 2\%$  of the binaries the progenitors will be sufficiently close in mass to burn helium at the same time, there would be sufficient double helium star binaries to evolve the relativistic binary pulsars.

The helium stars which are progenitors of PSR 1913+16 were estimated by Burrows & Woosley (1986) to have initial helium core mass of  $7 M_{\odot}$  which blew away, by wind, down to  $\sim 3.13 M_{\odot}$  before exploding, and ending up as neutron stars. Detailed evolution by Woosley et al. (1995) confirms that an initially  $7 M_{\odot}$  helium star ends up at  $3.20 M_{\odot}$  before exploding. Using standard formulae, we find that this corresponds to a ZAMS mass of  $\sim 24 M_{\odot}$ . Note that this is heavier than the progenitor of 1987A, where arguments have been made that the compact object went into a black hole (Brown & Weingartner, 1994). Based on the work by Woosley et al. (1995), Brown, Weingartner & Wijers (1996) show that stars substantially heavier than the

progenitor in 1987A can end up as neutron stars if they evolve in binaries, so there is no contradiction.

Let us begin with two O-stars, both of ZAMS masses close to  $24 M_{\odot}$ . The (slightly) heavier one will evolve first, and transfer mass to the companion in case B mass transfer, leaving a  $7 M_{\odot}$  helium core. Since the initial mass ratio  $q \sim 1$  is large, the transfer will be essentially conservative, with the mass of the secondary reaching  $\sim 40 M_{\odot}$ . The helium core of the secondary will not become appreciably larger than the  $7 M_{\odot}$  that would have evolved from the initial  $24 M_{\odot}$  star, because there will not be time for semiconvection to mix the additional hydrogen across the molecular weight barrier of the original core (Braun & Langer, 1995). In other words, there will not be rejuvenation of the secondary.

As the secondary evolves, the reverse mass transfer involves  $q \leq 1/6$ , so a common envelope should be formed. The two helium cores will then tighten their orbit, transferring energy to the hydrogen envelope by dynamical friction and by torquing the envelope up. In order to get a quantitative idea, we consider the envelope of the  $16 M_{\odot}$  supergiant, in red supergiant phase, of Bodenheimer & Taam (1984) which had a binding energy of  $1.8 \times 10^{48}$  ergs.

In the average helium burning time of  $\sim 5 \times 10^5$  yrs, the initially  $7 M_{\odot}$  helium core of the primary will have blown away down to  $5 M_{\odot}$  (Woosley et al., 1995), whereas that of the secondary will be at its initial mass of  $\sim 7 M_{\odot}$ . Both our total hydrogen core mass and the sum of helium core masses are  $\sim 3$  times those of Bodenheimer & Taam (1984). Gravitational energies of red supergiant envelopes generally go as  $GM^2/R$  ((Applegate, 1988), and private communication) which would mean using 9 times the  $1.8 \times 10^{48}$  ergs of Bodenheimer and Taam for the binding energy of the common envelope, although the radius  $R$  is very uncertain.

With this estimate, we equate

$$\alpha_{\text{CE}} \frac{GM_1 M_2}{a_0} \cong 1.6 \times 10^{49} \text{ ergs} \quad (1)$$

where  $a_0$  is the final separation, and where we neglect the initial binding energy. We take the efficiency  $\alpha_{\text{CE}}$  of the hydrodynamical coupling of the

orbital motion to the envelope to be 0.42 (Taam, 1996). We find

$$a_0 = 3.5 R_{\odot}. \quad (2)$$

The envelope and helium core masses we later need for PSR 1534+12 are only about half as large as those for 1913+16, implying about 1/4 of the envelope binding energy. Not surprisingly, since we have not put in any scaling with envelope radius  $R$ , our answer for  $a_0$  comes out to be about the same as Eq. (2).

Our estimates of  $a_0$  are of the same general size as those needed, at the end of common envelope evolution, for the helium-star, neutron-star separation, so our further numerical evolutionary calculations will apply to the standard scenario. Note that our  $a_0$ 's are comfortably larger than the sum of helium core radii, the radius of a  $7 M_{\odot}$  helium star being  $\sim 0.7 R_{\odot}$  (Van den Heuvel, 1994). Spiral-in at this stage is clearly avoided.

The evolution of PSR 1913+16 is intrinsically simpler than that of 1534+12. During the various burning stages, the radii of the  $7 M_{\odot}$  helium star progenitors do not become larger than  $\leq 1.3 R_{\odot}$  (Van den Heuvel, 1994). We shall cut off the evolution for orbits with eccentricities so high that the radius of the helium star during helium core burning does not fit inside the orbit. We do the same in the evolution of 1534+12. As we discuss, lighter helium stars, such as the progenitors of 1534+12, expand greatly in the supergiant phase, engulfing the orbit. We take no account of this.

For 1534+12, we will employ an initial helium star mass of  $3.5 M_{\odot}$ . From Woosley et al. (1993), this star blows away, by wind, down to  $\sim 2.1 M_{\odot}$  before it explodes. We are now faced with difficulties, because the extensive evolutions by Habets (1986a,b) for the lighter helium stars of mass less than  $4 M_{\odot}$  have been carried out without mass loss. The helium and carbon shell burning stages are particularly relevant, because, as noted, these lighter mass helium stars expand in red supergiant phase. Presumably we should look at Habets' calculation of a helium star closer to the final mass of  $2.1 M_{\odot}$ .

rather than the initial one of  $3.5 M_{\odot}$ , because the helium star mass will be closer to the final mass when the star evolves. Habets (1986b) finds that helium stars with masses of  $2-2.5 M_{\odot}$  expand up to radii  $> 200 R_{\odot}$  as they evolve, before beginning neon burning.

As the initially (slightly) heavier helium star evolves, it will transfer mass, in case BB mass transfer, to the companion helium star. It may explode as a type Ic supernova (Nomoto et al., 1994), leaving a neutron star remnant. The companion will evolve, then, so that the neutron star presumably goes into a common envelope. (Because of the extensive mass loss by wind, this is not, however, obvious.)

The outcome of this neutron-star, helium-star common envelope evolution is unclear. Presumably, under special circumstances analogous to those suggested earlier for the neutron star evolution in the hydrogen envelope, in some fraction of the cases the neutron star will survive spiral-in and will not go into a black hole. The total mass in the helium envelope is an order of magnitude less than that of the hydrogen envelope, so it is less likely that the neutron star goes into a black hole. However, because of the earlier common envelope evolution, of one type or other, the neutron star and helium star are spatially quite close to begin with, so spiral-in may be difficult to avoid. Van den Heuvel (1994) suggests that in only  $\sim 10\%$  of the cases spiral-in might be avoided. Common envelope numerical calculations have not been carried out for this case, so we do not know.

Note, however, that, because of mass transfer, the first helium star to explode will be the lighter one, giving the lighter neutron star. However, this neutron star can accrete mass both from the wind of the remaining helium star and, later, in the common envelope evolution. Thus, there is at least a chance of ending up with two such nearly equal masses as  $1.337 M_{\odot}$  and  $1.341 M_{\odot}$  as found in 1534+12. (As noted earlier, with double helium star progenitor, the first explosion will give the heavier neutron star, if there is no case BB mass transfer. This would be the case in 1913+16 where the pulsar is the heavier of the two neutron stars.)

That many of the neutron stars from the lighter mass helium stars do spiral in is suggested observationally by the fact that 1534+12 is the only close binary from the lighter mass progenitors. Given PSR 1913+16, with O-star progenitors of main sequence masses  $\sim 24 M_{\odot}$ , there should be at least an order of magnitude more neutron star binaries from the much more frequent lighter mass progenitors. It is, of course, dangerous to use statistics with a number of one. (PSR 2303+46 and possibly the newly discovered 1518+49 would also come from low mass helium stars.)

### 3. The evolution of the pulsar magnetic field

The pulsar magnetic field behavior is crucial for our calculations. Van den Heuvel (1994) has pointed out that single neutron stars, born with strong magnetic fields of  $10^{12}-5 \times 10^{12}$  G, spin down in a time

$$\tau_{sd} \sim 5 \times 10^6 \text{ yrs} \quad (3)$$

and then disappear into the graveyard of neutron stars. (The pulsation mechanism requires a minimum voltage  $V$  above the polar cap. This critical  $V$  can be attained if  $B_{12}/P^2 \geq 0.2$ , where  $B_{12} = B/10^{12}$  G and the period  $P$  is in seconds.) The relativistic binary pulsars PSR 1913+16 and 1534+12 have, however, weaker magnetic fields,  $B \sim 10^{10}$  G, and, therefore, emit less energy in magnetic dipole radiation. Van den Heuvel (1994) estimates their spin-down times to be

$$\begin{aligned} \tau_{sd} &= 10^8 \text{ yrs} && \text{PSR 1913 + 16} \\ \tau_{sd} &= 2.5 \times 10^8 \text{ yrs} && \text{PSR 1534 + 12} \\ \tau_{sd} &= 3 \times 10^7 \text{ yrs} && \text{PSR 2303 + 46}. \end{aligned} \quad (4)$$

We have included 2303+46. Its 1s pulse period and surface magnetic field,  $B \sim 8 \times 10^{11}$  G, suggest that it was never recycled and is probably the result of the second supernova in the system (Arzoumanian, 1995).

From Eq. (4) we note that the spin-down times of the relativistic pulsars, where the neutron stars are close together, are much longer than those of the

wide binary 2303+46 and of single neutron stars. This should give us a hint that the decrease in magnetic field, with accompanying lengthening of the spin-down time, is related to accretion. Indeed, Taam & Van den Heuvel (1986) have shown that empirically field decay is inversely correlated with mass accretion, although there is up to now no fundamental basis for this correlation. The empirical relationship between field decay and accretion has been modeled by Shibasaki et al. (1989)<sup>2</sup> by

$$B = \frac{B_0}{1 + \dot{M}t_{\text{acc}}/m_B}, \quad (5)$$

where  $t_{\text{acc}}$  is the time of accretion and  $m_B$  is a constant. For accretion at the Eddington limit, Brown (1995a,b) showed that this could be expressed as

$$B \cong 5 \times 10^{15} \text{ G yrs}/t_{\text{acc}} \quad (6)$$

for the times of accretion relevant here, where  $B_0$  was  $10^{12}$  G– $5 \times 10^{12}$  G. The form Eq. (6) fits the fields of the relativistic binary pulsars and of the millisecond pulsars. In fact, that this fits is easily seen. (It should not be used for shorter  $t_{\text{acc}}$ .) With an average helium burning time of  $5 \times 10^3$  yrs for the time of accretion for the relativistic binary pulsars (in our scenario of double helium star progenitor),  $B \cong 10^{10}$  G. For the millisecond pulsars, a time of  $10^7$  yrs is more appropriate (Cook et al., 1994), and this gives  $B \cong 5 \times 10^8$  G, again in the correct range. For accretion below the Eddington limit and the accretion times we use, this can be simplified to:

$$B \cong \frac{5 \times 10^{15} \text{ G yrs}}{\dot{M}t_{\text{acc}}/\dot{M}_{\text{Edd}}} \quad (7)$$

with  $\dot{M}_{\text{Edd}}$  the limit for hydrogen. This is the formula we shall use.

Accretion onto a rapidly rotating newly formed neutron star with strong magnetic field cannot immediately proceed because of the propeller effect (Illarionov & Sunyaev, 1975). However, a strong spin-down torque is exerted on the pulsar-plus

magnetosphere by the mass ejection. (See also Davidson & Ostriker (1973).)

#### 4. Binary pulsar evolution

Many of the evolutionary aspects we wish to discuss were considered by Hills (1983) and Wijers et al. (1992). [See also Yamaoka et al. (1993) and Tutukov (1988).] In addition to developing our double helium-star progenitor, we shall focus on the effects of accretion onto the neutron star from the helium star wind in the neutron-star, He-star binary phase which is common to both the standard scenario and to ours.

We evolve the binary pulsars through the three stages:

- o*) He-star + He-star binary
- i*) neutron-star + He-star binary
- f*) neutron-star + neutron-star binary.

We shall use the suffices *o* (for original), *i* (for intermediate), and *f* (for final) in labeling quantities applying to the particular stage. Effects of mass loss by wind from the He-star on the orbits will be taken into account. The fundamental role played by accretion (in stage *i*) is clear; we shall discuss it.

We believe the intermediate stage *i* to be most important in the evolution of the relativistic binary pulsars PSR 1913+16 and PSR 1534+12. The former has a relative semi-major axis of  $2.8 R_\odot$  and a period of 7.75 hr and the latter, a 10 hr period and similar semi-major axis. The only other neutron star binaries evolved as such are PSR 2303+46 with a 12d34 orbit and the newly discovered PSR J1518+4904 with an 8.5 day period. As we noted earlier, in 2127+11 in M15, the one neutron star probably picked up the other one later. In PSR 1820-11, a binary pulsar with a period of  $\sim 1$  yr, the companion may be a neutron star. We consider the interesting result to be that in two of the four neutron-star, neutron-star binaries observed to date, the two neutron stars are very close together. Those are the two relativistic binary pulsars.

Substantial accretion onto the neutron star can occur only in the He-star, neutron-star phase. Helium

<sup>2</sup>Shibasaki et al. show that the same amount of accretion that is necessary to bring the pulsar magnetic field down is sufficient to spin it up to the observed relatively short period.

stars are known to have high winds, with mass loss (Woosley et al., 1993)

$$\dot{M}_{\text{He}} \cong 5 \times 10^{-8} M_{\odot} \left( \frac{M_{\text{He}}}{M_{\odot}} \right)^{2.6} \text{ yr}^{-1}, \quad (8)$$

although the extrapolation of this  $\dot{M}$  to low mass He-stars is just an assumption. We begin by modeling PSR 1913+16, starting from a helium star mass of  $7 M_{\odot}$ . For such a relatively massive helium star, Eq. (8) for mass loss, which is determined from Wolf-Rayet stars, may be reasonable; it does not involve a large extrapolation downwards in mass. From Iben & Tutukov (1985), the relation between main sequence mass and helium core mass is

$$M_{\text{He}} \cong 0.08 M_{\text{ZAMS}}^{1.4} \quad (9)$$

so that a  $7 M_{\odot}$  helium star would correspond to a main sequence mass of  $\sim 24 M_{\odot}$ . Iron cores in the evolution of helium stars with mass loss (Woosley et al., 1993) tend to be substantially smaller than those developed without mass loss. Our helium stars, of course, have lost their hydrogen envelopes. The final helium star mass before explosion is found by Woosley et al. (1995) to be  $M_{\text{ex}} = 3.2 M_{\odot}$ . Note that the initial and final helium star masses are essentially those of Burrows & Woosley (1986).

The average helium-star mass for 1913+16 can be taken to be

$$\langle M_{\text{He}} \rangle \cong 5 M_{\odot}. \quad (10)$$

This gives

$$\begin{aligned} \dot{M}_{\text{He}} &= 3.3 \times 10^{-6} M_{\odot} \text{ yrs}^{-1} \\ &= 2.1 \times 10^{20} \text{ g s}^{-1}. \end{aligned} \quad (11)$$

The standard expression for Bondi–Hoyle–Lyttleton accretion theory onto an object of mass  $M$  is

$$\dot{M} = 2.2 \times 10^{29} \left( \frac{M}{M_{\odot}} \right)^2 \left( \frac{V_0}{V} \right)^3 \rho_{\infty} \text{ g s}^{-1}, \quad (12)$$

where  $\rho_{\infty}$  is the density in the neighborhood of the accreting object in cgs-units and  $V$  is the wind velocity. We employ

$$V_0 = 1000 \text{ km s}^{-1} \quad (13)$$

and take the mass  $M = 1.4 M_{\odot}$ . Using

$$\rho_{\infty} = \frac{\dot{M}_{\text{He}}}{4\pi a_i^2 V} = \frac{3.4 \times 10^{-11} \text{ g cm}^{-3}}{(a_i/R_{\odot})^2}, \quad (14)$$

we find

$$\dot{M} = \frac{1.5 \times 10^{19} \text{ g s}^{-1}}{(a_i/R_{\odot})^2} \quad (15)$$

for an assumed wind velocity of  $V = 1000 \text{ km s}^{-1}$ . For ionized helium, the Eddington limit is

$$\dot{M}_{\text{Edd}}(\text{He}) = 1.9 \times 10^{18} \text{ g s}^{-1}, \quad (16)$$

twice that of  $\dot{M}_{\text{Edd}}$  for hydrogen used in Eq. (7). We see that  $\dot{M}$  will exceed  $\dot{M}_{\text{Edd}}(\text{He})$  only for  $a_i \leq 2.8 R_{\odot}$ .

From Eq. (7) we find

$$B = 0.065 \times 10^{10} \left( \frac{a_i}{R_{\odot}} \right)^2 G, \quad (17)$$

where we have used the average helium burning time of  $5 \times 10^5$  yrs for a  $7 M_{\odot}$  helium core from Woosley et al. (1995). From Eqs. (3), (4) we see that 1913+16 is observable  $\sim 20$  times longer because its magnetic field has been decreased (by accretion). Therefore, in our Monte Carlo calculations, we weight the orbits in the helium-star, neutron-star stage  $i$  by the factor

$$F_i = \begin{cases} 20(6.2 R_{\odot}/a_i)^2 & \text{for } 2.8 R_{\odot} < a_i < 28 R_{\odot} \\ 1 & \text{for } a_i > 28 R_{\odot} \\ 98 & \text{for } a_i < 2.8 R_{\odot}, \end{cases} \quad (18)$$

the final factor saying that accretion at a rate higher than  $\dot{M}_{\text{Edd}}$  is treated as  $\dot{M}_{\text{Edd}}$ . Here,  $F_i$  is chosen to be 20 for  $B = 2.5 \times 10^{10} \text{ G}$ , the field of 1913+16, and to vary as  $B^{-1}$  since spin-down times vary inversely with  $B$ . (The overall pre-factor will be fixed by normalization.)

We shall also show results for PSR 1534+12 extrapolating the wind Eq. (8) down to the low masses  $M_{\text{in}} = 3.5 M_{\odot}$ ,  $\langle M_{\text{He}} \rangle \cong 2.5 M_{\odot}$ , and  $M_{\text{ex}} = 2.1 M_{\odot}$ , with

$$F_i = \begin{cases} 50(2R_\odot/a_i)^2 & \text{for } R_\odot < a_i < \sqrt{200}R_\odot \\ 1 & \text{for } a_i > \sqrt{200}R_\odot \\ 200 & \text{for } a_i < R_\odot. \end{cases} \quad (19)$$

We use an average helium-burning time of  $8 \times 10^5$  yrs (Habets, 1986b) for 1534+12. The extrapolation of the wind loss Eq. (8) to such low-mass helium stars is questionable. There is no observational guidance. Dimensional analysis (Langer, 1989) with normalization of the efficiency factor  $\alpha$  to  $15 M_\odot$  W-R stars would give roughly the wind loss Eq. (8), which is severely cut down for low mass He stars by the factor  $(M_{\text{He}}/M_\odot)^{2.6}$ . (Langer carries out the extrapolation to low-mass He stars.)

The common envelope during helium shell burning does not occur in the relatively massive  $7 M_\odot$  He stars used to evolve 1913+16. Because of wind loss, the helium mass at the time of explosion is only  $3.2 M_\odot$ , and the most probable  $a_i$  has drifted out  $\sim 0.5 R_\odot$  from the  $\sim 3.5 R_\odot$  at the time of  $M_{\text{He}} = 5 M_\odot$ , the average mass used in calculating the wind. This means that the accretion onto the neutron star at the time of explosion is only  $\sim 1/4$  of the time-averaged wind, which has been taken to be  $\dot{M}_{\text{Edd}}(\text{H})$ .

## 5. Results of the binary pulsar evolution

A detailed treatment of our explosion calculations is given in the appendix. We rely in part on the works of Hills (1983) and Wijers et al. (1992), but adapt them to our problem with double helium star progenitors. Details of how mass loss by wind, etc., is handled are also discussed in the appendix.

In our figures, we display results for the distribution of orbital parameters  $a$  and  $e$  at various stages of the evolution for both PSR 1913+16 and 1534+12. The parameters going into these calculations are, apart from various masses, the semi-major axis  $a_0$  of the initial He-star, He-star binary (reached by the common envelope evolution described in Section 2) and the parameter  $v_p$  characterizing the magnitude of the kick velocity. We take the initial eccentricity  $e_0 = 0$ , assuming the original orbit to

have been circularized by the common envelope evolution. The weighting applied at the intermediate stage corresponds to Eq. (18) for 1913+16 and to Eq. (19) for 1534+12, respectively.

The parameter  $v_p$  [and also the shape of  $P(v_k)$ ] is rather uncertain. Wijers et al. (1992) find that a wide range of  $v_p$  is acceptable. We consider two cases,  $v_p = 100 \text{ km s}^{-1}$  and  $v_p = 310 \text{ km s}^{-1}$  (the latter parameter corresponding to  $v_k = 350 \text{ km s}^{-1}$ ). It should not be surprising that our results depend on the choice of  $v_p$ , and we will discuss this effect below.

In Fig. 1, we display results for PSR 1913+16 using an initial separation of  $a_0 = 3 R_\odot$  and  $v_p = 310 \text{ km s}^{-1}$ . The masses appropriate for 1913+16 are  $7.0 M_\odot$  for the initial mass of a helium star,  $3.2 M_\odot$  for the mass at the time of explosion, and  $1.8 M_\odot$  for the mass loss in the explosion. The long-dashed line gives the distribution of helium-star, neutron-star binary separations immediately after the first explosion. As can be seen from the dashed-dotted curve, the binary is somewhat widened by the mass loss due to winds, as the stars drift apart. The short-dashed curve with highest peak builds in the weighting of Eq. (18), our "observability premium", to account for the longer visibility, once pulsar magnetic fields have been brought down.

Clearly our "observability premium" in the weighting Eq. (18) causes a large concentration of  $a_i$  in the region  $a_i \sim 3 R_\odot$  and more than compensates the drift apart because of wind. The distribution of  $a_f$  in the neighborhood of  $3 R_\odot$  is increased by our weighting, and there is little probability for large  $a_f$ .  $P_1$  gives the survival probability for the first explosion,  $P_2$  that for the survival of both explosions without taking the weighting factor Eq. (18) into account,  $P'_2$  the probability taking it into account. Roughly speaking, retrograde explosions, in which the kick velocity is in the backward hemisphere relative to the velocity of the orbit, survive. This results because kick velocities are roughly the same magnitude as orbital velocities in the binary pulsars.

Also shown in Fig. 1 are the distributions of intermediate and final eccentricities for the same set of input parameters.  $P(e)$  corresponds to an integration of  $P(a,e)$  over all  $a$ , whereas  $P_a(e)$  represents

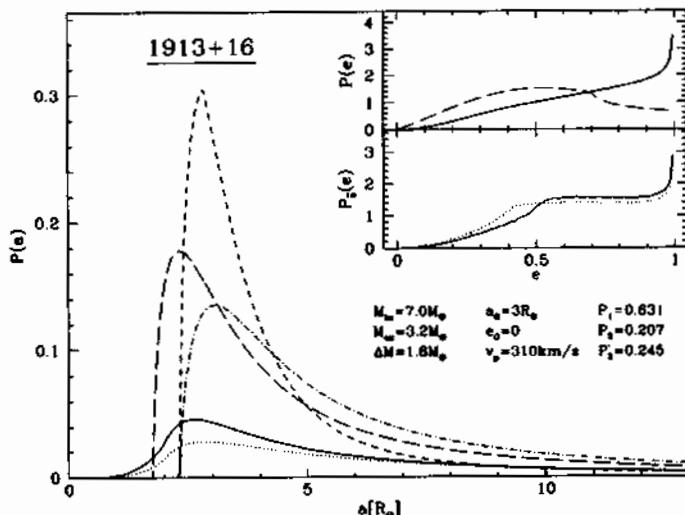


Fig. 1. Probability distribution of semi-major axes and eccentricities for PSR 1913+16 in various stages of the evolution. The input parameters are  $a_0 = 3R_\odot$ ,  $e_0 = 0$ , and  $v_p = 310 \text{ km s}^{-1}$ . Plotted are the distributions immediately after the first explosion (long dashes), prior to the second explosion (dot-short dash without weighting, short dashes with weighting), and after the second explosion (dotted line without weighting, solid line with weighting). The curves for  $P(e)$  correspond to an integration of  $P(a,e)$  over all  $a$ , whereas  $P_a(e)$  represents  $P(a = \bar{a}, e)$ , with  $\bar{a}$  being the value of  $a$  for which  $P(a)$  peaks. Since in our scenario the eccentricity is not affected by wind in the intermediate stage, we plot only one curve for  $P(e)$  in this stage. Also given are the survival probabilities after the first explosion ( $P_1$ ) and after two explosions ( $P_2$  without weighting,  $P'_2$  with weighting). The weighting of the intermediate orbits corresponds to Eq. (18).

$P(a = \bar{a}, e)$ , with  $\bar{a}$  being the value of  $a$  for which  $P(a)$  peaks. The curves for  $e$  are unexceptional, and relatively unaffected by our weightings. Note, however, the effect of our second survival criterion, Eq. (A.16), which suppresses large eccentricities in the intermediate stage. In the final stage, large eccentricities are preferred. PSR 1913+16 has an eccentricity of  $e = 0.617$ , which is supported by our findings. As in Wijers et al. (1992), intermediate orbits in which the pericenter would be inside the helium star are excluded.

In Fig. 2 we show similar results for PSR 1913+16 using an initial  $a_0$  of  $5R_\odot$  and the same value of  $v_p$ . The above remarks still apply qualitatively, the main difference being a shift of preferred final semi-major axes to somewhat larger values. The peaking is also less sharp. Clearly, an initial  $a_0$  closer to  $3R_\odot$  is favored for obtaining the observed orbits of the relativistic binaries, and  $a_0 = 3.5R_\odot$  of Eq. (2) will do fine.

In Fig. 3 we show results for PSR 1534+12 using  $a_0 = 3R_\odot$  and  $v_p = 310 \text{ km s}^{-1}$ . The initial masses are  $3.5M_\odot$ , the mass at the time of explosion is  $2.1M_\odot$ , and the mass loss in the explosion  $0.76M_\odot$ . The weighting applied at the intermediate stage is that of Eq. (19). The results are qualitatively similar to those of Fig. 1. However, 1534+12 has  $e_f = 0.274$  which is not favored by our results. Therefore, we repeat the calculation for a smaller value of the average kick velocity, using  $v_p = 100 \text{ km s}^{-1}$ . The results are displayed in Fig. 4. The resulting distributions of the final eccentricity are qualitatively altered by this smaller value of  $v_p$ , and we observe that the measured value of PSR 1534+12 is now well within the support of  $P(e_f)$ .

From the above results, it is clear that the common envelope evolution must bring the original  $a_0$  down to  $\sim 3-5R_\odot$ . From our estimates, sufficient energy is provided from the tightening of the double helium star orbit to expel the hydrogen envelope. The

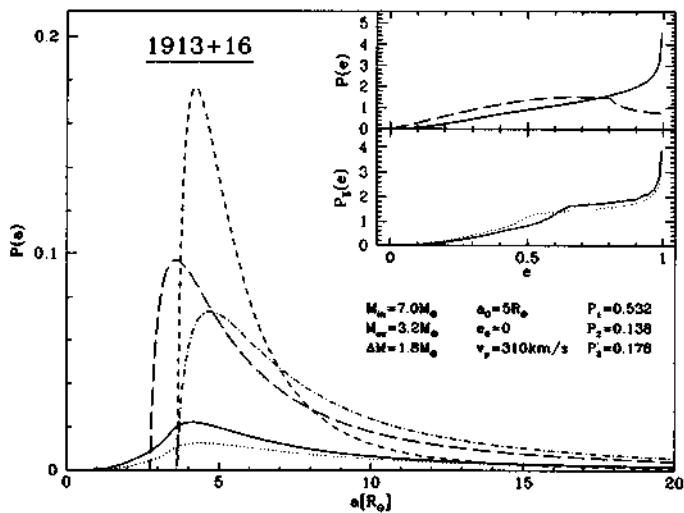


Fig. 2. Same as Fig. 1 but for initial parameters  $a_0 = 5 R_\odot$  and  $e_0 = 0$ .

accreted mass, which cannot cross the molecular weight barrier, will simply be expelled.

Let us remark at this stage that the Sutantyo (1975, 1992) scenario for the evolution of Her X-1 also avoids a neutron star going through the common

envelope of a large star. In this case an  $\sim 2 M_\odot$  main sequence star goes into the large O-star envelope, expelling it, leaving a He-star and the  $2 M_\odot$  one. In this sense, our scenario is not novel.

Our weighting factor, the “observability pre-

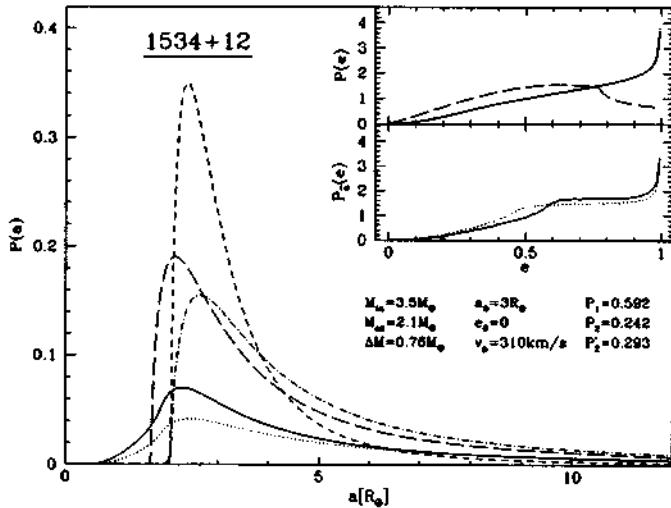


Fig. 3. Same as Fig. 1 but for masses appropriate for PSR 1534+12. The weighting is that of Eq. (19).

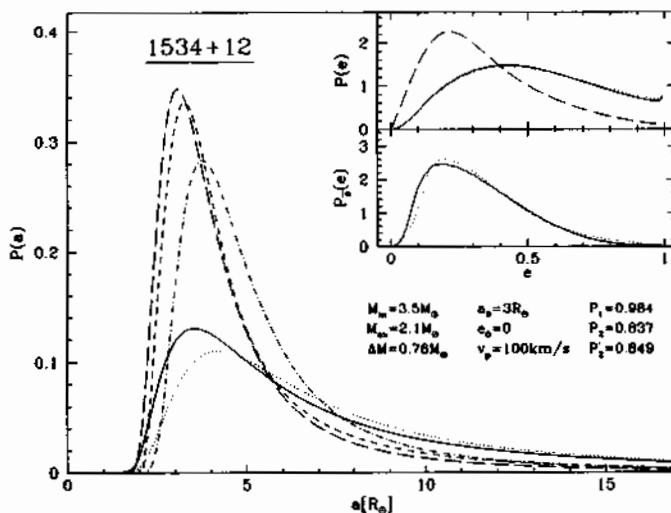


Fig. 4. Same as Fig. 3 but using  $v_p = 100 \text{ km s}^{-1}$ .

mium”, then somewhat sharpens the distribution of  $a_i$ , but chiefly incorporates the longer time of observability if the magnetic field is brought down by accretion. As noted earlier, this same accretion is needed to spin the pulsar up.

The probability of the helium star, neutron star binary surviving the last explosion is larger with inclusion of our “observability premium”, which favors small  $a_i$ , so that the gravitational field is large and, therefore, effective in holding the binary together. Since two relativistic binary pulsars are seen, single spun-up pulsars from breakup in the last explosion, with roughly the same period as the former, should exist with roughly equal probability. In fact, recently (Camilo et al., 1993) PSR J2235+1506, with a period of 59.7 ms has been found. The estimated magnetic field is also low,  $B_0 \leq 4 \times 10^9 \text{ G}$ , somewhat lower than that of 1913+16 and 1534+12.

Although our requirement of proximity in mass selects only  $\sim 2\%$  of the initial binaries in our double helium star scenario (Brown, 1995b), it is straightforward to show, by scaling to single pulsars, that our large observability premiums, Eqs. (18), (19), make it possible to produce enough relativistic

binary pulsars in our scenario to satisfy observational constraints.

Our scenario for the formation of neutron star binaries is represented in a schematic way in the accompanying movie (see Fig. 5 of the electronic version of this article, <http://www.elsevier.nl/locate/newast>).

## 6. Discussion

Because of the small number of relativistic binary pulsars, we have not carried out extensive statistical analyses. Our results should be evident from the figures here.

First of all, it is clear that the evolution of the two helium stars within the common hydrogen envelope must tighten the orbit to a separation of  $\sim 3-5 R_\odot$ . Because of the magnitude of the helium star masses, this will take place if the hydrogen common envelope of the two helium stars has binding energy  $\sim 1-2 \times 10^{49} \text{ ergs}$ .

Given the initial  $a_0$  in this range, the distribution of  $a_i$  in our evolution peaks up in the neighborhood of  $3 R_\odot$  where, according to Eq. (17), the pulsar

magnetic field will be brought down to  $B \sim 10^{10}$  G by the accretion during this helium-star, neutron-star stage. Note that this stage is shared by the standard scenario for evolution of binary pulsars. Therefore, our mechanism for bringing down the pulsar magnetic field would work there, too. As noted earlier, our scenario in which accretion onto the pulsar from the helium star is taken into account suggests an explanation for the Kulkarni (1992) "gap" in the distribution of pulsar magnetic field strengths. This "gap" is between the  $\geq 10^{10}$  G fields of the relativistic binary pulsars and those of the millisecond pulsars, which are below  $\sim 5 \times 10^8$  G. The average burning time of the helium star is  $\sim 5 \times 10^5$  yrs, which in the Taam & Van den Heuvel (1986) mechanism of field decay being correlated with mass accretion, is only long enough to bring the field of the pulsar in the relativistic binary down to  $\sim 10^{10}$  G, according to Eq. (6). (In our simulations of the relativistic binary pulsar evolution, the accretion in the neutron-star, helium-star stage is almost always near the Eddington limit.) On the other hand, in the standard scenario for speeding up the millisecond pulsars (Van den Heuvel, 1994), accretion continues for  $\geq 10^7$  yrs.

From our rough estimates, our scenario of double helium-star progenitors for the relativistic binary pulsars, and possibly also for 2303+46, gives about the right number for PSR 1913+16, which results from helium stars which are sufficiently heavy that there will not be a second common envelope stage of reverse case BB mass transfer. We would produce an order of magnitude too many light-mass neutron star binaries such as PSR 1534+12, and we suggest that most of these may disappear in spiral-in during the second common envelope stage.

As noted earlier, Terman & Taam (1996b) have been able to evolve 1913+16 so that the pulsar is the heavier of the two neutron stars, as observed. In their evolution, the primary and secondary do not interact before the supernova explosion of the primary. Thus, the secondary does not accrete mass, and it is unlikely that the progenitor system would be identified as a Be X-ray binary system. A nonconservative mass transfer scenario is discussed for PSR

1534+12, which is able to obtain the nearly equal neutron star masses. The survival of the binary neutron star systems requires that the progenitors must have, at one time, been long period ( $\sim 2-3$  yrs) binaries.

Chevalier (1995) has added effects of rotation to his 1993 scenario. His conclusion is that rotation does allow spiral-in of some neutron stars into stellar envelopes without their becoming black holes.

In the common envelope scenario of binary pulsar evolution, as modified by Taam, Terman and others, it does seem possible, by having the progenitor systems far enough apart that the common envelope evolution begins when the companion star becomes a supergiant, that some of the binary systems survive. Accretion onto the neutron star is included only schematically in these scenarios. In order to accrete the  $\sim 0.01 M_\odot$  necessary to spin up the neutron star (Terman & Taam, 1996b), accretion probably must occur at hypercritical rates, since the supergiant stage is short. It is then not clear why the accretion should stop with  $0.01 M_\odot$ , and the pulsar not be spun up more. Our view is that this scenario is possible, but that the description is incomplete without a detailed treatment of the accretion. We offer the scenario of double helium star progenitors as a possible alternative. Nature may well choose the Terman & Taam (1996b) track for 1913+16 and our track for 1534+12, where the masses of the two neutron stars are so very nearly equal.

#### Acknowledgments

We would like to thank Joe Weingartner, who began the binary pulsar evolution for us. We are grateful to Ralph Wijers for initial advice on how to go about the evolution and for useful criticism of the manuscript. Ron Taam and Jim Terman communicated prepublication results to us, and gave us helpful criticism. We are grateful to Peter Conti and Tony Moffat for instructions on W-R's. Hans Bethe gave many helpful suggestions to improve the manuscript.

## Appendix A

This appendix serves to discuss the more technical details of the explosion calculations and will also describe how mass loss due to wind evolution is taken into account. The former topic has been discussed in great detail by Hills (1983) and Wijers et al. (1992), and we will make use of their results wherever possible.

We are interested in the probability that a binary survives a supernova explosion of one of its components and in the probability distribution of the orbital parameters after the explosion. One has to decide whether it is possible to compute these quantities analytically or whether it is more economical to employ Monte Carlo methods. For one explosion, the former approach is possible in principle but leads to rather complicated integrations which have to be performed numerically (see Wijers et al., 1992). We are dealing with an additional explosion in which the orbital parameters prior to the explosion are not fixed but distributed according to the outcome of the first explosion (and modified by wind evolution and our additional weightings of narrow intermediate orbits). This gives rise to two additional integrations which can only be performed numerically, not to mention the complications arising from various special cases one has to consider. Hence, a complete analytical treatment of the problem appears to be highly inconvenient. Therefore, we will resort to Monte Carlo methods to obtain both survival probabilities and probability distributions of the orbital parameters after one and two explosions.

Let us now consider a single supernova explosion. We will use the index 0 to label orbital parameters prior to the explosion and omit the index when referring to parameters after the explosion. In a supernova explosion, a random kick velocity  $v_k$  is generated and imparted on the resulting neutron star. Following Hills (1983), we represent  $v_k$  in spherical coordinates  $(v_k, \theta, \phi)$ , where  $\theta$  denotes the angle between  $v_k$  and the orbital velocity  $v_0$  prior to the explosion. We choose the direction of the kick velocity to be random in direction, i.e.,  $P(\phi) = 1/2\pi$  and  $P(\theta) = (\sin\theta)/2$ . (Here and in the following,

functions  $P$  will always denote probability distributions of their arguments. Phase space factors are absorbed in the  $P$ 's.) We follow Wijers et al. (1992) in taking the distribution of kick velocity magnitudes to be  $P(v_k) \propto v_k^2 \exp(-v_k^2/v_p^2)$ . The parameter  $v_p$  is related to the average kick velocity by  $\bar{v}_k = 2v_p/\sqrt{\pi}$ .

Another element entering the determination of final orbital parameters is  $r$ , the separation of the two binary components. Unless we are dealing with a circular orbit (for which  $e_0 = 0$  and  $r = a_0$ ), we need the distribution of  $r$  which can be obtained from the equations of the ellipse,

$$r(\psi) = a_0(1 - e_0 \cos\psi) \quad (A.1)$$

$$t(\psi) = T(\psi - e_0 \sin\psi)/2\pi, \quad (A.2)$$

where  $T$  denotes the period. For our purposes, it is more convenient to use  $\psi$  as the independent parameter. It suffices to consider only one half of a period, i.e., the parameter  $\psi$  takes on values between 0 and  $\pi$ . Using  $P(t) = 2/T$ , we obtain for the distribution of  $\psi$

$$P(\psi) = \frac{1}{\pi}(1 - e_0 \cos\psi). \quad (A.3)$$

The semi-major axis after the explosion is determined by the change in total energy and given by (see Hills, 1983)

$$a = a_0 \frac{1 - \Delta M/M}{1 - (2a_0/r)(\Delta M/M) - (v^2 - v_0^2)/v_c^2}. \quad (A.4)$$

Here,  $M$  is the total mass of the binary before the explosion,  $\Delta M$  is the mass loss in the explosion,  $r$  is the separation of the two components at the time of the explosion,  $v_0$  and  $v$  are the orbital velocities before and after the explosion, and  $v_c = (GM/a_0)^{1/2}$ . Note that  $v_0$  and  $r$  are related by

$$v_0/v_c = [(2a_0/r) - 1]^{1/2}. \quad (A.5)$$

Since  $v = v_0 + v_k$ , we have

$$v^2 - v_0^2 = v_k^2 + 2v_k v_0 \cos\theta. \quad (A.6)$$

Eq. (A.4) gives rise to a survival condition which we will consider shortly. Let us first complete the determination of final orbital parameters by discussing how the new eccentricity is computed. In general, the eccentricity of the orbit is related to the angular momentum  $L$  by

$$1 - e^2 = \frac{L^2}{a\mu k}, \quad (\text{A.7})$$

where  $\mu = M_1 M_2 / (M_1 + M_2)$  is the reduced mass and  $k = GM_1 M_2$ . Using  $L = \mu \mathbf{r} \times \mathbf{v}$  we obtain

$$1 - e^2 = \frac{r^2 v_{\perp r}^2}{G(M_1 + M_2)a}. \quad (\text{A.8})$$

We have to determine  $v_{\perp r}$ , the component of the orbital velocity after the explosion perpendicular to  $\mathbf{r}$ . Let the original orbit lie in the  $xy$ -plane, centered at the origin. The  $z$ -component of the new orbital velocity is then just the  $z$ -component of the kick velocity  $\mathbf{v}_k$ , which in our representation is

$$v_z = v_k \sin\theta \sin\phi. \quad (\text{A.9})$$

This component is, of course, perpendicular to  $\mathbf{r}$ . It remains to determine  $v_{\perp xy}$ , the component of the orbital velocity in the  $xy$ -plane perpendicular to  $\mathbf{r}$ . We decompose the orbital velocity in the  $xy$ -plane into components parallel and normal to the original orbital velocity  $\mathbf{v}_0$ :

$$\mathbf{v}_{\parallel} = \mathbf{v}_0 + \mathbf{v}_k \cos\theta \quad (\text{A.10})$$

$$\mathbf{v}_{\perp} = \mathbf{v}_k \sin\theta \cos\phi. \quad (\text{A.11})$$

If  $\xi$  denotes the angle between  $\mathbf{r}$  and  $\mathbf{v}_{\parallel}$ , then  $v_{\perp xy} = v_{\parallel} \sin\xi + v_{\perp} \cos\xi$  so that

$$\begin{aligned} e_i &= [1 - \frac{r^2}{G(M - \Delta M)a} \{ (v_k \sin\theta \sin\phi)^2 \\ &\quad + [(v_0 + v_k \cos\theta) \sin\xi + v_k \sin\theta \cos\phi \cos\xi]^2 \}]^{1/2}. \end{aligned} \quad (\text{A.12})$$

The only information we still need is how  $\xi$  depends on  $r$ . Assuming that the orbit is traversed in the counter-clockwise direction and that we are only interested in that half of the orbit for which  $y \geq 0$ ,  $\xi$

is just the sum of the angles  $\alpha$  and  $\beta$  that  $\mathbf{r}$  and  $\mathbf{v}_{\parallel}$  make with the  $x$ -axis ( $0 \leq \alpha \leq \pi$ ,  $\pi/2 \geq \beta \geq -\pi/2$ , and  $0 \leq \xi_{\min} \leq \xi \leq \pi/2$ ). These angles are determined from geometry,

$$\tan\alpha = \frac{y}{x + a_0 e_0}, \quad (\text{A.13})$$

$$\tan\beta = -\frac{dy}{dx} = (1 - e_0^2) \frac{x}{y}. \quad (\text{A.14})$$

We now express  $x$  and  $y$  in terms of  $r$  and use  $\xi = \alpha + \beta$  to obtain after some algebra

$$\tan\xi(r) = \frac{\sqrt{1 - e_0^2}}{e_0 \sin\psi}. \quad (\text{A.15})$$

Eqs. (A.1), (A.12), (A.15) determine the eccentricity after the explosion.

Let us now return to the conditions for survival. The first condition follows from the requirement that the denominator in Eq. (A.4) be positive. For the first explosion, however, an additional criterion has to be taken into account. The result of the first explosion is a helium-star, neutron-star binary, and we require that the orbital parameters after the first explosion be such that the helium star does not overlap the resulting orbit. This translates to the condition that the pericenter of the orbit must be outside the helium star, i.e., the requirement that

$$a_i(1 - e_i) > R_{\text{He}}. \quad (\text{A.16})$$

This leads to a cut-off of large eccentricities after the first explosion as will be seen in our figures. The value of  $R_{\text{He}}$  is obtained from the relationship given by Wijers et al. (1992),

$$\frac{R_{\text{He}}}{R_{\odot}} = 0.22 \left( \frac{M_{\text{He}}}{M_{\odot}} \right)^{0.6}. \quad (\text{A.17})$$

The survival probability can be obtained analytically by integrating over all those kick velocities (magnitude and direction) that fulfill the two survival criteria. (For the second explosion, only the first criterion is needed due to the small size of neutron stars.) Unless we are dealing with a circular orbit (for which  $e_0 = 0$  and  $r = a_0$ ) we also have to

integrate over all possible values of  $r$  along the ellipse. The distribution of orbital parameters  $a$  and  $e$  after the explosion can also be computed analytically by inserting the appropriate  $\delta$ -functions in the integrals (again, see Wijers et al., 1992). However, the resulting expressions are rather cumbersome and can only be evaluated numerically.

In a Monte Carlo simulation, we avoid these difficulties by simply sampling the required results. Our method involves the following steps: (i) the independent variables ( $\psi, v, \theta, \phi$ , and, for the second explosion, also  $a_i$  and  $e_i$ ) are drawn from their respective distributions, (ii) the survival criteria are inspected and statistics on the survival probability are accumulated, (iii) in case of survival the new  $a$  and  $e$  are computed and their joint probability distribution is sampled. In the intermediate stage,  $a_i$  and  $e_i$  have to be drawn from the joint distribution obtained after the first explosion. It is precisely at this stage where our idea is incorporated: The distribution  $P(a_i, e_i)$  obtains an additional weighting towards smaller  $a_i$  according to the prescription of Eq. (18) or Eq. (19).

Plotted in our figures are the distributions  $P(a)$  and  $P(e)$  obtained from integrating  $P(a, e)$  over  $e$  and  $a$ , respectively, except for the distribution  $P_{\bar{a}}(e)$  which is  $P(a = \bar{a}, e)$ , with  $\bar{a}$  being the value of  $a$  where  $P(a)$  peaks.

A technical note on the side: The distribution of  $a$  after an explosion has the asymptotic form  $P(a, e) \sim 1/a^2$  for large  $a$ . For computational purposes, it is convenient to map  $a$  onto the interval (0,1) in such a way that the distribution of the new variable, say  $x$ , goes to a constant as  $a \rightarrow \infty$ . For  $a_i$  after the first explosion, this is accomplished by the mapping  $x_i = a_i^{\min}/a_i$ , where  $a_i^{\min} = a_0(1 - e_0)/2$  is the minimum value for  $a_i$  after the first explosion which can be obtained from Eq. (A.4). After the second explosion, a more convenient mapping is  $x_f = 1/(1 + a_f/\sigma)$ , where a good value for  $\sigma$  is  $3R_\odot$ .

We now discuss the implications of mass loss due to winds in the time between the first and the second supernova explosion. This mass loss is well described by the expression

$$\dot{M} = -\alpha M^\gamma, \quad (\text{A.18})$$

where  $\alpha > 0$  and  $\gamma = 2.6$  empirically (Woosley et al., 1993). This equation can be integrated immediately to yield

$$M(t) = M_0 [1 + (\gamma - 1)\alpha t M_0^{\gamma-1}]^{-1/(\gamma-1)} \quad (\text{A.19})$$

with  $M_0 = M(t=0)$ . Let the initial mass of a He-star be  $M_0$  and the mass at the time of the explosion  $M_T$ , where  $T$  is the He-burning time. We assume that at the time of explosion of the first He-star the age of the second He-star is  $T/2$ . This gives

$$M_{T/2} = [2/(M_0^{\gamma-1} + M_T^{\gamma-1})]^{1/(\gamma-1)} M_0 M_T \quad (\text{A.20})$$

for the mass of the second He-star at the time of explosion of the first. For PSR 1913+16, we use  $M_0 = 7M_\odot$  and  $M_T = 3.2M_\odot$  so that  $M_{T/2} = 4.2M_\odot$ , whereas for PSR 1534+12, we use  $M_0 = 3.5M_\odot$  and  $M_T = 2.1M_\odot$  so that  $M_{T/2} = 2.6M_\odot$ .

Subsequently, the second He-star loses some more mass over a time  $T/2$ . This mass loss leads to a change in the orbital parameters  $a_i$  and  $e_i$  and their corresponding probability distributions which has to be taken into account. Hills (1983) has given expressions for the change in semi-major axis and eccentricity due to sudden mass loss without kick velocity. We shall use his equations to derive differential equations describing continuous mass loss. Suppressing indices, we have for the change in  $a$

$$\frac{a + \Delta a}{a} = \frac{1 - \Delta M/M}{1 - (2a/r)(\Delta M/M)}. \quad (\text{A.21})$$

For an infinitesimal time interval  $dt$ , we write  $\Delta a = \dot{a} dt$  and  $\Delta M = -\dot{M} dt$  and expand the equation to first order in  $dt$ . We obtain

$$\frac{\dot{a}}{a} = -\frac{\dot{M}}{M} \left( \frac{2a}{r} - 1 \right). \quad (\text{A.22})$$

We shall make the assumption that only very little mass gets lost over one period so that it makes sense to approximate  $1/r$  in the above expression by the average  $\langle 1/r \rangle = 1/a$ . The resulting differential equation for  $a$  can now be integrated to yield

$$a(t) = a_0 \frac{M_0}{M(t)}, \quad (\text{A.23})$$

where the subscript 0 refers to  $t = 0$ . Hence, after a time  $T/2$ ,  $a_i$  has increased by a factor larger than 1. This increase has a simple effect on  $P(a_i, e_i)$ , but due to the mapping  $a_i = a_i^{\min}/x_i$ ,  $P(x_i, e_i)$  remains unaffected by this multiplicative constant. The only effect is a change in  $a_i^{\min}$  which gets multiplied by the same factor as  $a_i$ .

The change in eccentricity is given by

$$\frac{1 - (e + \Delta e)^2}{1 - e^2} = \frac{1 - (2a/r)(\Delta M/M)}{(1 - \Delta M/M)^2}. \quad (\text{A.24})$$

We again write  $\Delta e = \dot{e} dt$  and  $\Delta M = -\dot{M}dt$  and expand Eq. (A.24) to first order in  $dt$ . We obtain

$$\frac{2e\dot{e}}{1 - e^2} = 2\frac{\dot{M}}{M}\left(1 - \frac{a}{r}\right). \quad (\text{A.25})$$

Making the same approximation as before we replace  $1/r$  by  $\langle 1/r \rangle = 1/a$  and obtain  $\dot{e} = 0$ . While this result is due to an approximation, we take it as an indication that we do not have to be overly concerned with the change in the eccentricity. In short, we will assume that  $e_i$  does not change.

## References

- Applegate, J.H., 1988, *ApJ*, 329, 803  
 Arzoumanian, Z., 1995, Princeton University thesis, Radio Observations of Binary Pulsars, unpublished  
 Barba, R., & Niemela, V., 1994, IAU Circular No. 6099  
 Bhattacharya, D. & Van den Heuvel, E.P.J., 1991, *Phys. Rep.* 203, 1-124  
 Bodenheimer, P., & Taam, R.E., 1984, *ApJ*, 280, 771  
 Braun, H., & Langer, N., 1995, *A&A*, 297, 483  
 Brown, G.E., 1995a, *ApJ*, 440, 270  
 Brown, G.E., 1995b, *Phys. Rep.*, 256, 109  
 Brown, G.E., & Weingartner, J.C., 1994, *ApJ*, 436, 843  
 Brown, G.E., Weingartner, J.C., & Wijers, R.A.M.J., 1996, *ApJ*, 463, 297  
 Burrows, A., & Woosley, S.E., 1986, *ApJ*, 308, 680  
 Camilo, F., Nice, P.J., & Taylor, J.H., 1993, *ApJ*, 412, L37  
 Chevalier, R.A., 1993, *ApJ*, 411, L33  
 Chevalier, R.A., 1995, *ApJ*, 459, 322  
 Cook, G.B., Shapiro, S.L., & Teukolsky, S.A., 1994, *ApJ*, 423, L117  
 Darwin, G.H., 1908, "Scientific Papers", Vol. 2, (Cambridge: Cambridge Univ. Press) (original publication 1879)  
 Davidson, K. & Ostriker, J.P., 1973, *ApJ*, 179, 585  
 Doonan, C., & DeGreve, J.P., 1982, *Ap&SS*, 87, 357  
 Fomalont, E.B., Goss, W.M., Lyne, A.G., Manchester, R.N., & Justtanont, K., 1992, *MNRAS*, 258, 492  
 Fryer, C.L., & Benz, W., 1994, *BAAS*, 26, 1443  
 Habets, G.M.H.J., 1986a, *A&A*, 165, 95  
 Habets, G.M.H.J., 1986b, *A&A*, 167, 61  
 Harrison, P.A., Lyne, A.G., & Anderson, B., 1993, *MNRAS*, 261, 113  
 Hills, J.G., 1983, *ApJ*, 267, 322  
 Iben, I., & Tutukov, A.V., 1985, *ApJS* 58, 661  
 Illarionov, A.F. & Sunyaev, R.A., 1975, *A&A* 39, 185  
 Koenigsberger, G., Moffat, A.F.J., Auer, C.H., Drissen, L., & Seggewiss, W., 1994, *ApJ*, 436, 301  
 Kulkarni, S.R., 1992, *Phil. Trans. R. Soc. London A*, 341, 77  
 Langer, N., 1989, *A&A*, 210, 93  
 Lyne, A.G. & Lorimer, D.R., 1994, *Nature*, 369, 127  
 Nice, D.J., Sayer, R.W., & Taylor, J.H., 1996, *ApJ*, 466, L87  
 Nomoto, K.-I., Yamaoka, H., Pols, O.R., Van den Heuvel, E.P.J., Iwamoto, K., Kumagai, S., & Shigeyama, T., 1994, *Nature*, 371, 227  
 Phinney, E.S., & Sigurdsson, S., 1991, *Nature*, 349, 220  
 Pols, O., 1993, thesis, On the Evolution of Massive Close Binary Stars in Stellar Populations, University of Amsterdam  
 Shibasaki, N., Murakami, T., Shaham, J., & Nomoto, K., 1989, *Nature*, 342, 656  
 Sutanto, W., 1975, *A&A*, 41, 47  
 Sutanto, W., 1992, in: X-Ray Binaries and Recycled Pulsars, ed. E.P.J. Van den Heuvel & S.A. Rappaport (Dordrecht: Kluwer Acad. Publ.) 293  
 Taam, R.E., 1996, Proc. IAU Symposium 165, Compact Stars in Binaries, eds. J. van Paradijs, E.P.J. Van den Heuvel & E. Kuulkers (Dordrecht: Kluwer Acad. Publ.) 3-16  
 Taam, R.E., Bodenheimer, P., & Rózyczka, M., 1994, *ApJ*, 431, 247  
 Taam, R.E., & Van den Heuvel, E.P.J., 1986, *ApJ*, 305, 235  
 Terman, J.L., & Taam, R.E., 1996a, *ApJ*, 458, 692  
 Terman, J.L., & Taam, R.E., 1996b, *MNRAS*, 276, 1320  
 Terman, J.L., Taam, R.E., & Hernquist, L., 1994, *ApJ*, 422, 729  
 Tutukov, A.V., 1988, *Astrophysics* 28, 359  
 Van den Heuvel, E.P.J., 1976, Structure and Evolution of Close Binary Systems, 35, IAU, ed. P. Eggleton  
 Van den Heuvel, E.P.J., 1994, Interacting Binaries, Saas-Fee Course 22, Lecture Notes 1992, eds. H. Nussbaumer & A. Ott (Berlin: Springer-Verlag) 263  
 Van den Heuvel, E.P.J., & Van Paradijs, J., 1993 (Nov.), *Scientific American*, 38  
 Van Kerkwijk, M.H., Geballe, T.R., King, D.L., Van der Klis, M., Van Paradijs, J., 1996, *A&A*, in press

- Vrancken, J., DeGreve, J.P., Yungelson, L., & Tutukov, A.V., 1991, *A&A*, 249, 411
- Wijers, R.A.M.J., Van Paradijs, J., & Van den Heuvel, E.P.J., 1992, *A&A*, 261, 145
- Woosley, S.E., Langer, N., & Weaver, T.A., 1993, *ApJ*, 411, 823
- Woosley, S.E., Langer, N., & Weaver, T.A., 1995, *ApJ*, 448, 315
- Yamaoka, H., Shigeyama, T., & Nomoto, K., 1993, *A&A*, 267, 433
- Yungelson, L., & Tutukov, A.V., 1991, *Wolf-Rayet Stars in Galaxies*, ed. K. Van der Hucht & B. Hidayat (Dordrecht: Kluwer Acad. Publ.) 459

---

## Chapter 11

# Supernova Explosions, Black Holes and Nucleon Stars

G.E. Brown

Reprinted with permission from *Physikalische Blätter*, 53 (1997) 671-674

### Commentary by G.E. Brown

This popular article sketches the evidence for the dropping of the  $K^-$  effective mass with density found in GSI experiments. In the Commentary in our Appendix, the connection with heavy-ion experiments will be updated and covered in much more detail.

In this article it is suggested that we don't see binary pulsars which result from main sequence stars of ZAMS mass  $\lesssim 15M_\odot$  because in the pulsar, He-star stage a He star of mass  $\sim 3.7M_\odot$ , resulting from such a main sequence star, will expand in a red supergiant stage (reverse Case C mass transfer) and dump enough matter onto the pulsar to make it go into a black hole. (Fryer & Kalogera (1997) later developed this theme in much more detail, showing that in general, pulsars could only be saved from this fate because of the kick velocities neutron stars receive upon birth.)

The sentence "We suggest that the resulting large differences in the masses in 1913+16 pins down the main sequence progenitor mass to the transition mass in which the transition described occurs." should be deleted. We can evolve neutron-star binaries only in the double He-star scenario; i.e., only with naked He stars. These never skip convective carbon burning and there is no range in ZAMS masses where the size of the Fe core masses changes rapidly with mass.

### Reference

- (1) Fryer, C., and Kalogera, V. (1997), *Astrophysical Journal* **489**, 244.

This page is intentionally left blank

# Supernova Explosions, Black Holes and Nucleon Stars

G. E. Brown

In the collapse of large stars matter is formed at supranuclear densities. At a density  $\rho \sim 3\rho_0$ , where  $\rho_0$  is the nuclear matter density,  $K^-$ -mesons replace electrons, and go into a Bose condensate. This soft equation of state gives a maximum mass for neutron stars of  $\sim 1.5M_\odot$ . It is argued that the last nearby supernova, SN 1987A, went into a black hole. A new scenario of binary neutron star evolution is developed which results in the two masses in a given binary being nearly equal and all neutron star masses lying in a narrow band  $1.3M_\odot < M_{\text{NS}} < 1.5M_\odot$ .

## Supernova Explosions; Kaon Condensation

On February 23, 1987, we had our first "nearby" supernova explosion in 383 years. Neutrinos from the collapse of the star were detected in Kamiokande and the Pittsburgh Plate Glass mine, so we knew that the explosion was that of a large star. In fact, the star that exploded was Sanduleak 202-69, an  $\sim 18M_\odot$  blue supergiant, catalogued by Nick Sanduleak in 1969.

Many of our expectations of supernova explosions of large stars [1] were fulfilled, but we waited in vain to observe the birth of a neutron star. Neutron stars are expected to be born with high magnetic fields,  $B \sim 1-5 \times 10^{12}$  G, like the Crab Pulsar. Even if the magnetic field is pointed away from us, we would have expected to see the "plerion", a cloud illuminated by synchrotron radiation from charged particles flowing outwards along field lines. The plerion should be visible  $\sim 10^5$  years.

We didn't see anything, except the light from radiation coming from radioactive decays of the shells of cobalt, nickel, and iron isotopes blown off by the explosion.

We waited, expecting to see radiation from the neutron star. Observers kept telling us that the matter around the explosion was optically thick, that it would take time for it to clear out. Chevalier [2] and we [3] decided to see how much time. Using classical accretion theory and knowing the explosion kinematics, we could show that after  $\sim 1$  year, we should see light generated by ambient matter falling onto the neutron star, with a luminosity of

$$L \approx 4 \times 10^{38} \text{ ergs/sec} \sim 10^5 L_\odot. \quad (1)$$

The luminosity that we did see could be explained as originating in radioactive decays in the shells of cobalt, iron, etc. which were blown off. The total luminosity was  $L \approx 4 \times 10^{36}$  ergs/sec about 5 years following the explosion and can now be totally explained by the decay of the long-lived ( $\sim 50$  years) isotope of  $^{44}\text{Ti}$ .

So what became of the neutron star? Hans Bethe and I [4] proposed that it had gone into a black hole.

An upper limit of  $M = 1.56M_\odot$  for neutron star mass was derived by the following argument: From the amount of iron produced in the explosion [5] it was shown that the mass

of the compact object was  $\leq 1.56M_\odot$ . Then the argument that this compact object went into a black hole gave the upper limit on the neutron star mass. From this and other arguments we came to the conclusion that the maximum possible neutron star mass was

$$M_{\text{NS(max)}} \sim 1.5M_\odot. \quad (2)$$

Hans Bethe has raised a caveat about the determination eq. (2). In the case of SN 1987A, the explosion of a blue supergiant, a reverse shock arises when the primary shock hits the hydrogen envelope. This reverse shock is estimated to dump an additional  $0.1-0.15M_\odot$  of material onto the compact object, an hour or two following the initial explosion. Thus, it is possible that the neutron star lasted until this further material accreted; at least, nothing in the observation excludes this.

We will stick by the theoretical calculations which give  $M_{\text{NS(max)}} \sim 1.5M_\odot$ . We note that in the evolution of binary pulsars discussed later, there is no hydrogen envelope at the time of explosion; therefore, no reverse shock.

Meanwhile, developments in the theory of dense matter were being made. Kaplan and Nelson [6] showed that the  $K^-$ -meson experiences an attractive scalar, as well as vector, mean field interaction in nuclear matter. Whereas the  $K^-$ -meson has a mass

$$m_{K^-} = 495 \text{ MeV}, \quad (3)$$

in the middle of the  $^{56}\text{Ni}$  the study of kaonic atoms indicated that it was bound by  $200 \pm 20$  MeV [7]. In other words, the attraction from nuclear matter at a density  $\sim \rho_0$  was sufficient to greatly lower the mass. The  $\sim 200$  MeV turned out to be too large in magnitude; we now believe the attraction at density  $\rho_0$  to be somewhat less (see later). Nonetheless, this is sufficient to bring the  $K^-$ -mass down to

$$m_{K^-}^* \sim 200 \text{ MeV}, \quad (4)$$

Prof. Dr. Gerald E. Brown, Department of Physics, State University of New York at Stony Brook, Stony Brook, NY 11794-3800, USA – Talk given at the Magnus-Haus in Berlin on the occasion of receiving the Max Planck Medal.

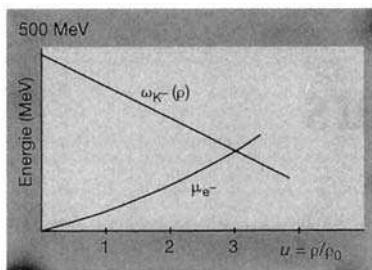


Fig. 1: Behavior of the  $K^-$ -energy (with the  $K^-$  at rest) and electron chemical potential with density.

by  $\rho \sim 3\rho_0$ , where  $\rho_0$  is the nuclear matter density. We have

$$m_{K^-}^* \sim m_{K^-}(1 - 0.2\rho/\rho_0) \quad (5)$$

for neutron rich matter.

Here we are discussing the energy of an antikaon at rest. For a moving antikaon we should divide the attraction into scalar and vector components.

Meanwhile, the electrons, which carry the negative electric charge in a normal neutron star, are squeezed together as the density increases. Since they are fermions, only one electron can occupy each state and the energies build up with electron density. The most energetic electron is at the top of the fermi sea, with fermi energy  $E_F \sim \mu_e$ , where  $\mu_e$  is the chemical potential. The electrons are highly degenerate. The situation is shown in Fig. 1.

Once the  $K^-$ -energy crosses the electron chemical potential, the energy of the neutron star can be lowered by replacing electrons by  $K^-$ -mesons. The  $K^-$ -mesons are bosons, and any number can go into the same state. In order to achieve the lowest energy they all go into the state of zero momentum, and form a Bose condensate. This lowers the energy of the compact object greatly, and produces a soft equation of state; i. e., the compact core offers less resistance to being squeezed.

The electrons change into  $K^-$ -mesons through the reaction

$$e^- \rightarrow K^- + \nu. \quad (6)$$

In the initial collapse and contraction of the inner core of the star, this cannot happen immediately, because all of the states available to the neutrino, up to the neutrino chemical potential  $\mu_\nu$ , are filled. The neutrinos formed in the collapse simply cannot get out in less than a few seconds (see below) because their mean free path is short, on the order of centimeters, and they cannot random walk out of the matter which is collapsing inwards. Basi-

cally, they are on a treadmill, which travels inward faster and faster, sweeping the neutrinos in with it. Thus, the  $K^-$ -mesons cannot replace the electrons, before the neutrinos from the decay find an empty state to go into, and this takes several seconds.

Thorsson, Prakash, and Lattimer [8] showed that for reasonable parameters a neutron star mass of only  $\lesssim 1.5M_\odot$  could be stabilized. Their scenario for the formation of the black hole is:

1. The central part of the neutron star implodes at time  $t=0$ .
2. In the 3–4 seconds after the center bounces, the resulting explosion blows off the outer shells of the star. The resulting matter forms the elements of carbon and heavier, as well as some hydrogen, helium, etc. This is the only known way of producing the heavy elements.
3. During the  $\sim 12$  seconds after bounce the original neutrinos leave. These were observed in Kamiokande and in the Pittsburgh Plate Glass mine in Ohio. As the neutrinos leave, kaon condensation takes place, the equation of state softens, and the pressure is no longer sufficient to support the neutron star; it goes into a black hole if the mass is  $\gtrsim 1.5M_\odot$ .

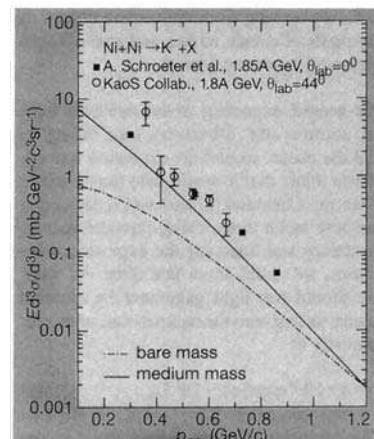


Fig. 2: The inclusive cross section for  $K^-$ -production as function of antikaon momentum in the nucleus-nucleus cms for  $K^-$ -mesons at  $\theta_{lab}=0^\circ$  for  $Ni + Ni$  at 1.85 GeV/Nucleon in comparison with the experimental data from Schröter, *et al.* [9]. Also shown is the preliminary data for  $Ni + Ni$  at 1.8 GeV/Nucleon from Seeger [10]. The dashed line is calculated for the bare antikaon mass, the solid line for the *in-medium* mass. The theoretical curves are from W. Cassing, E. L. Bratkovskaya, U. Mosel, S. Teis and A. Sibirtsev, Nucl. Phys. A, to be published. Confirmation that the theoretical fit to the data is better with a mean field potential that gives  $m_{K^-}^* = m_{K^-}(1 - 0.2\rho/\rho_0)$  has been obtained by G. Q. Li and collaborators (to be published).

Of course, extending the  $K^-$ -nucleus interaction from  $^{56}Ni$  to matter at densities  $\rho \sim 3\rho_0$  involved quite an extrapolation and considerable courage. Confirmation of this extrapolation from relativistic heavy ion reactions is now coming, and is of great importance. The cross section for subthreshold  $K^-$ -production in nucleus-nucleus collisions is very small, but if the  $K^-$ -mass is much decreased in medium, then the effective threshold drops by the binding energy. Experiments at GSI, shown in Fig. 2, show an order of magnitude greater  $K^-$ -production than would result if the  $K^-$ -mesons had their bare mass.

The in-medium  $K^-$ -mass for the solid line in Fig. 2 was taken to decrease with density as

$$m_{K^-}^* = m_{K^-}(1 - 0.2\rho/\rho_0). \quad (7)$$

In order to obtain the similar formula eq. (5) for neutron rich matter, one would need  $m_{K^-}^* \approx m_{K^-}(1 - 0.22\rho/\rho_0)$  for the nuclear matter in which the  $K^-$ -production measurements were made. Since the solid line in Fig. 2 lies below the experimental points, it can be seen that this parameterization would give a better fit to the experimental data.

Considerable activity to measure  $K^-$ -flow in relativistic heavy ion reactions is now going on to check the effects of mean fields, eq. (7).

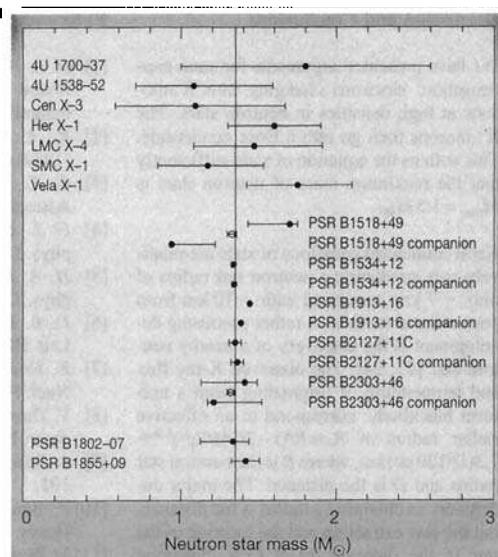
### Measured Neutron Star Masses, the Evolution of Binary Neutron Star Systems

Neutron star masses are measured only in binaries. In Fig. 3 we show the present status. The top seven, 4U 1700-37 – Vela X-1 are all neutron stars in a high mass X-ray binary. Those in the middle are neutron star binaries and the two lowest ones are neutron star, white dwarf binaries.

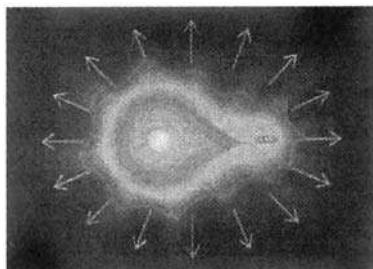
The lower limit on Vela X-1 has come down to  $1.37M_\odot$  due to new measurements [11]. The neutron stars, such as the one in Hercules X-1, have moved around in mass with little in the way of new measurements, just in the re-analysis of the data. It is doubtful that one can improve much on the accuracy of these. The neutron stars labeled PSR (pulsar) are all in either binary neutron stars or in neutron star, white dwarf binaries. In the case of the two wide neutron star binaries, PSR 1518+49 and PSR 2303+46, the individual masses cannot be measured with great accuracy. In these cases, the average star mass of the binary is shown as an open circle between the individual binary masses. My theory [12] would have the two masses nearly equal in a binary neutron star system.

As time has progressed, well measured neutron star masses have tended more and more to cluster between  $1.3M_\odot$  and  $1.5M_\odot$  (I as-

**Fig. 3:** Neutron star masses ( $M_{\odot}$ ). Compilation (January 1997) by Steve Thorsett, Princeton University.



**Fig. 4:** Dispersal of the common hydrogen envelope of the O-star, neutron star binary by dynamical friction as the neutron star spirals in through the envelope. The helium core is the solid white sphere in the center of the envelope. In reality, the neutron star is much smaller, in comparison with the helium star, than shown. ▼



sume that in the wide binaries the masses of the neutron stars within a given binary are nearly equal; see later.)

It is generally thought that neutron stars end up at masses  $\sim 1.4 M_{\odot}$  because of the way in which they evolve. The maximum (Chandrasekhar limit) white dwarf is  $1.44 M_{\odot}$ . However, the Chandrasekhar mass is a good bit smaller for the collapsing large star, because it is somewhat neutron rich. The number of electrons is equal to the number of protons, which is less than half the number of nucleons. Early on in the collapse, the electrons produce the pressure. Thus, the Chandrasekhar mass, which collapses homologously to form the core of the compact object is only  $\sim 1.1 M_{\odot}$  and the remaining mass must come from later accretion onto it. As noted above, in the case of 1987A the accretion seems to be sufficient to bring the mass up to  $\gtrsim 1.5 M_{\odot}$ .

In a binary, the two neutron stars come from two large O-star progenitors. In the conventional evolution there is first a mass transfer of the hydrogen envelope of the more massive to the less massive O-star and then the first O-star explodes, ending as a neutron

star. This leaves a neutron star circling the secondary O-star. When the secondary O-star evolves in the red giant stage after completing hydrogen core burning, the envelope extends beyond the neutron star (Fig. 4), and there is a period of common envelope evolution.

Dynamical friction from the hydrodynamic coupling of the motion of the spiraling-in neutron star in the common hydrogen envelope is supposed to remove the latter, leaving a helium star, neutron star binary. We have one example: Cygnus X-3.

The problem is that the neutron star also accretes matter as it spirals in. A rough calculation of the accreted matter, essentially dimensional analysis, gives [12]

$$\delta M < V^2 > \sim c E_B . \quad (8)$$

Here  $< V^2 >$  is the mean square velocity of the neutron star around the O-star, typically  $\sim 10^{14} \text{ cm}^2/\text{sec}^2$ , and  $E_B$  is the binding energy of the hydrogen envelope,  $\sim 10^{48} \text{ ergs}$ . The efficiency of hydrodynamic coupling of the neutron star motion to the envelope and the coefficient of dynamical friction give [12] the constant  $c \sim 2/3$ . One easily sees that

$$\delta M \sim 10^{34} \text{ ergs} = 5 M_{\odot} . \quad (9)$$

Thus, unless there are exceptional circumstances, such as a large amount of angular momentum [14] the neutron star will accrete enough matter to go into a black hole.

In order to avoid this fate, it was suggested [12] that the two O-stars should be very nearly equal, to within 4%, in mass. They would burn helium at the same time, the two helium

stars going into a common envelope as the O-stars evolved, expelling the envelope by dynamical friction. In this way, a neutron star never has to go through a hydrogen envelope. This scenario is very selective in needing nearly equal mass progenitors, but statistics showed that more than enough neutron star binaries could be formed this way [13,15]. One consequence of this scenario is that the neutron stars in a given binary should be nearly equal in mass. The masses in PSR 1534+12 are  $1.338 \pm 0.012$  and  $1.341 \pm 0.012 M_{\odot}$ . In PSR 1913+16 they are  $1.386 \pm 0.003$  and  $1.442 \pm 0.003 M_{\odot}$ . The latter is the pulsar, and it should have accreted some mass from the helium star wind during the helium star, neutron star binary stage in order to speed it up and to bring its magnetic field down. However, the amount accreted is estimated to be  $\lesssim 0.02 M_{\odot}$ , leaving  $0.04 M_{\odot}$  unaccounted for. Although this does not look large, it would imply an  $\sim 4 M_{\odot}$  difference in main sequence masses using standard relations between neutron star and helium star masses and helium star and main sequence masses. Woosley and collaborators have emphasized that there can be fluctuations in compact core masses, but the required magnitude seems large compared with the rather smooth behavior found by Brown, Weingartner and Wijers [16]. Just around main sequence mass  $19 M_{\odot}$  there is, however, a rapid change in masses, as emphasized by Weaver and Woosley [17]. This marks the transition between stars that burn carbon convectively in the core and those that burn it radiatively. In Woosley and Weaver (1995) [18] the compact core mass changes by 12.5% with change in main sequence mass from 18 to  $19 M_{\odot}$ . The precise main sequence mass at which this transition occurs depends upon the  $^{12}\text{C}(\alpha, \gamma)^{16}\text{O}$  reaction rate, which is not precisely known, and the transition could easily be moved up in mass, by a small change in reaction rate. Fig. 1 of Brown et al. [16] would reproduce the pulsar mass of  $1.442 M_{\odot}$  from a  $21 M_{\odot}$  main sequence star, for example. We suggest that the relatively large differences in the masses in 1913+16 pins down the main sequence progenitor mass to the transition mass in which the transition described occurs. This should set the  $^{12}\text{C}(\alpha, \gamma)^{16}\text{O}$  rate, as only a bit smaller than employed by Woosley and Weaver.

We believe that the individual neutron star masses of the wider binaries 1518+49 and 2303+46 are close to being equal, all of them  $\sim 1.3 M_{\odot}$ . This would follow within each binary from our evolution from double helium star binaries. In fact, neutron stars of masses below  $1.3 M_{\odot}$  would go into black holes, at least in our evolutionary scenario. The argument is as follows:

From Brown, Weingartner, and Wijers [16] (see Fig. 1) a  $1.3 M_{\odot}$  neutron star would evolve from an  $\sim 3.7 M_{\odot}$  helium star. (The

relativistic binary 1534+12 would evolve from an  $4M_{\odot}$  helium star, rather than the  $3.5M_{\odot}$  helium star employed by Wetzig and Brown [13] where less care was taken with the helium star masses.)

Now the opacity during helium burning increases as the temperature decreases. Somewhere between  $3M_{\odot}$  and  $4M_{\odot}$ , during helium shell burning the envelope of the helium star begins to expand in a red giant stage. In Habets [19] without inclusion of mass loss by wind, a  $3M_{\odot}$  helium star expands to  $\sim 10R_{\odot}$ . In Woosley, Langer, and Weaver [20], who include mass loss, this happens at  $4M_{\odot}$ . According to our latest estimates the Woosley et al. mass loss is at least twice too large. As a rough estimate, this expansion to  $\sim 10R_{\odot}$  would occur at a helium star mass of  $\sim 3.5M_{\odot}$ .

Habets [21] studied mass exchange of a  $2.5M_{\odot}$  helium star, during its red supergiant phase, to a  $17M_{\odot}$  O-star at  $84.2R_{\odot}$ . In 2860 years,  $0.3M_{\odot}$  was lost to the companion. In our case, we wish to replace the O-star by a neutron star at several  $R_{\odot}$ . Masses of neutron stars from helium stars in the mass range  $2-3.5M_{\odot}$  should lie in the range of  $1.2-1.3M_{\odot}$ . Addition of  $0.3M_{\odot}$  by accretion from the envelope of the companion helium star, during the neutron star, helium star binary stage, would send the neutron star into a black hole. (We envisage accretion at a hypercritical rate of  $>10^4$  times the Eddington limit, according to the outline of Brown [12].)

Although speculative and daring, our scenario would explain one salient feature of the neutron star masses. Until now, there is no real need observationally for neutron stars in binaries which resulted from progenitors with main sequence masses in the range  $8-15M_{\odot}$ . (The  $3.7M_{\odot}$  helium star corresponds to an  $\sim 15M_{\odot}$  main sequence star.) Yet the number of main sequence stars of masses in this range are relatively large, the Salpeter distribution function for main sequence masses going as

$$dN/dM = M^{-2.35}, \quad (10)$$

and one would expect neutron stars to predominantly have masses  $<1.3M_{\odot}$ . We remove them, by converting them into black holes. A rather complete discussion of neutron star masses is given by Timmes, Woosley, and Weaver [22] who have a different scenario. The work reported in this note corresponds to all supernova explosions in binary evolution being Type Ib (see Fig. 6c of [22]). We have chosen a particular scenario here. Almost all of our numbers came from S.E. Woosley and T. E. Weaver, however. Future observations may well require changes in our scenario.

## Discussion and Conclusions

We have presented arguments for kaon condensation: electrons changing into K-mesons at high densities in neutron stars. The K-mesons then go into a Bose condensate. This softens the equation of state sufficiently that the maximum mass of neutron stars is  $M_{\max} = 1.5M_{\odot}$ .

Kaon condensed equations of state are relatively soft, predicting a neutron star radius of only  $\sim 7$  km, compared with  $\sim 10$  km from conventional models. A rather promising development is the discovery of a nearby neutron star [23, 24]. The observed X-ray flux and temperature, if originating from a uniform blackbody, correspond to an effective stellar radius of  $R_s = R/(1-2GM/Rc^2)^{1/2} \approx 7.3(D/120\text{pc})\text{km}$ , where  $R$  is the neutron star radius and  $D$  is the distance. The major uncertainty in obtaining a radius is the distance, but the low extinction and the location of the star in the direction of R CrA molecular cloud limits the distance to  $D < 120$  pc. Upcoming parallax and proper motion observations [25] will pin down the radius of this star more accurately.

In order to avoid the first neutron star produced in the evolution of binary neutron stars from going into a black hole, a scenario in which the two O-star progenitors are very close in mass is proposed. In this case the two stars burn helium at the same time, expelling their hydrogen envelopes, and the neutron star can survive since it does not have to go through a common envelope. This results in nearly equal masses for the two neutron stars in a binary.

Helium stars of mass  $\lesssim 3.7M_{\odot}$  evolve after helium core burning, into a red supergiant stage. In this case of a helium star, neutron star binary, the stage preceding a double neutron star binary, the neutron star accretes the helium envelope during the RSG stage, and goes into a black hole. This results in a narrow band of neutron star masses  $1.3M_{\odot} < M_{\text{NS}} < 1.5M_{\odot}$ , coming from a band of main sequence masses  $15M_{\odot} < M_{\text{MS}} < 25M_{\odot}$ .

I would like to thank Stan Woosley for nearly all of the calculational results upon which my paper is based, although I am sure he will disagree with much of the interpretation. (Hopefully observations will decide this matter.) I am grateful to Steve Thorsett for his compilation of neutron star masses and useful information. Ralph Wijers has given much useful criticism. I wish to thank Josh Faber and Dr. C.-H. Lee for help with the manuscript. Supported by the U.S. Department of Energy under grant no. DE-FG02-88ER40388.

## References

- [1] H. A. Bethe and G. E. Brown, a general review of the theory of supernovae, *Scientific American*, May, 1985, p. 60.
- [2] R. A. Chevalier, *Astrophys. J.* **346** (1989) 847.
- [3] G. E. Brown and J. C. Weingartner, *Astrophys. J.* **423** (1994) 659.
- [4] G. E. Brown and H. A. Bethe, *Astrophys. J.* **423** (1994) 659.
- [5] H. A. Bethe and G. E. Brown, *Astrophys. J. Lett.* **445** (1995) L129.
- [6] D. B. Kaplan and A. E. Nelson, *Phys. Lett.* **B175** (1986) 57.
- [7] E. Friedman, A. Gal and C. J. Batty, *Nucl. Phys.* **A579** (1994) 518.
- [8] V. Thorsson, M. Prakash, and J. M. Lattimer, *Nucl. Phys.* **A572** (1994) 693.
- [9] A. Schröter et al., *Z. Phys.* **A350** (1994) 101.
- [10] P. Senger, *Strangeness 96*, Budapest, Heavy Ion Physics **4** (1996) 317.
- [11] M. H. van Kerwijk, T. R. Geballe, D. L. King, M. van der Klis and J. van Paradijs, *Astron. and Astrophys.* **314** (1996) 521.
- [12] G. E. Brown, *Astrophys. J.* **440** (1995) 270.
- [13] T. Wetzig and G. E. Brown, *New Astronomy*, **1** (1996) 17.
- [14] R. A. Chevalier, *Astrophys. J.* **459** (1996) 322.
- [15] G. E. Brown, *Phys. Rept.* **256** (1995) 109.
- [16] G. E. Brown, J. C. Weingartner and R. A. M. J. Wijers, *Astrophys. J.* **463** (1996) 297.
- [17] S. E. Woosley and T. E. Weaver, *Phys. Rept.* **227** (1993) 65.
- [18] S. E. Woosley and T. E. Weaver, *Astrophys. J. Suppl.* **101** (1995) 181.
- [19] G. M. H. J. Habets, *Astron. and Astrophys.* **161** (1986) 61.
- [20] S. E. Woosley, N. Langer and T. E. Weaver, *Astrophys. J.* **448** (1995) 315.
- [21] G. M. H. J. Habets, *Astron. and Astrophys.* **165** (1986) 95.
- [22] F. X. Timmes, S. E. Woosley and T. E. Weaver, *Astrophys. J.* **457** (1996) 834.
- [23] F. M. Walter, S. J. Wolk and R. Neuhäuser, *Nature*, **379** (1996) 233.
- [24] J. M. Lattimer and M. Prakash, to be published, 1997.
- [25] F. M. Walter, J. M. Lattimer, M. Prakash, L. Matthews, P. An and R. Neuhäuser, HST cycle 7 General Observer Proposal, 1997.

---

## Chapter 12

# Evolution of Binary Compact Objects That Merge

H.A. Bethe and G.E. Brown

Reprinted with permission from *Astrophysical Journal*, 506 (1998) 780–789  
Copyright (1998) by American Astronomical Society

### Commentary by G.E. Brown

At the end of January 1996, the day before I was to return to Stony Brook, Kip Thorne came to the office Hans and I shared in Kellogg Radiation Lab. at Caltech (Room # 1). He said that the production of gravitational waves that LIGO (The Laser Interferometer Gravitational Observatory) should measure from coalescing neutron stars had been estimated by several research workers theoretically, but that mergers of neutron stars and black holes had not received so much attention. We, Hans and I, were good at calculating things that had not been seen. Could we calculate the contribution for LIGO of neutron-star, black-hole mergers?

Since in Paper 5 I had already found that the standard scenario for evolving a neutron star binary resulted in a neutron-star, low-mass black-hole binary, I immediately knew that the latter would give at least an order-of-magnitude more mergings for LIGO than neutron-star binaries.

I said to Hans, "You're now calculated Roman numeral VI of your Theory of Supernovae. It's time to change." (After all, he was only 90.) I continued, "So we have a topic to work on next year at Caltech." He said, "Oh, no! I want to begin now." So after I got back to Stony Brook I sent him the 20-page paper of Meurs & van den Heuvel (1989). A few days later I received back from him a  $1\frac{1}{2}$ -page fax, in which he'd derived their main results. "I don't seem to have done it as accurately as they did, but I got the same results," he wrote.

I was at GSI in Germany during June of 1996 and I received results at a terrifying rate from Hans. Soon he had a factor of  $\sim 40$  increase over binary neutron star mergers for our low-mass black-hole, neutron-star systems. I panicked. Why hadn't we seen such systems? I asked my observer friends. I return to this point later in the Commentary in Paper 19, where we did estimate the probability that a low-mass black-hole binary would have been seen and found it to be small. The reader will discern that in going from Eq. (5.30) to

Eq. (5.31) in our Paper 12, a factor of 1.9 disappears. Between these two equations we have some discussion which exhibits the panic I felt. Hans, with his great confidence (even though he had never worked on binary evolution before), was unperturbed and perfectly happy with the factor of 1.9.

Starting with BBAL, Paper 1, I was generally the scribe of our works, first writing a free flow of consciousness, with Hans crossing out and inserting phrases into my handwritten draft. He was so excited about this paper, as well as about our recent Paper 22, that he wrote the first draft in his squiggly handwriting.

Hans gave a colloquium on this paper in Ramo Auditorium, a very large one, at Caltech. The auditorium was packed and his talk was carried by television to another room. Each piece of mathematics in this paper is done at senior level in a good high school, or at freshman mathematics level in a university. As Hans put each transparency onto the projector, there was a collective sigh of admiration from the myriad of students in the audience, who mostly poured over their computer screens the rest of the week and rarely did simple analytical work.

The referee of this paper was Peter Eggleton, the dean of British stellar evolvers. He wrote, "This is an interesting paper, making a useful point in an unusually simple and elegant manner. I recommend that it be published, but I would like to make a couple of points which the authors might be able to incorporate. I believe both of them help to reinforce the authors' point, that coalescing neutron-star binaries may be commoner than has been previously thought." Eggleton's points were incorporated. One of his points is the very useful Eq. (7.10) of this paper, giving the increase in mergings because of eccentricity. Just this dependence on eccentricity makes possible the mergings in binaries of high-mass black holes ejected from globular clusters (Portegies Zwart & McMillan 1999). Inclusion of these latter increases the expected number of LIGO-II mergings from our estimated two per month to several per week.

I said to Hans, after receiving Peter Eggleton's referee report, "I have never received such a good referee report before." "No," said Hans, "nor have I."

I leave the Commentary at this happy point, but will return to it for Paper 19, where I compare our simple analytical population synthesis with the computer-driven one of Simon Portegies Zwart.

## References

- (1) Meurs, E., and van den Heuvel, E.P.J. (1989), *Astronomy & Astrophysics* **226**, 88.
- (2) Portegies Zwart, S.F., and McMillan, S.L.W. (2000), *Astrophysical Journal* **528**, L17.

## EVOLUTION OF BINARY COMPACT OBJECTS THAT MERGE

HANS A. BETHE

Floyd R. Newman Laboratory of Nuclear Studies, Cornell University, Ithaca, NY 14853

AND

G. E. BROWN

Department of Physics and Astronomy, State University of New York, Stony Brook, NY 11794-3800

Received 1997 December 8; accepted 1998 June 1

### ABSTRACT

Beginning from massive binaries in the Galaxy, we evolve black hole–neutron star (BH-NS) binaries and binary neutron stars, such as the Hulse-Taylor system PSR 1913+16. The new point in our evolution is a quantitative calculation of the accretion of matter by a neutron star in a common-envelope evolution that sends it into a black hole. We calculate the mass of the latter to be  $\sim 2.4 M_{\odot}$ . The black hole fate of the first neutron star can only be avoided if the neutron star does not go through common-envelope evolution. This can be realized if the two massive binaries are sufficiently close in mass, not more than  $\sim 5\%$  apart, so that they burn helium at the same time. Then their common hydrogen envelope is expelled by the tightening in the double He star system, with attendant hydrodynamical coupling to the envelope. In this way, we obtain a rate of  $\lesssim 10^{-5}$  per yr per galaxy for production of the narrow neutron star binaries, such as Hulse-Taylor 1913+16 or Wolszczan 1934+12. This is in agreement with estimates based on the observed number of such systems extrapolated to the entire Galaxy, with beaming factors and corrections for the  $\sim 90\%$  of binary pulsars estimated to be unobservable. Our chief conclusion is that the production rate for BH-NS binaries (in which the neutron star is unrecycled) is  $\sim 10^{-4}$  per yr per galaxy, an order of magnitude greater than that of neutron star binaries. Not only should this result in a factor of  $\sim 10$  more mergings for gravitational wave detectors such as LIGO, but the signal should also be larger. We include some discussion of why BH-NS binaries have not been observed, but conclude that they should be actively searched for.

**Subject headings:** binaries: close — black hole physics — gravitation — stars: interiors — stars: neutron — stars: statistics

### 1. INTRODUCTION

The Laser Interferometer Gravitational-Wave Observatory (LIGO), an elaborate system of detectors, is being prepared to measure gravitational waves. It is generally believed that the best source of such waves is the merger of two compact stars, neutron stars (NS) or black holes (BH). It is therefore useful to predict the frequency of such mergers.

The current belief is that there are about  $10^{-5}$  mergers per galaxy per yr. We wish to show that the rate is about 10 times greater, on the order of  $10^{-4}$  per galaxy per yr.

We wish to show that a substantial fraction of supernovae in binary stars will in time lead into two compact stars that will ultimately merge. Our assumptions are listed in § 2. In § 3, we estimate the fraction of binaries containing two stars massive enough to become supernovae, provided that one of the stars is. We find this fraction to be about 50%.

The most likely way in which a binary of interest could be disrupted is by the kick that the neutron star receives in the two supernovae. We use recent observations by Cordes & Chernoff (1997) of pulsar velocities, and find that about 40% of binaries survive this kick (§ 2).

We find in § 4 that in 80% of the binaries of interest, the first supernova, in star A, occurs while the other star, star B, is still in its main sequence (MS). Subsequently, star B evolves into a giant. The neutron star, A, will accrete part of the envelope of B and thereby become a black hole. The rest of the envelope of B will be ejected, the energy for this being provided by A spiraling in (§ 4). We require that A stay

outside the He core of star B, so that we ultimately get two separate compact stars. This depends importantly on A becoming a black hole.

In the 20% of cases where B is already a giant when A becomes a supernova, the evolution is somewhat more complicated, as discussed in § 6. In the fraction of these cases where stars A and B burn helium at the same time, the first neutron star can escape common-envelope evolution and so avoid becoming a black hole. In these cases, a neutron star binary such as Hulse-Taylor 1913+16 can be formed, with total production probability for narrow NS binaries that merge of  $\sim 10^{-5}$  per yr per galaxy.

In § 7, we determine the maximum distance  $R$  between the two compact stars that will permit merger by gravitational-wave emission within a Hubble time. Essentially, all cases we discuss will permit such merger. We conclude in § 8 that  $\sim 1\%$  of all supernovae in binary stars ultimately will lead to mergers; thus, we predict  $10^{-4}$  mergers per galaxy per yr.

### 2. ASSUMPTIONS

We call the star that is initially heavier star A, the other star B. We denote initial masses by the subscript 0, so we have initial masses  $M_{A0}$ ,  $M_{B0}$ . We denote their ratio by  $q$ ; thus,

$$q = \frac{M_{B0}}{M_{A0}} \leq 1 . \quad (2.1)$$

Following Portegies Zwart & Yungelson (1998), we assume that  $q$  is distributed uniformly between 0 and 1. Likewise,

we also assume that  $\log a$  is uniformly distributed, where  $a$  is the semimajor axis of their orbit.

However, we assume different limits for  $a$  than Portegies Zwart & Yungelson (1998). Initially, both stars are massive main-sequence stars, with radius of at least  $3 R_{\odot}$ , so  $a > 6 R_{\odot} = 4 \times 10^6$  km. At the other end of the scale, we require  $a < 4 \times 10^9$  km. This corresponds to an orbital velocity of  $30 \text{ km s}^{-1}$ , the same as the Earth's orbital velocity around the Sun, and an orbital period of 25 yr. At a smaller orbital velocity (longer period), it would be difficult to recognize the stars as a binary. Then the fraction of binaries in a given interval of  $\ln a$  is

$$d\phi = \frac{d(\ln a)}{7}. \quad (2.2)$$

In fact, our lower limit is too low, since two zero-age main-sequence (ZAMS) early-type B stars of  $8 M_{\odot}$  would each fill their Roche lobes in an orbit of  $9 R_{\odot}$  rather than  $6 R_{\odot}$ . Thus,  $10^7$  km would be a more reasonable value. Furthermore, as P. Eggleton (1998, private communication) has also pointed out, our flat distribution over  $\log P$  is not well supported for massive stars. Garmany, Conti, & Massey (1980) did probably the best analysis so far and found a binarity of  $36\% \pm 7\%$  O stars in a magnitude-limited and declination-limited sample of 67 O stars. Longer periods are largely inaccessible to measurement because of the inherent instability of O stars. Although the observed binarity is smaller than our assumed 50%, the logarithmic interval in  $a$  is also smaller, so that the number of binaries in a given logarithmic interval is not much changed. Unfortunately, the interval  $a_i$ , which we find useful in equation (5.24), presumably includes some of the inaccessible binaries. In Garmany et al. (1980), there is a clustering of observed binaries in the period interval  $\sim 3\text{--}10$  days, as compared to the  $\sim 10\text{--}100$  days that we show below to be favorable for gravitational merging avoiding spiraling in. However, orbits widen during mass transfer, as we show below in equation (6.1), and the shorter period stars will mostly widen into our favorable interval. Thus, we believe that our flat distribution over  $\log P$  with the chosen interval and binarity somewhat underestimates the number of favorable binaries.

We assume that a star needs an initial mass of

$$M > M_s = 10 M_{\odot} \quad (2.3)$$

to become a supernova. Beyond this mass, we assume that the birthrate of stars of mass  $M$  is proportional to

$$\beta \sim M^{-n}, \quad (2.4)$$

where  $n$  is the Salpeter exponent, about 1.5. The rate of occurrence of supernovae (SNe) is the same as the birthrate. Therefore, if  $\alpha$  is the total rate of SNe, the rate of SNe in a mass interval  $dM$  is given by

$$da = \alpha n \left( \frac{M}{10 M_{\odot}} \right)^{-n} \frac{dM}{M}. \quad (2.5)$$

We assume the total rate of supernova (SN) events per galaxy to be

$$\alpha' = 2 \text{ per century} = 2 \times 10^{-2} \text{ yr}^{-1}. \quad (2.6)$$

We assume that half<sup>1</sup> of the stars are in binaries, so the rate of SN events in binaries is

$$\alpha = 1 \times 10^{-2} \text{ yr}^{-1}. \quad (2.7)$$

In the SN event, the resulting neutron star (NS) receives a kick. For the distribution of kicks, we take the measurements of Cordes & Chernoff (1997). They found that the distribution is well represented by the sum of two Gaussians with different standard dispersions  $\sigma$ :

$$\begin{aligned} 80\% \text{ have } \sigma_1 &= 175 \text{ km s}^{-1} \\ 20\% \text{ have } \sigma_2 &= 700 \text{ km s}^{-1}. \end{aligned} \quad (2.8)$$

If  $v$  is the orbital velocity of the star that is becoming a supernova, and  $U$  is the kick given to the supernova remnant, a rough approximation to the probability of the binary remaining together is given by<sup>2</sup>

$$\omega = \frac{v^2}{U^2 + v^2}. \quad (2.9)$$

If the kicks are distributed in a Gaussian, it is a fair approximation to replace  $U$  by  $\sigma$ , so

$$\omega' = \frac{v^2}{\sigma^2 + v^2}. \quad (2.10)$$

For a binary of total mass  $M = M_A + M_B$ , in a circular orbit of radius  $a$ , the orbital velocity is

$$v^2 = GMa^{-1}. \quad (2.11)$$

If  $\log a$  is uniformly distributed, then  $\log v^2$  is also, and

$$d\phi = \frac{d(v^2)}{7v^2}, \quad (2.12)$$

with limits on  $v^2$  as indicated below equation (2.1), i.e.,

$$30 \text{ km s}^{-1} < v < 1000 \text{ km s}^{-1}. \quad (2.13)$$

Then the fraction of all binaries surviving the first SN event is

$$\mathcal{W} = \int \omega' d\phi = \frac{1}{7} \int \frac{d(v^2)}{\sigma^2 + v^2} = \frac{1}{7} \ln \frac{\sigma^2 + v_{\max}^2}{\sigma^2 + v_{\min}^2}, \quad (2.14)$$

where  $v_{\max}$  and  $v_{\min}$  are the limits indicated in equation (2.13). Taking into account the kick distribution from equation (2.8),

$$\mathcal{W} = \frac{0.8}{7} \ln \frac{10^2 + 1.75^2}{0.3^2 + 1.75^2} + \frac{0.2}{7} \ln \frac{10^2 + 7^2}{0.3^2 + 7^2} = 0.43. \quad (2.15)$$

Note that the minimum velocity,  $30 \text{ km s}^{-1}$ , is unimportant. Changing this to  $10 \text{ km s}^{-1}$  alters the result in equation (2.15) by only 10%. However, the maximum orbital velocity, i.e., the minimum orbital radius, is relevant. The high kick velocities,  $\sigma_2$ , contribute only 0.03 of the 0.43 in equation (2.15). The rate of formation of binaries that survive the first supernova event is now 0.43 per century.

A small correction should be made for the fact that the

<sup>1</sup> With our smaller interval in  $a$ , we may be overestimating the binarity somewhat.

<sup>2</sup> Computer calculations with the program of Weitig & Brown (1996) show this approximation to be accurate to  $\leq 10\%$  in the region of appreciable survival probabilities. Survival with the Cordes & Chernoff (1997) parameterization is slightly greater than that with the formula of Portegies Zwart & Yungelson (1998).

hydrogen envelope in star A does not generally extend out to the larger  $a$  in our assumed interval. Weaver, Zimmerman, & Woosley (1978) find that for a  $15 M_{\odot}$  star (star A), the envelope reaches out to  $3.9 \times 10^8$  km in supergiant stage. Taking  $M_B = 10 M_{\odot}$ , we find the Roche lobe to be  $\sim \frac{1}{2}$  of the distance between A and B. Thus, Roche lobe overflow takes place only out to  $a \lesssim 1.2 \times 10^9$  km in this example, less than our  $4 \times 10^9$  km upper limit. However, relatively few systems at the higher  $a$  survive the kick velocities, and we estimate the correction to be  $\sim 10\%$ , which we neglect.

It should also be noted that we have assumed the semi-major axis distribution to be flat in  $\log a$  following mass transfer, whereas the empirical determinations of it are made before mass transfer. From rough estimates this seems to be a valid assumption. We neglect all binaries that do not survive the first phase of mass transfer and merge into single objects. From Table 4 of Portegies Zwart & Verbunt (1996), this is about 20%. We have not made this correction because in our calculations the close binaries do not survive spiraling in.

### 3. MASS TRANSFER

The more massive star A becomes a giant after time  $t_A$ . Its Roche lobe overflows and matter is transferred to star B. This continues until A is stripped of its hydrogen envelope and thus reduced to a He star. In the mass range we are considering,  $M_{A0}$  between 10 and perhaps  $50 M_{\odot}$ , the mass of the remaining He star is roughly 30% of the initial mass. Using a more accurate relation does not appreciably change our results. Denoting masses after this first transfer by subscript 1,

$$M_{A1} = 0.3 M_{A0}. \quad (3.1)$$

Only part of the mass lost by A will be attached to B. We adopt the estimate of Vrancken et al. (1991, hereafter VDYT) that the fraction attached to B is

$$\beta \approx \left( \frac{M_{B0}}{M_{A0}} \right)^2 = q^2. \quad (3.2)$$

(VDYT use the exponent 1.84). Thus, after transfer, the mass of B is

$$M_{B1} = (q + 0.7q^2)M_{A0} \equiv M_{A0}f(q). \quad (3.3)$$

For  $q = 1$ , the transfer is conservative, and  $f(q) = 1.7$ . For  $q < 0.68$ ,  $f(q) < 1$ .

We wish both stars to become SNe at some stage in their life. For star A, the condition is

$$M_{A0} > M_{\text{sup}} = 10 M_{\odot}. \quad (3.4)$$

It is here assumed that the minimum mass for a (Type II or Ib) supernova is  $10 M_{\odot}$ . For star B to become a supernova, we require that  $M_{B1} > 10 M_{\odot}$ ; in other words, that

$$f(q) > \frac{10 M_{\odot}}{M_{A0}}. \quad (3.5)$$

We have assumed that  $q$  is distributed uniformly between 0 and 1, so the probability that star B also becomes a SN, for any given  $M_{A0}$ , is  $1 - q_s$ , where

$$f(q_s) = \frac{10 M_{\odot}}{M_{A0}}. \quad (3.6)$$

Portegies Zwart & Verbunt (1996) give in their Table 4 results for other distributions in  $q$ .

The rate of supernovae in a given mass interval  $dM_{A0}$  is given by equation (2.5); hence, the birthrate of binaries that have supernovae in both stars is given by

$$\alpha'' = \alpha n \int_{10 M_{\odot}}^{50 M_{\odot}} \left( \frac{10 M_{\odot}}{M_{A0}} \right)^n \frac{dM_{A0}}{M_{A0}} [1 - q_s(M_{A0})]. \quad (3.7)$$

The upper limit has been set at  $50 M_{\odot}$  because above this mass stars tend to lose their envelope mass in a luminous blue variable stage, and become Wolf-Rayet stars. They cannot then participate in common-envelope evolution. We have evaluated the integral numerically, and find

$$\alpha'' = .50\alpha. \quad (3.8)$$

This value varies  $\lesssim 10\%$  with different reasonable Salpeter exponents (J. M. Lattimer 1997, private communication).

### 4. EVOLUTION

For masses near and above  $10 M_{\odot}$ , the lifetime of a star on the main sequence (MS) is approximately proportional to  $M^{-2}$ . Thus, when star A has come to the end of its MS life, after time  $t_A$ , star B will have gone through a fraction  $q^2$  of its MS lifetime.

The evolution time from the beginning of He burning to supernova is roughly 1/10 of the MS evolution time. This time is relevant for us because mass will generally be transferred during helium-core burning, the supergiant state in which the hydrogen envelope is greatly extended. During this time,  $0.1t_A$ , star B has a mass  $M_{A0}f(q)$ , so it will go through an added fraction  $0.1f^2(q)$  of its MS life. At the end of this time, the fraction of its MS life accomplished by star B is given by

$$g(q) = q^2 + 0.1f^2(q). \quad (4.1)$$

Table 1 gives values for the functions  $f(q)$  and  $g(q)$ . An approximate formula is

$$g(q) = 1.25q^2 \pm 0.02. \quad (4.2)$$

Of special interest is the value of  $q$  for which  $g(q) = 1$ . This is found to be

$$q_1 = 0.8897. \quad (4.3)$$

The evolution is different according to whether  $q < q_1$  or  $q > q_1$ . For  $q < q_1$ , we have case I, star B still in its main sequence when star A becomes a supernova. The evolution in this case will be treated in § 5. In case II,  $q > q_1$ , star B

TABLE 1  
MASS OF STAR B AFTER MASS TRANSFER

$q$	$f(q)$	$g(q)$	$1.25q^2$
0.95.....	1.582	1.153	1.128
0.9.....	1.467	1.025	1.012
0.89.....	1.444	1.001	0.990
0.85.....	1.356	0.906	0.903
0.8.....	1.248	0.796	0.800
0.7.....	1.043	0.599	0.613
0.6.....	0.852	0.433	0.450
0.5.....	0.675	0.296	0.313
0.3.....	0.363	0.103	0.112
0.1.....	0.107	0.011	0.013

NOTE.— $f(q) = M_{B1}/M_{A0}$ . The fraction  $g(q)$  of the main sequence accomplished by B when A goes supernova  $1.25q^2$  is shown for comparison with  $g(q)$ .

will be a giant when star A has its supernova event; this will be treated in § 6.

The probability for any  $M_{A0}$  that  $q > q_1$  is given by

$$1 - q_1 = 0.110, \quad (4.4)$$

so the rate of obtaining a case II binary is

$$\alpha(\text{II}) = 0.11\alpha, \quad (4.5)$$

which is about 20% of the total rate, given in equation (3.8).

### 5. CASE I: STAR B AS MAIN SEQUENCE

#### 5.1. Possible Appearance

As discussed in § 4, case I will hold if the initial mass ratio is

$$q = \frac{M_{B0}}{M_{A0}} < q_1 \approx 0.89. \quad (5.1)$$

As shown after equation (4.5), this applies to 80% of all SN binaries. After star A has had its SN event, and provided that the binary stays together (see § 2), we have a neutron star A and a MS star B of mass  $M_{B1} = M_{A0} f(q)$ , which must be greater than  $10 M_\odot$ , because we want B ultimately to become a SN. So, star B will be a B- or O-type star. The total lifetime of such stars is on the order of  $10^7$  yr. Star A will be a neutron star, but it will be a pulsar only for  $5 \times 10^6$  yr or less, because its spin will diminish by emitting pulsar radiation. In fact, the binaries 1259–63, with a Be star companion, and 0045–73, with a B star companion, contain radio pulsars. The former has a short 47 ms period and a double-pulse profile similar to the Crab; i.e., it is a young pulsar. The latter binary is in the Small Magellanic Cloud and is the most luminous binary radio pulsar known.

Since the pulsar is unrecycled, the expected number of these binaries should be compared with the single neutron star pulsar population, about 700 in number, with more uncataloged. This number should be multiplied by a factor of  $\sim \frac{1}{2}$  for binarity and a factor of 0.43 (eq. [2.15]) for breakup in the first explosion. This would leave the large number  $\sim 150$  if pulsars with massive companions were as easily observed as single pulsars. Of course, pulsars are predominantly produced in the Galactic disc, and because of the strong gravitational attraction of the dense matter in the disc, those with massive companions will be unable to move out of the disc. Stellar winds can also interfere with the radio pulses from these binaries, obscuring the narrower ones. Nonetheless, the factor necessary to reduce their observability is startlingly large. We return to this subject later.

#### 5.2. Structure of a Giant

A giant has a He core, containing some 25% of its mass, surrounded by an envelope consisting mostly of H. The envelope is usually in convection, so the entropy is constant. The particles, nuclei and electrons, are nonrelativistic and thus have  $\gamma = 5/3$ ; therefore, the envelope forms a polytrope of index  $n = 3/2$ . Applegate & Horner (1998) show that the binding energy of the envelope is

$$E_1 \cong \frac{2}{3} GM_B^2 R^{-1}, \quad (5.2)$$

where  $R$  is the outer radius. In this formula, the gravitational binding energy is decreased by 50% by the kinetic energy,  $E_1$  containing both effects.

After exhaustion of the core hydrogen, the radius  $R$  increases first slowly, then more rapidly, until it settles down (for masses in our range) at several times  $10^{13}$  cm.

#### 5.3. Accretion

Star A, beginning as a neutron star of mass  $1.4 M_\odot$ , sits at a distance  $r_A$  from the center of star B, waiting for the latter to expand into a giant. Its orbital period is a small fraction of the time required for B to evolve into a giant (on the order of  $10^5$  yr). When B reaches the radius  $r_A$ , material from B begins to accrete to A. The rate of accretion is given by the Bondi-Hoyle-Lyttleton theory,

$$\frac{dM_A}{dt} = \pi \rho v R_{ac}^2, \quad (5.3)$$

where  $\rho$  is the density of the B material,  $v$  is its velocity relative to A, and  $R_{ac}$  is the accretion radius,

$$R_{ac} = 2GM_A v^{-2}. \quad (5.4)$$

Here  $v$  is the velocity of the B material relative to A, which is essentially the orbital velocity of A around B. This is given by

$$v^2 = G(M_B + M_A)a^{-1}. \quad (5.5)$$

The energy loss is related to the accretion rate by (Iben & Livio 1993; Brown 1995)

$$\dot{E} = \frac{1}{2} c_d v^2 \dot{M}_A. \quad (5.6)$$

Here  $c_d$  is the drag coefficient (Shima et al. 1985),

$$c_d = 2 \ln(b_{\max}/b_{\min}) \approx 6, \quad (5.7)$$

for highly supersonic flow. (Iben & Livio 1993 use  $c_d = 8$ .) Here  $b_{\max}$  and  $b_{\min}$  are the maximum and minimum impact parameters of the envelope material relative to the neutron star. Note that if  $c_d$  decreases as the envelope moves toward corotation with the neutron star,  $c_d$  in equation (5.6) must be replaced (Brown 1995) by an effective coefficient of drag,

$$(c_d)_{\text{eff}} = c_d \left( \frac{\mathcal{M}^2 + 1}{\mathcal{M}^2} \right)^{3/2}$$

where  $\mathcal{M}$  is the Mach number  $v/c_s$ , with  $c_s$  the speed of sound. The drop in  $(c_d)_{\text{eff}}$  with decreasing Mach number is slower than in  $c_d$ .

The interaction energy of the two stars is

$$E = \frac{1}{2} GM_A M_B a^{-1}, \quad (5.8)$$

where  $a$  is the semimajor axis of the orbit. The term  $-E$  is the sum of the potential and kinetic (orbital) energy. The orbital velocity is given by equation (5.5). Generally,  $M_A \ll M_B$ , and we neglect  $M_A$  in equation (5.5). Inserting equation (5.5) into equation (5.6), we find

$$\dot{E} = \frac{1}{2} c_d GM_B a^{-1} \dot{M}_A. \quad (5.9)$$

Comparing equation (5.9) to equation (5.8), we see that we have two variables that depend on time,  $M_A$  and

$$Y = M_B a^{-1}. \quad (5.10)$$

Taking the derivative of equation (5.8) and inserting it into equation (5.9), we get

$$M_A \ddot{Y} + Y \dot{M}_A = c_d Y \dot{M}_A. \quad (5.11)$$

This can be integrated to give

$$\ln M_A = (c_d - 1)^{-1} \ln Y + \text{const.}, \quad (5.12)$$

$$M_A \propto Y^{1/(c_d - 1)} = Y^{1/5}. \quad (5.13)$$

The final energy of the binary is then

$$E_f = \frac{1}{2} GM_{A,i} Y_i \left( \frac{Y_f}{Y_i} \right)^{6/5}, \quad (5.14)$$

where the subscript *i* denotes the initial, *f* the final value of the respective quantities.

The binding energy  $E_f$  of star A to star B serves to expel the envelope of star B, whose initial binding energy is given by equation (5.2). But it is believed that only a fraction of  $E_f$  can serve to expel the envelope, which is usually assumed to be  $\frac{1}{2}$ ; hence,

$$E_f = 1.2GM_{B,i}^2 R^{-1} \quad (5.15)$$

is necessary to completely eject the envelope. The spiraling in begins when  $R$  reaches  $a_i$ . Spiraling in is fast compared to the expansion of the giant; therefore, we set

$$R = a_i. \quad (5.16)$$

Then equation (5.15) becomes

$$E_f = 1.2GM_{B,i} Y_i. \quad (5.17)$$

Inserting into equation (5.14) yields

$$\left( \frac{Y_f}{Y_i} \right)^{1/2} = 2.4 \frac{M_{B,i}}{M_{A,i}}. \quad (5.18)$$

Star A is initially a neutron star,  $M_{A,i} = 1.4$ . For star B, we assume  $M_{B,i} = 15$ . Then equation (5.18) yields

$$\frac{Y_f}{Y_i} = 15. \quad (5.19)$$

We use this first to find the result of accretion, with the help of equation (5.13),

$$\frac{M_{A,f}}{M_{A,i}} = 1.73, \quad (5.20)$$

$$M_{A,f} = 2.4 M_\odot. \quad (5.21)$$

Star A, by accretion, has become a black hole. The lower limit for a black hole, according to the argument of Brown & Bethe (1994), is  $M_{\min} = 1.5 M_\odot$ . For the range  $M_B = 10-20 M_\odot$ , the spread in  $M_{A,f}$  is  $2.25-2.5 M_\odot$ . To within a few percent accuracy,  $M_{A,f}/M_{A,i} = (M_{B,i}/M_\odot)^{0.2}$ , given the value we used for  $c_d$ .

Star B, by losing its envelope, becomes a He star. We estimate that

$$\frac{M_{B,f}}{M_{B,i}} = 0.3. \quad (5.22)$$

The size of the orbit is determined by equation (5.10),

$$\frac{a_i}{a_f} = \frac{M_{B,i}}{M_{B,f}} \frac{Y_f}{Y_i} = 50. \quad (5.23)$$

The final distance between the stars,  $a_f$ , should not be less than about  $10^{11}$ , so that the black hole A is comfortably outside the He star B. On the other hand, if the two stars are to merge within a Hubble time, equation (7.9) shows that  $a_f < 3.8 \times 10^{11}$ . Therefore, the initial distance of the two

stars, after the first mass exchange and the first supernova, should be

$$0.5 \times 10^{13} \text{ cm} < a_i < 1.9 \times 10^{13} \text{ cm}. \quad (5.24)$$

The assumption is made here that the final BH-NS orbit has the same  $a$  as the  $a_f$  of the BH-He star orbit. While the former will open out somewhat in the explosion of the He star, it will generally be quite eccentric. On average, this leads to a  $\sim 40\%$  increase in  $a_f$  for merger (J. M. Lattimer 1997, private communication), so there will be considerable cancellation between these neglected effects. (See our eq. [7.10].)

If the initial distribution of distances is  $da/7a$ , the probability of finding  $a$  between the limits of equation (5.24) is

$$P = 18\%. \quad (5.25)$$

The probability of the binary surviving the first explosion was found in equation (2.15) to be 0.43, so the combined probability is

$$PW = 18 \times 0.43 = 7.7\%. \quad (5.26)$$

Strictly speaking, the calculation of equation (2.15) should be made with the orbital velocity limits corresponding to the limits of  $a$ , given in equation (5.24), which are

$$\begin{aligned} v_{\max}^2 &= 7.3 \times 10^{14} (\text{cm s}^{-1})^2, \\ v_{\min}^2 &= 2.0 \times 10^{14} (\text{cm s}^{-1})^2. \end{aligned} \quad (5.27)$$

The result is then

$$PW = 8\%, \quad (5.28)$$

essentially unchanged.

The survival probability of the final He star explosion in the BH-He star binary should be similar to that found by Wettig & Brown (1996) for He star-NS binaries,

$$P \approx 50\%. \quad (5.29)$$

An important ingredient in this relatively high probability is the rather low mass of the He star before explosion. As is well known, the binary cannot stay together if its total mass after explosion is less than half the mass before explosion. The postexplosion mass is  $2.4 + 1.4 = 3.8 M_\odot$ , so the He star mass before explosion must be less than  $0.2 \times 3.8 - 2.4 = 5.2 M_\odot$ . Once the He stars have lost their H envelopes, they rapidly lose mass through wind, and for the relevant ZAMS masses they end up below  $4 M_\odot$ . For example, the initially  $7 M_\odot$  star of Woosley, Langer, & Weaver (1995) ends up at  $3.2 M_\odot$  before explosion. Then, if the SN kick is more or less opposite to the orbital velocity before explosion, the binary will stay together.

We now summarize the various factors entering into the BH-NS evolution:

$$\begin{aligned} \mathcal{R} &= 10^{-2} \times 0.50 \times 0.077 \times 0.5 \\ &= 1.9 \times 10^{-4} \text{ per galaxy per yr.} \end{aligned} \quad (5.30)$$

These factors follow from equations (2.7), (3.8), (5.26), and (5.29). Because of our possible overestimate of binarity, of the number of systems that undergo Roche lobe overflow, and other effects mentioned earlier, this may be somewhat too high. We feel that a reasonable lower limit is

$$\mathcal{R} = 10^{-4} \text{ per galaxy per yr.} \quad (5.31)$$

In fact, our rate per SN ( $\alpha'/2 + \alpha + \alpha''$ ) = 0.025 per yr is 0.004 to be compared to 0.0036 from Portegies Zwart &

Yungelson (1998), who find a SN rate of 0.015 per yr. Thus, the chief difference between our result in equation (5.31) and the  $\mathcal{R} = 5.3 \times 10^{-5}$  of these authors is due to the different assumed SN rate. This shows that the many effects neglected in our calculations but included in their detailed computer calculations are unimportant for the result.

Given our lower limit and the Wettig & Brown (1996)  $P = 50\%$ , we would have a rate of  $2 \times 10^{-4}$  per galaxy per yr for the formation of BH-He star binaries. Given an average He star lifetime of  $5 \times 10^5$  yr, this would give  $\sim 100$  BH-He star systems in the Galaxy. This is not far from the estimate of van den Heuvel (1995), who obtained 60 NS-He star systems, with a lower limit of  $2.5 M_{\odot}$  for the He star mass. (With our lower limit of  $2.2 M_{\odot}$  for the He stars that end up as neutron stars, his number would be  $\gtrsim 50\%$  greater.) Van den Heuvel (1995) discusses reasons why only one such system, Cyg X-3, is seen. If indeed the neutron stars would not become black holes, many of these neutron stars would be recycled by the He star wind, and this would lead to an overproduction, by a large factor, of recycled binary pulsars. We avoid this overproduction problem by having the neutron stars A become black holes.

Ergma & Yungelson (1997) find by means of population synthesis the same number (100) of BH-He star binaries in the Galaxy as we do. However, their massive black holes arise directly from stars of ZAMS masses  $> 30$  or  $50 M_{\odot}$ , whereas our low-mass black holes evolve from spirals in the hydrogen envelope of the compact object. (We have not evolved systems with massive black holes.) Ergma & Yungelson (1997) suggest a combination of reasons why, out of their many predicted systems, Cyg X-3 is the only example of such a system seen.

## 6. CASE II: STAR B AS GIANT

### 6.1. Evolution of Binary Neutron Stars

Following hydrogen-core burning, star B will expand in a red giant phase, which takes up  $\gtrsim 20\%$  of its lifetime. In the first half of the red giant phase, the temperature in the center is not high enough to burn helium. The core undergoes contraction, raising its temperature. Because of the concurrent temperature increase just outside the core, shell hydrogen burning begins and the envelope expands modestly. For a star of ZAMS mass  $16 M_{\odot}$ , Bodenheimer & Taam (1984) find that the radius increases out to  $\sim 4 \times 10^{12}$  cm. This first half of the red giant phase is not useful to this paper, because if the neutron star is met by the expanding giant envelope at  $R < 0.5 \times 10^{13}$  cm, from the condition given by equation (5.24), the neutron star will spiral into the core. Interestingly, we find that this occurs for all the standard high-mass X-ray binaries, such as SMC X-1, Cen X-3, LMC X-4, Vela X-1, and 4U 1538-52. The widest of these, Vela X-1, has  $a = 0.37 \times 10^{13}$  cm. Spiraling-in has been found for these objects in many numerical calculations in the literature. In order to avoid this spiraling-in, we need binaries wider than the HMXB's.

Particularly relevant for us is the next  $\sim 10\%$  of the lifetime of the star, the period of helium-core burning. During this (supergiant) stage, the star expands out to several times  $10^{13}$  cm. If the first-born neutron star is to escape common-envelope evolution, then the two massive stars in the binary must burn helium at the same time. As we outline below, they can then expel their hydrogen envelopes while burning helium. The neutron star produced later then has no hydro-

gen envelope. As noted earlier, stellar evolutionary times scale as  $M^{-2}$ , so that stars A and B must be within 5% of each other in mass if they are to burn helium at the same time. However, if one star goes supernova, the other will also. Thus, we have a rate of  $0.05\alpha$  (from eq. [2.7]), or  $\sim 10\% \alpha'$ .

The fraction of these binaries with very nearly equal masses that survive the spiraling-in will be roughly equal to the fraction of black holes that survive in the estimate given in our discussion in § 5.

If the two ZAMS masses are nearly equal, the two He stars tighten, expelling the common H envelope (Brown 1995). The binding energy of the H envelope that results from the two initial H envelopes is  $\sim 4$  times that of each individual envelope (see eq. [5.2]). If the He star were 4 times more massive than the compact object discussed in § 5 in common-envelope evolution, the final  $a_f$  would be the same as in § 5. Since this is not far from being true, the He stars will end up at roughly the same  $a_f$  as found earlier for the black hole and companion B. In the explosion of the first He star, Wettig & Brown (1996) found a survival probability of  $\sim 50\%$ , not far from the  $\mathcal{W}$  of equation (2.15). The survival probability in the explosion of the He star B in the He star-NS binary will be similar to that in the He star-BH binary. Thus, we can say that  $\sim 10\%$  of the original massive binaries will end up as binary pulsars. This same ratio to BH-NS binaries holds for the probability of merging during a Hubble time. Given our rate of  $10^{-4}$  per galaxy per yr for the latter binaries, we find  $10^{-5}$  per galaxy per yr for binary neutron stars. This may be an upper limit, because He stars with ZAMS masses of  $\leq 15 M_{\odot}$  expand in the He shell-burning phase and give the companion neutron star another chance to become a black hole (Brown 1997). Using unpublished evolutionary calculations of S. E. Woosley of He stars of mass  $< 4 M_{\odot}$  (roughly corresponding to  $15 M_{\odot}$  ZAMS) in which wind mass loss is included, Fryer & Kalogera (1997) show that only special conditions allow the pulsar in narrow NS binaries to avoid the envelope of these low-mass He stars. Otherwise, the pulsar accretes sufficient matter to become a black hole, in much the same way as was more crudely described by Brown (1997). This implies that most of the narrow binaries that evolved from He star masses of  $< 4 M_{\odot}$  in the double-He star scenario are low-mass BH-NS binaries. Because of the much lower total mass in the He envelope compared to the H envelope (Habets 1986, who does not include wind mass loss, gives the entire envelope of a  $2.5 M_{\odot}$  He star above the helium burning shell as  $\sim 0.8 M_{\odot}$ ), less mass will be accreted, so that the black hole will be less massive than the  $\sim 2.4 M_{\odot}$  given by equation (5.21). This will decrease our estimate of  $10^{-5}$  per galaxy per yr for merging binary neutron stars, probably by a factor of  $\sim 2$ . This also adds a class, formed at about the same rate as binary neutron stars, of low-mass BH-NS binaries in which the black hole is not much more massive than the neutron star. Again, our rate is in good agreement with the  $0.7 \times 10^{-5}$  found by Portegies Zwart & Yungelson (1998) in their detailed computer evolution.

Note that our rate of  $10^{-5}$  per galaxy per yr is essentially the same as the rate of  $\sim 8 \times 10^{-6}$  found by van den Heuvel & Lorimer (1996). These authors increased their estimated number of potentially observable binary pulsars by a beaming factor of 3 and a factor of 10 because 90% are estimated to be unobservable (Curran & Lorimer 1995). This factor of 30 is somewhat uncertain, so it may be useful

that we arrive at essentially the same rate directly from an evolutionary calculation.

For very nearly equal initial masses ( $q \sim 1$ ), the mass transfer is conservative. In this case, from angular momentum transfer, we can deduce that

$$\frac{a_f}{a_i} = \left( \frac{\mu_i}{\mu_f} \right)^2, \quad (6.1)$$

where

$$\mu = \frac{M_A M_B}{M_A + M_B}. \quad (6.2)$$

Since  $\mu_f < \mu_i$ , the final  $\mu_f$  being slightly smaller than  $M_{A,f}$ , the He core mass in A, the orbits open out in conservative mass transfer. The main effect of this is to shift the logarithmic interval of given binaries outward. But the magnitude of the logarithmic interval is unchanged, as is (by our assumption) the number of binaries.

Note that if  $q \sim 1$  initially, the initial near equality in the He core masses of A and B will not be changed by the mass transfer. Star B will not be rejuvenated by the transferred H envelope from A, because there is not time to cross the molecular weight barrier and convert the transferred H to He (Braun & Langer 1995). Thus, the near equality in He star masses in a given binary is conserved. This will lead to nearly equal neutron star masses in a binary, except for a small correction from the accretion by the first neutron star formed from the wind of its He star companion (Brown 1995). This explains why the neutron star masses in a given binary tend to be nearly equal. (In the standard scenario of evolution, where star B is rejuvenated by mass transfer, the companion neutron star B tends to be more massive than the pulsar A. This is the result found in the calculations of S. F. Portegies Zwart 1997, private communication.)

## 7. GRAVITATIONAL WAVES

In §§ 5 and 6 we have described how various types of compact binaries are formed. Once formed, they are subject to emission of gravitational waves. Shapiro & Teukolsky (1983) discuss gravitational waves and give a simple formula for the time required for a merger of the two stars in their equation (16.4.10),

$$T = \frac{5}{256} \frac{c^5}{G^3 M^2 \mu} R^4, \quad (7.1)$$

where  $M = M_A + M_B$  and

$$\mu = \frac{M_A M_B}{M_A + M_B}. \quad (7.2)$$

The masses here are the masses after both stars have become compact.

We are interested in the maximum initial distance permitted for the two stars to merge in a Hubble time, which we take to be

$$T_H = 10^{10} \text{ yr}. \quad (7.3)$$

Then,

$$R_{\max}^4 = 6.4 \frac{M_A M_B (M_A + M_B)}{M_\odot^3} R_{\text{Sh}}^2 c T_H, \quad (7.4)$$

where

$$R_{\text{Sh}} = \frac{2GM_\odot}{c^2} = 3.0 \text{ km} \quad (7.5)$$

is the Schwarzschild radius of the Sun.

Then,

$$R_{\max}^4 = R_0^4 \frac{M_A M_B (M_A + M_B)}{M_\odot^3}, \quad (7.6)$$

$$R_0^4 = 1.6 \times 10^{25} \text{ km}^4. \quad (7.7)$$

Taking the masses of case I (see § 5),

$$M_A = 2.4 M_\odot, \quad M_B = 1.4 M_\odot, \quad (7.8)$$

we get

$$R_{\max} = 3.8 \times 10^{11} \text{ cm} \approx 5 R_\odot. \quad (7.9)$$

However, we have shown in § 5 that after expulsion of the envelope, the distance between the two stars may be as low as  $1 \times 10^{11}$  cm, in which case the stars will merge in a small fraction of the Hubble time. The same is true of the Hulse-Taylor binary, as is well known.

As noted following equation (5.24), eccentricity in the final low-mass BH-NS binary leads to a value of  $a_f$  substantially larger than the  $3.8 \times 10^{11}$  cm given for merger in equation (7.9). In general, most of the final binaries will have  $e > 0.5$ , with a rapid rise just before  $e = 1$ . The rise occurs because preservation of the binary in the explosion is substantially greater if the kick velocity is opposite to the orbital velocity before explosion. In this case, the eccentricity  $\epsilon$  is large. The most favorable situation is when the orbital and kick velocities are equal in magnitude. (See figures in Wettig & Brown 1996.) P. Eggleton (1998, private communication) has kindly furnished us with a useful interpolation formula for the increase. The factor by which to multiply the  $T$  of equation (7.1), which refers to circular orbits, is

$$Z(e) \approx (1 - e^2)^{3.689 - 0.243e - 0.058e^2}. \quad (7.10)$$

This formula is accurate to about 1% for  $e \leq 0.99$ . Thus, if the initial eccentricity is 0.7, the time to shrink the orbit to zero is about 10% of the time required if the initial eccentricity were zero for the same initial period. The maximum  $a_f = 3.8 \times 10^{11}$  cm for circular orbits would be increased by the fourth root of the decrease in time; i.e., up to  $6.8 \times 10^{11}$  cm for this eccentricity. The maximum  $a_f$  in equation (5.24) would go up to  $3.4 \times 10^{13}$  cm, increasing the favorable logarithmic interval by  $\sim 40\%$ .

We wish to carry out a complete calculation of the distribution of eccentricities using the computer program of Wettig & Brown (1996) before assigning a definite number to the increase due to eccentricity. Note that this will affect our low-mass BH-NS binaries in the same way, leaving the ratio nearly unchanged. We note that Bloom, Sigurdsson, & Pols (1998), modifying the code created for binary evolution by Pols and Eggleton (see Pols 1994 and references therein) find an extremely low medium merging time of 2.1 Myr for binary neutron stars. This is more than 2 orders of magnitude smaller than the merging times of the narrow binaries 1913+16 and 1534+12, which are used to make estimates of the rate of mergers from observation. These binary neutron stars with very short lifetimes have understandably not been observed, but should be added in estimates made

from observations, such as those of van den Heuvel & Lorimer (1996).

#### 8. "OBSERVABILITY PENALTIES" FOR BLACK HOLE-NEUTRON STAR BINARIES

With two SNe per century, the formation rate of single pulsars is  $\sim 1.25 \times 10^{-2}$  per galaxy per yr. In fact, including binaries, as noted following equation (5.31), our total SN rate is 0.025 per yr. However, in only  $\sim 1/2$  of these cases is a neutron star formed, since ZAMS masses of  $\gtrsim 18 M_{\odot}$  evolve into black holes according to the Brown & Bethe (1994) scenario. (We believe that SN 1987A evolved into a low-mass black hole.) The single-pulsar formation rate of  $\sim 1.25 \times 10^{-2}$  per yr would seem to be the population to which our BH-NS binaries should be compared, since the neutron star in both cases is unrecycled. With a rate for the BH-NS binaries of  $10^{-4}$  per galaxy per yr, we should see several of these latter objects if this were true.

As noted in § 5, the binarity in NS-O/B binaries seems to severely inhibit observability, and it has been suggested that these binaries generally cannot get out of the Galactic disc, where their radio signals will be distorted by the dense matter.

The Hulse-Taylor pulsar is not far above the Galactic plane, at  $z = 0.26$  kpc. (Roughly half of the single pulsars are below this  $z$ , and half above.) The pulsar Wolszczan 1534 is at  $z = 0.51$  kpc, and the recently discovered 1518+49 is at 0.57 kpc, both high  $z$ . Both of these objects are very old, with estimated ages of 250 and 16,000 Myr. In all of these pulsars, the magnetic field is low,  $\sim 10^{10}$  G, so that their spin-down times are about 2 orders of magnitude greater than those of fresh pulsars. The unrecycled binary 2303+46 is at high  $z = 0.91$  kpc, presenting something of a mystery. According to our estimates of a production rate of  $10^{-2}$  for single pulsars and  $10^{-5}$  for binary pulsars, the latter corrected by a factor of  $\sim 100$  for longer observability, one might expect to see  $\sim 70$  binaries, given  $\sim 700$  single pulsars. Curran & Lorimer (1995), assuming the shape of the NS-NS luminosity function to be similar to that of normal pulsars, suggest that  $\sim 90\%$  of these binary pulsars are missed in current pulsar surveys. This would remove most of the discrepancy noted above. Pulsars in the Galactic plane are difficult to observe because the plane is full of dust and electrons that absorb or scatter electromagnetic radiation. This applies to single as well as binary pulsars. The difference is, however, that pulsars in a binary have greater difficulty in getting out of the Galactic plane. We suggest there is an "observability penalty" resulting from the fact that binary pulsars move more slowly out of the disc than single neutron stars because of their higher mass. This penalty would be greater for our BH-NS systems because they are more massive.

Furthermore, our BH-NS binaries will generally be narrow, with periods of  $\sim 2$ –18 hr, with more at the low end because of the  $a^{-1}$  distribution of binaries. Acceleration of the neutron star in its orbit in the binary will make it harder to find the signal by Fourier analysis.

We have not been able to quantify these penalties, and offer the above only as suggestions.

#### 9. CONCLUSION

We believe that at least 1% of massive binaries survive as binaries to ultimate merger. Taking the rate of SNe in

binaries to be 1 per century per galaxy, we find a merger rate of  $10^{-4}$  per yr per galaxy, 1 order of magnitude higher than previously believed.

The most important assumption for this conclusion is that the semimajor axes of binaries of heavy MS stars are distributed as  $da/a$ , and that this distribution extends out to  $a = 2 \times 10^8$  km (even farther once eccentricity in the final binaries is taken into account), or to orbital periods as long as 100 days (eq. [5.24]).

We suggest that most of the mergings, leading to gravitational waves, will be BH-NS binaries, rather than binary neutron stars. Earlier estimates are that the latter contribute  $\sim 10^{-5}$  per yr per galaxy. In our evolutionary scenario, binary neutron star systems only result from situations in which the first-born neutron star can escape the common-envelope evolution that otherwise sends it into a black hole. This can be realized if stars A and B are within  $\sim 5\%$  of each other in mass. Simple arguments show that this should result in  $\sim 10\%$  of the binaries, and scaling arguments show that about the same proportion of these selected binaries survive spiraling-in and end up close enough to merge, as in the BH-NS estimates. The 10 times higher rate for BH-NS mergers results, then, from the  $\sim 90\%$  of "failed" binary pulsar evolutions. This ratio is robust, given our results for the accretion onto the neutron star in the common envelope.

The reason for our high merger rate of  $10^{-4}$  is that only a moderate fraction of binaries get lost between the first supernova and the gravitational merger. In  $\sim 50\%$  of the cases, the second star B is heavy enough to end up in a supernova. About 40% survive the first supernova without splitting the binary. A small fraction are split by the second supernova kick. And another small fraction are lost because their separation, after both supernova events, is too great to permit merger by gravitational wave emission within a Hubble time.

A direct comparison with computer evolution is provided by case H of Portegies Zwart & Yungelson (1998). As noted following equation (5.31) and in § 6.1, for both binary pulsars and BH-NS binaries, agreement between our simple analytical evolution and their detailed numerical calculations is good,  $\lesssim$  a factor of 2 in the final results. If Portegies Zwart & Yungelson (1998) used the same assumptions for the rate of SNe, they would be even closer. Most importantly, our ratio of BH-NS mergings to binary neutron star mergings is nearly the same as theirs, 10. This indicates that the many detailed effects we leave out, e.g., widening of the orbits with mass exchange, are unimportant. We believe that our analytical work lends credence to the numerical work, and vice versa. It is much easier to assess changes that different effects would produce, so we believe it is worthwhile to have our simpler evolution.

Case H, which includes hypercritical accretion, was only one of eight cases studied by Portegies Zwart & Yungelson (1998). They remark that this case fails to reproduce the short-period binary pulsars. However, a further inclusion of the Wettig & Brown (1996) "observability premium" now underway by Portegies Zwart should improve this model.

In the standard scenario studied by Portegies Zwart & Yungelson (1998), without hypercritical accretion, masses of the pulsar progenitors are found to be substantially less than those of the companion progenitors (S. F. Portegies Zwart 1997, private communication). This is easily understood by rejuvenation of the companion following the

initial mass transfer to it. This runs contrary to observations; the pulsar is more massive than the companion neutron star in 1913+16, and the two are nearly equal in mass in 1534+12. The near equality in masses follows from our double He star scenario, the somewhat greater pulsar mass in 1913+16 pinning down the progenitor mass (Brown 1997).

A partial comparison can be made with the work of Iben, Tutukov, & Yungelson (1995), who used an extensive numerical scenario program developed at the Institute of Astronomy in Moscow several years ago to carry out in much more detail than given here the evolution of binary compact objects. Iben et al. (1995) obtain a rate of  $3 \times 10^{-4}$  per yr for merging neutron stars. With the inclusion of hypercritical accretion in the common-envelope evolution, 90% of these would become merging BH-NS binaries, to be compared with our rate of  $1.9 \times 10^{-4}$  per yr (our eq. [5.30]).

Their birthrate for high-mass X-ray binaries (our neutron stars with massive companions) is  $2.6 \times 10^{-3}$  per yr. Our rate is  $10^{-2} \times 0.58 \times 0.43 = 2.5 \times 10^{-3}$ , with kicks included. Their other results are difficult to compare with ours because of different basic assumptions.

Wettig & Brown (1996) suggested an observability premium for the longer spin-down time of binary pulsars once their magnetic field had been brought down (by accretion from the He star wind during the He star-NS stage in the Wettig & Brown evolution). We suggest an observability penalty for our higher mass BH-NS binaries, as a consequence of their greater difficulty in getting out of the Galactic disc. In the Galactic disk, they are more difficult to observe. Still, we believe that some of our BH-NS binaries should be observable, and we hope that our work is a challenge to observers.

Our results should be important for LIGO. At the same time, LIGO becomes essential for testing our ideas about massive binaries. Our rates in the Galaxy of  $10^{-4}$  per yr for BH-NS binaries<sup>3</sup> and  $10^{-5}$  per yr for NS-NS binaries extrapolate to rates per unit volume in the universe of  $8 \times 10^{-7} \text{ Mpc}^{-3}$  per yr for BH-NS binaries and  $8 \times 10^{-8} \text{ Mpc}^{-3}$  per yr for NS-NS binaries (Phinney 1991), which means that to see one event per yr, LIGO must look out to a distance of 70 Mpc for BH-NS and 150 Mpc for NS-NS binaries.

Because of their larger masses, the BH-NS binaries can be seen at greater distances by LIGO than the NS-NS binaries. Kip Thorne informs us that LIGO's first long gravitational wave search in 2002-2003 is expected to see BH-NS binaries (assuming masses of 2.4 and  $1.4 M_{\odot}$ ) to a distance

<sup>3</sup> Here and in the following discussion, BH means low-mass black hole. We are presently preparing a paper that estimates the rate for high-mass BH-NS binaries, such as would result from Cyg X-1.

of about 35 Mpc (too short by a factor of 2 according to our lower limit; see eq. [5.31]), and NS-NS to 25 Mpc (too short by a factor of 6). However, enhancements of the initial LIGO interferometers, planned for 2004, should reach out beyond 70 Mpc for BH-NS, bringing them into view, and other planned enhancements should reach NS-NS soon thereafter (see Thorne 1998).

LIGO will measure each binary's chirp mass  $M_{\text{chirp}} = \mu^{3/5} M^{2/5}$  to an accuracy of a few tenths of a percent (Poisson & Will 1995). We predict a bimodal distribution for these measurements; the BH-NS systems should display chirp masses concentrated near  $1.6 M_{\odot}$ , while the NS-NS systems should concentrate near  $1.2 M_{\odot}$ . We predict a ratio of event rates in LIGO of about 20 heavier systems to each lighter one (a factor of 10 from the rate per unit volume; a factor of 2 from seeing BH-NS systems farther than NS-NS).

Mergers of binary neutron stars and BH-NS have been considered in many papers as progenitors for gamma-ray bursters. If one of the neutron stars in the former does not collapse into a black hole, the radiated neutrinos will deposit their energy in lifting baryons from the strong gravitational potential instead of powering a relativistically expanding pair-plasma fireball. This can be avoided if one of the compact objects is a black hole, in which case baryons drop through the event horizon, carrying their binding energy with them (Ruffert & Janka 1998). This leaves a low baryon contamination of the fireball, which cannot be larger than  $\sim 10^{-5} M_{\odot}$ . The black hole must have a mass of less than  $5 M_{\odot}$ ; otherwise the neutron star would be swallowed before disruption (Rees 1997). Our many low-mass BH-NS binaries offer a heretofore unexpected number of possible progenitors.

The authors would like to thank Simon F. Portegies Zwart for many useful discussions and for providing us with unpublished material. We thank Lev Yungelson for correspondence and for sending us the Ergma & Yungelson paper in advance of publication. We are also grateful to C.-H. Lee for using the Wettig & Brown (1996) code to check a number of analytical calculations. We would like to thank Zaven Arzoumanian and Steve Thorsett for much advice about observability, and David Chernoff for giving us the results of Cordes & Chernoff (1997) long before publication. One of us (G. E. B.) is grateful to Madappa Prakash for many helpful conversations. Jim Lattimer checked many of our calculations and gave many helpful criticisms. G. E. B. was partially supported by the U. S. Department of Energy under grant DE-FG02-88ER40388. Kip Thorne helped us to relate our results more quantitatively to LIGO. We are grateful to Peter Eggleton for advice on the distribution of massive binaries and for the interpolation equation (7.10) for the effect of eccentricity.

#### REFERENCES

- Applegate, J. H., & Horner, J. 1998, in preparation
- Bloom, J. S., Sigurdsson, S., & Pols, O. R. 1998, MNRAS, in press
- Bodebeimer, P., & Taam, R. E. 1984, ApJ, 280, 771
- Braun, H., & Langer, N. 1995, A&A, 297, 771
- Brown, G. E. 1995, ApJ, 440, 270
- . 1997, Phys. Blatt., 53, 671
- Brown, G. E., & Bethe, H. A. 1994, ApJ, 423, 659
- Cordes, J. M., & Chernoff, D. F. 1997, preprint, astro-ph/970738
- Curran, S. J., & Lorimer, D. R. 1995, MNRAS, 276, 347
- Ergma, E., & Yungelson, L. R. 1997, A&A, in press
- Fryer, C., & Kalogera, V. 1997, ApJ, 489, 244
- Garnavich, C. D., Conti, P. S., & Massey, P. 1980, ApJ, 242, 1063
- Habets, G. M. H. J. 1986, A&A, 165, 95
- Iben, I., & Livio, M. 1993, PASP, 105, 1373
- Iben, I., Tutukov, A. V., & Yungelson, L. R. 1995, ApJS, 100, 217
- Phinney, E. S. 1991, ApJ, 380, L17
- Poisson, E., & Will, C. M. 1995, Phys. Rev. D, 52, 848
- Pols, O. R. 1994, A&A, 290, 119
- Portegies Zwart, S. F., & Verbunt, F. 1996, A&A, 309, 179
- Portegies Zwart, S. F., & Yungelson, L. R. 1998, A&A, 332, 173
- Rees, M. J. 1997, preprint, astro-ph/9701162
- Ruffert, M., & Janka, H.-J. 1998, in Proc. 4th Huntsville Gamma-Ray Burst Symposium, ed. C. A. Meegan, R. D. Preece, & T. M. Kochut (New York: AIP), 793
- Shapiro, S. L., & Teukolsky, S. A. 1983, Black Holes, White Dwarfs and Neutron Stars (New York: Wiley), 477

- Shima, E., Matanda, T., Takeda, H., & Sawada, K. 1985, MNRAS, 217, 367  
Thorne, K. S. 1998, in Black Holes and Relativistic Stars, ed. R. M. Wald  
(Chicago: Univ. Chicago Press), 41  
Van den Heuvel, E. P. J. 1995, J. Astrophys. Astron., 16, 255  
Van den Heuvel, E. P. J., & Lorimer, D. R. 1996, MNRAS, 283, L37  
Vrancken, J., DeGreve, J. P., Yungelson, L., & Tutukov, A. 1991, A&A,  
249, 411 (VDYT)  
Weaver, T. A., Zimmerman, G. B., & Woosley, S. E. 1978, ApJ, 225, 1021  
Wetzig, T., & Brown, G. E. 1996, New Astron., 1, 17  
Woosley, S. E., Langer, N., & Weaver, T. A. 1995, ApJ, 448, 315

---

## Chapter 13

# Contribution of High-Mass Black Holes to Mergers of Compact Binaries

H.A. Bethe and G.E. Brown

Reprinted with permission from *Astrophysical Journal*, 517 (1999) 318–327  
Copyright (1999) by American Astronomical Society

### **Commentary**

It was natural to continue the work on the merging of low-mass black-hole, neutron-star binaries to that of high-mass black-hole, neutron-star binaries. One advantage is that the most visible, continuously shining such object is Cygnus X-1, with  $\gtrsim 10M_{\odot}$  black hole and  $\sim 17M_{\odot}$  O-star companion (Herrero *et al.* 1995). On the other hand it is the *only* continuously shining such binary, except for LMC X-1 and LMC X-3 in the Large Magellanic Clouds, about which we have nothing to say. The O-star companion in Cyg X-1 is near its Roche Lobe, with an estimated time, Eq. (56), that it will be bright of  $2.7 \times 10^5$  yrs.

In Sec. 2 of this paper we repeat our estimate of compact object masses which will end up as neutron stars, low-mass black holes and high-mass black holes, always the same (which does not mean that they are correct, but does mean that throughout the course of the work summarized in this book, no observation has forced us to change our estimate).

Our chief conclusion is that the primary in a massive binary system must have a mass  $\gtrsim 80M_{\odot}$  in order for it to go into a high-mass black hole. This is a much higher mass than found earlier in the literature, but it should be remembered that this is in a binary. In Papers 20 and 21 we establish that only  $20M_{\odot}$  is required for a single star to go into a black hole. In this paper a factor  $\sim 3$  too high He-star wind losses were used, which could make this conclusion questionable. A better job is done in Paper 21 using more correct wind losses, with essentially the same conclusion.

We manage in this paper to increase the factor of  $\sim 20$  of the ratio of observable low-mass black-hole, neutron-star mergers to binary neutron star mergers by a factor of  $\sim 1.3$ , giving a total factor of  $\sim 26$  for black-hole, neutron-star mergers over those of binary neutron stars. Thus, if advanced LIGO, planned to be completed in  $\sim 2006$ , is estimated to see three mergers per year of binary neutron stars, then it should see  $\sim$  one merger per week of a black hole, low- or high-mass, and a neutron star. As we noted in the Commentary to Chapter 12, a binary high-mass black hole should contribute several times more.

In detailed calculation in Eq. (37) we incorrectly used the separation of the black hole and the O-star as  $a_1 = 17R_\odot$ . In fact, this is the Roche Lobe, and the correct separation ( $a_i = 33R_\odot$  before the explosion that formed the black hole,  $a_f \simeq 40R_\odot$  now) is used in Paper 21.

### Reference

- (1) Herrero, A., Kudritzki, R.P., Gabler, R., Vilchez, J.M., and Gabler, A. (1995), *Astronomy & Astrophysics* **297**, 556.

## CONTRIBUTION OF HIGH-MASS BLACK HOLES TO MERGERS OF COMPACT BINARIES

HANS A. BETHE

Floyd R. Newman Laboratory of Nuclear Studies, Cornell University, Ithaca, New York 14853

AND

G. E. BROWN

Department of Physics and Astronomy, State University of New York, Stony Brook, New York 11794-3800

Received 1998 June 16; accepted 1998 December 28

### ABSTRACT

We consider the merging of compact binaries consisting of a high-mass black hole and a neutron star. From stellar evolutionary calculations that include mass loss, we estimate that a zero-age main sequence (ZAMS) mass of  $\gtrsim 80 M_{\odot}$  is necessary before a high-mass black hole can result from a massive O star progenitor. We first consider how Cyg X-1, with its measured orbital radius of  $\sim 17 R_{\odot}$ , might evolve. Although this radius is substantially less than the initial distance of two O stars, it is still so large that the resulting compact objects will merge only if an eccentricity close to unity results from a high kick velocity of the neutron star in the final supernova explosion. We estimate the probability of the necessary eccentricity to be  $\sim 1\%$ ; i.e., 99% of the time the explosion of a Cyg X-1-type object will end as a binary of compact stars, which will not merge in Hubble time (unless the orbit is tightened in common envelope evolution, which we discuss later). Although we predict  $\sim 7$  massive binaries of Cyg X-1 type, we argue that only Cyg X-1 is narrow enough to be observed, and that only Cyg X-1 has an appreciable chance of merging in Hubble time. This gives us a merging rate of  $\sim 3 \times 10^{-8} \text{ yr}^{-1}$  in the galaxy, the order of magnitude of the merging rate found by computer-driven population syntheses, if extrapolated to our mass limit of  $80 M_{\odot}$  ZAMS mass for high-mass black hole formation. Furthermore, in both our calculation and in those of population syntheses, almost all of the mergings involve an eccentricity close to unity in the final explosion of the O star. From this first part of our development we obtain only a negligible contribution to our final results for mergers, and it turns out to be irrelevant for our final results. In our main development, instead of relying on observed binaries, we consider the general evolution of binaries of massive stars. The critical stage is when the more massive star A has become a black hole and the less massive star B is a giant reaching out to A. We then have a common envelope, and we expect hypercritical accretion to star A. Star A will accept a small fraction of the mass of the envelope of star B, but it will plunge deep into star B while expelling the envelope of star B. We expect that star B can at least be in the mass range  $15 \sim 35 M_{\odot}$ , while the black hole A has a mass of  $10 M_{\odot}$ . About 20% of the binaries of this type are found to end up in a range of orbital radii favorable for merging; i.e., outside of the relevant Roche lobes, but close enough so that these final binaries of compact objects will merge in Hubble time. The narrow black hole O star orbits do not seem to be found in population syntheses, because in them mergers happen almost completely as a result of kick velocities. In the exception (case H of Portegies Zwart & Yungelson, which includes hypercritical accretion), common envelope evolution is more effective and we are in agreement with their results. We find that the high-mass black hole neutron star systems contribute substantially to the predicted observational frequency of gravitational waves. We discuss how our high-mass black hole formation can be reconciled with the requirements of nucleosynthesis, and we indicate that a bimodal distribution of masses of black holes in single stars can account, at least qualitatively, for the many transient sources that contain high-mass black holes.

*Subject headings:* binaries: close — black hole physics — gravitation — stars: neutron

### 1. INTRODUCTION

The supernova (SN) community once believed that stars above a certain mass, about  $30\text{--}40 M_{\odot}$  zero-age main sequence (ZAMS), will collapse into a massive black hole (MBH) of mass of order  $10 M_{\odot}$ . The argument was that in these stars the mantle was bound with binding energy well above  $10^{51}$  ergs so that the SN shock was not strong enough to expel it.

Whereas this may be true for single stars, Woosley, Langer, & Weaver (1995) showed that in binaries, where the hydrogen envelope of the primary star has been transferred to the companion in Roche lobe overflow (RLOF), the evolution of the resulting “naked” He star (i.e., the star without a hydrogen envelope) led to a substantially smaller pre-SN

core than that of a single star with a hydrogen envelope. A comparison of compact core masses from naked He stars and those evolved by Woosley & Weaver (1995) for single stars is shown in Figure 1, taken from Brown, Weingartner, & Wijers (1996). Detailed reasons for the great difference in the evolution of “clothed” and naked He cores are given in Woosley et al. (1995).

Stars with ZAMS masses  $\gtrsim 40 M_{\odot}$  lose their masses by strong winds, whether in binaries or not, and become Wolf-Rayet (W-R) stars. In an earlier paper, Woosley, Langer, & Weaver (1993) investigated ZAMS masses of 35, 40, 60, and  $85 M_{\odot}$ . In those up through  $60 M_{\odot}$ , the hydrogen envelope was blown off early enough for the He cores to evolve as naked cores, and compact core masses were around  $1.5 M_{\odot}$ .

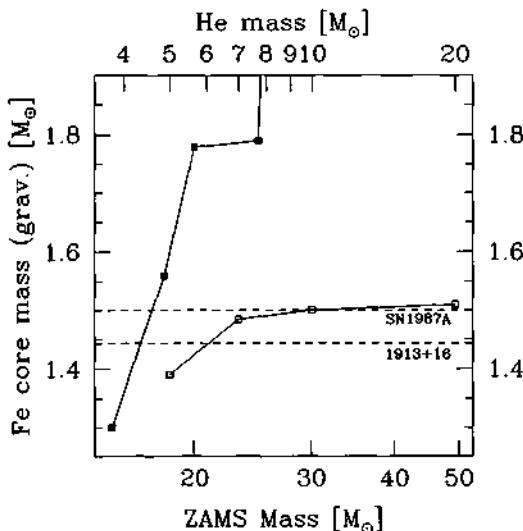


FIG. 1.—Comparison of the compact core masses resulting from the evolution of single stars (filled squares), case B of Woodsley & Weaver (1995), and naked helium stars (Woodsley et al. 1995) with masses equal to the corresponding He core mass of single stars. The horizontal dashed lines indicate the mass of the heaviest known well-measured pulsar and the maximum mass of a neutron star.

(gravitational). In fact, with inclusion of extensive mass loss the lower line in Figure 1, which heads just above  $1.5 M_{\odot}$  for the higher ZAMS masses, gave the Woodsley et al. (1993) correspondence of Fe core mass to ZAMS mass for single stars of masses  $35-60 M_{\odot}$ . Thus rapid mass loss by wind in this mass region that removes the H envelope before appreciable He core burning begins leaves a He core, which burns as a naked core. (In the case of the massive stars, the relation shown in Figure 1 between Fe core mass and He core mass no longer holds because of large wind losses in the latter.) In the case of the  $85 M_{\odot}$  star, some hydrogen envelope remained during an appreciable part of the He burning, so the He core burned as a (partially) clothed core, and the compact core was more massive, in the range  $1.7-2.0 M_{\odot}$ , depending on the  $^{12}\text{C}(\alpha, \gamma)^{16}\text{O}$  burning rate. For the very massive stars, the He core burns as if clothed at least part of the time.

## 2. THE COMPACT STAR

Thorsson, Prakash, & Lattimer (1994) and Brown & Bethe (1994) have studied the compact core after the collapse of an SN, assuming reasonable interactions between hadrons. Initially, the core consists of neutrons, protons, and electrons, and a few neutrinos. It has been called a proto-neutron star. It is stabilized against gravity by the pressure of the Fermi gases of nucleons and leptons, provided its mass is less than a limiting mass  $M_{pc}$  (proto-compact) of  $\sim 1.8 M_{\odot}$ .

If the assembled core mass is greater than  $M_{pc}$ , there is no stability and no bounce; the core collapses immediately into a black hole. It is reasonable to take the core mass to be equal to the mass of the Fe core in the pre-SN, and we shall make this assumption, although small corrections for fall-

back in the later SN explosion can be made as in Brown, Weingartner, & Wijers (1996). If the center collapses into a black hole, the outer part of the star has no support and will also collapse. We then get a massive black hole containing the entire mass of the pre-SN star, perhaps of order  $\lesssim 10 M_{\odot}$ . (At least in binaries the companion star will have removed the hydrogen envelope by either RLOF or by common envelope evolution, depending on the mass of the companion, and wind will diminish the He core before the SN explosion.)

If the mass of the core is less than  $M_{pc}$ , the electrons will be captured by protons,

$$p + e^- \rightarrow n + \nu, \quad (1)$$

and the neutrinos will diffuse out of the core. This process takes of order 10 s, as has been shown by the duration of the neutrino signal from SN 1987A. The result is a neutron star with a small concentration of protons and electrons. The Fermi pressures of the core are chiefly from the nucleons, with small correction from the electron. On the other hand, the nucleon energy is increased by the symmetry energy; i.e., by the fact that we now have nearly pure neutrons instead of an approximately equal number of neutrons and protons. Thorsson et al. (1994) have calculated that the maximum mass of the neutron star  $M_{ns}$  is still about  $1.8 M_{\odot}$ ; i.e., the symmetry energy compensates the loss of the Fermi energy of the leptons. Corrections for thermal pressure are small (Prakash et al. 1997).

The important fact is that the 10 s of neutrino diffusion from the core give ample time for the development of a shock, which expels most of the mass of the progenitor star.

But this is not the end of the story. The neutrons can convert into protons plus  $K^-$  mesons,

$$n \rightarrow p + K^-. \quad (2)$$

Since the density at the center of the neutron star is very high, the energy of the  $K^-$  is very low, as confirmed by Li, Lee, & Brown (1997) using experimental data. By this conversion the nucleons can again become half neutrons and half protons, thereby saving the symmetry energy needed for pure neutron matter. The  $K^-$ , which are bosons, will condense, saving the kinetic energy of the electrons they replace. The reaction (eq. [2]) will be slow, since it is preceded by

$$e^- \rightarrow K^- + \nu, \quad (3)$$

which is actually effected in the star by reactions such as  $e^- + p \rightarrow K^- + p + \nu$ , with strangeness breaking. (Times are long enough for chemical equilibrium to be realized.) It becomes energetically advantageous to replace the fermionic electrons by the bosonic  $K^-$  at higher densities. Initially, the neutrino states in the neutron star are filled up to the neutrino chemical potential with trapped neutrinos, and it takes some seconds for them to leave the star. These must leave before new neutrinos can be formed from the process (eq. [3]). Thorsson et al. (1994) have calculated that the maximum mass of a star in which reaction (2) has gone to completion is

$$M_{np} \approx 1.5 M_{\odot}, \quad (4)$$

where the lower suffix "np" denotes their nearly equal content of neutrons and protons, although we continue to

use the usual name of neutron star. This is the maximum mass of neutron stars, which is to be compared with the masses determined in binaries. The masses of 19 neutron stars determined in radio pulsars (Thorsett & Chakrabarty 1999) are consistent with this maximum mass.

The core mass  $M_c$  formed by the collapse of SN must, therefore, be compared to the two limiting masses,  $M_{pc}$  and  $M_{np}$ . If

$$(1) M_c > M_{pc}, \quad (5)$$

we get a high-mass black hole of mass essentially equal to the full mass of the pre-SN star. If

$$(2) M_{pc} > M_c > M_{np}, \quad (6)$$

we get a low-mass black hole of mass  $M_c$ . Only if

$$(3) M_c < M_{np} \quad (7)$$

do we get a neutron (more precisely, "nucleon") star from the SN. Only in this case can we observe a pulsar. In cases (2) and (3), we can see an SN display. In case (1), only initial neutrinos from electrons captured in the collapse before  $M_c$  becomes greater than  $M_{pc}$  can be observed, but no light would reach us.

We tentatively choose the lower limit of ZAMS mass for making MBHs to be  $80 M_\odot$ . On the other hand, it is believed that ZAMS above  $100 M_\odot$  do not exist because of excessive formation of electron pairs. So we assume that the range of  $80\text{--}100 M_\odot$  is available.

### 3. RATE OF FORMATION

We are interested in massive binaries containing one star of ZAMS mass between  $80$  and  $100 M_\odot$ . As in Bethe & Brown (1998), we start from the assumption that there is one SN per century per galaxy in a binary. Assuming also that  $10 M_\odot$  is required for a star to end up as an SN of Type II (or Ib or Ic), the formation of binaries of  $M > 10 M_\odot$  is also  $10^{-2} \text{ yr}^{-1} \text{ galaxy}^{-1}$ . We assume a Salpeter function with index  $n = 1.5$ ; then the fraction of such stars between  $80$  and  $100 M_\odot$  is

$$f = 8^{-3/2} - 10^{-3/2} = 1.26 \times 10^{-2}. \quad (8)$$

So the rate of formation is

$$\alpha_1 = 1.26 \times 10^{-4} \text{ yr}^{-1} \text{ galaxy}^{-1}. \quad (9)$$

We require that this star A be accompanied by a companion B of  $M > 10 M_\odot$ . Assuming that the distribution of mass of the companion is  $dq$ , with

$$q = M_B/M_A, \quad (10)$$

and having assumed  $M_A = 90 M_\odot$  on average, we need  $q > \frac{1}{9}$ , which has a probability

$$1 - q = 0.9, \quad (11)$$

hence a formation rate

$$\alpha_2 = (1 - q)\alpha_1 = 1.13 \times 10^{-4} \text{ yr}^{-1}. \quad (12)$$

In order to observe strong X-rays from the MBH, the distance  $a$  between star A (the MBH) and star B (an O or B

star) should not be too large, let us say less than  $150 R_\odot$ . On the other hand, it must not be less than  $30 R_\odot$ , because otherwise the two stars would merge already at this stage of evolution. Assuming as in Bethe & Brown (1998) a distribution  $da/7a$ , the probability that  $a$  falls in the desired limits is

$$p = 7^{-1} \ln(150/30) = 0.23, \quad (13)$$

giving for the probability of formation (per galaxy)

$$\alpha_3 = 0.23\alpha_2 = 2.6 \times 10^{-5} \text{ yr}^{-1}. \quad (14)$$

Although we take the same logarithmic distribution as in Bethe & Brown (1998), we think of the lower limit as greater than  $2 \times 10^7$  km so that the more massive stars we deal with here lie inside their Roche lobes, and the upper limit as less than  $2 \times 10^{10}$  km, although the latter is uncertain because O star binaries probably will not be recognized for such a large separation (Garmany, Conti, & Massey 1980). High-mass transfers from the O star occur only when it nearly fills its Roche lobe. It is then bright for a time (Massevich et al. 1979)

$$\tau = 2.7 \times 10^5 \text{ yr}, \quad (15)$$

so the expected number of strong X-ray-emitting binaries in the galaxy is

$$N = \alpha_3 \tau = 7. \quad (16)$$

Only one such binary has been observed, Cyg X-1, at a distance of 2.5 kpc. It might be thought that X-ray binaries at larger distances have escaped detection because of absorption of X-rays in the highly ionized galactic medium. But, in fact, X-rays of energy greater than 3 keV should be seen throughout the galaxy, penetrating even the Galactic disc. On the other hand, we are probably seeing the closest binary of Cyg X-1 type, with orbital separation only about half of the initial separation in the double O star progenitor. Thus other such objects may well have substantially lower luminosities. The O star in such a binary would probably be in the disc and might not be observed at greater distances, so that the X-rays could not be associated with the binary. In any case, it is clear that in the disc there is only one bright Cyg X-1 type object, a fact we use below.

In the Magellanic clouds with less than 1/10 the mass of the galaxy, two high-mass black hole binaries, LMC X-3 with a B star companion and LMC X-1, where the donor is probably an O star, have been observed. We believe the high incidence of high-mass black holes in the LMC to be not only a consequence of the large amount of star formation there, but also of the low metallicity compared with solar. We return to a discussion of high-mass black hole formation in § 8.

### 4. MERGERS

In the case of low-mass black holes, the rate of mergers is limited chiefly by the disruption of the binary due to the recoil that star B experiences when it goes SN. For MBHs this is of little concern since the orbital velocity of the binary before the SN process is already of order  $600 \text{ km s}^{-1}$  (see below), comparable to the higher recoil velocities, so that the recoil is unlikely to disrupt the binary.

Instead of this, the concern for MBHs is that two stars of the binary may be too far apart so they will not merge

during Hubble time (assumed to be  $10^{10}$  yr). Note that the progenitor binary must have a separation greater than  $30 R_\odot$  in order that the two O stars do not fill their Roche lobes. At present, the separation in Cyg X-1 is only slightly more than half this, which is permissible because now only the O star needs to be in its Roche lobe, and the black hole has negligible radius. But as we show below, even this small separation is still large enough to substantially cut down the probability of merging within Hubble time. Because of rapid mass loss in such massive stars of ZAMS 80–100  $M_\odot$ , there is no well-established procedure for calculating their evolution in binaries. Starting from a separation greater than  $30 M_\odot$ , mass exchange by the very massive star with the lower mass companion when the former is on the main sequence (case A mass transfer) could decrease the separation substantially, perhaps causing the stars to coalesce. However, the resulting binary would then widen (roughly back to its original separation) as the W-R, which remains after mass transfer from the very massive star, loses mass by wind. We do not have any evidence of the net change (S. Portegies Zwart 1998, private communication).

Because of the uncertainties in the evolutionary scenario, we adopt an empirical approach beginning from the measured separation  $a_1$  in Cyg X-1. The requirement that the two stars merge in Hubble time then sets a lower limit on the eccentricity  $e$  of the final neutron star black hole orbit:

$$a_2(1 - e^2) < a_0, \quad (17)$$

where  $a_2$  is the semimajor axis of the orbit after the SN event, and  $e$  is the eccentricity of the orbit;  $a_0$  will be calculated in § 5.

The left-hand side of equation (17) is related to the angular momentum  $J$  of the post-SN orbit. In fact,

$$J = a_2 V_3(1 - e^2)^{1/2}, \quad (18)$$

where  $V_3$  is the orbital velocity if the orbit were circular. (We do not need to consider the mass of the system.) By Newton's laws,

$$a_2 V_3^2 = G(M_A + M_n), \quad (19)$$

where  $M_A$  is the mass of the MBH and  $M_n$  is that of the neutron star resulting from the SN of star B. Hence

$$J^2 = G(M_A + M_n)a_2(1 - e^2). \quad (20)$$

On the other hand, immediately after the SN the distance between stars A and B is still  $a_1$ , and the relative velocity is

$$V_2 = V_1 + Q, \quad (21)$$

where  $Q$  is the recoil velocity due to the SN. The angular momentum is

$$\begin{aligned} J &= a_1 \times V_2, \\ J^2 &= a_1^2 V_2^2 \sin^2 \psi, \end{aligned} \quad (22)$$

where  $\psi$  is the angle between  $V_2$  and  $a_1$ .

This angle can be calculated, but it is complicated. We now make the approximation of replacing  $\sin^2 \psi$  by

$$\langle \sin^2 \psi \rangle = \frac{2}{3}. \quad (23)$$

Then, equating equations (20) and (22) and using equation (17),

$$a_1^2 V_2^2 < \frac{3}{2} G(M_A + M_n)a_0. \quad (24)$$

Equation (23) involves a double approximation: (1)  $\psi$  is uniformly distributed over the sphere, and (2) only the average of  $\sin^2 \psi$  is used instead of calculating the limit on  $a_1 V_2$  separately for each possible direction of  $Q$  and then averaging the results. We believe that a correct calculation would give a somewhat higher limit on  $a_1 V_2$ .

The product  $a_1 V_2^2$  can be calculated assuming that the orbit of the binary is circular before the SN event, an assumption that is probably very nearly correct. Then, similar to equation (19),

$$a_1 V_1^2 = G(M_A + M_B), \quad (25)$$

where  $M_B$  is the mass of star B, the companion of the MBH, before the SN event. In Cyg X-1, it is believed that the mass of the MBH is  $M'_A = 10 M_\odot$ , and that of the O star companion is  $M'_B = 17 M_\odot$  (Herrero et al. 1995). After RLOF of the O star, these masses happen to be just exchanged, assuming conservative mass transfer, so that immediately before the SN

$$M_A = 17 M_\odot, \quad M_B = 10 M_\odot. \quad (26)$$

However, we shall use "standard" values

$$M_A = M_B = 10 M_\odot. \quad (27)$$

The SN changes star B to a neutron star,  $M_n = 1.4 M_\odot$ , while star A remains unchanged. We now have

$$a_1^2 V_1^2 = G(M_A + M_n)a_1. \quad (28)$$

Dividing equation (24) by equation (28),

$$\frac{V_2^2}{V_1^2} < \frac{3}{2} \left( \frac{M_A + M_n}{M_A + M_B} \right) \frac{a_0}{a_1} \equiv B. \quad (29)$$

Usually,  $a_1 > a_0$ , so  $V_2 < V_1$ . From equation (21), this means that the recoil velocity  $Q$  of the SN must be approximately opposite to the pre-SN orbital velocity  $V_1$  and of the same order of magnitude. Setting

$$Q \cdot V_1 = -QV\mu, \quad (30)$$

$$Q/V_1 = \omega, \quad (31)$$

thus equation (29) becomes

$$1 + \omega^2 - 2\omega\mu < B \quad (32)$$

$$1 - \mu < \frac{B - (1 - \omega)^2}{2\omega}. \quad (33)$$

Since B is, in general, considerably smaller than 1, we thus get the condition

$$|1 - \omega| < B^{1/2}, \quad (34)$$

<sup>1</sup> Later we shall show that common envelope evolution of the black hole in the evolving O star is more likely.

stating once more that  $Q$  must be of the same order as  $V_1$ . The fraction of the total solid angle available to  $Q$  is  $(1 - \mu)/2$ .

Now, according to equation (28),

$$V_1^2 = \frac{2}{3} \times \frac{10^{-7} \times 20 \times 2 \times 10^{33}}{a_1} = 2.7 \times 10^{15} \times \frac{10^{12}}{a_1}, \quad (35)$$

$$V_1 = 520 \text{ km s}^{-1} \left( \frac{10^{12}}{a_1} \right)^{1/2}. \quad (36)$$

In Cyg X-1, the distance between the two stars is

$$a_1 = 17 R_\odot = 1.2 \times 10^{12} \text{ cm}. \quad (37)$$

So  $a_1$  is likely to be of order  $10^{12}$ . This means  $V_1$  is of order  $500 \text{ km s}^{-1}$ , and  $Q$  is likewise. Now Cordes & Chernoff (1997) have found that the recoil velocities of pulsars are distributed bimodally with  $\sigma = 175$  and  $700 \text{ km s}^{-1}$ , respectively. The smaller velocity Gaussian is negligible for our  $Q$ . For the larger  $\sigma$ , we set

$$x \equiv Q/\sigma. \quad (38)$$

The Gaussian distribution is

$$\int dx = 2\pi^{-1/2} x^2 dx e^{-x^2}. \quad (39)$$

We assume  $V_1 = \sigma$  for the Gaussian with the greater  $\sigma$ , and then for this Gaussian  $\omega = x$ , and the probability is (cf. eq. [33])

$$\int \frac{1}{4x} [B - (1 - x)^2] dx \approx \frac{2e^{-1}}{3\sqrt{\pi}} B^{3/2} = 0.14 B^{3/2}. \quad (40)$$

For large  $B$ , substitute  $0.9B$  for  $B^{3/2}$ . Considering that only 20% of pulsar recoils are in the larger  $\sigma$  Gaussian, the fraction of binaries containing massive black holes that merge during Hubble time is

$$\phi = 0.03 B^{3/2}, \quad (41)$$

with  $B$  given by equation (29). Inserting our values for  $M_A$ ,  $M_B$ , and  $M_*$ ,

$$B = 0.85 a_0/a_1. \quad (42)$$

It remains to calculate  $a_0$ .

## 5. NUMERICAL MERGERS

If the orbit is circular, a merger will take place within Hubble time if the initial distance is less than  $R_{\max}$ , with  $R_{\max}$  given by equation (7.7) of Bethe & Brown (1998):

$$R_{\max}^4 = 1.6 \times 10^{25} \text{ km}^4 \frac{M_A M_B (M_A + M_B)}{M_\odot^3}. \quad (43)$$

$R$  is the same as the semimajor axis introduced in § 4, so  $a_0 = R_{\max}$ . If the orbit is eccentric, the merger time for a given major axis is diminished by a factor (P. Eggleton 1998, private communication)

$$Z(e) = (1 - e^2)^{3.689 - 0.243e - 0.058e^2}. \quad (44)$$

Therefore, the condition on the initial semimajor axis,  $a_2$ , is

$$a_2^4 Z(e) < R_{\max}^4. \quad (45)$$

Taking the fourth root,

$$a_2 Z^{1/4} < R_{\max}, \quad (46)$$

$Z^{1/4}$  is almost  $1 - e^2$ , so the left-hand side of equation (46) is almost  $a_2(1 - e^2)$ , thus

$$a_2(1 - e^2) < R_{\max}(1 - e^2)^{0.0777 + 0.0608e + 0.0195e^2}. \quad (47)$$

The last factor depends only slightly on the eccentricity. So we replace in it  $e^2$  by an average of 0.5. This makes that factor

$$0.5^{0.130} = 0.914. \quad (48)$$

Inserting in equation (43)  $M_A = 10 M_\odot$ ,  $M_B = 1.4 M_\odot$ ,

$$R_{\max} = 7.1 \times 10^6 \text{ km}. \quad (49)$$

The upper limit of  $a_2(1 - e^2)$  is the definition of  $a_0$ , according to equation (17). Thus the right-hand side of equation (47) equals  $a_0$ , and we have

$$a_0 = 6.5 \times 10^6 \text{ km}, \quad (50)$$

and equation (29) becomes

$$B = 5.1 \times 10^6 \text{ km}/a_1. \quad (51)$$

The measured distance between the two stars in Cyg X-1 is  $1.2 \times 10^7 \text{ km}$ . If this distance does not change in RLOF, then  $a_1 = 1.2 \times 10^7 \text{ km}$ ,

$$B = 0.42, \quad (52)$$

$$\phi = 0.009. \quad (53)$$

This estimate can be checked from Figure 4a of Kalogera (1996). In our situation, her parameters are  $\beta \approx 0.6$ ,  $\xi \approx 0.9$ . Using the Eggleton formula (eq. [44]), we find that  $e \gtrsim 0.76$  for a merger. The average survival probability is  $\sim 0.3$  for these eccentricities, and our  $\xi$  is  $\sim 0.3$ , so that the final probability is  $\frac{1}{3} \times 0.24 \times 0.3 \simeq 0.014$ , close to our  $\phi$ .

Multiplying this by the formation rate  $\alpha_3$ , equation (14), we find for the rate of mergers

$$\alpha_3 \phi = 2 \times 10^{-7} \text{ yr}^{-1}. \quad (54)$$

So this is a rare event, as we might expect. In fact, equation (54) is probably still an overestimate because we predicted seven X-ray-emitting binaries in the galaxy (eq. [16]), and we see only Cyg X-1. In our assumed distribution of these objects out to  $150 R_\odot$ , the others would be expected to have substantially larger  $a_1$  than Cyg X-1, in which case their merger rate would be negligible. Thus a possible guess of the rate of mergers is not much larger than  $1/7$  of equation (54).

We compare our result with results from population synthesis:

- First, we note that the smallness of the rate of mergers results from the separation at the black hole O star stage being too large for merger (for spherical orbits) without a high eccentricity from kick velocities. Whereas Portegies Zwart & Yungelson (1998) obtain a (high-mass) black hole neutron star birthrate of  $18.7 \times 10^{-8} \text{ yr}^{-1}$  in the galaxy without kick velocities, their merger rate is zero for this case. Lipunov, Postnov, & Prokhorov (1997) find an order

of magnitude greater merger rate with inclusion of kick velocities than without.

2. Portegies Zwart & Yungelson (1998)<sup>2</sup> use  $40 M_{\odot}$  for their ZAMS mass for the black hole limit, and we compare with those for  $40 M_{\odot}$  from Lipunov et al. (1997). Increasing the mass limit to  $80 M_{\odot}$  and taking into account the factor  $1 - q$  would decrease the results of these authors by a factor of  $\sim 5$ . Thus Portegies Zwart & Yungelson (1998) would have a merging rate of  $2 \times 10^{-7} \text{ yr}^{-1}$ , and Lipunov et al. (1997) would have one of  $\sim 4 \times 10^{-8} \text{ yr}^{-1}$  for our mass limit of  $80 M_{\odot}$ . Our rate obtained by taking  $\frac{1}{2}$  of equation (54) is

$$R = 3 \times 10^{-8} \text{ yr}^{-1}, \quad (55)$$

about the same as the latter. We believe that our simple considerations give some understanding of the results obtained in the latter population synthesis, and we shall return to a discussion of the work of Portegies Zwart & Yungelson (1998) later.

We can also try an "observational estimate" of the merging of Cyg X-1 type objects, although we have only one such object near its Roche lobe. Assuming, as in equation (15), Cyg X-1 to be bright for a time  $\tau = 2.7 \times 10^5 \text{ yr}$ , the birth rate of such bright objects is

$$\tau^{-1} = 3.7 \times 10^{-6} \text{ yr s}^{-1}. \quad (56)$$

Given the merging probability of the resulting black hole neutron star binary  $\phi = 0.009$  from equation (53), we arrive at a merging rate of  $4 \times 10^{-8} \text{ yr}^{-1}$  for the Galaxy, in rough agreement with equation (55). This may be an underestimate since Cyg X-1 will probably go into stable mass transfer after losing substantial mass to the black hole. LMC X-3 with roughly equal black hole and B star masses (Kuiper et al. 1997) may already be in stable mass transfer. Thus the  $2.7 \times 10^5 \text{ yr}$  may well be an underestimate.

Although the last two sections give results for mergers that are negligible compared to our final results in § 7, once hypercritical accretion is included in our common envelope evolution we believe that they are instructive in that they match rather well the results of population synthesis, especially those of Lipunov et al. (1997). In Portegies Zwart & Yungelson (1998), (generally unstable) mass transfer starts on the dynamical timescale of the donor, changing along the way to the thermal timescale, and the last part of the mass is sometimes even transferred on nuclear timescales (S. Portegies Zwart 1998, private communication). Their initial separations of the O star binaries are nearly double our separations. (We take ours before substantial main-sequence burning, which increases the stellar radius by a factor of  $\sim 2$ .) Thus their common envelope evolution does tighten the orbits, but clearly not enough for a substantial merger of the compact objects evolved later. The exception is case H of Portegies Zwart & Yungelson (1998), the case similar to ours with inclusion of hypercritical accretion, which gives a merger rate  $\sim 35$  times that of cases B–G. We develop in the next section why hypercritical accretion is so effective.

<sup>2</sup> We compare with their case H in which hypercritical accretion was included in the next section.

## 6. COMMON ENVELOPE EVOLUTION

In the last two sections we have addressed the evolution of massive binaries in which the primary evolves into a high-mass black hole in an empirical way. We match rather well the population synthesis results of Lipunov et al. (1997). In particular, these authors find that introduction of kick velocities increases their merging rate by an order of magnitude. Portegies Zwart & Yungelson (1998) find zero mergers without kick velocities. We believe this to result from the fact that their black hole O star binaries lie too far apart for the resulting binary of compact objects to merge in Hubble time without the increase in merger distance given by an eccentricity close to unity; i.e., these authors have very little (or no) merging from circular black hole neutron star orbits, except for their case H that includes hypercritical accretion, which we discuss below.

Clearly the population syntheses of Lipunov et al. (1997) do not include efficient common envelope evolution that brings the compact object (high-mass black hole) close to the He core of the companion star.

We now develop our scenario for common envelope evolution, which hypercritical accretion makes particularly plausible, in that we have quantitative control over the crucial quantities. As the envelope of the giant expands to meet the compact object one can see from Rasio & Livio (1996) that the common envelope evolution begins rather quickly. Soon after the beginning of mass transfer the compact object creates a tidal bulge in the evolving companion, transferring angular momentum to the companion. As the giant companion loses mass, the isentropic envelope responds by expanding. The compact object plunges into the companion, the chief loss in orbital energy occurring in  $\lesssim 1 \text{ yr}$  in time (Terman, Taam, & Hernquist 1995; Rasio & Livio 1996).

In the literature, common envelope evolution is thought to occur when the two stars involved differ in mass more than a factor of 2–3. This condition arises because at higher mass ratios the mass-receiving star will then not be able to accept the large amount of mass transferred to it on the short thermal timescale of the companion. This thermal timescale is (van den Heuvel 1994)

$$\tau_{th} = \frac{GM^2}{RL} \simeq \frac{3 \times 10^7}{(M/M_{\odot})^2}. \quad (57)$$

In fact, the factor of 2–3 in  $q$  initially came from a factor of  $\sim 10$  in  $\tau_{th}$  from the work of Kippenhahn & Meyer-Hofmeister (1977) who considered case A mass transfer (during main-sequence evolution). In case B mass transfer, the transfer takes place as the giant traverses the Herzprung gap, which is much more quickly than in case A, so it is not clear why  $q$  could not be substantially reduced and still have common envelope evolution.

The situation with hypercritical accretion is different from either case A or case B (RLOF) mass transfer. Initially, our  $\sim 10 M_{\odot}$  black hole is met by the expanding red giant or supergiant envelope. The black hole accretes some of the matter and transfers enough energy to the remaining matter to expel it. The convective envelope has constant entropy and must expand in order to replace the accreted and expelled matter. The black hole drops in gravitational potential closer to the He core of the companion in order to furnish the necessary energy to expel most of the matter. This whole process happens very rapidly. (In our case of

hypercritical accretion, an accretion disc will be set up around the black hole, which cannot immediately accept the matter because of the high angular momentum of the latter. We assume the viscosity to be high enough so that angular momentum, but little mass, will be advected outward. Observed masses of  $\sim 7 M_{\odot}$  of high-mass black holes in transient sources discussed by Brown, Lee, & Bethe 1999 substantiate this scenario.)

Hypercritical accretion sets in when the envelope density reaches  $\sim 10^{-9} \text{ cm}^{-3}$ . From equations (5.3) and (5.7) of Bethe & Brown (1998), we can show that the contribution to the coefficient of dynamical friction  $c_d$  from matter inside the accretion radius  $R_{sc}$ , i.e., from matter that can accrete onto the compact object, is  $c_d = 2$ . Now the total coefficient of dynamical friction is  $c_d = 6-8$  (Ruffert 1994; Ruffert & Arnett 1994). The remainder of  $c_d$  comes chiefly from the wake, at greater distances than  $R_{sc}$ . Thus for

$$c_d \gg 2, \quad (58)$$

we encounter highly nonconservative mass transfer, and it is plausible that common envelope evolution ensues for a wider range of  $q$  than usually thought. In fact, as we show below, out of the  $\sim 20 M_{\odot}$  hydrogen envelope of a  $30 M_{\odot}$  companion star, only  $\sim 3 M_{\odot}$  is accepted by the black hole.

From the above argumentation we believe that  $c_d \gg 2$  strongly favors common envelope evolution. (For  $c_d = 2$ , the mass transfer is nearly conservative.) The companion star should have mass less than  $35-40 M_{\odot}$ , because for higher masses it loses mass in an LBV phase. It should have somewhat higher mass than the black hole, say,  $15 M_{\odot}$  in our case, so that mass transfer is unstable throughout the transfer of envelope mass, but the final transfer is so rapid that it will overshoot into the range of stable mass transfer. As noted, Cyg X-1 is probably included in our interval of masses favorable for common envelope evolution.

Having established the plausibility of common envelope evolution for the black hole in the expanding H envelope of an O star with ZAMS mass of  $15-35 M_{\odot}$ , we now carry out this common envelope evolution for a typical O star mass of  $20 M_{\odot}$  following Bethe & Brown (1998). The He core is  $6 M_{\odot}$ .

We choose the coefficient of dynamical friction  $c_d$  to be 6. In the Bethe & Brown (1998) notation,  $M_A$  is the compact object mass,  $M_B$  the companion mass, and

$$Y = M_B a^{-1}. \quad (59)$$

The initial and final  $Y$  are related by

$$\left(\frac{Y_f}{Y_i}\right)^{1+1/(c_d-1)} = \frac{2.4 M_{B_i}}{M_{A_i}}, \quad (60)$$

depending only on the initial masses. The ratio of final to initial black hole masses is given by

$$\frac{M_{A_f}}{M_{A_i}} = \left(\frac{Y_f}{Y_i}\right)^{1/(c_d-1)} = \left(\frac{Y_f}{Y_i}\right)^{1/5} = 1.3, \quad (61)$$

so the final black hole mass is  $13 M_{\odot}$ , for an initial  $10 M_{\odot}$ . We find

$$\frac{a_i}{a_f} = \frac{M_{B_i}}{M_{B_f}} \frac{Y_f}{Y_i} = 12, \quad (62)$$

where  $M_{B_f}$  is the companion He core mass of  $6 M_{\odot}$ .

## 7. MERGERS AFTER COMMON ENVELOPE

The most important result of the calculation at the end of § 6 is that  $a_i/a_f$  is large, of the order of 12. In equation (62), there are two factors:  $Y_f/Y_i$  is fairly large, about 4;  $M_{B_f}/M_{B_i}$  is also fairly large, the ratio of the entire mass of the star to its He core taken to be  $(0.3)^{-1}$  by Bethe & Brown (1998).

Thus the radius of the orbit shrinks by a large factor in the common envelope. We know the range of the final radius  $a_f$ ; its maximum is given by the condition that the merger should occur within Hubble time. According to equation (49), it would be about  $7 \times 10^6 \text{ km}$  if the orbit were circular. Eccentricity may raise this about 30%,

$$a_f^{\max} \approx 9 \times 10^{11} \text{ cm}. \quad (63)$$

The minimum  $a_f$  is some multiple of the radius of the He star, which star B becomes after removal of its hydrogen envelope. The He star radius may be about  $5 \times 10^{10} \text{ cm}$ , and we estimate

$$a_f^{\min} = (1.5-3) \times 10^{11} \text{ cm}, \quad (64)$$

so  $a_f^{\max}/a_f^{\min} = 3 \sim 6$ . We assume the distribution of  $a$  to be  $da/da_f$ , so the probability of having  $a_f$  in the permitted range is

$$p = \frac{\ln 3}{7} \sim \frac{\ln 6}{7} = 0.15 \sim 0.25. \quad (65)$$

The center of the useful range is at about  $5 \times 10^{11} \text{ cm}$ . Multiplying by  $a_i/a_f = 12$ , we get

$$a_i \lesssim 10^{13} \text{ cm}. \quad (66)$$

This is too large a distance to permit appreciable X-rays to come from the black hole. If the two compact stars are to merge by gravitational waves within Hubble time, their precursors (a black hole and an O or B star) presumably cannot be observed as emitting X-rays.

Conversely, objects like Cyg X-1 probably will not lead to an observable gravitational wave merger. When the O star expands as a giant and goes into a common envelope with the black hole, the orbit will contract to the extent that the black hole falls into the He core of the giant; there is a merger, but its gravitational waves will be too low a frequency to be observed. Thus the discussions in §§ 4 and 5 are actually irrelevant, given our scenario for common envelope evolution.

Returning to our discussion up to equation (65), we must examine the values of the initial mass of star B. It should be greater than  $15 M_{\odot}$  to permit common envelope evolution, but less than  $35 M_{\odot}$  so as to avoid undue mass loss by wind. With  $M_{A_0} = 90 M_{\odot}$ , this means a range of  $q$  of 0.2. Multiplying this by  $\alpha_1$  in equation (9) and by the mean of  $p$  in equation (65), the rate of merger of a massive black hole with another compact star is

$$\alpha_4 \approx (4-6) \times 10^{-6} \text{ yr}^{-1} \text{ galaxy}^{-1}. \quad (67)$$

This is in remarkably good agreement with the results for case H with hypercritical accretion of Portegies Zwart & Yungelson (1998) if we decrease their number by a factor of 5 to take into account our greater high-mass black hole

mass limit; namely, with this decrease they would have  $R = 7 \times 10^{-6} \text{ yr}^{-1} \text{ galaxy}^{-1}$ .

The great uncertainty is in  $\alpha_1$ , because we do not know very well the minimum ZAMS mass that leads to a massive black hole; thus  $\alpha_1$  is likely to be uncertain by a factor of 2 either way. Together with  $p$  in equation (65), the uncertainty of merger is about a factor 3, so

$$\alpha_4 = 1 \sim 10 \times 10^{-6} \text{ yr}^{-1} \text{ galaxy}^{-1}. \quad (68)$$

This rate is much smaller than that of mergers of small black holes with neutron stars,  $\alpha_s \approx 10^{-4}$ , as one might expect from the difficulty of forming massive black holes. But the signal-to-noise ratio in the gravitational wave detector depends on the chirp mass as  $(M_{\text{chirp}})^{2.5}$ . The chirp mass is

$$M_{\text{chirp}} = \mu^{0.6} M^{0.4}, \quad (69)$$

where  $\mu$  is the effective mass and  $M$  is the total system mass. We have  $M_{\text{chirp}} \sim 3.3 M_\odot$  for our high-mass black hole neutron star binaries, as compared with  $\sim 1.6 M_\odot$  for the low-mass black hole neutron star binaries. Thus effectively the high-mass black hole merging rate should be multiplied by a factor of  $\sim 6$ . The two types can be distinguished by the chirp mass.

Star B, after its SN event, will be either a neutron star or a small black hole of mass slightly over  $1.5 M_\odot$ , and it has no occasion to accrete extra mass. Thus we estimate that signals from massive black hole mergers with neutron stars should have a frequency about one-third of those from smaller black holes. Taking the central value of  $5 \times 10^{-6}$  from equation (67) and multiplying it by 6, we obtain an effective increase of  $\sim 30\%$  over the merging rate of  $10^{-4} \text{ yr}^{-1}$  for the low-mass black hole neutron star binaries.

#### 8. OTHER EFFECTS OF HYDROGEN CLOTHING

We have seen that the He core of stars evolves quite differently according to whether it is clothed with an H envelope or not. We believe that this difference may explain problems in some recent investigations.

##### 8.1. Formation of Black Holes with Light Companions

Portegies Zwart, Verbunt, & Ergma (1997) have discussed the formation of black holes in low-mass X-ray binaries. Many of these transient X-ray sources have been discussed recently with black hole masses in the probable range of  $\sim 6-7 M_\odot$ . Assuming a lower limit of  $\sim 40 M_\odot$  for black hole formation, Portegies Zwart et al. (1997) find a rate of formation much too low.

In their evolution, the black hole originates from the more massive component (star A) of the binary. The less massive component (star B) will spiral into star A when star A becomes a giant and touches star B. They show that star B will survive this spiral-in only if its original distance from star A is at least several hundred solar radii. Otherwise, star B will spiral into the He core of star A, will merge with star A, and thus be lost. In order for star A to have a radius of several hundred  $R_\odot$ , it must have completed He core burning. This means that star A has burned He while clothed with most of its H envelope. Consequently, the upper curve in Figure 1 applies; the massive star A essentially burns as a single star. In this case stars of ZAMS mass as low as  $20-25 M_\odot$  can go into a black hole. We suggest

this as a possible solution to the Portegies Zwart et al. (1997) underproduction of high-mass black holes.

Our scenario also suggests an explanation of why the transient black holes are generally accompanied by low-mass companions. As the massive progenitor star A of the black hole evolves as a giant, RLOF will generally transfer its H envelope to a massive companion (if one exists) during hydrogen shell burning or early in the He core-burning phase. After RLOF, the He core of the primary will burn as a naked He core, with its possible fate as either a low-mass black hole or a neutron star (Brown et al. 1996), but not as a high-mass black hole. Therefore, a high-mass black hole will generally not have a massive companion, except in the relatively rare cases.

If, on the other hand, star B has small mass, it can accept only very little mass from star A in RLOF. Instead, interaction of the two stars will wait until star A becomes a supergiant and its surface reaches star B. Then in the common envelope star B spirals in as described by Portegies Zwart et al. (1997). The He core of star A, mass about one-third of the original ZAMS mass, evolves as described by Portegies Zwart et al. (1997), finally collapses as an SN, and leaves a remnant of mass approximately equal to the He core mass,  $6-7 M_\odot$ .

##### 8.2. Cyg X-3

Cyg X-3 may be a progenitor of a binary of two low-mass black holes. We believe the envelope of the O star progenitor of the compact object (star A) in Cyg X-3 was probably lifted off in RLOF. This star A then had an SN event making it into a neutron star. When star B, the present He star, became a giant, star A would spiral in, which would have converted the neutron star into a low-mass black hole (if it was not one already). The He star B in Cyg X-3 is certainly burning as naked. Our estimate is that the ZAMS mass of the progenitor of this He star is about the maximum mass that does not go into an LBV stage (because the envelope must have been used up in common envelope evolution), say,  $35 M_\odot$ . From our Figure 1, the most likely fate of such a star with an H envelope lifted off in RLOF is a low-mass black hole, but a neutron star of mass  $\sim 1.5 M_\odot$  cannot be excluded.

##### 8.3. Nucleosynthesis

The galactic ratio of oxygen to iron depends on the ZAMS mass above which single stars evolve into high-mass black holes (cut-off mass), and therefore do not return matter to the galaxy. Oxygen is chiefly produced in quiescent burning before the SN explosion, and the amount is roughly proportional to ZAMS mass. Fe, on the other hand, is produced explosively in the SN explosion, and the amount is roughly independent of ZAMS mass, possibly decreasing slightly as the latter increases. Tsujimoto et al. (1997) have recently used the [O/Fe] ratio and observations in metal poor stars to determine a cutoff mass of

$$M_{\text{cutoff}} = 50 \pm 10 M_\odot. \quad (70)$$

Tsujimoto et al. (1997) remark that the influence of the metallicity dependence of stellar wind losses may be significant, but they do not take it into account. We believe, however, that the larger wind losses of stars with solar metallicity have important effects. Over a wide range of ZAMS masses, between  $\sim 35-40 M_\odot$  and  $\sim 80 M_\odot$  accord-

ing to Woosley et al. (1993), the winds remove the H envelope sufficiently rapidly that the He cores evolve as naked. Thus the cores evolve into either low-mass black holes or neutron stars, their (gravitational) core masses following the lower line heading slightly above  $1.5 M_{\odot}$  in Figure 1. Brown et al. (1996) found that the primary in 1700–37 probably evolved into a low-mass black hole (i.e., it does not pulse), whereas 1223–62 is known to contain a neutron star. In both cases, the ZAMS mass of the primary was found to be  $\sim 40 M_{\odot}$ , so this approximately locates the ZAMS mass corresponding to  $M_{\text{ap}}$  of § 2 in the case of solar metallicity. (As noted earlier, we estimate the ZAMS mass corresponding to  $M_{\text{pc}}$  to be  $\sim 80 M_{\odot}$ .) It may well be that for metal-poor stars with much weaker winds the hydrogen envelope is not removed rapidly enough for their He cores to evolve as naked. In this case the limit (eq. [70]) for ZAMS masses giving core mass  $M_{\text{pc}}$  could be more appropriate.

We have not attempted to carry out a calculation of nucleosynthesis in our scenario, but we wish to point out features of our scenario that will tend to increase the  $M_{\text{cutoff}}$ , above which element production for the galaxy ceases.

Disregarding mass loss, single stars above a certain mass  $M_{\text{min}}$  evolve into high-mass black holes. From the upper line in Figure 1, one might take the minimum He mass to be  $6-8 M_{\odot}$ , corresponding to a ZAMS mass of  $20-25 M_{\odot}$ . But as the mass of the single star increases, there is big mass loss before the star ever reaches the SN stage. So with extensive mass loss, Woosley et al. (1993) find that ZAMS mass  $35 M_{\odot}$  leads, after SN, to a compact object of about  $1.5 M_{\odot}$ , thus a low-mass black hole. Thus high-mass black holes are only formed by single stars in a limited mass range, from  $20-25 M_{\odot}$  to about  $35-40 M_{\odot}$ . Stars in this intermediate mass range, of width  $10-20 M_{\odot}$ , do not return their matter, especially their Fe, to the galaxy. This group of stars is in addition to the stars above  $M_{\text{cutoff}}$ . To compensate for this fact,  $M_{\text{cutoff}}$  must be raised above the Tsujimoto et al. (1997) value. Because the abundance of stars decreases with increasing mass, the raise must be more than  $10-20 M_{\odot}$ . Thus we suggest that  $M_{\text{cutoff}}$  may be as high as the  $80 M_{\odot}$ , which we used for other reasons in § 3.

We note from the Woosley et al. (1993) calculations that naked He cores evolve into less massive carbon/oxygen cores than clothed ones. Thus we see that inclusion of mass loss will tend to move the mass above which all nucleosynthesis ceases even higher. Therefore, we believe that inclusion of the mentioned effects may decrease the apparent discrepancy between  $M_{\text{cutoff}}$  of equation (70) and our ZAMS mass of  $80 M_{\odot}$  for making MBHs.

We return to the higher observed incidence of high-mass black holes (LMC X-1 and LMC X-3) in the LMC. Suppose that the limit of equation (70) applies for the lower metallicity of the LMC; i.e., stars with ZAMS masses  $\sim 40-80 M_{\odot}$  in the LMC would not experience as large a wind loss as those in the disc because of the lower metallicity in the LMC. This could help explain why two high-mass X-ray binaries containing high-mass black holes are observed in the LMC, and only Cyg X-1 is seen in the Galactic disc.

## 9. CONCLUSIONS

Our chief result is that those massive binary systems in which the primary is sufficiently massive,  $\gtrsim 80 M_{\odot}$ , to go into a high-mass black hole can contribute importantly to observable gravitational waves upon merger of the final

binary of compact objects. These mergers come chiefly from companion O stars in the range of  $15-35 M_{\odot}$  ZAMS mass.

We argue that the situation created by hypercritical accretion is favorable to common envelope evolution in this range of companion masses, and that the timescale of this evolution is very fast.

We show that population syntheses do not have a common envelope with efficiency comparable with ours, possibly because they generally do not include hypercritical accretion. Certainly they do not have as effective a tightening of orbits as we have. The exception is case H of Portegies Zwart & Yungelson (1998), which does include hypercritical accretion. For the same mass limit for black holes, their results are in good agreement with ours. At first sight it may seem surprising that we match so well the Portegies Zwart & Yungelson (1998) work, provided they include hypercritical accretion, as we do. Their calculation includes various phases of mass transfer before forming the binary consisting of a neutron star and a black hole. However, results depend only on the ratio of the logarithmic interval favorable for gravitational merger with avoidance of coalescence to the total logarithmic interval over which binaries are distributed. Whereas the favorable logarithmic interval is shifted around in the various mass transfers, its magnitude is unchanged because the only scale is the radial separation of the two objects.

Our final merger rate of  $\sim (4-6) \times 10^{-6} \text{ yr}^{-1}$  increases the Bethe & Brown (1998) rate of gravitational waves by a factor of  $\sim 1.3$ , largely because of the higher chirp mass with high-mass black holes.

This important effect comes in spite of a much higher mass limit ZAMS  $80 M_{\odot}$  for evolution into high-mass black holes in binaries. We justify this high-mass limit from results for the evolution of naked He stars by Woosley et al. (1993). Namely, the more massive primary in binaries has its mass lifted off early, either RLOF or in LBV stage for the stars with ZAMS masses greater than  $35-40 M_{\odot}$ . The resulting naked W-R or He star is deprived of the H envelope, which normally “insulates” He cores during their burning, so the convective carbon-burning stage is not skipped, as it is in single stars of ZAMS mass  $\gtrsim 20 M_{\odot}$ . The great entropy loss during the long duration of this stage results in a low-mass compact object.

We show that the high-mass limit for evolution into high-mass black holes in binaries is consistent with nucleosynthesis, because single stars with ZAMS masses in the range of  $20-35$  or  $40 M_{\odot}$  evolve into high-mass black holes without return of matter to the galaxy.

We also indicate that we could evolve enough transient sources, binaries of a high-mass black hole, and a low-mass main-sequence star.

There is interest in the merging of binaries composed of two high-mass black holes (Brady et al. 1998). With our mass limit of  $80 M_{\odot}$  for high-mass black holes, the initial separation of the two massive O star progenitors must be greater than  $40 R_{\odot}$  for both of them to lie inside their Roche lobes. Case A mass transfer between two nearly equal mass stars can only widen the binary, since  $a_f/a_i = (\mu_f/\mu_i)^2$ , and  $\mu$  will decrease in the transfer. Stars of ZAMS mass  $\gtrsim 40 M_{\odot}$  lose their H envelopes rapidly in the LBV phase, so a common envelope phase in which the orbit is tightened is excluded. The W-R stars that result after loss of the H will lose most of their mass by wind, down to  $\sim 10-15 M_{\odot}$  before going into black holes. With spherically symmetric

mass loss,  $a_f/a_i = M_i/M_f \sim 4$ . Thus the separation of the two final black holes will be several times the initial greater than  $40 R_\odot$ , which precludes merging in Hubble time. We are in agreement with Portegies Zwart & Yungelson (1998) on this point.

We believe that in Brown & Bethe (1998) together with this paper we have given a consistent description of the evolution of most binaries of compact stars containing a black hole.

We would like to thank the referees for extremely helpful criticism. We are grateful to Simon Portegies Zwart and Ralph Wijers for advice. We would like to thank Chang-Hwan Lee for useful advice and help, and Fred Walter for information on observability. We are grateful to Kip Thorne for discussion of the signal-to-noise ratio in LIGO. G. E. B. was partially supported by the U.S. Department of Energy under grant DE-FG02-88ER40388.

## REFERENCES

- Bethe, H. A., & Brown, G. E. 1998, *ApJ*, 506, 780  
 Brady, P. R., Creighton, J. D. E., & Thorne, K. S. 1998, *Phys. Rev. D*, 58, 061501  
 Brown, G. E., & Bethe, H. A. 1994, *ApJ*, 423, 659  
 Brown, G. E., Lee, C.-H., & Bethe, H. A. 1999, *New Astr.*, submitted  
 Brown, G. E., Weingartner, J. C., & Wijers, R. A. M. J. 1996, *ApJ*, 463, 297  
 Cordes, J. M., & Chernoff, D. F. 1997, *ApJ*, 482, 971  
 Germany, C. D., Conti, P. S., & Massey, P. 1980, *ApJ*, 242, 1063  
 Herrero, A., Kudritzki, R. P., Gabler, R., Vilchez, J. M., & Gabler, A. 1995, *A&A*, 297, 556  
 Kalogera, V. 1996, *ApJ*, 471, 352  
 Kippenhahn, R., & Meyer-Hofmeister, E. 1977, *A&A*, 54, 539  
 Kuiper, L., van Paradijs, J., & van der Klis, M. 1998, *A&A*, 203, 79  
 Li, G. Q., Lee, C.-H., & Brown, G. E. 1997, *Nucl. Phys. A*, 625, 372  
 —. 1997b, *Phys. Rev. Lett.*, 79, 5214  
 Lipunov, V. M., Postnov, K. A., & Prokhorov, M. E. 1997, *MNRAS*, 288, 245  
 Masevich, A. G., Popova, E. I., Tutukov, A. V., & Yungelson, L. R. 1979, *Ap&SS*, 62, 451  
 Portegies Zwart, S. F., Verbunt, F., & Ergma, E. 1997, *A&A*, 321, 207  
 Portegies Zwart, S. F., & Yungelson, L. R. 1998, *A&A*, 332, 173  
 Prakash, M., Bombaci, I., Prakash, M., Ellis, P. J., Lattimer, J. M., & Knopren, R. 1997, *Phys. Rep.*, 280, 1  
 Rasio, F. A., & Livio, M. 1996, *ApJ*, 471, 366  
 Ruffert, M. 1994, *ApJ*, 427, 351  
 Ruffert, M., & Arnett, D. 1994, *A&AS*, 106, 505  
 Terman, J. L., Taam, R. E., & Hernquist, L. 1995, *ApJ*, 445, 367  
 Thorsett, S. E., & Chakrabarty, D. 1999, *ApJ*, 512, 288  
 Thorsson, V., Prakash, M., & Lattimer, J. M. 1994, *Nucl. Phys. A*, 572, 693  
 Tsujimoto, T., Yoshii, Y., Nomoto, K., Matteuci, F., Thielemann, F.-K., & Hashimoto, M. 1997, *ApJ*, 483, 228  
 van den Heuvel, E. P. J. 1994, in *Saas Fe Adv. Course 22*, ed. H. Nussbaumer & A. Orr (Berlin: Springer), 263  
 Woosley, S. E., Langer, N., & Weaver, T. A. 1993, *ApJ*, 411, 823  
 —. 1995, *ApJ*, 448, 315  
 Woosley, S. E., & Weaver, T. A. 1995, *ApJS*, 101, 181

## Chapter 14

# The Formation of High-Mass Black Holes in Low-Mass X-Ray Binaries

G.E. Brown, C.-H. Lee and H.A. Bethe

Reprinted with permission from *New Astronomy*, 4 (1999) 313–323

Copyright (1999) by Elsevier Science B.V.

### Commentary

By the time we worked on this paper, we knew that naked He stars blew away, and that not enough was left to form high-mass black holes. Ed van den Heuvel kept emphasizing that He wind losses employed in the evolution were too large by a factor of 2 to 4. He argued as if there were a nearly linear relation between wind loss and amount of matter removed. From our experience with stellar evolution we knew that in most such cases there are large feedbacks. None the less, it was comforting to have the work of Wellstein & Langer (1999), where they showed that final He cores were not much larger when Langer's earlier too large wind loss rate was cut down by a factor of 3. Cutting down further did produce substantially larger final He cores.

Whereas we see only one continuously shining Cygnus X-1 in the Galaxy (plus LMC X-1 and LMC X-3 in the Large Magellanic Cloud), we have an estimated  $\sim 2000$  high-mass ( $\sim 7M_{\odot}$ ) black holes with low-mass main sequence companion in the soft X-ray transient sources. We have many more with subgiant companions, many that will not be observed until the companion evolves further, close enough to dump matter on the accretion disk of the black hole.

From our experience with naked He stars, we saw that we had to keep them clothed during He core burning, so that they didn't lose much mass by wind. They can then be made naked by mass transfer to the companion star following the He core burning (Case C mass transfer) because they then have only  $\sim 10^4$  yrs to live, and not much mass can be lost by wind in this short time. Consequently, the conclusion of Paper 14 is that mass must be transferred from the massive black hole progenitor only after the He core burning is finished. In other words, this progenitor has to be tricked into evolving as a single star, its contact with the companion coming only shortly before its death.

This means that the massive star will have evolved through its supergiant (He core burning stage) before matter overflows its Roche Lobe. Then, by that time, a main sequence companion must be at just the right distance to receive the overflow; this means  $a \sim 1500R_{\odot}$ , the Roche Lobe of the massive star being at  $\sim \frac{2}{3}a$ . Since the binding energy of the envelope of the massive star goes as  $1/a$ , this binding energy is very small, so that the envelope can be removed by the drop in gravitational energy of an  $\sim 1M_{\odot}$  main sequence star as it moves inwards in common envelope evolution with the massive star from  $\sim 1500R_{\odot}$  to the much smaller Roche Lobe of the He star which results when the H envelope is removed from the massive star. In this way one could understand why all of the main sequence companions of the black holes in the transient sources were of nearly the same low masses,  $0.5-1M_{\odot}$ . We develop the further evolution of these binaries in Paper 20.

In Eq (3.1) of the present paper we chose an interval of ZAMS masses of  $20M_{\odot}$  to  $35M_{\odot}$  for progenitors of the high-mass black hole.

In the work we discussed here, we wondered whether the star had to be clothed by hydrogen during all of the He core burning, or only during some fraction of it, in order for a high-mass black hole to result. In Paper 20 it will be seen that if it is clothed by hydrogen during some part of He core burning, then it will be clothed during all of it, at least with present stellar evolution.

#### Reference

- (1) Wellstein, S., and Langer, N. (1999), *Astronomy & Astrophysics* **350**, 148.



ELSEVIER

## New Astronomy

New Astronomy 4 (1999) 313–323

[www.elsevier.nl/locate/newast](http://www.elsevier.nl/locate/newast)

# The formation of high-mass black holes in low-mass X-ray binaries

G.E. Brown<sup>a,1</sup>, C.-H. Lee<sup>a,2</sup>, Hans A. Bethe<sup>b,3</sup><sup>a</sup>*Department of Physics and Astronomy, State University of New York, Stony Brook, NY 11794-3800, USA*<sup>b</sup>*Floyd R. Newman Laboratory of Nuclear Studies, Cornell University, Ithaca, NY 14853, USA*

Received 8 December 1998; accepted 13 April 1999

Communicated by E.P.J. van den Heuvel

## Abstract

We suggest that high-mass black holes; i.e., black holes of several solar masses, can be formed in binaries with low-mass main-sequence companions, provided that the hydrogen envelope of the massive star is removed in common envelope evolution which begins only after the massive star has finished He core burning. © 1999 Elsevier Science B.V. All rights reserved.

PACS: 97.60.Lf; 97.80.Jp

Keywords: Black hole physics; X-rays: bursts; Binaries

## 1. Introduction

In this note we suggest that high-mass black holes; i.e., black holes of several solar masses, can be formed in binaries with low-mass main-sequence companions, provided that the hydrogen envelope of the massive star is removed in common envelope evolution which begins only after the massive star has finished He core burning. That is, the massive star is in the supergiant stage, which lasts only  $\sim 10^4$  yr, so effects of mass loss by He winds are small. Since the removal of the hydrogen envelope of the massive star occurs so late, it evolves essentially as a single star, rather than one in a binary. Thus, we can use evolutionary calculations of Woosley & Weaver (1995) of single stars.

Using the Brown & Bethe (1994) upper limit of

$\sim 1.8 M_{\odot}$  for the (gravitational) compact core mass that can evolve into a low-mass black hole, we find that high-mass black holes can be formed in the collapse of stars with ZAMS mass  $\geq 20 M_{\odot}$ . We somewhat arbitrarily take the upper limit for the evolution of the so-called transient sources to be  $\sim 35 M_{\odot}$  ZAMS mass. Mass loss by winds in stars sufficiently massive to undergo the LBV (luminous blue variable) stage may seriously affect the evolution of stars of ZAMS  $> 35-40 M_{\odot}$ , but we need calculations with improved mass loss rates before discussing these quantitatively. Both Portegies Zwart et al. (1997) and Ergma & Van den Heuvel (1998) have suggested that roughly our chosen range of ZAMS masses must be responsible for the transient sources. We believe that the high-mass black hole limit of ZAMS mass  $\sim 40 M_{\odot}$  suggested by Van den Heuvel & Habets (1984) and later revised to  $\geq 50 M_{\odot}$  (Kaper et al., 1995) applies to massive stars in binaries, which undergo RLOF (Roche Lobe Overflow) early in their evolution. We will not

<sup>1</sup>E-mail: popence@nuclear.physics.sunysb.edu<sup>2</sup>E-mail: chlee@silver.physics.sunysb.edu<sup>3</sup>E-mail: dusty@heph.cornell.edu

pursue this here because calculations with improved He-star mass loss rates by wind, now being carried out, are necessary before quantitative results can be obtained.

The most copious high-mass black holes of masses  $\sim 6\text{--}7 M_{\odot}$  have been found in the transient sources such as A0620-00. These have low-mass companions, predominantly of  $\lesssim 1 M_{\odot}$ , such as K- or M-stars. In the progenitor binaries the mass ratios must have been tiny, say  $q \sim 1/25$ . Normally such small ratios are thought to be rare; e.g. in binary evolution the companion distribution is often taken as  $dq$ , implying a very low probability of such a binary.

In this note we follow the evolutionary scenario for the black hole binary of De Kool et al. (1987). We show that the reason for this small  $q$ -value lies in the common envelope evolution of the binary. The smaller the companion mass, the greater the radius  $R_g$  the giant must reach before its envelope meets the companion. This results because the orbit of a low-mass companion must shrink by a large factor in order to expel the envelope of the giant, hence the orbit must initially have a large radius. (Its final radius must be just inside its Roche Lobe, which sets a limit to the gravitational energy it can furnish.)

A large radius  $R_g$  in turn means that the primary star must be in the supergiant stage. Thus it will have completed its He core burning while it is still "clothed" with hydrogen. This prevents excessive mass loss so that the primary retains essentially the full mass of its He core when it goes supernova. We believe this is why K- and M-star companions of high-mass black holes are favored.

We find that the black holes in transient sources can be formed from stars with ZAMS masses in the interval  $20\text{--}35 M_{\odot}$ . The black hole mass is only slightly smaller than the He core mass, typically  $\sim 7 M_{\odot}$ .

## 2. Development

Our evolutionary scenario is essentially the same as that of De Kool et al. (1987) for the black hole binary A0620-00. We apply this scenario somewhat schematically for the range of ZAMS  $20\text{--}35 M_{\odot}$  for the massive star. In Fig. 1 we show results of calculations of Fe core masses of single stars by

Woosley & Weaver (1995). Somewhere around ZAMS mass  $20 M_{\odot}$  the Fe core exceeds the Brown & Bethe (1994) limit of  $\sim 1.8 M_{\odot}$  for low-mass black-hole formation, so we take this as the beginning of high-mass black hole formation. The detailed behavior of this curve should not be taken seriously, but the large increase around  $20\text{--}25 M_{\odot}$  is of importance, as we discuss in detail later.

We wish to carry out a population synthesis so that we can estimate the number of transient sources. We make roughly the same assumptions as Bethe & Brown (1998); Bethe & Brown (1999), our massive star  $M_B$  lying in mass somewhere in between the O,B star progenitors of binary neutron stars and the progenitor of the massive black hole in Cyg X-1. Our low-mass companion is a main sequence star of mass  $\sim 1 M_{\odot}$ . Thus, the ratio

$$q = \frac{M_{A_i}}{M_{B_i}} \quad (2.1)$$

is very small, and there will be great uncertainty in the initial number of binaries for such a small  $q \sim 1/25$ . We take the distribution as  $dq$ . The distribution in  $q$  is unknown for such low-mass companions as are involved here, but our results will show that the flat distribution in  $q$  is not unreasonable. We assume  $\log a$ , where  $a$  is the semi-major axis of the orbit, to be uniformly distributed, over a total logarithmic interval of 7. Thus, the fraction of binaries in a given interval of  $\ln a$  is

$$d\phi = \frac{d(\ln a)}{7}. \quad (2.2)$$

We take  $\alpha$  the supernova rate to be  $\sim 0.02$  per galaxy per year, somewhat larger than the 0.015 given by Cappellaro et al. (1997) and somewhat smaller than the 0.025 assumed by Bethe & Brown (1998). The rate of supernovae (SN) is the same as the rate of birth of massive stars  $M > 10 M_{\odot}$ , so

$$d\alpha = \alpha n \left( \frac{M}{10 M_{\odot}} \right)^{-n} \frac{dM}{M}, \quad (2.3)$$

with  $n$  the Salpeter exponent which we take to be 1.5. We furthermore assume a binarity of 0.5. Once again, for our very small  $q$  values this is uncertain.

We thus evolve as typical a ZAMS  $25 M_{\odot}$  (star B) with companion  $\sim 1 M_{\odot}$  main sequence star (star A) as typical progenitor of the transient X-ray sources.

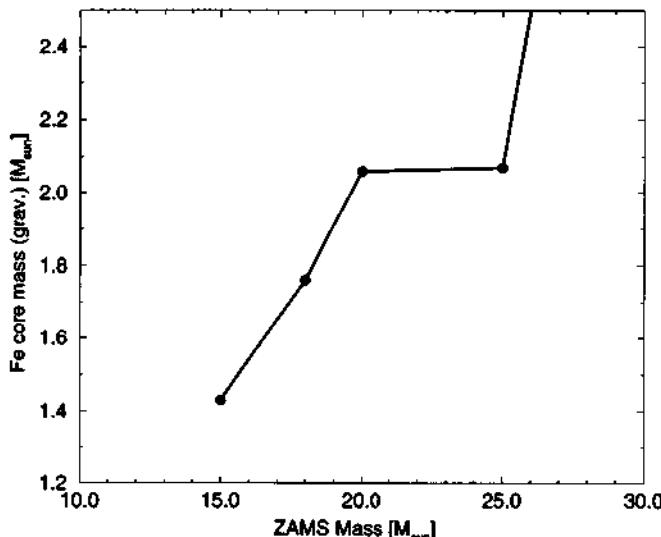


Fig. 1. Comparison of the compact core (Remnant) masses resulting from the evolution of single stars (filled symbols), case of solar metallicity of Woosley & Weaver (1995).

The common envelope evolution can be done as in Bethe & Brown (1998). With  $M_{B_i} = 25 M_\odot$  and neglect of the accretion onto the main sequence mass  $M_A$ , we find from Bethe & Brown

$$\left(\frac{Y_f}{Y_i}\right)^{1.2} = \frac{1.2}{\alpha_{ce}} \frac{M_{B_i}}{M_A}, \quad (2.4)$$

where  $Y = M_B/a$ . Here the coefficient of dynamical friction  $c_d$  was taken to be 6, in the range of the 6–8 for supersonic flow with Mach number 3–10 (Ruffert, 1994; Ruffert & Arnett, 1994). The result is relatively insensitive to  $c_d$ , the exponent 1.2 resulting from  $1 + 1/(c_d - 1)$ .

Thus, in our case

$$\frac{Y_f}{Y_i} = 17 \left( \frac{\alpha_{ce} M_A}{M_\odot} \right)^{-0.83} = 30 \left( \frac{0.5 M_\odot}{\alpha_{ce} M_A} \right)^{0.83}. \quad (2.5)$$

We expect  $\alpha_{ce} = 0.5$ , under the assumption that the kinetic energy of the expelled envelope is equal to that it originally possessed in the massive star, but it could be smaller. From this we obtain

$$\frac{a_i}{a_f} = \frac{M_{B_i} Y_f}{M_{B_i} Y_i} = 90 \left( \frac{0.5 M_\odot}{\alpha_{ce} M_A} \right)^{0.83}, \quad (2.6)$$

where we have taken the He star mass  $M_{B_i}$  to be 1/3

of  $M_{B_i}$ . In order to survive spiral-in, the final separation  $a_f$  must be sufficient so that the main sequence star lies at or inside its Roche Lobe, about  $0.2a_f$  if  $M_A = M_\odot$ . This sets  $a_f \sim 5R_\odot = 3.5 \times 10^{11}$  cm and

$$a_i = 3.15 \left( \frac{0.5}{\alpha_{ce}} \right)^{0.83} \times 10^{13} \text{ cm}, \quad (2.7)$$

~2 A.U. This exceeds the radius of the red giant tip in the more numerous lower mass stars in our interval, so the massive star must generally be in the supergiant phase where it meets the main sequence star, i.e., the massive star must be beyond He core burning. E.g., the red giant tip (before the He core burning) for a  $20 M_\odot$  star is at  $0.96 \times 10^{13}$  cm, for a  $25 M_\odot$  star,  $2.5 \times 10^{13}$  cm (Schaller et al., 1992). These numbers are, however, somewhat uncertain. Notice that decreasing  $\alpha_{ce}$  will increase  $a_i$ . Decreasing  $M_A$  has little influence, because with the smaller stellar radius the minimum  $a_f$  will decrease nearly proportionately. Note that neglect of accretion onto the main sequence star would change the exponent 0.83 to unity, so accretion is unimportant except in increasing the final mass.

Now a ZAMS  $25 M_\odot$  star ends up at radius  $6.7 \times 10^{13}$  cm ( $\sim 2a_i$ ) following He shell burning (Weaver

et al., 1978). Thus the interval between  $a_i$  and  $6.7 \times 10^{13}$  cm is available for spiral-in without merger<sup>4</sup> so that a fraction

$$\frac{1}{7} \ln \left( \frac{6.7}{3.15(0.5/\alpha_{ce})^{0.83}} \right) \approx 0.11 \quad (2.8)$$

of the binaries survive spiral-in, but are close enough so that the main sequence star is encountered by the evolving H envelope of the massive star. The He core burning will be completed before the supergiant has moved out to  $\sim 2$  A.U., so binaries which survive spiral-in will have He cores which burn as "clothed", namely as in single stars.

### 3. Birth rate of transient sources

Given our assumptions in Section 2, the fraction of supernovas which arise from ZAMS stars between  $20$  and  $35 M_\odot$  is

$$\frac{1}{2^{3/2}} - \frac{1}{3.5^{3/2}} = 0.20, \quad (3.1)$$

where we have assumed the mass  $10 M_\odot$  is necessary for a star to go supernova. A Salpeter function with index  $n = 1.5$  is assumed here. Our assumption that the binary distribution is as  $dq$  is arbitrary, and gives us a factor  $1/25$  for a  $1 M_\odot$  companion. Note, however, that had we included higher mass companions, the change in the final projected number of transient sources would not be order of magnitude, because the hydrogen burning time goes inversely as mass squared. Thus, for supernova rate 2 per century, our birth rate for transient sources in the Galaxy is

$$2 \times 10^{-2} \times 0.5 \times 0.11 \times 0.20 \times 0.04 \\ \simeq 8.8 \times 10^{-6} \text{ yr}^{-1}, \quad (3.2)$$

where 0.5 is the assumed binarity, 0.11 comes from Eq. (2.8), and the final (most uncertain) factor 0.04 results from a distribution flat in  $q$  and an assumed  $1 M_\odot$  companion star.

In order to estimate the number of transient sources with black holes in the Galaxy, we should know the time that a main sequence star of mass  $\sim 1 M_\odot$  transfers mass to a more massive companion. For a main-sequence donor, the mass transfer rate is  $\sim 10^{-10} M_\odot \text{ yr}^{-1}$ , almost independent of donor mass (Verbunt & Van den Heuvel, 1995). As mass is transferred, the mass of the donor decreases and with it the radius of the donor. Quite a few low-mass X-ray binaries have X-ray luminosities that imply accretion rates in excess of  $10^{-10} M_\odot \text{ yr}^{-1}$ , leading to suggestions of additional mechanisms for loss of angular momentum from the binary, to increase mass transfer. Verbunt & Zwaan (1981) estimate that magnetic braking can boost the transfer of mass in a low-mass binary. We somewhat arbitrarily adopt an effective mass transfer rate of  $10^{-9} \text{ yr}^{-1}$  for main sequence stars and  $10^{-8} \text{ yr}^{-1}$  for the two systems that have subgiant donors (V404 Cyg and XN Sco 94). In order to estimate the number of high-mass black hole, main sequence star binaries in the Galaxy we should multiply the birth rate (3.2) times the  $10^9 \text{ yr}$  required, at the assumed mass loss rate, to strip the main sequence star, obtaining 8800 as our estimate. Not all of these will fill their Roche Lobes. Those that do not may not now be visible, but will be later, as they begin evolving as one of those with subgiant donors. The fact that two of eight observed binaries are subgiants, although the lifetime of the latter is two orders of magnitude less than the main sequence lifetime, suggests that some fraction of our 8800 estimated binaries do not fill their Roche Lobes. From the observed black-hole transient sources Wijers (1996) arrives at 3000 low-mass black hole sources in the Galaxy, but regards this number as a lower limit. Beginning from this, which we regard as an observational estimate, we note that the two subgiants in Table 1 involve the more massive F,G and A stars. These indicate an  $\sim 100$  times greater population of unevolved main sequence stars in this range which lie quietly inside their Roche Lobes. Thus, including the quiescent binaries might give as many as  $(2/8) \times 100 \times 3000$  or  $\sim 10^5$  additional binaries, suggesting that our above estimate may be an order of magnitude too low. Estimates of the number of transient sources are very uncertain, but it is clear that there are orders of magnitude more of them than of the Cygnus X-1

<sup>4</sup>Note that envelope removal does not occur at the Roche Lobe on the thermal time scale  $\tau_{th}$  but at the low-mass star since the remaining lifetime of the giant is  $\sim 10^4$  yr, much shorter than  $\tau_{th}$ .

type objects with high-mass black hole and massive star companion.

If we assume that ZAMS masses  $\sim 10\text{--}18 M_{\odot}$  evolve into a neutron star, we should have  $\sim 3$  times more neutron stars than high-mass black holes (see Eq. (3.1)). The upper limit follows from our belief that SN (1987A) with progenitor  $\sim 18 M_{\odot}$  ZAMS went into a low-mass black hole, following the scenario of Brown & Bethe (1994). On the basis of a Monte Carlo calculation using the kick velocities of Cordes & Chernoff (1998) we find that  $\sim 1/2$  of the binaries containing He-star, low-mass main sequence companion (with  $M = 1 M_{\odot}$ ) will be disrupted in the explosion. Thus, we find only a slightly higher birth rate for LMXBs with neutron stars than with black holes, although the numbers could be equal to within our accuracy. The LMXBs with neutron stars tend to be much brighter than those with black holes, indicating an order of magnitude greater transfer rate. With the correspondingly shorter main-sequence lifetime, this would give us several hundred LMXBs

with neutron stars, a factor of several greater than the observed number,  $\sim 130$ . Given the lifetime of  $\sim 10^8$  yr of a LMXB and accretion at roughly the Eddington rate of  $\sim 10^{-9}\text{--}10^{-8} M_{\odot} \text{ yr}^{-1}$ , it is reasonable that some neutron stars accrete a reasonable fraction of a solar mass  $M_{\odot}$  (Van den Heuvel, 1995). We expect the masses of these to exceed the Brown & Bethe (1994) limit of  $1.5 M_{\odot}$  for maximum neutron star mass, and evolve into low-mass black holes. The low-mass black-hole, low-mass main sequence star systems would not be seen.

#### 4. Estimated masses of the black holes in transient sources

As we showed below in Eq. (2.7), the He core of the massive star will in general be uncovered only after He core burning is completed. The remaining time for He burning (in a shell) will be short, e.g., for a  $20 M_{\odot}$  ZAMS star it is only  $1.4 \times 10^4$  yr

Table 1

Parameters of suspected black hole binaries with measured mass functions (Wijers (1996), Chen et al. (1997), Bailyn et al. (1998), Orosz et al. (1998), Bailyn, private communication). N means nova, XN means X-ray nova. Numbers in parenthesis indicate errors in the last digits

X-ray names	Other name(s)	Compan. type $q (M_{\text{opt}}/M_X)$	$P_{\text{orb}}$ (d) $K_{\text{opt}}$ ( $\text{km s}^{-1}$ )	$f(M_X) (M_{\odot})$ $i$ (deg)	$M_{\text{opt}} (M_{\odot})$ $M_X (M_{\odot})$	$(l,b)$ $d$ (kpc)
Cyg X-1	V1357 Cyg	O9.7Iab	5.5996	0.25(1)	33(9)	(73.1, + 3.1)
1956 + 350	HDE 226868		74.7(10)		16(5)	2.5
LMC X-3		B3Ve	1.70	2.3(3)		(273.6, - 32.1)
0538 - 641			235(11)		5.6-7.8	55
LMC X-1		O7-9III	4.22	0.14(5)		(280.2, - 31.5)
0540 - 697			68(8)			55
XN Mon 75	V616 Mon	K4 V	0.3230	2.83-2.99	0.53-1.22	(210.0, - 6.5)
A 0620 - 003	N Mon 1917	0.057-0.077	443(4)	37-44 (*)	9.4-15.9	0.66-1.45
XN Oph 77	V2107 Oph	K3 V	0.5213	4.44-4.86	0.3-0.6	(358.6, + 9.1)
H 1705 - 250			420(30)	60-80	5.2-8.6	5.5:
XN Vul 88	QZ Vul	K5 V	0.3441	4.89-5.13	0.17-0.97	(63.4, - 3.1)
GS 2000 + 251		0.030-0.054	520(16)	43-74	5.8-18.0	2
XN Cyg 89	V404 Cyg	K0 IV	6.4714	6.02-6.12	0.57-0.92	(73.2, - 2.2)
GS 2023 + 338	N Cyg 1938, 1959	0.055-0.065	208.5(7)	52-60	10.3-14.2	2.2-3.7
XN Mus 91		K5 V	0.4326	2.86-3.16	0.41-1.4	(295.0, - 6.1)
GS 1124 - 683		0.09-0.17	406(7)	54-65	4.6-8.2	3.0
XN Per 92		M0 V	0.2127(7)	1.15-1.27	0.10-0.97	(197.3, - 11.9)
GRO J0422 + 32		0.029-0.069	380.6(65)	28-45	3.4-14.0	
XN Sco 94		F5-G2	2.6127(8)	2.64-2.82	1.8-2.5	(345.0, + 2.2)
GRO J1655 - 40		0.33-0.37	227(2)	67-71	5.5-6.8	3.2
XN	MX 1543-475	A2 V	1.123(8)	0.20-0.24	1.3-2.6	(330.9, + 5.4)
4U 1543 - 47			124(4)	20-40	2.0-9.7	9.1(11)

(\*) A much higher inclination for A0620 has been claimed by Haswell et al. (1993) of up to  $i = 70$ . In this case, the lower limits on the component masses would be  $M_X > 3.8$  and  $M_{\text{opt}} > 0.22$ .

(Schaller et al., 1992). Therefore the mass loss by wind after uncovering the He core will not be large, and when the star finally becomes a supernova, its mass will be almost equal to the He core of the original star. The latter can be calculated from

$$M_{\text{He}} = 0.08(M_{\text{ZAMS}})^{1.4}. \quad (4.1)$$

so for ZAMS masses  $20\text{--}35 M_{\odot}$ ,  $M_{\text{He}}$  will lie in the interval  $\sim 5.3\text{--}11.6 M_{\odot}$ . In fact, the lower limit looks a bit small, because the He core of the  $\sim 18 M_{\odot}$  progenitor of 1987A is generally taken as  $\sim 6 M_{\odot}$ .

Bailyn et al. (1998) find the black hole masses in transient sources to be clustered about  $\sim 7 M_{\odot}$ , except for V404 Cyg which has a higher mass. This is in general agreement with our scenario, because most of the black holes will come from the more numerous stars of ZAMS mass not far from our lower limit of  $\sim 20 M_{\odot}$ . Two points are important to note:

1. Not much mass can have been lost by wind. Naked He stars have rapid wind loss. However in our scenario the He star is made naked only during He shell burning and therefore does not have much time ( $\leq 10^4$  yr) to lose mass by wind.
2. There are good reasons to believe that the initial He core will be rotating (cf. Mineshige et al., 1993). The way in which the initial angular momentum affects the accretion process has been studied by Mineshige et al. (1997) for black hole accretion in supernovae. In general accretion discs which are optically thick and advection dominated are formed. The disc is hot and the produced energy and photons are advected inward rather than being radiated away. The disc material accretes into the black hole at a rate of  $> 10^6 \dot{M}_{\text{Edd}}$  for the first several tens of days. Angular momentum is advected outwards. Our results show that little mass is lost, because the final  $\sim 7 M_{\odot}$  black hole masses are not much less massive than the He core masses of the progenitors, and some mass is lost by wind before the core collapses. The latter loss will not, however, be great, because there is not much time from the removal of the He envelope until the collapse.

Accretion of the He into the black hole will differ

quantitatively from the above, but we believe it will be qualitatively similar. The fact that the helium must be advected inwards and that little mass is lost as the angular momentum is advected outwards is extremely important to establish. This is because angular momentum, essentially centrifugal force, has been suggested by Chevalier (1996) to hold up hypercritical accretion onto neutron stars in common envelope evolution. (Chevalier (1993) had first proposed the hypercritical accretion during this evolutionary phase to turn the neutron stars into black holes, the work followed up by Brown (1995) and Bethe & Brown (1998).) However, once matter is advected onto a neutron star, temperatures  $\geq 1$  MeV are reached so that neutrinos can carry off the energy. The accreted matter simply adds to the neutron star mass, evolving into an equilibrium configuration. Thus, this accretion does not differ essentially from that into a black hole. In either case of neutron star or black hole an accretion disc or accretion shock, depending on amount of angular momentum, but both of radius  $\sim 10^{11}$  cm, is first formed, giving essentially the same boundary condition for the hypercritical accretion in either case, black hole or neutron star. Thus, the masses of the black holes in transient sources argue against substantial inhibition of hypercritical accretion by jets, one of the Chevalier (1996) suggestions.

Measured mass functions, which give a lower limit on the black hole mass are given in Table 1. Only GRO J0422+32 and 4U 1543-47 have a measured mass function  $\leq 3 M_{\odot}$ . Results of Callanan et al. (1996) indicate that the angle  $i$  between the orbital plane and the plane of the sky for GRO J0422+32 is  $i < 45^\circ$ , and recent analysis by Orosz et al. (1998) indicate that the angle  $i$  for 4U 1543-47 is  $20^\circ < i < 40^\circ$ . So both GRO J0422+32 and 4U 1543-47 also contain high-mass black holes.

## 5. General discussion

There is agreement (Portegies Zwart et al., 1997; Ergma & Van den Heuvel, 1998) that in order to make enough transient sources the progenitors of the black holes must begin at relatively low masses,  $\sim$  ZAMS  $20 M_{\odot}$ . We take the upper limit somewhat arbitrarily to be  $\sim 35 M_{\odot}$  about where rapid mass

loss occurs and stars may enter into the LBV phase. Our upper limit is relatively unimportant since most of the stars considered will lie near the lower limit.

Based on the observations of Kaper et al. (1995) that the companion is a hypergiant, Ergma & Van den Heuvel (1998) argue that the progenitor of the neutron star in 4U1223-62 must have a ZAMS mass  $\gtrsim 50 M_{\odot}$ . Brown et al. (1996), by similar argumentation, arrived at  $\sim 45 M_{\odot}$ , but then had the difficulty that 4U1700-37, which they suggested to contain a low-mass black hole appeared to evolve from a lower mass star than the neutron star in 1223. Wellstein & Langer (1999) suggest the alternative that in 1223 the mass occurs in the main sequence phase (Case A mass transfer) which would be expected to be quasi conservative. They find that the progenitor of the neutron star in 1223 could come from a mass as low as  $26 M_{\odot}$ . This is in agreement with Brown et al. (1996) for conservative mass transfer (their Table 1), but these authors discarded this possibility, considering only RLOF (Case B mass transfer) in which case considerable mass would be lost.

Wellstein & Langer (1999) are in agreement with Brown et al. (1996) that 4U1700-37 should come from a quite massive progenitor. Conservative evolution here is not possible because of the short period of 3.4 days (Rubin et al., 1996). The compact object mass is here  $1.8 \pm 0.4 M_{\odot}$  (Heap & Corcoran, 1992). Brown et al. (1996) suggest that the compact object is a low-mass black hole. The upper mass limit for these was found by Brown & Bethe (1994) to be  $\sim 1.8 M_{\odot}$ , as compared with an upper limit for neutron star masses of  $\sim 1.5 M_{\odot}$ . Thus, there seems to be evidence for some ZAMS masses of  $\sim 40$ – $50 M_{\odot}$  ending up as low-mass compact objects, whereas we found that lower mass stars in the interval from  $\sim 20$ – $35 M_{\odot}$  ended up as high-mass black holes. In this sense we agree with Ergma & Van den Heuvel (1998) that low-mass compact object formation “is connected with other stellar parameters than the initial stellar mass alone”. We suggest, however, following Brown et al. (1996) that stars in binaries evolve differently from single stars because of the different evolution of the He core in binaries resulting from RLOF in their evolution. Namely, “naked” He cores evolve to smaller final compact objects than “clothed” ones.

In fact, this different evolution of binaries was found by Timmes et al. (1996). They pointed out that stars denuded of their hydrogen envelope in early RLOF in binaries would explode as Type Ib supernovae. They found the resulting remnant gravitational mass following explosion to be in the interval of  $1.2$ – $1.4 M_{\odot}$ , whereas in exploding stars of all masses with hydrogen envelope (Type II supernova explosion) they found a peak at about  $1.28 M_{\odot}$ , chiefly from stars of low masses and another peak at  $1.73 M_{\odot}$  more from massive stars. Our Fe core masses in Fig. 1 come from essentially the same calculations, but the “Remnant” masses of Woosley & Weaver (1995) are somewhat greater than those used by Timmes et al. (1996). In fact, the differences between the masses we plot and those of Timmes et al. come in the region  $\sim 1.7$ – $1.8 M_{\odot}$  (gravitational). This is just in the Brown & Bethe (1994) range for low-mass black holes. It may be that some of the stars with low-mass companions evolve into low-mass black holes. Presumably these would give lower luminosities than the high-mass black holes, although at upper end of the mass range we discuss 4U1700-37 seems to be an example of such a system. Of course here the high luminosity results from the high mass loss rate of the giant companion. There are substantial ambiguities in fallback, etc., from the explosion. Our point in this paper is that most of the higher mass single stars  $20 go into high mass black holes. (The Brown & Bethe (1994) limit for low-mass black hole formation is  $\sim 1.5$ – $1.8 M_{\odot}$  gravitational, but there is some give and take in both lower and upper limit. Also the stars are not all the same. In particular different metallicities will give different wind losses.)$

In determining the upper mass limit for which a low-mass compact object can result from binary evolution with RLOF (Case B mass transfer) the Brown et al. (1996) scenario must be revised because it is now realized that He-star wind loss rates employed by Woosley et al. (1995) were too high. With lower rates the differences in behavior between “clothed” and “naked” He core evolutions will be diminished because much of the difference arises from the mass losses and how it affects the convective  $^{12}\text{C}$  burning as we discuss below.

In the mass range  $\sim 20$ – $35 M_{\odot}$ , the compact objects resulting from naked He stars (Type Ib SN

explosions) are sufficiently far below the maximum neutron star mass that they will still remain neutron stars when better (lower) He star mass loss rates are employed. Indeed, for our  $25 M_{\odot}$  star, Wellstein & Langer (1999) find that halving the mass loss rate lowers the mass of the CO core by only 2.5%. However, at about the upper end of this range of masses, the outcome may, indeed, be changed. We will somewhat arbitrarily focus on the mass region  $\sim 40 M_{\odot}$ , studying how lower mass loss rates might affect the outcome.

Until recently W-R wind loss rates were taken from observed winds which originated chiefly from free-free scattering. These depend quadratically on density. Because of "clumpiness" in the winds, the mass loss rate was overestimated. Polarization measurements of the Thomson scattering, which depend linearly on the wind, give substantially lower mass rates, in approximate agreement with the rates that would be deduced from the observed rate of increase in orbital periods for spherical mass loss

$$\frac{\dot{M}}{M} = \frac{P}{2P}. \quad (5.1)$$

In V444 Cygni  $P = 0.202 \pm 0.018 \text{ s yr}^{-1}$  was obtained by Khaliullin et al. (1984) and  $M_{\text{WR}} = 9.3 \pm 0.5 M_{\odot}$  by Marchenko et al. (1994), resulting in

$$\dot{M}_{\text{dyn}} = 1 \times 10^{-5} M_{\odot} \text{ yr}^{-1}. \quad (5.2)$$

This is to be compared with the

$$\dot{M} = 0.75 \times 10^{-5} M_{\odot} \text{ yr}^{-1} \quad (5.3)$$

obtained by St.-Louis et al. (1993). In later work Moffat & Robert (1994) arrive at a mean of  $(0.7 \pm 0.1) \times 10^{-5} M_{\odot} \text{ yr}^{-1}$ . The mass loss rate employed by Woosley et al. (1995) was that of Langer (1989). Specifically,  $\dot{M} = -kM^{2.5}$ , with  $M$  in  $M_{\odot}$  and  $\dot{M}$  in  $M_{\odot} \text{ yr}^{-1}$  and  $k = 6 \times 10^{-8}$  as long as the carbon surface mass fraction does not exceed 0.02 and  $k = 10^{-7}$  afterwards. Choosing an average of  $k = 8 \times 10^{-8}$  we find the WLW rate to be  $2.1 \times 10^{-5} M_{\odot} \text{ yr}^{-1}$ , a factor of 2 larger than  $\dot{M}_{\text{dyn}}$ . Given the many uncertainties in our estimate, we feel the range of two to three times less than the Langer (1989) mass loss rate to be reasonable. It should also be remembered that stars vary substantially in metallicity, and that there will be a range of variation even

of those in the Galactic disc. Of course stars in the metal poor Magellanic clouds should have substantially lower winds.

In the region of ZAMS  $\sim 40 M_{\odot}$  calculation of Wellstein & Langer (1999) show that both the final He core mass and C/O core mass increase  $\sim 23\%$  in the mass loss rate is halved. Thus, reducing the mass loss rate by a factor of 2 does not increase the final mass by a similar factor. The main reason is that if the mass loss is reduced, then the He stars remain somewhat more massive, thus also more luminous and therefore have a higher mass loss rate than had their mass been reduced earlier. There are also other, less important, feedbacks.

In the case of  $\sim 60 M_{\odot}$  stars, the He core and C/O cores increase  $\sim 31\%$  if the mass loss rate is halved and a factor of  $\sim 1.8\text{--}1.9$  if it is cut by  $1/4$ . In fact, in this case, a reasonable interpolation formula for these lower mass loss rates is

$$\frac{M}{M_0} = (1.33)^{1/2f}, \quad (5.4)$$

where  $M_0$  was calculated with the Langer (1989) rates and  $f$  is the fractional decrease in winds from these rates. Wellstein & Langer (1999) also evolve a  $36 M_{\odot}$  He core, which would come from an  $\sim 85 M_{\odot}$  ZAMS star, with  $f = 1/4$  of the Langer (1989) wind loss rates. Their final  $7.5 M_{\odot}$  He star mass is not enough for Cyg X-1. However, Woosley et al. (1995) with the Langer (1989) mass loss rate found a  $9.71 M_{\odot}$  He core in evolving a ZAMS  $85 M_{\odot}$  star with mass loss. In the latter case the WNL phase lasted  $> 1/3$  of the W-R phase; i.e., the He core initially had some hydrogen envelope. This illustrates the order of uncertainties that may come in the evolution of very massive stars, arising from mass loss in the LBV phase.

Although the final He core and C/O core scale by roughly the same factor the Fe core, crucial for the compact object masses, is not expected to do the same. Firstly, rather trivially, even if the baryon number Fe core did scale in the same way, one would expect corrections downwards in core mass from binding energy correction. These decreases in gravitational mass would be greater for higher mass cores, so the Fe cores would scale somewhat less rapidly than the C/O cores.

More important for the mass of the Fe core is the

ZAMS mass at which the convective carbon burning is skipped, because, as seen in the single stars in Fig. 1, a big jump in Fe core masses occurs here.

The convective carbon burning phase (when it occurs) is extremely important in presupernova evolution, because this is the first phase in which a large amount of entropy can be carried off in  $\nu\bar{\nu}$ -pair emission, especially because this phase is of long duration. The reaction in which carbon burns is  $^{12}\text{C}(\alpha, \gamma)^{16}\text{O}$  (other reactions like  $^{12}\text{C} + ^{12}\text{C}$  would require excessive temperatures). The cross section of  $^{12}\text{C}(\alpha, \gamma)^{16}\text{O}$  is still not accurately determined; the lower this cross section the higher the temperature of the  $^{12}\text{C}$  burning, and therefore the more intense the  $\nu\bar{\nu}$  emission. With the relatively low  $^{12}\text{C}(\alpha, \gamma)^{16}\text{O}$  rates<sup>5</sup> determined both directly from nuclear reactions and from nucleosynthesis by Weaver & Woosley (1993), the entropy carried off during  $^{12}\text{C}$  burning in the stars of ZAMS mass  $\leq 18 M_{\odot}$  is substantial. The result is rather low-mass Fe cores for these stars, which can evolve into neutron stars. Note that in the literature earlier than Weaver & Woosley (1993) often large  $^{12}\text{C}(\alpha, \gamma)^{16}\text{O}$  rates were used, so that the  $^{12}\text{C}$  was converted into oxygen and the convective burning did not have time to be effective. Thus its role was not widely appreciated.

Of particular importance is the ZAMS mass at which the convective carbon burning is skipped. In the Woosley & Weaver (1995) calculations this occurs for single stars at ZAMS mass  $19 M_{\odot}$ , but with a slightly lower  $^{12}\text{C}(\alpha, \gamma)^{16}\text{O}$  rate it might come at  $20 M_{\odot}$  or higher (Brown, 1997). As the progenitor mass increases, it follows from general polytropic arguments that the entropy at a given burning stage increases. At the higher entropies of the more massive stars the density at which burning occurs is lower, because the temperature is almost fixed for a given fuel. Lower densities decrease the rate of the triple- $\alpha$  process which produces  $^{12}\text{C}$  relative to the two-body  $^{12}\text{C}(\alpha, \gamma)^{16}\text{O}$  which produces oxygen. Therefore, at the higher entropies in the more massive stars the ratio of  $^{12}\text{C}$  to  $^{16}\text{O}$  at the end of He burning is lower. The star skips the long convective

carbon burning and goes on to the much shorter oxygen burning. Oxygen burning goes via  $^{16}\text{O} + ^{16}\text{O}$  giving various products, at very much higher temperature than  $^{12}\text{C}(\alpha, \gamma)^{16}\text{O}$  and much faster. Since neutrino cooling during the long carbon-burning phase gets rid of a lot of entropy of the core, skipping the convective carbon burning phase leaves the core entropy higher and the final Chandrasekhar core fatter.

In Fig. 1 the large jump in compact object mass in single stars at ZAMS mass  $\sim 19 M_{\odot}$  is clearly seen. From our discussion in the last section we see that this is just about at the point where our Fe core mass goes above  $\sim 1.8 M_{\odot}$  and, therefore, above this mass one would expect single stars to go into high-mass black holes. Arguments have been given that SN (1987A) with progenitor ZAMS mass of  $\sim 18 M_{\odot}$  evolved into a low-mass black hole (Brown & Bethe, 1994). We believe from our above arguments and Fig. 1 that soon above the ZAMS mass of  $\sim 18 M_{\odot}$ , single stars go into high-mass black holes without return of matter to the Galaxy. Thus, the region of masses for low-mass black hole formation in single stars is narrow. The precise upper mass limit is not clear, but certainly in the range of  $\gtrsim 20 M_{\odot}$  ZAMS. This is in agreement with Wellstein & Langer (1999) who find a minimum black hole progenitor mass for single stars of  $21 M_{\odot}$ .

Thus far our discussion has been chiefly about single stars, in which the He burns while clothed by a hydrogen envelope.

In binary evolution, if the hydrogen envelope is removed before the helium core burning in either Case A or Case B (RLOF) mass transfer, the resulting "naked" He star burns quite differently from a "clothed" one. Weaver & Woosley (1993) find that the convective carbon burning tends to be skipped when the central  $^{12}\text{C}$  abundance at the end of helium core burning is less than  $\sim 15\%$ . In Woosley et al. (1995) with the Langer (1989) helium mass loss rate, this central  $^{12}\text{C}$  abundance is 34% for a ZAMS  $40 M_{\odot}$  star. With mass loss rate decreased to half, it is still 33% (Wellstein & Langer, 1999) hardly changed. However, as noted earlier, the CO mass for a ZAMS  $40 M_{\odot}$  star is increased from  $2.33 M_{\odot}$  to  $2.87 M_{\odot}$  with halved mass loss rate. Provisionally Wellstein & Langer (1999) have suggested that the magnitude of the CO mass chiefly

<sup>5</sup>Weaver & Woosley use  $S(E) = S(300) = 170$  keV barn, remarkably close to the  $169 \pm 55$  keV barn arrived at by Barnes (1995). Given the large uncertainty (stemming chiefly from that in the E2 rate), the good agreement may be somewhat accidental.

determines the fate of the star, and that the  $2.87 M_{\odot}$  CO core could go into a black hole. However, halving the mass loss rate of the naked He star decreases the central carbon abundance at the end of He core burning hardly at all. Thus, there will still be a long period of  $^{12}\text{C}$  convective core burning. As seen from Fig. 1 for less massive stars in the mass region  $\sim 20 M_{\odot}$ , this can easily lower the compact core mass by  $\sim 0.5 M_{\odot}$ . On the other hand, for each  $1 M_{\odot}$  added to the He envelope at the envelope at the time of SN explosion, an additional  $\sim 3 \times 10^{49}$  erg is needed to expel it in either a prompt or delayed supernova explosion. This is  $\sim 2\%$  of the SN explosion energy, so would work towards formation of a high-mass black hole.

Given the increase in He and CO cores from the decreased mass loss rates, we estimate that stars of ZAMS masses  $\sim 40 M_{\odot}$  which lose their H envelopes by wind will end up as low-mass black holes, 1700-37 being an example. The possible mass ranges in which this can happen cannot be determined until the dynamic evolution of the CO cores formed with the lowered He-star wind rates of Wellstein & Langer (1999) is carried out.

It seems clear that with the lower metallicity in the LMC and consequently lower mass loss rates that the  $^{12}\text{C}$  convective core burning will not be skipped so the Fe cores in the  $\geq 40 M_{\odot}$  region of ZAMS masses will be larger, and will evolve into black holes. This may help to explain why there are two HMXB's with high-mass black holes, LMC X-1 and LMC X-3, in the LMC, whereas Cyg X-1 is the only clear example in the disc.

## 6. Conclusion

We have shown that it is likely that single stars in the range of ZAMS masses  $\sim 20\text{--}35 M_{\odot}$  evolve into high-mass black holes without return of matter to the Galaxy. This results because at mass  $\sim 20 M_{\odot}$  the convective carbon burning is skipped and this leads to substantially more massive Fe cores. Even with more realistic reduced mass loss rates on He stars, it is unlikely that stars in this mass range in binaries evolve into high-mass black holes, because the progenitor of the compact object when stripped of its hydrogen envelope in either Case A (during main

sequence) or Case B (RLOF) mass transfer will burn as a “naked” He star, ending up as an Fe core which is not sufficiently massive to form a high-mass black hole.

In the region of ZAMS mass  $\sim 40 M_{\odot}$ , depending sensitively on the rate of He-star wind loss, the fate of the primary in a binary may be a low-mass black hole. We are unable to pin down the limit for high-mass black hole formation until better mass loss rates are determined.

In our estimates we have assumed the Brown & Bethe (1994) estimates of  $1.5 M_{\odot}$  for maximum neutron star mass and  $1.5\text{--}1.8 M_{\odot}$  for the range in which low-mass black holes can result.

In our evolution of the transient sources using Case C (during He shell burning) mass transfer, almost the entire He core will collapse into a high-mass black hole, explaining the more or less common black hole mass of  $\sim 7 M_{\odot}$  for these objects, with the possible exception of V404 Cyg where the mass may be greater. Our evolution gives an explanation for the seemingly large gap in black-hole masses, between the  $\geq 1.5 M_{\odot}$  for the black hole we believe was formed in 1987A and the  $\sim 1.8 M_{\odot}$  black hole we suggest in 1700-37 and the  $\sim 7 M_{\odot}$  in the transient sources.

We note that following the removal of the H envelope by Case C mass transfer, the collapse inwards of the He envelope into the developing black hole offers the Collapsar scenario for the most energetic gamma ray bursters of Woosley (1993) and MacFadyen & Woosley (1999). Especially if our higher estimate of  $\sim 10^5$  of the high-mass black-hole, main-sequence star binaries is roughly correct, a simple estimate (Brown et al., 1999) shows that they would be the largest population of possible progenitors.

## Acknowledgements

We dedicate this work to Dave Schramm, a giant among astrophysicists. One of the authors (G.E.B.) invited him to edit the astrophysics section of Physics Reports and enjoyed his collegiality for many years. We are especially grateful to S. Wellstein and N. Langer for communicating their results with

modified mass loss rates to us before publication. We would like to thank Charles Bailyn for useful discussions and, especially, for putting the data in Table 1 together for us. We would like to thank Stan Woosley, who not only provided us with most of the results we used, but also clarified their interpretation in several communications. We also thank Adam Burrows for a clear explanation of the consequences of skipping the convective carbon burning. We wish to thank Simon Portegies Zwart for several helpful discussions. We were supported by the U.S. Department of Energy under Grant No. DE-FG02-88ER40388.

## References

- Bailyn, C.D., Jain, R.K., Coppi, P., & Orosz, J.A., 1998, *ApJ*, 499, 367.
- Barnes, C.A., 1995, *Nucl. Phys. A*, 588, 295c.
- Bethe, H.A. & Brown, G.E., 1998, *ApJ*, 506, 780.
- Bethe, H.A. & Brown, G.E., 1999, *ApJ*, 517, 318.
- Brown, G.E., 1995, *ApJ*, 440, 270.
- Brown, G.E., 1997, *Phys. Bl.*, 53, 671.
- Brown, G.E. & Bethe, H.A., 1994, *ApJ*, 423, 659.
- Brown, G.E., Wijers, R.A.M.J., Lee, C.-H., Lee, H.K., Bethe, H.A., 1999, *astro-ph/9905337*.
- Brown, G.E., Weingartner, J.C., & Wijers, R.A.M.J., 1996, *ApJ*, 463, 297.
- Callanan, P.J., et al. 1996, *ApJ*, 461, 351.
- Cappellaro, E., et al. 1997, *A&A*, 322, 431.
- Chen, W., Shrader, C.R., & Livio, M., 1997, *ApJ*, 491, 312.
- Chevalier, R.A., 1993, *ApJ*, 411, L33.
- Chevalier, R.A., 1996, *ApJ*, 459, 322.
- Cordes, J.M. & Chernoff, D.F., 1998, *ApJ*, 505, 315.
- De Kool, M., Van den Heuvel, E.P.J., & Pylyser, E., 1987, *A&A*, 183, 47.
- Ergma, E. & Van den Heuvel, E.P.J., 1998, *A&A*, 331, L29.
- Haswell, C.A., Robinson, E.L., Horne, K., Stiening, R.F., & Abbott, T.M.C., 1993, *ApJ*, 411, 802.
- Heap, S.R. & Corcoran, M.F., 1992, *ApJ*, 387, 340.
- Kaper, L., Lamers, H.J.G.L.M., Van den Heuvel, E.P.J., & Zuiderwijk, E.J., 1995, *A&A*, 300, 446.
- Khaliullin, K.F., Khaliullina, A.I., & Cherepashchuk, A.M., 1984, *Sov. Astron. Lett.*, 10, 250.
- Langer, N., 1989, *A&A*, 220, 135.
- MacFadyen, A. & Woosley, S.E., 1999, *ApJ*, accepted.
- Marchenko, S.V., Moffat, A.F.J., & Koenigsberger, G., 1994, *ApJ*, 422, 810.
- Mineshige, S., Nomoto, K., & Shigeyama, T., 1993, *A&A*, 267, 95.
- Mineshige, S., Nomura, H., Hirose, M., Nomoto, K., & Suzuki, T., 1997, *ApJ*, 489, 227.
- Moffat, A.F.J. & Robert, C., 1994, *ApJ*, 421, 310.
- Orosz, J.A., et al. 1998, *ApJ*, 499, 375.
- Portegies Zwart, S.F., Verbunt, F., & Ergma, E., 1997, *A&A*, 321, 207.
- Rubin, B.C., et al. 1996, *ApJ*, 459, 259.
- Ruffert, M., 1994, *ApJ*, 427, 351.
- Ruffert, M. & Arnett, D., 1994, *A&AS*, 106, 505.
- St.-Louis, N., Moffat, A.F.J., Lapointe, L., Efimov, Y.S., Shakhskoy, N.M., Fox, G.K., & Pirola, V., 1993, *ApJ*, 410, 342.
- Schaller, G., Schaerer, D., Meynet, G., & Maeder, A., 1992, *A&AS*, 96, 269.
- Timmes, F.X., Woosley, S.E., & Weaver, T.A., 1996, *ApJ*, 457, 834.
- Van den Heuvel, E.P.J., 1995, *JApA*, 16, 255.
- Van den Heuvel, E.P.J. & Habets, G.M.H.J., 1984, *Natur*, 309, 598.
- Verbunt, F. & Van den Heuvel, E.P.J., 1995, in *X-ray Binaries*, Cambridge Univ. Press, Eds. W.H.G. Lewin, J. van Paradijs & E.P.J. Van den Heuvel, p. 457.
- Verbunt, F. & Zwaan, C., 1981, *A&A* 100, L7.
- Weaver, T.A. & Woosley, S.E., 1993, *Phys. Rep.*, 227, 65.
- Weaver, T.A., Zimmerman, G.B., & Woosley, S.E., 1978, *ApJ*, 225, 1021.
- Wellstein, S. & Langer, N., 1999, to be published.
- Wijers, R.A.M.J., 1996, *Evolutionary Processes in Binary Stars*, Eds. R.A.M.T. Wijers et al., Kluwer, Dordrecht, pp. 327–344.
- Woosley, S.E., 1993, *ApJ*, 405, 273.
- Woosley, S.E., Langer, N., & Weaver, T.A., 1995, *ApJ*, 448, 315.
- Woosley, S.E. & Weaver, T.A., 1995, *ApJS*, 101, 181.

This page is intentionally left blank

## Chapter 15

# Evolution of Black Holes in the Galaxy

G.E. Brown, C.-H. Lee, R.A.M.J. Wijers and H.A. Bethe

Reprinted with permission from *Physics Reports*, 333–334 (2000) 471–504

Copyright (2000) by Elsevier Science B.V.

### **Commentary**

Dave Schramm had died tragically in an airplane crash. We decided to summarize our work to date in this review article, which is part of the Schramm Memorial Volume of *Physics Reports*.

This page is intentionally left blank



ELSEVIER

## PHYSICS REPORTS

Physics Reports 333–334 (2000) 471–504

[www.elsevier.com/locate/physrep](http://www.elsevier.com/locate/physrep)

## Evolution of black holes in the galaxy

G.E. Brown<sup>a,\*</sup>, C.-H. Lee<sup>a</sup>, R.A.M.J. Wijers<sup>a</sup>, H.A. Bethe<sup>b</sup>

<sup>a</sup>*Department of Physics & Astronomy, State University of New York, Stony Brook, NY 11794-3800, USA*

<sup>b</sup>*Floyd R. Newman Laboratory of Nuclear Studies, Cornell University, Ithaca, NY 14853, USA*

### Abstract

In this article we consider the formation and evolution of black holes, especially those in binary stars where radiation from the matter falling on them can be seen. We consider a number of effects introduced by some of us, which are not traditionally included in binary evolution of massive stars. These are (i) hypercritical accretion, which allows neutron stars to accrete enough matter to collapse to a black hole during their spiral-in into another star. (ii) The strong mass loss of helium stars, which causes their evolution to differ from that of the helium core of a massive star. (iii) The direct formation of low-mass black holes ( $M \sim 2M_{\odot}$ ) from single stars, a consequence of a significant strange-matter content of the nuclear-matter equation of state at high density. We discuss these processes here, and then review how they affect various populations of binaries with black holes and neutron stars.

We have found that hypercritical accretion changes the standard scenario for the evolution of binary neutron stars: it now usually gives a black-hole, neutron-star (BH-NS) binary, because the first-born neutron star collapses to a low-mass black hole in the course of the evolution. A less probable double helium star scenario has to be introduced in order to form neutron-star binaries. The result is that low-mass black-hole, neutron star (LBH-NS) binaries dominate the rate of detectable gravity-wave events, say, by LIGO, by a factor  $\sim 20$  over the binary neutron stars.

The formation of high-mass black holes is suppressed somewhat due to the influence of mass loss on the cores of massive stars, raising the minimum mass for a star to form a massive BH to perhaps  $80M_{\odot}$ . Still, inclusion of high-mass black-hole, neutron-star (HBH-NS) binaries increases the predicted LIGO detection rate by another  $\sim 30\%$ ; lowering of the mass loss rates of Wolf-Rayet stars may lower the HBH mass limit, and thereby further increase the merger rate.

We predict that  $\sim 33$  mergers per year will be observed with LIGO once the advanced detectors planned to begin in 2004 are in place.

Black holes are also considered as progenitors for gamma ray bursters (GRB). Due to their rapid spin, potentially high magnetic fields, and relatively clean environment, mergers of black-hole, neutron-star

\* Corresponding author.

E-mail addresses: [popenoe@nuclear.physics.sunysb.edu](mailto:popenoe@nuclear.physics.sunysb.edu) (G.E. Brown), [chlee@silver.physics.sunysb.edu](mailto:chlee@silver.physics.sunysb.edu) (C.-H. Lee), [rwijers@astro.sunysb.edu](mailto:rwijers@astro.sunysb.edu) (R.A.M.J. Wijers).

binaries may be especially suitable. Combined with their 10 times greater formation rate than binary neutron stars this makes them attractive candidates for GRB progenitors, although the strong concentration of GRBs towards host galaxies may favor massive star progenitors or helium-star, black-hole mergers.

We also consider binaries with a low-mass companion, and study the evolution of the very large number of black-hole transients, consisting of a black hole of mass  $\sim 7M_{\odot}$  accompanied by a K or M main-sequence star (except for two cases with a somewhat more massive subgiant donor). We show that common envelope evolution must take place in the supergiant stage of the massive progenitor of the black hole, giving an explanation of why the donor masses are so small. We predict that there are about 22 times more binaries than observed, in which the main-sequence star, somewhat more massive than a K- or M-star, sits quietly inside its Roche Lobe, and will only become an X-ray source when the companion evolves off the main sequence.

We briefly discuss the evolution of low-mass X-ray binaries into millisecond pulsars. We point out that in the usual scenario for forming millisecond pulsars with He white-dwarf companions, the long period of stable mass transfer will usually lead to the collapse of the neutron star into a black hole. We then discuss Van den Heuvel's "Hercules X-1 scenario" for forming low-mass X-ray binaries, commenting on the differences in accretion onto the compact object by radiative or semiconvective donors, rather than the deeply convective donors used in the earlier part of our review.

In Appendix A we describe the evolution of Cyg X-3, finding the compact object to be a black hole of  $\sim 3M_{\odot}$ , together with an  $\sim 10M_{\odot}$  He star. In Appendix B we do the accounting for gravitational mergers and in Appendix C we show low-mass black-hole, neutron-star binaries to be good progenitors for gamma ray bursters. © 2000 Elsevier Science B.V. All rights reserved.

PACS: 97.60.Lf; 97.60. – s; 98.70.Rz

**Keywords:** Black hole physics – stars; Binaries; Close – gamma rays; Bursts – stars; Neutron – gravity waves – accretion

## 1. Introduction

The fate of massive stellar cores, both in single and binary stars, has many observable consequences, both for what types of compact object may be found in what type of binary, and for the formation rates of all types of compact-object binary. We have discussed various aspects of this problem in previous works, and here give an overview of all these together, applying the same set of principles to all and obtaining a consistent picture of the evolution of massive stars and binaries.

The best-known compact-object (i.e., neutron star or black hole) binaries are the binary neutron stars. They are key testing grounds of general relativity, and the usually favored gravity-wave source for LIGO. Until recently the theoretical formation rate of binary neutron stars gave at least one order-of-magnitude higher rate than was arrived at empirically by extrapolation from observed binary neutron stars. Because there are few binary neutron stars, and even fewer dominate the empirical estimates, the latter are frequently revised. The recent doubling of the estimated distance to PSR 1534 + 12 [2] has lowered the empirical birth rate significantly, widening the gap.

A solution to this discrepancy comes from combining the strange-matter equation of state, which results in a relatively low maximum mass for neutron stars, with hypercritical accretion [1]. In the standard scenario the first neutron star formed spirals into the other star, in a phase of

common-envelope evolution. Bethe and Brown [1] argued that when a neutron star spirals into a red giant with a deeply convective envelope, it accretes matter at a very high rate of up to  $1M_{\odot} \text{ yr}^{-1}$ . Photons are trapped in the flow and carried adiabatically inwards to the surface of the neutron star [3]. The latter is heated to  $T \sim 1 \text{ MeV}$ , temperatures at which neutrino emission can carry off the thermal energy. Hence, the Eddington limit of  $\dot{M}_{\text{Edd}} \sim 1.5 \times 10^{-8} M_{\odot} \text{ yr}^{-1}$  does not apply. As a result, the neutron star accretes about a solar mass of material and collapses to a low-mass black hole. Only if the two stars are initially so close in mass that at the time of spiral-in the first supernova has not yet exploded (i.e. the object that spirals in is still a helium star) a binary neutron star is formed. The sum total of binary neutron stars and black-hole, neutron-star binaries is almost the same as what was found for binary neutron stars in previous estimates, but now the binary neutron stars are only a small fraction of the total. The result is that an order of magnitude more black-hole, neutron-star binaries than binary neutron stars are formed. Together with the fact that the black holes are somewhat more massive than neutron stars, this implies that binaries with black holes should play an important part in mergers producing gravitation waves. They may also be good candidates for producing gamma-ray bursts.

No low-mass black-hole, neutron-star binaries have been observed. This is due to the fact that the one neutron star in them is unrecycled, hence is observable for only a short time. The rarer binary neutron stars, like PSR 1913 + 16, do have a long-lived recycled pulsar, which more than offsets their lower formation rate and makes them dominate the observed population.

We do observe high-mass black holes in Cyg X-1 and in soft X-ray transients. In the former, the black hole is of  $\gtrsim 10M_{\odot}$  [4]. The companion O-star is near its Roche Lobe, and its wind is continuously feeding the black hole, which shines through X-ray emission. In addition to Cyg X-1, high-mass black holes are seen in the LMC in LMC X-3 and perhaps LMC X-1. Much more copious are the transient sources, with black holes of mass  $M_{\text{BH}} \sim 7M_{\odot}$ , most of which flare up only occasionally with long quiescent times between flare ups. Wijers [5] estimated  $\sim 3000$  of these in the Galaxy. That is, these are the numbers that are presently operative. Remarkable about the transient sources with unevolved donors is that the main sequence star is K- or M-star, less massive than our sun. Brown et al. [6] explain this in terms such that higher-mass donors can also participate in the formation of such binaries containing a high-mass black hole, but will end up in the evolution further away from the black hole so that they can pour matter on the latter only when they evolve in subgiant or giant stage. Thus, there are a large factor estimated to be  $\sim 22$  more of those binaries which will not be seen until the main sequence star evolves [6]. The mechanism describing the evolution of the transient sources required the massive progenitor of the black hole to carry out core helium burning as if it were a single star, i.e., before having its H envelope removed in RLOF by its main sequence companion. An interval of  $\sim 20\text{--}35M_{\odot}$  ZAMS was estimated for the progenitors of the high-mass black hole. Consequently, this same interval of single stars, not in binary, would be expected to end up as high-mass black holes. In the formation of these high-mass black holes, most of the helium envelope of the progenitor must drop into the black hole in order to form their high mass, so little matter is returned to the Galaxy.

This brings us to the intriguing matter of SN 1987A which we believe did go into a black hole, but after the supernova explosion which did return matter to the Galaxy. The progenitor of SN 1987A was known to have ZAMS mass  $\sim 18M_{\odot}$ . This leads us to the interesting concept of low-mass black holes with delayed explosion, which result from the ZAMS mass range  $\sim 18\text{--}20M_{\odot}$ , although the precise interval is uncertain. The delayed explosion mechanism has

been elucidated by Prakash et al. [7]. The range of ZAMS masses of single stars in which neutron stars are formed is thus only  $\sim 10\text{--}18M_{\odot}$ .

The absence of matter being returned to the Galaxy in the ZAMS mass range  $\sim 20\text{--}35M_{\odot}$  impacts on nucleosynthesis, especially in the amount of oxygen produced. Bethe and Brown [8] suggested that matter was again returned to the Galaxy by stars in the ZAMS range  $\sim 35\text{--}80M_{\odot}$ . In this case, the progenitor was stripped of H envelope in an LBV phase, and the naked He star was suggested to evolve to a low-mass black hole, with return of matter to the galaxy before its formation in a delayed explosion, or to a neutron star. Thus, elements like oxygen were produced in a bimodal distribution of ZAMS masses  $M \lesssim 20M_{\odot}$  and  $35M_{\odot} \lesssim M \lesssim 80M_{\odot}$ .

The Bethe and Brown [8] suggestion was based on naked He stars evolved by Woosley et al. [9], who used a too-large wind loss rate for He stars. Wellstein and Langer [10] have evolved naked He stars with lower rates, in which case the final He envelope is somewhat larger. However, the central carbon abundance following core He burning is high  $\sim 33\%$ . With this abundance, the stars will not skip the convective carbon burning stage in their evolution, and according to the arguments of Brown et al. [6] would still be expected to end up as low-mass compact objects, in which case matter would be returned to the Galaxy. This matter will not, however, be settled until the CO cores evolved with lowered He-star wind loss rates by Wellstein and Langer have been burned further up to the Fe core stage, so the Bethe and Brown [8] bimodal mass region for nucleosynthesis should be viewed as provisional.

In Section 2, we discuss the maximum mass of neutron stars and the processes that determine which range of initial stellar masses gives rise to what compact object, and how mass loss in naked helium stars changes those ranges. Then we describe the Bethe and Brown [1] scenario for the evolution of massive binary stars, and especially their treatment of common-envelope evolution and hypercritical accretion (Section 3). We then discuss a few specific objects separately, first binary neutron stars (Section 4), then Cyg X-1 and its ilk (Section 5) and the black-hole transients (Section 6). Then we comment briefly on how our results would affect the evolution of low-mass X-ray binaries with neutron stars (Section 7) and summarize our conclusions (Section 8). The discussion of Cyg X-3 and the possible implications of neutron-star, black-hole binaries for gravity waves and gamma-ray bursts are in Appendices A–C.

## 2. The compact star

Thorsson et al. [11] and Brown and Bethe [12] have studied the compact core after the collapse of a supernova, assuming reasonable interactions between hadrons. Initially, the core consists of neutrons, protons and electrons and a few neutrinos. It has been called a proto-neutron star. It is stabilized against gravity by the pressure of the Fermi gases of nucleons and leptons, provided its mass is less than a limiting mass  $M_{\text{PC}}$  (proto-compact) of  $\sim 1.8M_{\odot}$ .

If the assembled core mass is greater than  $M_{\text{PC}}$  there is no stability and no bounce; the core collapses immediately into a black hole. It is reasonable to take the core mass to be equal to the mass of the Fe core in the pre-supernova, and we shall make this assumption, although small corrections for fallback in the later supernova explosion can be made as in Brown et al. [13]. If the center collapses into a black hole, the outer part of the star has no support (other than centrifugal force from angular momentum) and will also collapse.

If the mass of the core is less than  $M_{PC}$ , the electrons will be captured by protons



and the neutrinos will diffuse out of the core. This process takes of order of 10 s, as has been shown by the duration of the neutrino signal from SN 1987A. The result is a neutron star, with a small concentration of protons and electrons. The Fermi pressures of the core are chiefly from the nucleons, with small correction from the electrons. On the other hand, the nucleon energy is increased by the symmetry energy, i.e., by the fact that we now have nearly pure neutrons instead of an approximately equal number of neutrons and protons. Thorsson et al. [11] have calculated that the maximum mass of the neutron star  $M_{NS}$  is still about  $1.8M_\odot$ , i.e., the symmetry energy compensates the loss of the Fermi energy of the leptons. Corrections for thermal pressure are small [14].

The important fact is that the 10 s of neutrino diffusion from the core give ample time for the development of a shock which expels most of the mass of the progenitor star.

But this is not the end of the story. The neutrons can convert into protons plus  $K^-$  mesons,



This is short-hand for the more complicated interaction  $N + e^- \rightarrow N' + K^- + \nu$  where  $N$  is a nucleon. The neutrinos leave the star. The times are sufficiently long that chemical equilibrium is assured. Since the density at the center of the neutron star is very high, the energy of the  $K^-$  is very low, as confirmed by Li et al. [15] using experimental data. By this conversion the nucleons can again become half neutrons and half protons, thereby saving the symmetry energy needed for pure neutron matter. The  $K^-$ , which are bosons, will condense, saving the kinetic energy of the electrons they replace. The reaction equation (2) will be slow, since it is preceded by



(with the reaction equation (2) following) as it becomes energetically advantageous to replace the fermionic electrons by the bosonic  $K^-$ 's at higher densities. Initially, the neutrino states in the neutron star are filled up to the neutrino chemical potential with trapped neutrinos, and it takes some seconds for them to leave the star. These must leave before new neutrinos can be formed from the process equation (3). Thorsson et al. [11] have calculated that the maximum mass of a star in which reaction equation (2) has gone to completion is

$$M_{NP} \simeq 1.5M_\odot , \quad (4)$$

where the lower suffix NP denotes their nearly equal content of neutrons and protons, although we continue to use the usual name “neutron star”. This is the maximum mass of neutron stars, which is to be compared with the masses determined in binaries. The masses of 19 neutron star masses determined in radio pulsars [16] are consistent with this maximum mass.

The core mass  $M_C$  formed by the collapse of a supernova must therefore be compared to the two limiting masses,  $M_{PC}$  and  $M_{NP}$ . If

$$(I) \quad M_C > M_{PC} , \quad (5)$$

we get a high mass black hole. If

$$(II) \quad M_{\text{PC}} > M_{\text{C}} > M_{\text{NP}}, \quad (6)$$

we get a low-mass black hole, of mass  $M_{\text{C}}$ . Only if

$$(III) \quad M_{\text{C}} < M_{\text{NP}} \quad (7)$$

do we get a neutron (more precisely, “nucleon”) star from the SN. Only in this case can we observe a pulsar. In cases (II) and (III) we can see a supernova display. In case (I) we receive only initial neutrinos from electrons captured in the collapse before  $M_{\text{C}}$  becomes greater than  $M_{\text{PC}}$  but no light would reach us. (Except perhaps if the new black hole rotates rapidly enough to power an explosion, a mechanism proposed by MacFadyen and Woosley [93] for gamma-ray bursts.)

Woosley et al. [17] evolve massive stars with mass loss. For stars in the ZAMS mass range  $\sim 20\text{--}30M_{\odot}$ , mass loss is relatively unimportant and since  $M_{\text{PC}} \gtrsim 1.8M_{\odot}$  for this range, we find from the earlier calculation of Woosley and Weaver [18] that most of the single stars in this range will go into high-mass black holes. Evolution of these stars in binaries is another matter. Timmes et al. [19], Brown et al. [13], and Wellstein and Langer [10] find that substantially smaller core masses result if the hydrogen envelope is taken off in RLOF so that the helium star is naked when it burns. Thus, stars of ZAMS masses  $\sim 20\text{--}35M_{\odot}$  in such binaries evolve into low-mass compact cores, black hole or neutron star. Woosley et al. [17] used helium-star wind loss rates which were too high by a factor  $\sim 2\text{--}3$ , but lower wind losses give only slightly larger He cores in the ZAMS mass range  $\sim 20\text{--}35M_{\odot}$  [10] so our above conclusion is unlikely to be reversed.

On the other hand, the fate of single stars in the ZAMS mass range  $\sim 35\text{--}80M_{\odot}$  is uncertain. In the published Woosley et al. [17] work with too high mass loss rate, so much matter is blown away, first in LBV stage and later in W-R stage that low-mass compact objects, black-hole or neutron-star, result [13]. Bethe and Brown [8] attribute this to the fact that convective carbon burning is not skipped in these stars. In this stage a lot of entropy can be removed by  $v\bar{v}$  emission, so that a low entropy, and therefore small, core results. In this range, Wellstein and Langer [10] find central  $^{12}\text{C}$  abundances of 33–35% following He core burning, more than double the  $\sim 15\%$  required for convective carbon core burning. Therefore, we believe that this range of stars will still go into low-mass compact objects, even though their final He cores are substantially larger because of the lower, more correct, He-star wind mass loss rates used by Wellstein and Langer [10]. However, this problem cannot be considered as settled until the Wellstein and Langer CO cores are burned up to the Fe core stage. We will therefore not discuss the evolution of Cyg X-1 like objects, high-mass black holes accompanied by sufficiently massive giant companion so that they shine continuously in X-rays. It is not clear to us whether LMC X-3, with a high-mass black hole and a B-star companion of roughly equal mass, has a history more like Cyg X-1 or like the transient black-hole binaries which we discuss below.

Bethe and Brown [8] took  $80M_{\odot}$  as the lower mass limit for high-mass black-hole formation in binaries which experience RLOF, i.e., in those for which helium core burning proceeds in a naked helium star. Because of our above discussion, we believe this mass limit may be too high, so that the contributions from high-mass black-hole, neutron-star binaries were, if anything, underestimated in their work. However, we will not know until the CO cores obtained with better He-star mass loss rates are evolved further.

### 3. Evolution of binary compact objects

We summarize the Bethe and Brown [1] evolution of binary compact objects, paying special attention to their common envelope evolution. In particular, we shall show that their schematic evolution should be applicable to donors with deeply convective envelopes, whereas for non-convective or semiconvective envelopes, such as encountered in the evolution of low-mass X-ray binaries, their common envelope evolution would not be expected to apply.

We call the star that is initially heavier star A, the other star B. We denote initial masses by subscript i, so we have masses  $M_{A,i}$ ,  $M_{B,i}$ . We denote their ratio by  $q$ ; thus

$$q = M_{B,i}/M_{A,i} \leq 1 . \quad (8)$$

Following Portegies Zwart and Yungelson [20], we assume that  $q$  is distributed uniformly between 0 and 1. Likewise, we also assume that  $\ln a$  is uniformly distributed, where  $a$  is the semi-major axis of their orbit.

However, we assume different limits for  $a$  than Portegies Zwart and Yungelson [20]. Initially, both stars are massive main sequence stars, with radius at least  $3R_\odot$ , so  $a > 6R_\odot = 4 \times 10^6$  km. At the other end of the scale, we require  $a < 4 \times 10^9$  km. We assume that 50% of all stars are binaries with separations in this range (stars in wider binaries would evolve as if they were single). Then the fraction of binaries in a given interval of  $\ln a$  is

$$d\phi = d(\ln a)/7 . \quad (9)$$

We assume that a star needs an initial mass of

$$M > M_s = 10M_\odot \quad (10)$$

in order to go supernova. Therefore, if  $\alpha$  is the total rate of SNe, the rate of SNe in mass interval  $dM$  is given by

$$d\alpha = \alpha n \left( \frac{M}{10M_\odot} \right)^{-n} \frac{dM}{M} , \quad (11)$$

where we have used a power-law initial mass function with  $n = 1.5$  (close to the Salpeter value  $n = 1.35$ ). The birth rate of supernova systems was taken to be

$$\alpha = 0.02 \text{ yr}^{-1} \quad (12)$$

in the Galaxy. By a supernova system we mean a single star that goes supernova (i.e., has  $M_{\text{ZAMS}} > 10M_\odot$ ) or a close binary containing at least one such star (close here means within the separation range mentioned above). Bethe and Brown [1] find that if the primary is massive enough to go supernova, then there is an  $\sim 50\%$  chance for the secondary to also go supernova. This was calculated for a distribution flat in  $q = M_{B,i}/M_{A,i}$ . Therefore, the supernova rate in our notation would be  $1.25\alpha = 0.025 \text{ yr}^{-1}$ .

Using the Cordes and Chernoff [21] distribution of kick velocities, 43% of the binaries were found to survive the first explosion. Thus, at this stage, we are left with a birth rate of

$$R = 0.02 \times \frac{1}{2} \times \frac{1}{2} \times 0.43 \simeq 2 \times 10^{-3} \text{ per yr} \quad (13)$$

for the formation of binaries consisting of a neutron star with a companion massive enough to go supernova ( $M > 10M_{\odot}$ ). The lifetime of such systems is the companion lifetime of  $\sim 10^7$  yr, but star A will be a pulsar for only  $\sim 5 \times 10^6$  yr because it will spin down electromagnetically until it is no longer observable. From these numbers we estimate the number of such systems to be  $\sim 10^4$  in the Galaxy.

Since the pulsar is unrecycled, the expected number should be compared with the detected population of active radio pulsars in the galaxy, about  $10^3$ . This number should be multiplied by a factor of 1/2 for binarity, a further factor of 1/2 for a binary in which both stars can go supernova and the 0.43 for survival of the first explosion. This would leave the large number  $\sim 10^2$  if pulsars with massive companions were as easily detected as single pulsars. In fact, only 2 are observed; PSR 1259-63 with a Be-star companion and PSR 0045-73 with a B-star companion. Stellar winds interfere with the radio pulses from these binaries, obscuring the narrower ones. Doppler shifts also make these difficult to observe. Nevertheless, the factor necessary to reduce their observability is large. We return to the subject later.

At this stage we have an  $\sim 1.4M_{\odot}$  neutron star with O or B-star companion. We take the latter to have mass  $\sim 15M_{\odot}$ . The giant has a He core containing some 30% of its mass, surrounded by an envelope consisting mainly of H. We take the envelope to be deeply convective,<sup>1</sup> so the entropy is constant. The particles, nuclei and electrons, are nonrelativistic and thus have  $\gamma = 5/3$ . Therefore, the envelope forms a polytrope of index  $n = 3/2$ . Applegate [26] shows that the binding energy of the envelope is

$$E \simeq 0.6GM_B^2R^{-1}, \quad (14)$$

where  $R$  is the outer radius. In this formula the binding energy is decreased 50% by the kinetic energy,  $E$ , containing both effects.

The major difference of the Bethe and Brown calculations and of case H of Portegies Zwart and Yungelson [20] compared with other work is the use of hypercritical accretion. In a series of papers, Chevalier [27,28] showed that once  $\dot{M}$  exceeded  $\sim 10^4M_{\text{Edd}}$ , the photons were carried inwards in the adiabatic inflow, onto the neutron star. The surface of the latter was heated sufficiently that energy could be carried off by neutrino pairs. Brown [3] reproduced Chevalier's results in analytical form. The idea has a much longer history: Colgate [73] showed already in 1971 that if neutrinos carry off the bulk of the energy, accretion can proceed at a much greater rate than Eddington. In 1972 Zeldovich et al. [74], before the introduction of common envelope evolution, used hypercritical accretion of a cloud onto a neutron star. Bisnovatyi-Kogan and Lamzin [75] and Chevalier [27] pointed out that during the common envelope phase of binary evolution, photons would be trapped and accretion could occur at much higher rates, and that neutron stars that go through this phase generally will go into black holes.

We begin by considering the work done by the neutron star on the envelope matter that it accretes. This will turn out to be only a fraction of the total work, the rest coming from the

<sup>1</sup> The assumption that the envelope is deeply convective is essential for our later treatment of common envelope evolution with hypercritical accretion. Recent developments with nonconvective or semiconvective donors show that the accretion rate is also highly super-Eddington, but still significantly less [22–24]. For very massive donors the rate is always highly super-Eddington.

production of the wake, but it illustrates simply our procedure. Taking the neutron star to beat rest, the envelope matter is incident on it with the Keplerian velocity  $v$ . The rate of accretion is given by the Bondi–Hoyle–Lyttleton theory

$$\frac{dM_A}{dt} = \pi \rho v R_{ac}^2 , \quad (15)$$

where  $\rho$  is the density of the B material,  $v$  is its velocity relative to the neutron star A, and  $R_{ac}$  is the accretion radius

$$R_{ac} = 2GM_A v^{-2} . \quad (16)$$

The rate of change of momentum  $P$  is

$$\frac{dP}{dt} = v \frac{dM_A}{dt} , \quad (17)$$

the matter being brought to rest on the neutron star, and this is equal to the force  $F$ . Consequently, the rate at which the neutron star does work in the material is

$$\dot{E} = Fv = v^2 dM_A/dt . \quad (18)$$

Inclusion of the work done in creating the wake involves numerical calculations [29–31] with the result that the coefficient of the right-hand side of Eq. (18) is changed, i.e.,

$$\dot{E} = \left(\frac{c_d}{2}\right)v^2 \frac{dM_A}{dt} , \quad (19)$$

with  $c_d \sim 6\text{--}8$  for our supersonic flow. It is, in fact, very important that the wake plays such a large role, in that its the fact that  $c_d/2 > 1$  (we consider  $c_d/2$  to be  $\gg 1$ ) that makes our later common envelope evolution strongly nonconservative, the proportion of the total H-envelope mass accreted onto the neutron star being relatively small.

In Eq. (19)  $v^2$  is the velocity of the B (giant) material relative to A, the neutron star. This is given by

$$v^2 = G(M_A + M_B)a^{-1} . \quad (20)$$

The interaction energy of A and B is

$$E = \frac{1}{2}GM_A M_B a^{-1} . \quad (21)$$

Since we know  $M_{B,i}$  and  $M_{B,f}$ , the initial mass of B and the mass of its He core, our unknown is  $a_f$ . We can obtain it by considering

$$Y = M_B a^{-1} \quad (22)$$

as one variable,  $M_A$  as the other. Differentiating Eq. (21) we have

$$\dot{E} = \frac{1}{2}G(\dot{M}_A Y + M_A \dot{Y}) , \quad (23)$$

whereas combining Eqs. (19) and (20) and neglecting  $M_A$  with respect to  $M_B$ , we have

$$\dot{E} = G \left( \frac{c_d}{2} \right) Y M_A . \quad (24)$$

Thus, Eqs. (23) and (24) are equal, so we have

$$\frac{\dot{M}_A}{M_A} = \frac{1}{(c_d - 1)} \frac{\dot{Y}}{Y} , \quad (25)$$

which can be integrated to give

$$M_A \propto Y^{1/(c_d - 1)} = Y^{1/5} \quad (26)$$

where we have chosen  $c_d = 6$  [29]. The final energy is then

$$E_f = \frac{1}{2} G M_{A,i} Y_i (Y_f/Y_i)^{6/5} . \quad (27)$$

The binding energy  $E_f$  of star A to star B serves to expel the envelope of star B, whose initial binding energy is given by Eq. (14). Mass transfer begins at the Roche Lobe which lies at  $\sim 0.6a_i$  for the masses involved. However, star B expands rapidly in red giant stage before the mass transfer can be completed. To keep the numbers easy to compare with Bethe and Brown [1], we use their approximation of starting spiral-in when the giant's radius equals the orbital separation rather than the Roche-lobe radius. Since for the large mass ratios considered here,  $R_L/a \sim 0.5$  for the giant, this implies we require  $E_f$  of Eq. (27) to be about twice the binding energy (Eq. (14)), i.e.

$$E_f = \frac{0.6}{\alpha} G \frac{M_{B,i}^2}{a_i} = 1.2 G \frac{M_{B,i}^2}{a_i} . \quad (28)$$

(We set the common-envelope efficiency,  $\alpha$ , to 0.5.) The ejected material of B is, therefore, released with roughly the thermal energy it had in the envelope; in other words, the thermal energy content of the star is not used to help expel it. Inserting Eq. (28) into Eq. (27) yields

$$(Y_f/Y_i)^{1/2} = 2.4 M_{B,i}/M_{A,i} . \quad (29)$$

Star A is initially a neutron star,  $M_{A,i} = 1.4M_\odot$ . For star B we assume  $M_{B,i} = 15M_\odot$ . Then Eq. (29) yields

$$Y_f/Y_i = 15 . \quad (30)$$

We use this to find the result of accretion, with the help of Eq. (26),

$$M_{A,f}/M_{A,i} = 1.73 \quad (31)$$

or

$$M_{A,f} = 2.4M_\odot . \quad (32)$$

This is well above any of the modern limits for neutron star masses, so we find that the neutron star has gone into a black hole.

Our conclusion is, then, that in the standard scenario for evolving binary neutron stars, if the giant is deeply convective, accretion in the common envelope phase will convert the neutron star into a black hole.

Star B, by losing its envelope, becomes a He star. We estimate that

$$M_{B,f}/M_{B,i} \simeq 0.3 . \quad (33)$$

The size of the orbit is determined by Eq. (22),

$$a_i/a_f = M_{B,i}/M_{B,f} \quad Y_f/Y_i = 50 . \quad (34)$$

The final distance between the stars  $a_f$  should not be less than about  $10^{11}$  cm, so that the He star (mass  $M_{B,f}$ ) fits within its Roche lobe next to the black hole of mass  $M_{A,f}$ . Bethe and Brown [1] showed that if the black hole and the neutron star resulting from the explosion of star B are to merge in a Hubble time, then  $a_f < 3.8 \times 10^{11}$  (for circular orbits; correction for eccentricity will be given later). Therefore, the initial distance of the two stars, after the first mass exchange and the first supernova should be

$$0.5 \times 10^{13} \text{ cm} < a_i < 1.9 \times 10^{13} \text{ cm} . \quad (35)$$

If the initial distribution of distances is  $da/7a$ , the probability of finding  $a$  between the limits of Eq. (35) is

$$P = 18\% . \quad (36)$$

As noted earlier, 43% of the binaries survive the first explosion, so the combined probability is now

$$P = 8\% \quad (37)$$

for the survivors falling in the logarithmic interval in which they survive coalescence, but are narrow enough to merge in a Hubble time. Our final result, following from a birth rate of  $10^{-2}$  binaries per year in which one star goes supernova, half of which have both stars going supernova, is

$$R = 10^{-2} \times 0.5 \times 0.08 \times 0.5 = 2 \times 10^{-4} \text{ yr}^{-1} \quad (38)$$

in the Galaxy. The final factor of 0.5 is the survival rate of the He-star, neutron star binary, calculated by Monte Carlo methods. Bethe and Brown [1] quoted  $10^{-4} \text{ yr}^{-1}$ , or half of this rate, in order to take into account some effects not considered by them in which the binary disappeared (e.g., Portegies Zwart and Verbunt [57]).

Our final rate is, then,

$$R = 10^{-4} \text{ yr}^{-1} \text{ galaxy}^{-1} . \quad (39)$$

Using our supernova rate of 0.025 per year, which includes the case where both stars in the binary go supernova, we can convert this birth rate to 0.004 per supernova for comparison with other work. Portegies Zwart and Yungelson [20] in their case H, which included hypercritical accretion, got 0.0036 per supernova, within 10% of our value. Thus, the chief difference between our result in Eq. (39) and the  $R = 5.3 \times 10^{-5}$  of these authors is due to the different assumed SN rate.

In our above estimates we have assumed the second neutron star to be formed to have a circular orbit of the same  $a$  as its He-star progenitor. However, eccentricity in its orbit leads to a value of  $a_f$  substantially larger than the  $3.8 \times 10^{11}$  cm used above as the maximum separation for merger. In general, most of the final binaries will have  $e > 0.5$ , with a heavy peak in the distribution close to  $e = 1$ . The rise occurs because preservation of the binary in the explosion is substantially greater if

the kick velocity is opposite to the orbital velocity before explosion. In this case the eccentricity  $e$  is large. The most favorable situation is when the orbital and kick velocities are equal in magnitude. (See the figures in Wettig and Brown [32].) Eggleton [33] has kindly furnished us with a useful interpolation formula for the increase. The factor by which to multiply the time for merger in circular orbits, is

$$Z(e) \approx (1 - e^2)^{3.689 - 0.243e - 0.058e^2}. \quad (40)$$

This formula is accurate to about 1% for  $e \leq 0.99$ . Thus, if the initial eccentricity is 0.7, the time to shrink the orbit to zero is about 10% of the time required if the initial eccentricity were zero for the same initial period. The maximum  $a_t = 3.8 \times 10^{11}$  cm for circular orbits would be increased by the fourth root of the decrease in time, i.e., up to  $6.8 \times 10^{11}$  cm for this eccentricity. The maximum  $a_i$  in Eq. (35) would go up to  $3.4 \times 10^{13}$  cm, increasing the favorable logarithmic interval by  $\sim 40\%$ . We have not introduced this correction because it is of the same general size as the uncertainty in the supernova rate. However, this correction gives us some comfort that our final numbers are not unreasonably large.

If we produce an order of magnitude more low-mass black-hole, neutron-star binaries than binary neutron stars, the obvious question is why we have not seen any. The neutron star in this object is “fresh” (unrecycled) so it would spin down into the graveyard of neutron stars in  $\sim 5 \times 10^6$  yr. The two relativistic binary pulsars we do see 1913 + 16 and 1534 + 12 have been recycled, have magnetic fields  $B \sim 10^{10}$  G, two orders of magnitude less than a fresh pulsar, and will therefore be seen for about 100 times longer than an unrecycled neutron star. So even with a 10 times higher birth rate, we should see 10 times fewer LBH-NS binaries than NS-NS binaries. Furthermore, the binary with black hole will have a somewhat higher mass, therefore greater Doppler shift, and therefore be harder to detect. In view of the above, it is reasonable that our low-mass black-hole, neutron-star binaries have not been observed, but they should be actively looked for.

We should also calculate the rate of coalescences of the black hole with the He star. These have been suggested by Fryer and Woosley [34] as candidate progenitors for the long time gamma-ray bursters. Note that they will occur for a range of  $0.04 \times 10^{13}$  cm  $< a_i < 0.5 \times 10^{13}$  cm, a logarithmic interval double that of Eq. (35). Thus, the black-hole, He-star coalescence has a probability

$$P = 36\%. \quad . \quad (41)$$

Furthermore, this situation does not have the 50% disruption in the final explosion, so the black-hole, He-star coalescences occur with a total rate of 4 times that of the black-hole, neutron-star mergers.

There has been much discussion in the literature of the difficulties in common envelope evolution. We believe our model of deeply convective giants and hypercritical accretion offers an ideal case. Of course, the initiation of the common envelope evolution requires some attention, but it can be modeled in a realistic way [35]. As the giant evolves across its Roche lobe, the compact object creates a tidal bulge in the giant envelope, which follows the compact object, torquing it in. As the convective giant loses mass, the envelope expands in order to keep entropy constant. In Bondi-Hoyle-Lyttleton accretion, a density  $\rho_\infty \sim 10^{-13}$  g cm $^{-3}$  is sufficient with wind velocities  $\sim 1000$  km s $^{-1}$  in order to give accretion at the Eddington rate. Thus to achieve  $\dot{M} \sim 10^8 M_{\text{Edd}} \sim 1 M_\odot \text{ yr}^{-1}$  we need  $\rho \sim 10^{-5}$  g cm $^{-3}$  which is found at  $0.9 R$ , where  $R$  is the

radius of the giant. At this rate of accretion, angular momentum, etc., are hardly able to impede it appreciably. The total mass accreted onto the compact object is  $\sim 1M_{\odot}$ , so the common envelope evolution has dynamical time of years. As noted earlier, it is nonconservative.

#### 4. Evolution of binary neutron stars

Since the standard scenario of evolution of binary compact objects ends up with low-mass black-hole, neutron-star binaries, another way must be found to evolve neutron star binaries. In the double He-star scenario was suggested by Brown [3] and developed further by Wettig and Brown [32] the neutron star avoids going through common envelope with a companion star. In this way the neutron star can avoid being converted into a black hole by accretion. For two giants to burn He at the same time, they must be within  $\sim 5\%$  of each other in mass, the helium burning time being  $\sim 10\%$  of the main sequence lifetime, and stellar evolution time going roughly with the inverse square of the mass. With a flat mass ratio distribution, this happens in 5% of all cases, making the ratio of NS-NS to NS-LBH binaries 1:20. However, when the primary becomes an LBH, only half the secondaries will be massive enough to form a NS, whereas for the very close mass values of the double-He scenario this factor 2 loss does not occur. Thus, binary neutron stars should be formed 10% as often as low-mass black-hole, neutron-star binaries. This 10% is nearly model independent because everything else roughly scales.

The scenario goes as in Wettig and Brown [32]. The primary O-star evolves transferring its H-envelope to the companion. Often, this would lead to ‘rejuvenation’ of the secondary, i.e. its evolution would restart also from the ZAMS with the now higher total mass, and it would make a much heavier core. However, here the core of the secondary has evolved almost as far as the primary’s core, so the core molecular weight is much higher than that of the envelope. This prevents convection in the core from extending into the new envelope to make the bigger core, so no rejuvenation takes place [36]. Since  $q \sim 1$ , the first mass transfer is nearly conservative. The second is not, so the two He-cores then share a common H envelope, which they expel, while dropping to a lower final separation  $a_f$ .

Following the explosion of the first He star, the companion He-star pours wind matter onto the pulsar, bringing the magnetic field down and spinning it up [3,32]. The end result is two neutron stars of very nearly equal mass, although wind accretion can change the mass 2–3%.

The above scenario ends for He-star masses greater than  $4$  or  $5M_{\odot}$ , corresponding to ZAMS masses greater than  $\sim 16$  or  $18M_{\odot}$ . However, less massive He stars evolve in the He shell-burning stage, and a further mass transfer (Case C) can take place. The transfer of He to the pulsar can again bring about a black hole, which Brown [37] very roughly estimates to occur in  $\sim 50\%$  of the double-neutron star binaries. This is roughly consistent with results of Fryer and Kalogera [38]. Taking a rate of  $R = 10^{-4}$  per year per galaxy for the low-mass black-hole, neutron-star binaries, we thus arrive at a birth rate of

$$R \simeq 5 \times 10^{-6} \text{ per year per galaxy} \quad (42)$$

for binary neutron-star formation. However, the black holes formed in the He shell burning evolution will not have accreted much mass and will have about the same chirp mass as binary neutron stars (see below) for gravitational merging.

Our best guess values, Eqs. (39) and (42), thus give an  $\sim 20$  to 1 ratio for formation of low-mass black-hole, neutron-star binaries to binary neutron stars. The former are better progenitors for gravitational waves from mergers because of their higher masses and they have many advantages as progenitors of gamma-ray bursters [39]. Note that our estimated rate of  $R = 5 \times 10^{-6}$  per galaxy per year for binary neutron star formation is consistent with the empirical rates discussed in our introduction.

## 5. High-mass black-hole O/B-star binaries

We will be brief in our review of these, because we believe the evolution of these objects such as Cyg X-1, LMC X-1 and LMC X-3 to be less well understood than the low-mass black-hole, neutron-star binaries. Evolutionary calculations now proceeding by Alexander Heger, using the CO cores evolved by Wellstein and Langer [10] should clarify this situation substantially.

Bethe and Brown [8] arrived at a limit of ZAMS mass  $80M_{\odot}$  for stars in binaries to go into high-mass black holes (unless Case C mass transfer takes place as we discuss in our next section). This limiting mass is much higher than other workers have used. It was based on calculations of Woosley et al. [17] and was so high because of very high mass loss rates used by these authors. With more correct lower rates the limiting mass may come down, so the Bethe and Brown evolution should be viewed as giving a lower limits to the number of high-mass black-hole, O/B-star binaries. Their estimated birth rate of about  $3 \times 10^{-5}$  per galaxy per year does agree reasonably well with the fact that only one such system is known in the Galaxy. However, since even with a twice larger separation the accretion rate of the black hole from the fast wind of the O star becomes small, it is possible that substantially more systems with somewhat wider orbits exist undetected, and that Cyg X-1 is the only one presently in the (very short) phase of incipient Roche lobe overflow when it is bright. Bethe and Brown [8] found this narrowness of the Cyg X-1 orbit ( $40R_{\odot}$  according to Herrero et al. [4]) to be puzzling: the massive stars in the progenitor binary initially had to fit within their Roche Lobes, therefore a separation of at least double the current  $17R_{\odot}$  was needed. And most evolutionary effects from then on, such as wind mass loss or supernova-like mass loss, would tend to widen the orbit. Of course, the orbit could be narrowed in Case A mass transfer (i.e. during the main sequence) since the progenitor of the black hole was more massive than the present donor, but it could not become so narrow that the present donor filled its Roche lobe, and would widen again once the mass ratio became reversed and widen further due to wind loss after the whole primary envelope was lost.

In any case, a binary as narrow as Cyg X-1 would coalesce in the common envelope evolution once the O-star companion of the massive black-hole goes into red giant phase, according to the Bethe and Brown [8] estimates. Since the black hole in Cyg X-1 has mass  $\gtrsim 10M_{\odot}$  and is probably the most massive black hole in a binary observed in the Galaxy, in the Fryer and Woosley [34] model where the black hole “eats” the W.-R. companion, such a coalescence should produce the most energetic long-lasting gamma-ray burster. We are unable to evaluate the probability of Cyg X-1 like objects merging following common envelope evolution because we have been unable to understand why Cyg X-1, before common envelope evolution, is so narrow. The LBV, RSG, and WNL stages of W.-R. development are not quantitatively understood.

After the main sequence star in a Cyg X-1-like object explodes and becomes a neutron star, according to Bethe and Brown [8] the binary will eventually merge. They estimated the contribution to the merger rate of these systems to be  $(4\text{--}6) \times 10^{-6} \text{ yr}^{-1} \text{ galaxy}^{-1}$ , however with considerable uncertainty due to the fact that the evolution of Cyg X-1 itself is uncertain. Lowering the mass limit for black-hole formation by having lower mass loss rates would increase this number (e.g. a limit of  $40M_{\odot}$  would increase the merger rate by a factor 5).

## 6. The formation of high-mass black holes in low-mass X-ray binaries

### 6.1. General

Crucial to our discussion here is the fact that single stars evolve very differently from stars in binaries that lose their H-envelope [6,10,13,19] either on the main sequence (Case A) or in the giant phase (Case B). However, stars that transfer mass or lose mass after core He burning (Case C) evolve, for our purposes, as single stars, because the He core is then exposed too close to its death for wind mass loss to significantly alter its fate. Single stars above a ZAMS mass of about  $20M_{\odot}$  skip convective carbon burning following core He burning, with the result, as we shall explain, that their Fe cores are substantially more massive than stars in binaries, in which H-envelope has been transferred or lifted off before He core burning. These latter “naked” He stars burn  $^{12}\text{C}$  convectively, and end up with relatively small Fe cores. The reason that they do this has to do chiefly with the large mass loss rates of the “naked” He cores, which behave like W.-R.’s. Unfortunately, in calculation until recently, substantially too large mass loss rates were used, so we cannot pin limits down quantitatively. In this section we will deal with the ZAMS mass range  $\sim 20\text{--}35M_{\odot}$ , in which it is clear that many, if not most, of the single stars go into high-mass black holes, whereas stars in binaries which burn “naked” He cores go into low-mass compact objects. In this region of ZAMS masses the use of too-high He-star mass loss rates does not cause large effects [6].

The convective carbon burning phase (when it occurs) is extremely important in pre-supernova evolution, because this is the first phase in which a large amount of entropy can be carried off in  $v\bar{v}$ -pair emission, especially if this phase is of long duration. The reaction in which carbon burns is  $^{12}\text{C}(\alpha, \gamma)^{16}\text{O}$  (other reactions like  $\text{C} + \text{C}$  would require excessive temperatures). The cross section of  $^{12}\text{C}(\alpha, \gamma)^{16}\text{O}$  is still not accurately determined; the lower this cross section the higher the temperature of the  $^{12}\text{C}$  burning, and therefore the more intense the  $v\bar{v}$  emission. With the relatively low  $^{12}\text{C}(\alpha, \gamma)^{16}\text{O}$  rates determined both directly from nuclear reactions and from nucleosynthesis by Weaver and Woosley [42], the entropy carried off during  $^{12}\text{C}$  burning in the stars of ZAMS mass  $\sim 10\text{--}20M_{\odot}$  is substantial. The result is rather low-mass Fe cores for these stars, which can evolve into neutron stars. Note that in the literature earlier than Weaver and Woosley [42] often large  $^{12}\text{C}(\alpha, \gamma)^{16}\text{O}$  rates were used, so that the  $^{12}\text{C}$  was converted into oxygen and the convective burning did not have time to be effective. Thus its role was not widely appreciated.

Of particular importance is the ZAMS mass at which the convective carbon burning is skipped. In the Woosley and Weaver [18] calculations this occurs at ZAMS mass  $19M_{\odot}$  but with a slightly lower  $^{12}\text{C}(\alpha, \gamma)^{16}\text{O}$  rate it might come at  $20M_{\odot}$  or higher [37]. As the progenitor mass increases, it follows from general polytropic arguments that the entropy at a given burning stage increases. At the higher entropies of the more massive stars the density at which burning occurs is lower, because

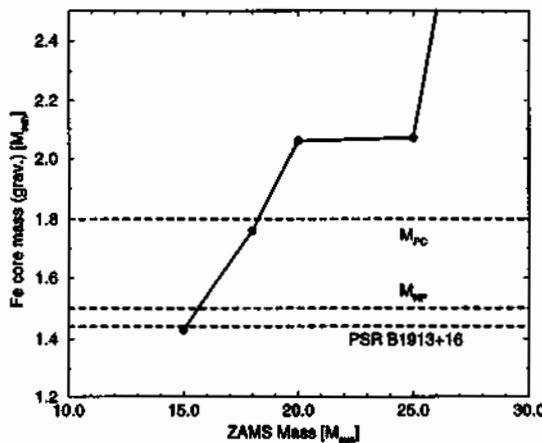


Fig. 1. Compact core masses resulting from the evolution of single stars, Case B of solar metallicity of Woosley and Weaver [18]. The horizontal dashed lines indicate the mass of the heaviest known well-measured pulsar [16], the maximum mass of a neutron star, and our estimate of  $M_{pc}$  (proto-compact), the maximum compact core mass for which matter can be returned to the galaxy.

the temperature is almost fixed for a given fuel. Lower densities decrease the rate of the triple- $\alpha$  process which produces  $^{12}\text{C}$  relative to the two-body  $^{12}\text{C}(\alpha, \gamma)^{16}\text{O}$  which produces oxygen. Therefore, at the higher entropies in the more massive stars the ratio of  $^{12}\text{C}$  to  $^{16}\text{O}$  at the end of He burning is lower. The star skips the long convective carbon burning and goes on to the much shorter oxygen burning. Oxygen burning goes via  $^{16}\text{O} + ^{16}\text{O}$  giving various products, at very much higher temperature than  $\text{C}(\alpha, \gamma)$  and much faster. Since neutrino cooling during the long carbon-burning phase gets rid of a lot of entropy of the core, skipping this phase leaves the core entropy higher and the final Chandrasekhar core fatter.

In Fig. 1 the large jump in compact object mass in single stars at ZAMS mass  $\sim 19M_{\odot}$  is clearly seen. From our discussion in Section 2 we see that this is just at the point where the Fe core mass goes above the proto-compact mass of  $\sim 1.8M_{\odot}$  and, therefore, above this mass one would expect single stars to go into high-mass black holes. Arguments have been given that SN 1987A with progenitor ZAMS mass of  $\sim 18M_{\odot}$  evolved into a low-mass black hole [12]. We believe from our above arguments and Fig. 1 that just above the ZAMS mass of  $\sim 20M_{\odot}$ , single stars go into high-mass black holes without return of matter to the Galaxy. Thus, the region of masses for low-mass black hole formation in single stars is narrow, say  $\sim 18\text{--}20M_{\odot}$  (although we believe it to be much larger in binaries).

Thus far our discussion has been chiefly about single stars, in which the He burns “clothed” by a hydrogen envelope. In this case the convective helium core grows in stars as time passes. In the “naked” He cores, in which the H envelope has been lifted off in RLOF or driven off by wind either before or early in the He burning the temperature and the entropy will be slightly lower, because the insulating layer is gone, so it is not surprising that their carbon abundance is large. Furthermore, the core mass continually decreases because of mass loss by wind. In fact, even for the naked  $20M_{\odot}$  He core, corresponding to ZAMS mass  $45M_{\odot}$ , the central carbon abundance was  $\sim 33\%$  at the end of He core burning [18] whereas only  $\sim 15\%$  is necessary for convective carbon burning

[42]. For lower mass He stars the  $^{12}\text{C}$  abundance was, of course, larger. Even with He-star wind mass loss rates reduced by half, Wellstein and Langer [10] find a central carbon abundance of  $\gtrsim 1/3$  at the end of He core burning all the way up through  $60M_{\odot}$  stars, so it is clear that convective carbon burning will take place. Unfortunately, the cores have not yet been evolved past the CO stage. Thus, in the range of ZAMS masses up to  $\geq 60M_{\odot}$ , if the H envelope is lifted off early in the core He burning phase, the convective carbon burning will take place after the He burning.

By ZAMS mass  $\sim 40M_{\odot}$ , where stars evolve into WR stars almost independent of whether they have a companion, the ultimate fate of the compact core is uncertain: Brown et al. [13] suggest that 1700-37, with a progenitor of about  $40M_{\odot}$  went into a low-mass black hole. This would seem to indicate that the H-envelope of such massive stars is blown off in an LBV phase rapidly enough that the He core again burns as “naked”. In any case,  $^{12}\text{C}$  is burned convectively following He core burning, so the resulting Fe core should be small.

We believe that our discussion earlier in this section indicates that single stars in the region of ZAMS masses  $\sim 20\text{--}35M_{\odot}$  end up as high-mass black holes. We can obtain the high-mass black holes, according to our above discussion, if we make the He-stars burn with “clothing”, i.e., lift their H-envelope off only following He core burning. Thus, the evolving massive star should meet the companion main sequence star only following He core burning (in the supergiant stage). By then its radius  $R$  is several hundred  $R_{\odot}$ , and its binding energy  $0.6GM^2/R$ , very small because of the large  $R$ . In order to see effects of matter stripped off from the main sequence companion in the transient sources, we want it to end up close to the black hole. Because of its low binding energy the supergiant envelope will be expelled by a relatively small binding energy of the companion,  $\frac{1}{2}M_{\text{A}}M_{\text{B},i}/a_{\text{f}}$  where  $a_{\text{f}}$  is the distance between black hole and companion. In order to make  $a_{\text{f}}$  small the mass  $M_{\text{A}}$  of the companion must be small. (More massive main sequence stars will spiral in less far, hence end up further from the black hole, and not fill their Roche Lobes. However, when they evolve in subgiant or giant phase they will fill it.) Both Portegies Zwart et al. [43] and Ergma and Van den Heuvel [44] have suggested that roughly the above region of ZAMS masses must be responsible for the  $\sim 7M_{\odot}$  black holes in the transient X-ray sources in order to form enough such sources. Our scenario is essentially the same as that of de Kool et al. [45] for the black hole binary A0620-00. We refer to this work for the properties of the K-star companion, stressing here the evolutionary aspects of the massive black hole progenitor.

## 6.2. Calculation

We now calculate the common envelope evolution following the formalism of Section 3. Here  $M_{\text{A}}$  is the mass of the main sequence companion,  $M_{\text{B}}$  that of the massive black hole progenitor. The ratio

$$q = M_{\text{A},i}/M_{\text{B},i} \quad (43)$$

is very small and there is great uncertainty in the initial number of binaries for such a small  $q \sim 1/25$ . We again take the distribution as  $dq$ , and again assume  $\ln a$  to be uniformly distributed over a logarithmic interval of 7. Again, the fraction of binaries in a given interval is

$$d\phi = d(\ln a)/7. \quad (44)$$

We evolve a typical  $25M_{\odot}$  star (B) with a companion  $\sim 1M_{\odot}$  main sequence star (star A) as the progenitor of the transient X-ray sources. The common envelope evolution can be done as in Section 3. With  $M_{B,i} = 25M_{\odot}$  and neglect of the accretion onto the main sequence mass  $M_A$ , we find from Bethe and Brown [1]

$$\left(\frac{Y_f}{Y_i}\right)^{1.2} = \frac{1.2}{\alpha_{ce}} \frac{M_{B,i}}{M_A}, \quad (45)$$

where  $Y = M_B/a$ . Here the coefficient of dynamical friction  $c_d$  was taken to be 6. The result is relatively insensitive to  $c_d$ , the exponent 1.2 resulting from  $1 + 1/(c_d - 1)$ .

Thus, in our case

$$\frac{Y_f}{Y_i} = 17 \left( \frac{\alpha_{ce} M_A}{M_{\odot}} \right)^{-0.83} = 30 \left( \frac{0.5 M_{\odot}}{\alpha_{ce} M_A} \right)^{0.83}. \quad (46)$$

We expect  $\alpha_{ce} \simeq 0.5$ , under the assumption that the thermal energy of the expelled envelope is equal to that it originally possessed in the massive star (i.e. that it is not used as extra energy to help remove the envelope), but it could be smaller. From this we obtain

$$\frac{a_i}{a_f} = \frac{M_{B,i} Y_f}{M_{B,f} Y_i} = 90 \left( \frac{0.5 M_{\odot}}{\alpha_{ce} M_A} \right)^{0.83}, \quad (47)$$

where we have taken the He star mass  $M_{B,f}$  to be 1/3 of  $M_{B,i}$ . In order to survive spiral-in, the final separation  $a_f$  must be sufficient so that the main sequence star lies at or inside its Roche Lobe, about  $0.2a_f$  if  $M_A = M_{\odot}$ . This sets  $a_f \sim 5R_{\odot} = 3.5 \times 10^{11}$  cm and

$$a_i = 3.15 \left( \frac{0.5}{\alpha_{ce}} \right)^{0.83} \times 10^{13} \text{ cm}, \quad (48)$$

which is about 2 AU. This exceeds the radius of the red giant tip in the more numerous lower mass stars in our interval, so the massive star must generally be in the *supergiant* phase when it meets the main sequence star, i.e., the massive star must be beyond He core burning. E.g., the red giant tip (before the He core burning) for a  $20M_{\odot}$  star is at  $0.96 \times 10^{13}$  cm, for a  $25M_{\odot}$  star,  $2.5 \times 10^{13}$  cm [46]. These numbers are, however, somewhat uncertain. Notice that decreasing  $\alpha_{ce}$  will increase  $a_i$ . Decreasing  $M_A$  has little influence, because with the smaller stellar radius the minimum  $a_f$  will decrease nearly proportionately. Note that neglect of accretion onto the main sequence star would change the exponent 0.83 to unity, so accretion is unimportant except in increasing the final mass.

Now a ZAMS  $25M_{\odot}$  star ends up at radius  $6.7 \times 10^{13}$  cm ( $\sim 2a_i$ ) following He shell burning [47]. Thus, the interval between  $a_i$  and  $6.7 \times 10^{13}$  cm is available for spiral-in without merger so that a fraction

$$\frac{1}{7} \ln \left( \frac{6.7}{3.15(0.5/\alpha_{ce})^{0.83}} \right) \simeq 0.11 \quad (49)$$

of the binaries survive spiral-in, but are close enough so that the main sequence star is encountered by the evolving H envelope of the massive star. The He core burning will be completed before the supergiant has moved out to  $\sim 2$  AU, so binaries which survive spiral-in will have He cores which burn as “clothed”, namely as in single stars.

Given our assumptions in Section 3, the fraction of supernovas which arise from ZAMS stars between  $20$  and  $35M_{\odot}$  is

$$1/2^{3/2} - 1/3.5^{3/2} = 0.20, \quad (50)$$

where we have assumed the mass  $10M_{\odot}$  is necessary for a star to go supernova. A Salpeter function with index  $n = 1.5$  is assumed here. Our assumption that the binary distribution is as  $dq$  is arbitrary, and gives us a factor  $1/25$  for a  $1M_{\odot}$  companion. Thus, for supernova rate 2 per century, our birth rate for transient sources in the Galaxy is

$$2 \times 10^{-2} \times 0.5 \times 0.11 \times 0.20 \times 0.04 \simeq 8.8 \times 10^{-6} \text{ yr}^{-1} \quad (51)$$

where  $0.5$  is the assumed binarity,  $0.11$  comes from Eq. (49), and the final (most uncertain) factor  $0.04$  results from a distribution flat in  $q$  and an assumed  $1M_{\odot}$  companion star.

In order to estimate the number of transient sources with black holes in the Galaxy, we should know the time that a main sequence star of mass  $\sim 1M_{\odot}$  transfers mass to a more massive companion. This depends on the angular-momentum loss rate that drives the mass transfer. A guaranteed loss mechanism for close binaries is gravitational radiation, which for a main-sequence donor gives a mass transfer rate of  $10^{-10}M_{\odot} \text{ yr}^{-1}$ , almost independent of donor mass [48]. As mass is transferred, the mass of the donor decreases and with it the radius of the donor. Quite a few low-mass X-ray binaries have X-ray luminosities that imply accretion rates in excess of  $10^{-10}M_{\odot} \text{ yr}^{-1}$ , leading to suggestions of additional mechanisms for loss of angular momentum from the binary, to increase mass transfer. Verbunt and Zwaan [49] estimate that magnetic braking can boost the transfer of mass in a low-mass binary. We somewhat arbitrarily adopt an effective mass transfer rate of  $10^{-9}M_{\odot} \text{ yr}^{-1}$  for main sequence stars. In order to estimate the number of high-mass black hole, main sequence star binaries in the Galaxy we should multiply the birth rate equation (51) by the  $10^9 \text{ yr}$  required, at the assumed mass loss rate, to strip the main sequence star, obtaining 8800 as our estimate. From the observed black-hole transient sources, Wijers, [5] arrives at 3000 low-mass black hole sources in the Galaxy, but regards this number as a lower limit. With the uncertainties in formation rate and lifetime, the agreement between the two numbers is as good as may be expected.

### 6.3. Observations

We believe that there are many main sequence stars more massive than the  $\lesssim 1M_{\odot}$  we used in our schematic evolution, which end up further away from the black hole and will fill their Roche Lobe only during subgiant or giant stage. From our earlier discussion, we see that a  $2M_{\odot}$  main sequence star will end up about twice as far from the black hole as the  $1M_{\odot}$ , a  $3M_{\odot}$  star, three times as far, etc. Two of the 9 systems in our Table 1 have subgiant donors (V404 Cyg and XN Sco). These have the longest periods, 6.5 and 2.6 days and XN Sco is suggested to have a relatively massive donor of  $\sim 2M_{\odot}$ . It seems clear that these donors sat inside their Roche Lobes until they evolved off the main sequence, and then poured matter onto the black hole once they expanded and filled their Roche Lobe. For a  $2M_{\odot}$  star, the evolutionary time is about a percent of the main-sequence time, so the fact that we see two subgiants out of nine transient sources means that

**Table 1**

Parameters of suspected black hole binaries with measured mass functions [5,50–54]. N means nova, XN means X-ray nova. Numbers in parentheses indicate errors in the last digits

X-ray names	Other name(s)	Compan. type	$P_{\text{orb}}$ (d)	$f(M_X)$ ( $M_\odot$ )	$M_{\text{opt}}$ ( $M_\odot$ )	$(l, b)$
		q ( $M_{\text{opt}}/M_X$ )	$K_{\text{opt}}$ (km s $^{-1}$ )	i (deg)	$M_X$ ( $M_\odot$ )	d (kpc)
Cyg X-1 1936 + 350	V1357 Cyg HDE 226868	O9.7Iab	5.5996 74.7(10)	0.25(1)	33(9) 16(5)	(73.1, + 3.1) 2.5
LMC X-3 0538-641		B3Ve	1.70 235(11)	2.3(3)	5.6–7.8	(273.6, – 32.1) 55
LMC X-1 0540-697		O7–9III	4.22 68(8)	0.14(5)		(280.2, – 31.5) 55
XN Mon 75 A0620-003	V616 Mon N Mon 1917	K4 V 0.057–0.077	0.3230 443(4)	2.83–2.99 37–44°	0.53–1.22 9.4–15.9	(210.0, – 6.5) 0.66–1.45
XN Oph 77 H1705-250	V2107 Oph	K3 V	0.5213 420(30)	4.44–4.86 60–80	0.3–0.6 5.2–8.6	(358.6, + 9.1) 5.5:
XN Vul 88 GS 2000 + 251	QZ Vul	K5 V 0.030–0.054	0.3441 520(16)	4.89–5.13 43–74	0.17–0.97 5.8–18.0	(63.4, – 3.1) 2
XN Cyg 89 GS 2023 + 338	V404 Cyg N Cyg 1938, 1959	K0 IV 0.055–0.065	6.4714 208.5(7)	6.02–6.12 52–60	0.57–0.92 10.3–14.2	(73.2, – 2.2) 2.2–3.7
XN Mus 91 GS 1124-683		K5 V 0.09–0.17	0.4326 406(7)	2.86–3.16 54–65	0.41–1.4 4.6–8.2	(295.0, – 6.1) 3.0
XN Per 92 GRO J0422 + 32		M0 V 0.029–0.069	0.2127(7) 380.6(65)	1.15–1.27 28–45	0.10–0.97 3.4–14.0	(197.3, – 11.9)
XN Sco 94 GRO J1655-40		F5-G2 0.33–0.37	2.6127(8) 227(2)	2.64–2.82 67–71	1.8–2.5 5.5–6.8	(345.0, + 2.2) 3.2
XN 4U1543-47	MX1543-475	A2 V	1.123(8) 124(4)	0.20–0.24 20–40	1.3–2.6 2.0–9.7	(330.9, + 5.4) 9.1(11)
XN Vel 93		K6-M0 0.137 ± 0.015	0.2852 475.4(59)	3.05–3.29 ~ 78	0.50–0.65 3.64–4.74	

\*A much higher inclination for A0620 has been claimed by Haswell et al. [55] of up to  $i = 70$ . In this case, the lower limits on the component masses would be  $M_X > 3.8$  and  $M_{\text{opt}} > 0.22$ .

many more of these massive donors are sitting quietly well within their Roche Lobes. Indeed, we could estimate from the relative time, that there are  $2/9 \times 100 = 22$  times more of these latter quiet main sequence stars in binaries.

Amazingly, this factor 22 almost cancels the 1/25 we had for the interval in  $q$  over which the donors contribute. This is not coincidental. Essentially any mass donor, at least almost up to the  $25M_{\odot}$  progenitor of the black hole, can give rise to a common envelope phase. The BH progenitor crosses the Herzsprung gap very quickly, in a time in which the companion can hardly accept its mass. (The ratio of  $q \lesssim 1/4$  for common envelope evolution was determined by Kippenhahn and Meyer-Hofmeister [56] for case A mass transfer.) Thus, one can expect essentially all companions, up to  $q \lesssim 1$ , to go into common envelope evolution and contribute. Beginning from Wijers' empirical estimate we would thus have  $(2/9) \times 100 \times 3000 = 6.7 \times 10^4$  binaries with high-mass black holes and main-sequence companions. This number is determined, as shown above, chiefly by the number of observed systems with subgiant donors.

If we assume that ZAMS masses  $\sim 10\text{--}18M_{\odot}$  evolve into a neutron star, we should have  $\sim 3$  times more neutron stars than high-mass black holes (see Eq. (50)). The range follows from our belief that SN 1987A with progenitor  $\sim 18M_{\odot}$  ZAMS went into a low-mass black hole, following the scenario of Brown and Bethe [12]. On the basis of a Monte Carlo calculation using the kick velocities of Cordes and Chernoff [21] we find that  $\sim 1/2$  of the binaries containing He-star, low-mass main sequence companion (with  $M \simeq 1M_{\odot}$ ) will be disrupted in the explosion. Thus, we find only a slightly higher birth rate for LMXBs (low mass X-ray binaries) with neutron stars than with black holes, although the numbers could be equal to within our accuracy. With comparable lifetimes (since the donor masses and mass transfer rates are comparable), this would give us one to a few thousand LMXBs with neutron stars, much above the total number of observed LMXBs ( $\sim 130$ ). Indeed, from Table 6 of Portegies Zwart and Verbunt [57] one sees that their estimated empirical birth rate for low-mass X-ray binaries is  $2 \times 10^{-7} \text{ yr}^{-1}$ , whereas in either theoretical evolution including kick velocities they obtain  $4 \times 10^{-6} \text{ yr}^{-1}$ . This factor of 20 discrepancy is by far the greatest between theoretical and empirical rates in their table, and supports our point that many of the neutron stars must have disappeared along the way. Alternatively, a large number of LMXBs with neutron stars could be transients as well (like e.g. Aql X-1). Just at the present there are new developments in the evolution of low-mass X-ray binaries, which we shall shortly summarize in Section 7.

As we showed below Eq. (48), the He core of the massive star will in general be uncovered only after He core burning is completed. The remaining time for He burning (in a shell) will be short, e.g., for a  $20M_{\odot}$  ZAMS star it is only  $1.4 \times 10^4 \text{ yr}$  [46]. Therefore the mass loss by wind after uncovering the He core will not be large, and when the star finally becomes a supernova, its mass will be almost equal to the He core of the original star. The latter can be calculated from

$$M_{\text{He}} \simeq 0.10 (M_{\text{ZAMS}})^{1.4} \quad (52)$$

so for ZAMS masses  $20\text{--}35M_{\odot}$   $M_{\text{He}}$  will lie in the interval  $\sim 7\text{--}14M_{\odot}$ .

Bailyn et al. [51] find the black-hole masses in transient sources to be clustered about  $\sim 7M_{\odot}$ , except for V404 Cyg which has a higher mass. This is in general agreement with our scenario, because most of the black holes will come from the more numerous stars of ZAMS mass not far

from our lower limit of  $\sim 20M_{\odot}$ . Two points are important to note:

- (1) Not much mass can have been lost by wind. Naked He stars have rapid wind loss. However, in our scenario the He star is made naked only during He shell burning and therefore does not have much time ( $\lesssim 10^4$  yr) to lose mass by wind.
- (2) There are good reasons to believe that the initial He core will be rotating [58]. The way in which the initial angular momentum affects the accretion process has been studied by Mineshige et al. [59] for black-hole accretion in supernovae. In general, accretion discs which are optically thick and advection dominated are formed. The disc is hot and the produced energy and photons are advected inward rather than being radiated away. The disc material accretes into the black hole at a rate of  $> 10^6 \dot{M}_{\text{Edd}}$  for the first several tens of days. Angular momentum is advected outwards. Our results show that little mass is lost, because the final  $\sim 7M_{\odot}$  black hole masses are not much less massive than the He core masses of the progenitors, and some mass is lost by wind before the core collapses. The latter loss will not, however, be great, because there is not much time from the removal of the He envelope until the collapse.

Accretion of the He into the black hole will differ quantitatively from the above, but we believe it will be qualitatively similar. The fact that the helium must be advected inwards and that little mass is lost as the angular momentum is advected outwards is extremely important to establish. This is because angular momentum, essentially centrifugal force, has been suggested by Chevalier [28] to hold up hypercritical accretion onto neutron stars in common envelope evolution. (Chevalier [27] had first proposed the hypercritical accretion during this evolutionary phase to turn the neutron stars into black holes, the work followed up by Brown [3] and Bethe and Brown [1].) However, once matter is advected onto a neutron star, temperatures  $\gtrsim 1$  MeV are reached so that neutrinos can carry off the energy. The accreted matter simply adds to the neutron star mass, evolving into an equilibrium configuration. Thus, this accretion does not differ essentially from that into a black hole. In either case of neutron star or black hole an accretion disc or accretion shock, depending on amount of angular momentum, but both of radius  $\sim 10^{11}$  cm, is first formed, giving essentially the same boundary condition for the hypercritical accretion for either black hole or neutron star. Thus, the masses of the black holes in transient sources argue against substantial inhibition of hypercritical accretion by jets, one of the Chevalier suggestions [28].

Measured mass functions, which give a lower limit on the black hole mass are given in Table 1. Only GRO J0422 + 32 and 4U 1543-47 have a measured mass function  $\lesssim 3M_{\odot}$ . Results of Callanan et al. [60] indicate that the angle  $i$  between the orbital plane and the plane of the sky for GRO J0422 + 32 is  $i < 45^\circ$ , and recent analysis [52] indicate that the angle  $i$  for 4U 1543-47 is  $20^\circ < i < 40^\circ$ . So both GRO J0422 + 32 and 4U 1543-47 also contain high-mass black holes.

Based on the observations of Kaper et al. [61] that the companion is a hypergiant, Ergma and Van den Heuvel [44] argue that the progenitor of the neutron star in 4U1223-62 must have a ZAMS mass  $\gtrsim 50M_{\odot}$ . Brown et al. [13], by similar argumentation, arrived at  $\sim 45M_{\odot}$ , but then had the difficulty that 4U1700-37, which they suggested contains a low-mass black hole, appeared to evolve from a lower mass star than the neutron star in 1223. Wellstein and Langer [10] suggest the alternative that in 1223 the mass occurs in the main-sequence phase (Case A mass transfer), which would be expected to be quasi conservative. They find that the progenitor of the neutron star in 1223 could then come from a mass as low as  $26M_{\odot}$ . This is in agreement with Brown et al. [13]

for conservative mass transfer (their Table 1), but these authors discarded this possibility, considering only Case B mass transfer in which case considerable mass would be lost.

Wellstein and Langer [10] are in agreement with Brown et al. [13] that 4U1700-37 should come from a quite massive progenitor. Conservative evolution here is not possible because of the short period of 3.4 days [62]. The compact object mass is here  $1.8 \pm 0.4 M_{\odot}$  [63]. Brown et al. [13] suggest that the compact object is a low-mass black hole. The upper mass limit for these was found by Brown and Bethe [12] to be  $\sim 1.8 M_{\odot}$ , as compared with an upper limit for neutron star masses of  $\sim 1.5 M_{\odot}$ . Thus, there seems to be evidence for some ZAMS masses of  $\sim 40\text{--}50 M_{\odot}$  ending up as low-mass compact objects, whereas we found that lower mass stars in the interval from  $\sim 20$  to  $35 M_{\odot}$  ended up as high-mass black holes. In this sense we agree with Ergma and Van den Heuvel [44] that low-mass compact object formation “is connected with other stellar parameters than the initial stellar mass alone”. We suggest, however, following Brown et al. [13] that stars in binaries evolve differently from single stars because of the different evolution of the He core in binaries resulting from RLOF in their evolution. Namely, “naked” He cores evolve to smaller final compact objects than “clothed” ones.

In fact, this different evolution of binaries was found by Timmes et al. [19]. They pointed out that stars denuded of their hydrogen envelope in early RLOF in binaries would explode as Type Ib supernovae. They found the resulting remnant gravitational mass following explosion to be in the interval of  $1.2\text{--}1.4 M_{\odot}$ , whereas in exploding stars of all masses with hydrogen envelope (Type II supernova explosion) they found a peak at about  $1.28 M_{\odot}$ , chiefly from stars of low masses and another peak at  $1.73 M_{\odot}$  more from massive stars. Our Fe core masses in Fig. 1 come from essentially the same calculations, but the “Remnant” masses of Woosley and Weaver [18] are somewhat greater than those used by Timmes et al. [19]. In fact, the differences between the masses we plot and those of Timmes et al. come in the region  $\sim 1.7\text{--}1.8 M_{\odot}$  (gravitational). This is just in the Brown and Bethe [12] range for low-mass black holes. It may be that some of the stars with low-mass companions evolve into low-mass black holes. Presumably, these would give lower luminosities than the high-mass black holes, although at upper end of the mass range we discuss 4U1700-37 seems to be an example of such a system. Of course, here the high luminosity results from the high-mass loss rate of the giant companion. There are substantial ambiguities in fallback, etc., from the explosion. Our point in this paper is that most of the higher mass single stars  $20\text{--}35 M_{\odot}$  go into high-mass black holes. (The Brown and Bethe [12] limit for low-mass black-hole formation is  $\sim 1.5\text{--}1.8 M_{\odot}$  gravitational, but there is some give and take in both lower and upper limit. Also the stars are not all the same. In particular, different metallicities will give different wind losses.)

## 7. Evolution of low-mass X-ray binaries

We shall briefly point out new developments in the evolution of low-mass X-ray binaries. These were foreseen in the excellent review by Van den Heuvel [64], and there has been substantial development in this field lately.

Low-mass X-ray binaries are considered to be progenitors of recycled pulsars with helium white dwarf companions. In order to bring the magnetic fields of the latter down to  $\sim 10^8$  G and to speed them up to their final period, Van den Heuvel and Bitzarakis [65] had the neutron star accreting  $\sim 0.5 M_{\odot}$  from the main-sequence donor. More detailed recent calculations by Tauris and

Savonije [66] find that if the initial orbital period is below  $\sim 30$  days with a main sequence donor of  $\sim 1M_{\odot}$  which undergoes stable mass transfer with the neutron star, the mass of the latter is increased up to  $\sim 2M_{\odot}$  if the amount of material ejected as a result of propeller effect or disk instabilities is insignificant. This presents a problem for us because the Brown and Bethe [12] mass limit for neutron stars is  $1.5M_{\odot}$ . From this limit, we would say that these neutron stars in low-mass X-ray binaries would have gone into black holes.

A way out of this problem was suggested by Van den Heuvel [64], which is called the evolution of Her X-1 type X-ray binaries (see especially the Appendix of Van den Heuvel [64]). In this case a radiative donor more massive than the neutron star pours matter in unstable mass transfer across the Roche Lobe onto the neutron star. This mass transfer can occur onto the accretion disc by as much as  $\sim 10^4 \dot{M}_{\text{Edd}}$ , if  $\dot{M}_{\text{Edd}} \sim 1.5 \times 10^{-8} M_{\odot} \text{ yr}^{-1}$  is accreted onto the neutron star, since the Eddington limit goes linearly with  $R$  and the radius of the disc can be  $\sim 10^{10} \text{ cm}$ . The advection-dominated inflow-outflow solution (ADIOS) of Blandford and Begelman [23] suggests that the binding energy of the matter released at the neutron star can carry away mass, angular momentum and energy from the gas accreting onto the edge of the accretion disc provided the latter does not cool too much. In this way, the binding energy of a gram of gas at the neutron star can carry off  $\sim 10^3 \text{ g}$  of gas at the edge of the accretion disc. Such radiatively-driven outflows are suggested by King and Begelman [22] to enable common envelope evolution to be avoided. Tauris and Savonije [66] have carried out a detailed evolution of low-mass X-ray binaries with  $P_{\text{orb}} > 2$  days using computer programs based on Eggleton's, which for radiative and semiconvective donors follow, in at least a general way, the above ideas. For a deeply convective donor a short phase of rapid mass loss may reach a rate as large as  $10^4 \dot{M}_{\text{Edd}}$  while the mass of the donor drops to well below the neutron star mass. Although rates  $> 10^4 \dot{M}_{\text{Edd}}$  would be hypercritical for spherical accretion, somewhat higher rates survive hypercritical accretion provided angular momentum is taken into account [28]. The important point is that the donor mass can be brought down sufficiently far before stable mass transfer at a rate  $\lesssim \dot{M}_{\text{Edd}}$  sets in, so that the neutron star can avoid accreting sufficient mass to send it into a black hole. It is not clear what percentage of the neutron stars will survive black-hole fate. Our rough estimates in Section 6 indicate that only a small fraction need to do so.

For even more massive donors ( $2-6M_{\odot}$ ) which are either radiative or semiconvective, work by Tauris et al. [24] indicates that the low-mass X-ray binaries with C/O white-dwarf (CO) companions can be made in much the same way. In an earlier paper, Van den Heuvel [67] had suggested that these binaries would originate from donor stars on the asymptotic giant branch. In order to evolve these, he needed an efficiency  $\alpha > 1$ , i.e., sources additional to those included in our earlier common envelope evolution, such as mass loss by instabilities in the AGB, dissociation energy, etc., have to participate in helping to remove the envelope of the donor star.

King and Ritter [25] have computed a scenario for Cyg X-2 with an initial donor mass of  $\sim 3.6M_{\odot}$ . Currently, the donor has a mass of  $0.5-0.7M_{\odot}$  and a large radius, about  $7R_{\odot}$ . About  $2M_{\odot}$  must have been lost in super-Eddington accretion, roughly along the lines sketched above. More massive donors can lead to relatively more massive white-dwarf companions, which will be C/O white dwarfs.

In fact, the present situation is that no circular NS-CO<sub>c</sub><sup>2</sup> binaries which went through common envelope evolution seem to be observed, the alternative Tauris et al. [24] evolution which avoids

<sup>2</sup> The lower suffix c (e) denotes the circular (eccentric) binaries.

common envelope evolution being preferred. This presents a real dilemma for the standard scenario of common envelope evolution. It seems clear [68] that in the binary B2303 + 46 the companion to the pulsar is a C/O white dwarf. B2303 + 46 is an eccentric binary NS-CO<sub>e</sub>, indicating that the neutron star was formed last. This is confirmed by the unrecycled field strength of the pulsar  $B = 8 \times 10^{11}$  G. Cases have made that the recently discovered J1141-6545 [69] and B1820-11 [70] are also NS-CO<sub>e</sub> binaries.

On the other hand, evolutionary calculations show that formation probability of NS-CO<sub>e</sub> binaries through common envelope evolution is  $\gtrsim 50\%$  as probable as of NS-CO<sub>c</sub> binaries [71]. In this evolution the pulsar magnetic moment will be recycled, brought down at least a factor of 100 [3] and possibly even further, down to the empirical values of  $\sim 5 \times 10^8$  G found in the NS-CO<sub>c</sub> binaries. The lowering of the magnetic fields increases the time of observation by a factor of  $\sim 100$  or of 2000, depending on whether the theoretical or empirical magnetic field is used. Since we fairly certainly observe at least one NS-CO<sub>e</sub> binary, we should see either 100 or 2000 NS-CO<sub>e</sub> binaries which have gone through common envelope evolution. We certainly do not see anything like this, at most the 5 that had earlier been attributed to common envelope evolution, and probably none. Brown et al. [71] remove at least most of this discrepancy by showing that with the introduction of hypercritical accretion the neutron star in common envelope evolution with the evolving main sequence companion goes into a black hole.

## 8. Discussion and conclusion

Our chief new point in the evolution of binaries of compact objects is the use of hypercritical accretion in common envelope evolution, although the idea of hypercritical accretion is not new (Section 3). Chevalier [28] discussed the possibility that angular momentum might hinder hypercritical accretion. In his treatment of the accretion disc, he assumed gas pressure to dominate, in order to raise the temperature sufficiently for neutrinos to be emitted. This entailed a tiny viscosity, characterized by  $\alpha \lesssim 10^{-6}$  in the  $\alpha$ -description. More reasonable values of  $\alpha$  are  $\sim 0.1$ .

Bethe et al. [72] have shown that for larger  $\alpha$ 's,  $\alpha \sim 0.01$ –1, the disc pressure is radiation dominated, and they find a simple hypercritical advection-dominated accretion flow (HADAF) of matter onto the neutron star.

The Bethe et al. HADAF appears to reproduce the Armitage and Livio [77] numerical two-dimensional hydro solution. These latter authors suggest that jets will prevent hypercritical accretion by blowing off the accreting matter. At such high rates of accretion  $\sim 1M_{\odot} \text{ yr}^{-1}$  the Alfvén radius is, however, close to the neutron star surface, and we believe that this will effectively, shut down any magnetically driven jets.

In Section 7 we discussed the advection of a rotating He envelope into a black hole. We believe that two possibilities exist. Phinney and Spruit suggest [80] that the magnetic turbulence is strong enough to keep the He envelope in corotation with the core of the star until shortly before it evolves into a black hole. Then not much angular momentum would have to be advected away in order to let the matter accrete. Alternatively, magnetic turbulence is strong enough so that angular momentum can be carried away from a rapidly rotating He core; then the matter can accrete. From the measured masses of  $\sim 7M_{\odot}$  we know that most of the He core must fall into the black hole, so

one of these scenarios should hold. Both favor high magnetic turbulence, lending credence to the Chevalier suggestion we quoted.

### Note added in proof

The evolutionary calculations by Alexander Heger, referred to in Section 5: High-mass black-hole O/B-star binaries, evolving the CO cores of Wellstein and Langer [10] have now been completed. These CO cores were evolved with lower, more correct, He-star wind losses than had been used by Woosley, Langer, and Weaver [17]. A factor 2–3 reduction is now favored. The Heger calculations show that the scenario outlined in Section 5, particularly the limit of  $\sim 80M_{\odot}$  ZAMS mass for a star in a binary to go into a black hole, is unchanged by the factor of 2–3 lower He wind loss rates. A paper by Brown et al. [40] summarizing these results and suggesting a scenario for the evolution of high mass X-ray black hole binaries is now in preparation.

The discrepancy between eccentric and circular NS-CO binaries discussed at the end of Section 7 has been increased by details of observations on the eccentric PSR J1141-6545 [41]. This pulsar is found to have a characteristic age  $\tau_c = 1.4 \times 10^6$  yr, inferred surface magnetic field strength  $B = 1.3 \times 10^{12}$  G. The total mass of the system is  $2.300 \pm 0.012M_{\odot}$ . Arguments are given that “the companion is probably a massive white dwarf, which formed prior to the birth of the pulsar. Since the companion to the pulsar has not yet been observed optically, there is a small chance that J1141-6545 is a double neutron star system of nearly equal masses. However,  $\sim 1.15M_{\odot}$  is substantially smaller than any other well measured neutron star masses. If J1141-6545 is confirmed as an eccentric (NS-CO<sub>c</sub>) binary, then one would expect to see an additional  $\sim 70$  circular (NS-CO<sub>c</sub>) ones, because of the much longer time the latter can be observed due to their low fields.

Finally, we believe that there have been important developments in the theory of gamma ray bursters. First of all, a supernova origin for the black hole in Nova Sco 1994 (GRO J1655-40) [76] has been observed. The atmosphere of the companion F-star (see our Table 1) has a large excess of  $\alpha$ -particle nuclei, especially  $^{32}\text{S}$ . In ordinary supernova explosions little of this element but in the highly-energetic explosions called hypernovae, which accompany gamma ray bursters, much of the latter is produced. Following the Israelian et al. suggestion that a hypernova explosion took place in the formation of the black hole, Brown et al. [78] suggested that Nova Sco 1994 was a relic of a gamma ray burster. This theme was then developed in Ref. [79] which showed that the binary progenitors of the transient black hole sources were also good progenitors of gamma ray bursters.

### Acknowledgements

G.E.B. and C.H.L. wish to acknowledge support from the U.S. Department of Energy under Grant No. DE-FG02-88ER40388. We thank C. Bailyn for his help with Table 1 and discussions on black-hole transients.

### Appendix A. Common envelope evolution of Cygnus X-3

The closeness of the compact object in Cyg X-3 to its  $\sim 10M_{\odot}$  companion helium star bears witness to an earlier stage of common envelope evolution. Although the mass of the He star has not been measured, the star is similar to V444 Cygni, the mass of which is  $9.3 \pm 0.5M_{\odot}$  [81]. For example, from the period change its mass loss rate would be  $\dot{M}_{\text{dyn}} = 0.6 \times 10^{-5}(M_{\text{He}}/10M_{\odot})M_{\odot} \text{ yr}^{-1}$  [82], whereas that of V444 Cygni is  $\dot{M}_{\text{dyn}} = 1 \times 10^{-5}M_{\odot} \text{ yr}^{-1}$  [83] indicating an  $M_{\text{He}} \sim 10M_{\odot}$ . Mass loss rates cannot easily be obtained from W.-R. winds because of large nonlinear effects which necessitate corrections for “clumpiness”. However, polarization measurement of the Thomson scattering, which depend linearly on the wind, give a mass loss rate of  $\sim \dot{M} = 0.75 \times 10^{-5}M_{\odot} \text{ yr}^{-1}$  [84], roughly compatible with the period change. In agreement with many other authors we take  $M_{\text{He}} = 10M_{\odot}$  in Cyg X-3.

Here we evolve a massive O-star binary with initial ZAMS masses of  $33M_{\odot}$  and  $23M_{\odot}$  as possible progenitor for Cyg X-3. In red giant phase the  $33M_{\odot}$  star will transfer its H envelope to the  $23M_{\odot}$  companion, leaving a He star of

$$M_{\text{He}} = 0.1M_{\text{ZAMS}}^{1.4} = 13M_{\odot}. \quad (\text{A.1})$$

With efficiency of mass transfer assumed to go as  $q^2$ , about half of the  $20M_{\odot}$  H-envelope will be accepted by the companion, which then becomes a rejuvenated  $33M_{\odot}$  star. The He core of the primary then explodes, going into a  $1.5M_{\odot}$  compact object, neutron star or low-mass black hole. After the companion  $33M_{\odot}$  star evolves, the binary will go into common envelope evolution. Eq. (29) can be written

$$\left(\frac{Y_f}{Y_i}\right) = \left(\frac{2.4M_{\text{B},i}}{M_{\text{A},i}}\right)^{(c_d - 1)/c_d} \quad (\text{A.2})$$

where we again take  $c_d = 6$ . With  $M_{\text{B},i} = 33M_{\odot}$  and  $M_{\text{A},i} = 1.5M_{\odot}$ ,

$$Y_f/Y_i = 27. \quad (\text{A.3})$$

The compact object mass scales as

$$M_{\text{A}} \propto Y^{1/(c_d - 1)} = Y^{1/5} \quad (\text{A.4})$$

so that

$$M_{\text{A},f} = 2.9M_{\odot} \quad (\text{A.5})$$

and the final compact object is certainly a black hole, in agreement with Cherepaschchuk and Moffat [85] and Ergma and Yungelson [86]. We believe our evolution here to show that this  $\sim 3M_{\odot}$  black hole is about the most massive that can be formed in common envelope evolution by accretion onto a low-mass compact object, since our  $33M_{\odot}$  companion is near to the ZAMS mass range that will lose mass in an LBV phase, unsuitable for common envelope evolution, so it cannot be made much more massive. We next find

$$a_i/a_f = M_{\text{B},i}/M_{\text{B},f} Y_f/Y_i \simeq 70. \quad (\text{A.6})$$

For an  $a_f \sim 3.5R_{\odot}$  this gives

$$a_i \sim 250R_{\odot} \quad (\text{A.7})$$

comfortably within the red-giant range.

Following Ergma and Yungelson [86] we calculate the accretion rate as

$$\dot{M}_{\text{acc}} = 0.14 \left( \frac{M_{\text{BH}}}{M_{\odot}} \right)^2 v_{1000}^{-4} P_{\text{hr}}^{-4/3} \left( \frac{M_{\odot}}{M_{\text{tot}}} \right)^{2/3} \dot{M}_{\text{wind}} . \quad (\text{A.8})$$

Here  $v_{1000}$  is the wind velocity in units of  $1000 \text{ km s}^{-1}$  and  $P_{\text{hr}}$  is the period in hours.<sup>3</sup> For  $\dot{M}_{\text{wind}}$  we, as Ergma and Yungelson, take  $\dot{M}_{\text{dyn}}$ . These authors take  $v_{1000} = 1.5$ , essentially the result of Van Kerkwijk et al. [82]. An earlier estimate by Van Kerkwijk et al. [87] was  $v_{1000} = 1$ . We believe that the  $v_{\text{wind}}$  to be used here may be different from the (uncertain) measured terminal wind velocities, because the velocity near the compact object is substantially less. Therefore, we take  $v_{1000} = 1$ . Taking  $\dot{M}_{\text{wind}} = \dot{M}_{\text{dyn}}$  we obtain

$$\dot{M}_{\text{acc}} = 2.2 \times 10^{-7} M_{\odot} \text{ yr}^{-1} . \quad (\text{A.9})$$

This is to be compared with

$$\dot{M}_{\text{Edd}} = 4\pi c R / \kappa_{\text{es}} = 2.6 \times 10^{-8} (M_{\text{BH}}/M_{\odot}) M_{\odot} \text{ yr}^{-1} , \quad (\text{A.10})$$

where  $\kappa_{\text{es}} = 0.2 \text{ g/cm}^2$  for He accretion. Our result is in fair agreement with Ergma and Yungelson [86], who find  $\dot{M}_{\text{Edd}} \sim 2.3 \times 10^{-7} M_{\odot} \text{ yr}^{-1}$  for a  $10M_{\odot}$  black hole. The presence of jets in Cyg X-3 argues for super-Eddington rates of accretion, which we find.

Cherepaschchuk and Moffat [85] estimated the total luminosity of Cyg X-3 to be  $L_{\text{bol}} \sim 3 \times 10^{39} \text{ erg}$ . The efficiency of black-hole accretion varies as

$$0.057 < \varepsilon < 0.42 \quad (\text{A.11})$$

for a black hole at rest to a (maximally rotating) Kerr black hole. We expect the black hole to be spun up by accretion from the wind or accretion disc. Taking an intermediate  $\varepsilon = 0.2$ , we find

$$L = 2.5 \times 10^{39} \text{ erg s}^{-1} \quad (\text{A.12})$$

in rough agreement with the Cherepaschchuk and Moffat value.

Cyg X-3 is often discussed as the “missing link” in binary pulsar formation. In fact, because of its high He star mass, upon explosion of the latter, it most probably will break up. But it should be viewed as “tip of the iceberg” [64], in that there must be a great many more such objects with lower mass He stars which are not seen. We have shown, in Section 3 however, that these objects are more likely to contain a black hole than a neutron star.

In our evolutionary scenario, the He star progenitor has about the same ZAMS mass as that of the primary. Thus, the fate of the “naked” He star should be the same low-mass compact object, neutron star or low-mass black hole that resulted from the explosion of the primary.

<sup>3</sup> Through a slip, the two factors preceding  $\dot{M}_{\text{wind}}$  appear in the denominator in [86], although we confirm that they carried out their calculations with the correct formula.

## Appendix B. Implications for LIGO

Our results that there are 10 times more<sup>4</sup> black hole, neutron star binaries than binary neutron stars has important results for LIGO, the detection rates of which were based on the  $\sim 10^{-5}$  per year per galaxy rates of merging for the latter. The combination of masses which will be well determined by LIGO is the chirp mass

$$M_{\text{chirp}} = \mu^{3/5} M^{2/5} = (M_1 M_2)^{3/5} (M_1 + M_2)^{-1/5}, \quad (\text{B.1})$$

where  $M = M_1 + M_2$  is the total system mass. The chirp mass of a NS–NS binary, with both neutron stars of mass  $1.4M_\odot$ , is  $1.2M_\odot$ . A  $10^{-5}$  birth rate implies a rate of  $3 \text{ yr}^{-1}$  out to 200 Mpc [88]. Kip Thorne informs us that LIGO's first long gravitational-wave search in 2002–2003 is expected to see binaries with  $M_{\text{chirp}} = 1.2M_\odot$  out to 21 Mpc.

The chirp mass corresponding to the Bethe and Brown [1] LMBH-NS binary with masses  $2.4M_\odot$  and  $1.4M_\odot$  is  $1.6M_\odot$ . Including an  $\sim 30\%$  increase in the rate to allow for high-mass black-hole, neutron-star mergers (which should be regarded as a lower limit because of the high-mass limit of  $80M_\odot$  used by Bethe and Brown for going into a HMBH) gives a 26 times higher rate than Phinney's estimate for NS–NS mergers. These factors are calculated from the signal-to-noise ratio, which goes as  $M_{\text{chirp}}^{5/6}$  and then cubing it to obtain the volume of detectability. We then predict a ratio of  $3 \times (21/200)^3 \times 26 = 0.09 \text{ yr}^{-1}$  for 2003, rather slim. The enhanced LIGO interferometer planned to begin in 2004 should reach out beyond 150 Mpc for  $M_{\text{chirp}} = 1.2M_\odot$ , increasing the detection rate to  $3 \times (150/200)^3 \times 26 = 33 \text{ yr}^{-1}$ . We therefore predict that LIGO will see more mergers per month than NS–NS mergers per year.

## Appendix C. Binary contributions to gamma-ray bursters

The sheer numbers of black-hole, neutron-star binaries should dominate the mergers for gravitational waves, which could be detected by LIGO [39]. For gamma-ray bursts, the presence of an event horizon eases the baryon pollution problem, because energy can be stored in the rotational energy of the black hole, and then released into a cleaner environment via the Blandford–Znajek magnetohydrodynamic process.

Binaries containing a black hole, or single black holes, have been suggested for some time as good progenitors for gamma-ray bursts [89–92,34,93]. Reasons for this include the fact that the rest mass of a stellar mass black hole is comparable to what is required to energize the strongest GRB. Also, the horizon of a black hole provides a way of quickly removing most of the material present in the cataclysmic event that formed it. This may be important because of the baryon pollution problem: we need the ejecta that give rise to the GRB to be accelerated to a Lorentz factor of 100 or more, whereas the natural energy scale for any particle near a black hole is less than its mass. Consequently, we have a distillation problem of taking all the energy released and putting

<sup>4</sup> Actually about 20 times more if we include the binaries in which the pulsar goes into a black hole in the He shell burning evolution. However, these will have masses not very different from the binary neutron stars so we do not differentiate them.

it into a small fraction of the total mass. The use of a Poynting flux from a black hole in a magnetic field [94] does not require the presence of much mass, and uses the rotation energy of the black hole, so it provides naturally clean power.

As a neutron star in a binary moves nearer to a black hole companion, it is distorted into a torus around the latter. Most of the torus matter enters the black hole from the last stable Keplerian orbit of  $R = 6GM_{\text{BH}}/c^2$ , carrying considerable angular momentum. In the process the black hole is spun up until it rotates with some fraction of the speed of light. A magnetic field which originates from the neutron star, but which could have been enhanced by differential rotation is anchored in the remaining part of the torus, the accretion disc.

When a rapidly rotating black hole is immersed in a magnetic field, frame dragging twists the field lines near the hole, which causes a Poynting flux to be emitted from near the black hole. This is the Blandford–Znajek mechanism [94]. The source of energy for the flux is the rotation of the black hole. The source of the field is the surrounding accretion disk or debris torus. We showed [95] that at most 9% of the rest mass of a rotating black hole can be converted to a Poynting flux, making the available energy for powering a GRB

$$E_{\text{BZ}} = 1.6 \times 10^{53} (M/M_{\odot}) \text{ erg}. \quad (\text{C.1})$$

The power depends on the applied magnetic field:

$$P_{\text{BZ}} \sim 6.7 \times 10^{50} B_{15}^2 (M/M_{\odot})^2 \text{ erg s}^{-1} \quad (\text{C.2})$$

(where  $B_{15} = B/10^{15}$  G). This shows that modest variations in the applied magnetic field may explain a wide range of GRB powers, and therefore of GRB durations. There has been some recent dispute in the literature whether this mechanism can indeed be efficient [96] and whether the power of the BH is ever significant relative to that from the disk [97]. The answer in both cases is yes, as discussed by Lee et al. [95].

The issue, therefore, in finding efficient GRB sources among black holes is to find those that spin rapidly. There are a variety of reasons why a black hole might have high angular momentum. It may have formed from a rapidly rotating star, so the angular momentum was there all along ('original spin', according to Blandford [98]); it may also have accreted angular momentum by interaction with a disk ('venial spin') or have formed by coalescence of a compact binary ('mortal spin'). We shall review some of the specific situations that have been proposed in turn.

Neutron star mergers are among the oldest proposed cosmological GRB sources [99–101], and especially the neutrino flux is still actively studied as a GRB power source [102]. However, once the central mass has collapsed to a black hole it becomes a good source for BZ power, since it naturally spins rapidly due to inheritance of angular momentum from the binary [103]. Likewise BH-NS binaries [104] will rapidly transfer a large amount of mass once the NS fills its Roche lobe, giving a rapidly rotating BH [105]. The NS remnant may then be tidally destroyed, leading to a compact torus around the BH. It is unlikely that this would be long-lived enough to produce the longer GRB, but perhaps the short ( $t \lesssim 1$  s) ones could be produced [106]. However, mass transfer could stabilize and lead to a widening binary in which the NS lives until its mass drops to the minimum mass of about  $0.1M_{\odot}$ , and then becomes a debris torus [107]. By then, it is far enough away that the resulting disk life time exceeds 1000 s, allowing even the longer GRB to be made. Thus BH-NS and NS-NS binaries are quite promising. They have the added advantage that their environment

is naturally reasonably clean, since there is no stellar envelope, and much of the initially present baryonic material vanishes into the horizon.

In addition to the mergers from compact objects, Fryer and Woosley [34] suggested that GRBs could originate from the coalescence of low-mass black hole and helium-star binaries in the Bethe and Brown [1] scenario. From Eq. (35) we see that binaries survived in the initial range of  $0.5 \times 10^{13} \text{ cm} < a_i < 1.9 \times 10^{13} \text{ cm}$ . Inside that range for  $0.04 \times 10^{13} \text{ cm} < a_i < 0.5 \times 10^{13} \text{ cm}$  the low-mass black hole coalesces with the core. Hence, using a separation distribution flat in  $\ln a$ , coalescences are more common than low-mass black-hole, neutron-star binaries by a factor  $\ln(0.5/0.04)/\ln(1.9/0.5) = 1.9$ . In Bethe and Brown [1] the He star compact-object binary was disrupted  $\sim 50\%$  of the time in the last explosion, which we do not have here. Thus, the rate of low-mass black-hole, He-star mergers is 3.8 times the formation rate of low-mass black-hole, neutron-star binaries which merge, or  $R = 3.8 \times 10^{-4} \text{ yr}^{-1}$  in the Galaxy.

In Table 2 we summarize the formation rates of GRBs and gravity waves from the binaries considered in this review.

Because gamma-ray bursts have a median redshift of 1.5–2 [108]), and the supernova rate at that redshift was 10–20 times higher than now, the gamma-ray burst rate as observed is higher than one expects using the above rates. However, for ease of comparison with evolutionary scenarios we shall use the GRB rate at the present time (redshift 0) of about 0.1 GEM. (Wijers et al. [108] found a factor 3 lower rate, but had slightly underestimated it because they overestimated the mean GRB redshift; see Ref. [106] for more extensive discussions of the redshift dependence.) An important uncertainty is the beaming of gamma-ray bursts: the gamma rays may only be emitted in narrow cones around the spin axis of the black hole, and therefore most GRBs may not be seen by us. An upper limit to the ratio of undetected to detected GRB is 600 [109], so an upper limit to the total required formation rate would be 60 GEM. We may have seen beaming of about that factor or a bit less in GRB 990123 [110], but other bursts (e.g. 970228, 970508) show no evidence of beaming in the afterglows (which may not exclude beaming of their gamma rays). At present, therefore, any progenitor with a formation rate of 10 GEM or more should be considered consistent with the observed GRB rate.

An exciting possibility for the future will be to receive both gravitational-wave and gamma-ray burst signals from the same merger, with attendant detailed measurement, which would give witness to them arising from the same binary.

Because we dealt in this review with binaries, we did not explain one popular model of GRBs, the Woosley Collapsar model [92]. In this model a black hole is formed in the center of a rotating

**Table 2**  
Summary of the formation rates of various sources of gamma-ray bursts (GRB) or gravity waves (GW) from the binaries considered in this review. L(H)BH means low- (high-)mass black hole

Object	GRB	GW	Rate [GEM*]
NS-NS merger	X	X	10
NS-BH merger	X	X	100
WR star-LBH merger	X		380

\*GEM means Galactic Events per Megayear; rates are quoted for redshift 0.

W.-R. star. The outer matter can then be accreted into the neutron star, spinning it up. If, however, magnetic turbulence is sufficient to keep the envelope of the progenitor in corotation with the core until a few days before collapse of the latter, as suggested by Phinney and Spruit [80] the He envelope could not furnish enough angular momentum to the black hole for the latter to drive the necessary jets (see the end of Section 8).

## References

- [1] H.A. Bethe, G.E. Brown, *Astrophys. J.* 506 (1998) 780.
- [2] I.H. Stairs, Z. Arzoumanian, F. Camilo, A.G. Lyne, D.J. Nice, J.H. Taylor, S.E. Thorsett, A. Wolszczan, *Astrophys. J.* 505 (1998) 352.
- [3] G.E. Brown, *Astrophys. J.* 440 (1995) 270.
- [4] A. Herrero, R.P. Kudritzki, R. Gabler, J.M. Vilchez, A. Gabler, *Astron. Astrophys.* 297 (1995) 556.
- [5] R.A.M.J. Wijers, in: R.A.M.J. Wijers et al. (Eds.), *Evolutionary Processes in Binary Stars*, Vol. 327, Kluwer Acad. Publ., Dordrecht, 1996.
- [6] G.E. Brown, C.-H. Lee, H.A. Bethe, *New Astron.* 4 (1999) 313.
- [7] M. Prakash, J.R. Cooke, J.M. Lattimer, *Phys. Rev. D* 52 (1995) 661.
- [8] H.A. Bethe, G.E. Brown, *Astrophys. J.* 517 (1999) 318.
- [9] S.E. Woosley, N. Langer, T.A. Weaver, *Astrophys. J.* 448 (1995) 315.
- [10] S. Wellstein, N. Langer, *Astron. Astrophys.* 350 (1999) 148.
- [11] V. Thorsson, M. Prakash, J.M. Lattimer, *Nucl. Phys. A* 572 (1994) 693.
- [12] G.E. Brown, H.A. Bethe, *Astrophys. J.* 423 (1994) 659.
- [13] G.E. Brown, J.C. Weingartner, R.A.M.J. Wijers, *Astrophys. J.* 463 (1996) 297.
- [14] M. Prakash, I. Bombaci, M. Prakash, P.J. Ellis, J.M. Lattimer, R. Knorren, *Phys. Rep.* 280 (1997) 1.
- [15] G.Q. Li, C.-H. Lee, G.E. Brown, *Nucl. Phys. A* 625 (1997) 372; *Phys. Rev. Lett.* 79 (1997) 5214.
- [16] S.E. Thorsett, D. Chakrabarty, *Astrophys. J.* 512 (1999) 288.
- [17] S.E. Woosley, N. Langer, T.A. Weaver, *Astrophys. J.* 411 (1993) 823.
- [18] S.E. Woosley, T.A. Weaver, *Astrophys. J. Suppl. Ser.* 101 (1995) 181.
- [19] F.X. Timmes, S.E. Woosley, T.A. Weaver, *Astrophys. J.* 457 (1996) 834.
- [20] S.F. Portegies Zwart, L.R. Yungelson, *Astron. Astrophys.* 332 (1998) 173.
- [21] J.M. Cordes, D.F. Chernoff, *Astrophys. J.* 482 (1997) 971.
- [22] A.R. King, M.C. Begelman, *Astrophys. J.* 519 (1999) L169.
- [23] R.D. Blandford, M.C. Begelman, *Mon. Not. R. Astron. Soc.* 303 (1999) L1.
- [24] T. Tauris, E.P.J. van den Heuvel, G.J. Savonije, *Astrophys. J.* 530 (2000) L93.
- [25] A.R. King, H. Ritter, *Mon. Not. R. Astron. Soc.* 309 (1999) 253.
- [26] J.H. Applegate, Columbia Univ. Preprint, 1998.
- [27] R.A. Chevalier, *Astrophys. J.* 411 (1993) L33.
- [28] R.A. Chevalier, *Astrophys. J.* 459 (1996) 322.
- [29] E. Shima, T. Matsuda, H. Takeda, K. Sawada, *Mon. Not. R. Astron. Soc.* 217 (1985) 367.
- [30] M. Ruffert, *Astron. Astrophys. Suppl. Ser.* 106 (1994) 505.
- [31] M. Ruffert, D. Arnett, *Astrophys. J.* 427 (1994) 351.
- [32] T. Wettig, G.E. Brown, *New Astron.* 1 (1996) 17.
- [33] P. Eggleton, Private communication, 1998.
- [34] C.L. Fryer, S.E. Woosley, *Astrophys. J.* 502 (1998) L9.
- [35] F.A. Rasio, M. Livio, *Astrophys. J.* 471 (1996) 366.
- [36] H. Braun, N. Langer, *Astron. Astrophys.* 297 (1995) 771.
- [37] G.E. Brown, *Phys. Bl.* 53 (1997) 671.
- [38] C. Fryer, B. Kalogera, *Astrophys. J.* 489 (1997) 244.
- [39] G.E. Brown, R.A.M.J. Wijers, C.-H. Lee, H.K. Lee, H.A. Bethe, *Astrophys. J. Lett.* (1999) submitted; astro-ph/9905337.

- [40] G.E. Brown, A. Heger, N. Langer, C.-H. Lee, S. Wellstein, H.A. Bethe, in preparation.
- [41] V.M. Kaspi, A.G. Lyne, R.N. Manchester, F. Crawford, F. Camilio, J.F. Bell, N. D'Amico, I.H. Stairs, N.P.F. McKay, D.G. Morris, A. Possenti, in preparation.
- [42] T.A. Weaver, S.E. Woosley, Phys. Rep. 227 (1993) 65.
- [43] S.F. Portegies Zwart, F. Verbunt, E. Ergma, Astron. Astrophys. 321 (1997) 207.
- [44] E. Ergma, E.P.J. van den Heuvel, Astron. Astrophys. 331 (1998) L29.
- [45] M. de Kool, E.P.J. van den Heuvel, E. Pylyser, Astron. Astrophys. 183 (1987) 47.
- [46] G. Schaller, D. Schaefer, G. Meynet, A. Maeder, Astron. Astrophys. Suppl. Ser. 96 (1992) 269.
- [47] T.A. Weaver, G.B. Zimmerman, S.E. Woosley, Astrophys. J. 225 (1978) 1021.
- [48] F. Verbunt, E.P.J. van den Heuvel, in: W.H.G. Lewin, J. van Paradijs, E.P.J. van den Heuvel (Eds.), X-ray Binaries, Cambridge Univ. Press, Cambridge, 1995, p. 457.
- [49] F. Verbunt, C. Zwaan, Astron. Astrophys. 100 (1981) L7.
- [50] W. Chen, C.R. Shrader, M. Livio, Astrophys. J. 491 (1997) 312.
- [51] C.D. Bailyn, R.K. Jain, P. Coppi, J.A. Orosz, Astrophys. J. 499 (1998) 367.
- [52] J.A. Orosz et al., Astrophys. J. 499 (1998) 375.
- [53] A.V. Filippenko, D.C. Leonard, T. Matheson, W. Li, E.C. Moran, A.G. Riess, PASP, 111 (1999) 969.
- [54] C.D. Bailyn, Private communication, 1998.
- [55] C.A. Haswell, E.L. Robinson, K. Horne, K., R.F. Stiening, T.M.C. Abbott, Astrophys. J. 411 (1993) 802.
- [56] R. Kippenhahn, E. Meyer-Hofmeister, Astron. Astrophys. 54 (1977) 539.
- [57] S. Portegies Zwart, F. Verbunt, Astron. Astrophys. 309 (1996) 179.
- [58] S. Mineshige, K. Nomoto, T. Shigeyama, Astron. Astrophys. 267 (1993) 95.
- [59] S. Mineshige, H. Nomura, M. Hirose, K. Nomoto, T. Suzuki, Astrophys. J. 489 (1997) 227.
- [60] P.J. Callanan, M.R. Garcia, J.E. McClintock, P. Zhao, R.A. Remillard, F. Haberl, Astrophys. J. 461 (1996) 351.
- [61] L. Kaper, H.J.G.L.M. Lamers, E.P.J. van den Heuvel, E.J. Zuiderwijk, Astron. Astrophys. 300 (1995) 446.
- [62] B.C. Rubin et al., Astrophys. J. 459 (1996) 259.
- [63] S.R. Heap, M.F. Corcoran, Astrophys. J. 387 (1992) 340.
- [64] E.P.J. van den Heuvel, J. Astrophys. Astron. 16 (1995) 255.
- [65] E.P.J. van den Heuvel, O. Bitzaraki, Astron. Astrophys. 297 (1995) L41.
- [66] T. Tauris, G.J. Savonije, Astron. Astrophys. 350 (1999) 928.
- [67] E.P.J. van den Heuvel, Astron. Astrophys. 291 (1994) L39.
- [68] M.H. Van Kerkwijk, S.R. Kulkarni, Astrophys. J. 516 (1999) L25.
- [69] T.M. Tauris, T. Sennels, Astron. Astrophys. 355 (2000) 236.
- [70] S.F. Portegies Zwart, L.R. Yungelson, Mon. Not. R. Astron. Soc. 309 (1999) 26.
- [71] G.E. Brown, C.-H. Lee, S.F. Portegies Zwart, H.A. Bethe, New Astron. (1999) submitted, astro-ph/9911130.
- [72] G.E. Brown, C.-H. Lee, H.A. Bethe, Astrophys. J. (2000) in press, astro-ph/9909132.
- [73] S.A. Colgate, Astrophys. J. 163 (1971) 221.
- [74] Ya.B. Zeldovich, L.N. Ivanova, D.K. Nadezhin, Sov. Astron. Astrophys. J. 16 (1972) 209.
- [75] G.S. Bisnovatyi-Kogan, S.A. Lamzin, Sov. Astron. 28 (1984) 187.
- [76] G. Israelian, R. Rebolo, G. Basri, J. Casares, E.L. Martin, Nature 401 (1999) 142.
- [77] P. Armitage, M. Livio, Astrophys. J. 532 (2000) 540.
- [78] G.E. Brown, C.-H. Lee, H.K. Lee, H.A. Bethe, astro-ph/9911458, Proc. 5th Huntsville Gamma Ray Burst Symposium, Oct. 18–22, 1999.
- [79] G.E. Brown, C.-H. Lee, R.A.M.J. Wijers, H.K. Lee, G. Israelian, H.A. Bethe, astro-ph/0003361.
- [80] E.S. Phinney, H. Spruit, Nature 393 (1998) 139.
- [81] S.V. Marchenko, A.F.J. Moffat, G. Koenigsberger, Astrophys. J. 422 (1994) 810.
- [82] M.H. van Kerkwijk, T.R. Geballe, D.L. King, M. van der Klis, J. van Paradijs, Astron. Astrophys. 314 (1996) 521.
- [83] K.F. Khalilullin, A.I. Khalilullina, A.M. Cherepashchuk, Sov. Astron. Lett. 10 (1984) 250.
- [84] N. St.-Louis, A.F.J. Moffat, L. Lapointe, Y.S. Efimov, N.M. Shaklovskoy, G.K. Fox, V. Pirola, Astrophys. J. 410 (1993) 342.
- [85] A.M. Cherepashchuk, A.F.J. Moffat, Astrophys. J. 424 (1994) L53.
- [86] E. Ergma, L.R. Yungelson, Astron. Astrophys. 333 (1998) 151.

- [87] M.H. van Kerkwijk, *Astron. Astrophys.* 276 (1993) L9.
- [88] E.S. Phinney, *Astrophys. J.* 380 (1991) L17.
- [89] B. Paczyński, *Acta Astron.* 41 (1991) 257.
- [90] B. Paczyński, *Astrophys. J.* 494 (1998) L45.
- [91] R. Mochkovitch, M. Hernanz, J. Isern, X. Martin, *Nature* 361 (1993) 236.
- [92] S.E. Woosley, *Astrophys. J.* 405 (1993) 273.
- [93] A. MacFadyen, S.E. Woosley, *Astrophys. J.* 524 (1999) 262.
- [94] R.D. Blandford, R.L. Znajek, *Mon. Not. R. Astron. Soc.* 179 (1977) 433.
- [95] H.K. Lee, R.A.M.J. Wijers, G.E. Brown, *Phys. Rep.* 325 (2000) 83.
- [96] L.-X. Li, *Phys. Rev. D* 61 (2000) 084016.
- [97] M. Livio, G.I. Ogilvie, J.E. Pringle, *Astrophys. J.* 512 (1999) L100.
- [98] R.D. Blandford, Talk at Santa Barbara meeting on GRBs, 15–19 March 1999.
- [99] D. Eichler, M. Livio, T. Piran, D.N. Schramm, *Nature* 340 (1989) 126.
- [100] J. Goodman, A. Dar, S. Nussinov, *Astrophys. J.* 314 (1987) L7.
- [101] B. Paczyński, *Astrophys. J.* 308 (1986) L43.
- [102] M. Ruffert, H.-T. Janka, in: C.A. Meegan, R.D. Preece, T.M. Koshut (Eds.), *Proceedings of the 4th Huntsville Gamma-Ray Bursts Symposium*, Huntsville, Alabama, September 15–20, 1997, New York, 1998, AIP Conference Proceedings, p. 428, 793.
- [103] M.J. Rees, P. Mészáros, *Mon. Not. R. Astron. Soc.* 258 (1992) L41.
- [104] J.M. Lattimer, D.N. Schramm, *Astrophys. J.* 192 (1974) L145.
- [105] W. Kluzniak, W.H. Lee, *Astrophys. J.* 494 (1998) L53.
- [106] C.L. Fryer, S.E. Woosley, D.H. Hartmann, *Astrophys. J.* 526 (1999) 152.
- [107] S.F. Portegies Zwart, *Astrophys. J.* 503 (1998) L53.
- [108] R.A.M.J. Wijers, J.S. Bloom, J.S. Bagla, P. Natarajan, *Mon. Not. R. Astron. Soc.* 294 (1998) L13.
- [109] P. Mészáros, M.J. Rees, R.A.M.J. Wijers, *New Astron.* 4 (1999) 303.
- [110] S.R. Kulkarni et al., *Nature* 398 (1999) 389.

---

## Chapter 16

# The Blandford-Znajek Process as a Central Engine for a Gamma-Ray Burst

H.K. Lee, R.A.M.J. Wijers and G.E. Brown

Reprinted with permission from *Physics Reports*, 325 (2000) 83–114

Copyright (2000) by Elsevier Science B.V.

### Commentary by G.E. Brown

Hyun Kyu Lee from Hanyang University had come to Stony Brook in 1998 to spend his sabbatical with us. It had seemed clear to me that the varying times of the longer bursts, from a few up to some thousands of seconds, and the vast amount of energy  $\sim 10^{52}$  ergs delivered by gamma-ray bursts could come only from the rotational energy of black holes. In the Blandford-Znajek (BZ) mechanism, the extraction of rotational energy is envisioned in the form of Poynting flux winds.

There had been many criticisms of different parts of the Blandford-Znajek mechanism. Hyun Kyu ran into almost all of the critics at meetings such as a joint Italian-Korean one, all over the world during the next year, and set matters right in discussion. What emerged is this Physics Report.

A substantial improvement in the Blandford-Znajek mechanism, in our view, has been carried out since the publication of our Physics Report, in several papers by Maurice van Putten, and it is summarized in his Physics Report (2001). An unsatisfactory feature of the BZ mechanism was that the power in the Poynting flux to run the gamma-ray burster was delivered into the “loading region”. The latter was never clearly defined. Basically, it was a “dumpster” for the Poynting flux. The load resistance was estimated by Lovelace, MacAuslan and Barns (1979), by Macdonald and Thorne (1982) and by Phinney (1983), all estimates assuming that there exists a distant load whose resistance is equal to that of the impedance of the black hole, to maximize the power output.

In order to power a gamma-ray burst, it is necessary to put sufficient energy into a hot, fairly compact, baryon poor fireball sufficient to raise its temperature well above the electron-positron pair creation threshold (Goodman 1986; Paczyński 1986). Sufficient energy is transported by the Poynting vector in the BZ process to do this, but, first of all, it is not clear what the loading area is — somehow a region in which the energy is deposited — and it is not clear how the Poynting flux is converted to electron-positron pairs.

Van Putten, in the approximation of cylindrical symmetry for the core of the jet, identifies a piece of flux tube some distance from the black hole in which there is differential rotation, so that energy in the form of electron-positron pairs is delivered in this region.

In the force-free limit, the electric field (in Gaussian unit) satisfies

$$\vec{E} + \frac{\vec{v}}{c} \times \vec{B} = 0. \quad (16.1)$$

The associated charge density near the axis of rotation of a black hole satisfies

$$n_e = \rho = -\frac{\tilde{\Omega} \cdot \vec{B}}{2\pi c}. \quad (16.2)$$

This is easily derived from

$$\vec{\nabla} \cdot \vec{E} = 4\pi\rho \quad (16.3)$$

and

$$\vec{v} = \tilde{\Omega} \times \vec{r}, \quad (16.4)$$

where  $\Omega$  is the angular frequency of rotation. In (16.1)–(16.4) we have tacitly assumed working in a frame of reference, whose time coordinate is orthogonal to the space-like coordinates. In the vicinity of a rotating black hole, this corresponds to working in a frame of reference in co-rotation with a zero angular momentum observer. Thus,  $\Omega$  refers the relative angular velocity of the magnetosphere with respect to a local zero angular momentum observer.

Equation (16.2) gives a current density at horizon and infinity

$$j = -\frac{\tilde{\Omega} \cdot \vec{B}}{2\pi}, \quad (16.5)$$

and the total current in the open flux tube becomes

$$I = -\Omega A_\phi, \quad (16.6)$$

where  $A_\phi$  is the magnetic stream function defined by the net horizon magnetic flux  $\Phi_H^e$ :

$$\Phi_H^e = 2\pi A_\phi \approx 4\pi B \left( \frac{GM}{c^2} \right)^2. \quad (16.7)$$

The angular velocity as seen by a distant observer will differ from that relative to a zero angular momentum observer due to frame-dragging. Frame-dragging reaches the angular velocity  $\Omega_H$  of the black hole on the horizon, and vanishes at infinity. Current continuity implies that the angular velocity of rotation  $\Omega_-$  of the outer section of the flux tube at infinity must be equal in magnitude to  $\Omega_H - \Omega_+$ :

$$\Omega_- = \Omega_H - \Omega_+, \quad (16.8)$$

where  $\Omega_+$  denotes the angular velocity of the inner section near the horizon of the black hole. This guarantees that the (negative) current of electrons going outwards from the flux tube is equal to the positive current going into the black hole, i.e., current conservation. Van Putten also shows from a topological argument that

$$\Omega_- = \Omega_D, \quad (16.9)$$

where  $\Omega_D$  is the angular frequency of the torus (disk).

From Faraday's law, Eq. (16.1), the potential difference away from the gap, taken to be a cylinder in differential rotation, is

$$\Delta V = \frac{1}{c}(\Omega_+ - \Omega_-)A_\phi, \quad (16.10)$$

giving a delivered power

$$\begin{aligned} P &= I\Delta V = I \times \frac{1}{c}(\Omega_+ - \Omega_-)A_\phi^2 \\ &= \frac{1}{c} \Omega_D(\Omega_H - 2\Omega_D)A_\phi^2. \end{aligned} \quad (16.11)$$

The energy that can be got out of the black hole rotational energy is, of course, determined by the Kerr geometry, and this is similar as found by Blandford and Znajek. The difference, however, is that the power (16.11) is carried by particle outflow.

## References

- (1) Goodman, J. (1986), *Astrophysical Journal* **308**, L47.
- (2) Paczyński, B.P. (1986), *Astrophysical Journal* **308**, L43.
- (3) Lovelace, R.V.E., MacAuslan, J., and Barns, M. (1979), in *Experimental Gravitation*, ed. B. Bertatti (Academic Press, N.Y.), p. 361.
- (4) MacDonald, D.A., and Thorne, K.S. (1982), *Monthly Notices of the Royal Astronomical Society* **198**, 345.
- (5) Phinney, E.S. (1983), Ph.D. diss., Univ. of Cambridge.
- (6) Van Putten, M.H.P.M. (2001), *Physics Reports* **345**, 1.

This page is intentionally left blank

# THE BLANDFORD-ZNAJEK PROCESS AS A CENTRAL ENGINE FOR A GAMMA-RAY BURST

**Hyun Kyu LEE<sup>a</sup>, R.A.M.J. WIJERS<sup>b</sup>, G.E. BROWN<sup>b</sup>**

<sup>a</sup>*Department of Physics, Hanyang University, Seoul 133-791, South Korea*

<sup>b</sup>*Department of Physics and Astronomy, State University of New York at Stony Brook, Stony Brook,  
NY 11794-3800, USA*



AMSTERDAM – LAUSANNE – NEW YORK – OXFORD – SHANNON – TOKYO



ELSEVIER

# The Blandford-Znajek process as a central engine for a gamma-ray burst

Hyun Kyu Lee<sup>a,\*</sup>, R.A.M.J. Wijers<sup>b</sup>, G.E. Brown<sup>b</sup>

<sup>a</sup>Department of Physics, Hanyang University, Seoul 133-791, South Korea

<sup>b</sup>Department of Physics and Astronomy, State University of New York at Stony Brook, Stony Brook, NY 11794-3800, USA

Received June 1999; editor: G.E. Brown

## Contents

1. Introduction	86	Acknowledgements	104
2. Overview of the proposed central engine for GRB	87	Appendix A. Rotating black hole	104
3. Blandford-Znajek process	89	Appendix B. Electromagnetic fields in vacuum around rotating black hole	106
4. Evolution of a black hole via the Blandford-Znajek process	92	Appendix C. Axial-symmetric force-free magnetosphere	107
5. Magnetic field and force-free plasma	94	C.1. Force-free magnetosphere	107
6. Rotation of the black hole	97	C.2. Energy and angular momentum outflow	108
7. Magnetized accretion disks	100	Appendix D. Power from rotating black hole	111
8. Conclusion	103	Note added in proof	113
		References	113

## Abstract

We investigate the possibility that gamma-ray bursts are powered by a central engine consisting of a black hole with an external magnetic field anchored in a surrounding disk or torus. The energy source is then the rotation of the black hole, and it is extracted electromagnetically via a Poynting flux, a mechanism first proposed by Blandford and Znajek (*Mon. Nat. R. Astron. Soc.* 179 (1997) 433) for AGN. Our reanalysis of the strength of the Blandford-Znajek power shows that the energy extraction rate of the black hole has been underestimated by a factor 10 in previous works. Accounting both for the maximum rotation energy of the hole and for the efficiency of electromagnetic extraction, we find that a maximum of 9% of the rest mass of the hole can be converted to a Poynting flow, i.e. the energy available to produce a gamma-ray burst is  $1.6 \times 10^{53} (M/M_{\odot})$  erg for a black hole of mass  $M$ . We show that the black holes formed in a variety of gamma-ray burst scenarios probably contain the required high angular momentum. To extract the energy from a black hole in the required time of  $\lesssim 1000$  s a field of  $10^{15}$  G near the black hole is needed. We give an example of a disk-plus-field structure that both delivers the required field and makes the Poynting flux from

\*Corresponding author.

E-mail address: hklee@hepth.hanyang.ac.kr (H.K. Lee)

the hole dominate that of the disk. Thereby we demonstrate that the Poynting energy extracted need not be dominated by the disk, nor is limited to the binding energy of the disk. This means that the Blandford–Znajek mechanism remains a very good candidate for powering gamma-ray bursts. © 2000 Elsevier Science B.V. All rights reserved.

*PACS:* 98.70.Rz; 97.60.Lf

*Keywords:* Gamma-ray bursts; Black hole; Accretion disk

---

## 1. Introduction

Gamma-ray bursts presently provide great excitement in astronomy and astrophysics as optical observations by way of many instruments give considerable detail of the history of each burst. We are concerned here with the prodigious energy in each burst, the estimate for GRB971214 being  $\gtrsim 3 \times 10^{53}$  erg [1], although this could be diminished if considerable beaming is involved in the central engine, as we will discuss.

Amazingly,  $2 \times 10^{54}$  erg is just the rest mass energy of our sun, so it seems immediately clear that the central engine for the GRB must be able to extract a substantial fraction of the rest mass energy of a compact object, neutron star or black hole, and convert it into energy of GRB.

The second criterion for the central engine is that it must be able to deliver power over a long interval up to  $\sim 1000$  s, since some GRBs last that long although other GRBs last only a fraction of a second. It must also be able to account for the vast diversity in pulses, etc., or, alternatively, one must have a number of diverse mechanisms.

We believe the need to deliver power over the long-time found in some bursters to be the most difficult requirement to fulfill, since the final merger time of the compact objects is only a fraction of a second and it is difficult to produce a high-energy source of, e.g.,  $v\bar{v}$ -collisions that goes on for more than 2 or 3 s.

For many years mergers of binary neutron stars were considered to be likely sources for the GRBs. The estimated merger rate in our Galaxy of a few GEM<sup>1</sup> is of the right order for the occurrence of GRBs. The possible problem with binary mergers might be the ejected materials during the merging processes. Not more than  $\sim 10^{-5} M_{\odot}$  of baryons can be involved in the GRB, since it would not be possible to accelerate a higher mass of nucleons up to the Lorentz factors  $\Gamma \sim 100$  needed with the energies available.

We find the merger of a neutron star with a black hole to be a particularly attractive mechanism. The baryon number “pollution” problem can be solved by the main part of the baryons going over the event horizon. In the Blandford-Znajek mechanism [2] we wish to invoke, a substantial proportion of the rotational energy of the black hole, which will be sent into rapid rotation by swallowing up the neutron star matter, can be extracted through the Poynting vector. The rate of extraction is proportional to the square of magnetic field strength,  $B^2$ , as we shall discuss, so that power can be furnished over varying times, depending upon the value of  $B$ . With substantial beaming, we estimate that  $B \sim 10^{15}$  G would be sufficient to power the most energetic GRBs with  $\sim 10^{53}$  erg.

Recently at least three magnetars, neutron stars with fields  $B \sim 10^{14}$ – $10^{15}$  G, have been observed. Their visible lifetime is only a few thousand years, because neutron stars slow down by emission of magnetic dipole radiation and join the “graveyard” where they no longer emit pulses. The time of observability is proportional to  $B^{-1}$ , so the number of these high magnetic field stars may not be an order of magnitude less than the garden variety  $10^{12}$  G neutron stars. It is also possible that existing magnetic fields can be increased by the dynamo effect.

Failed supernovae were suggested by Woosley [3] as a source of GRB. In this case the black hole would be formed in the center of a massive star, and surrounding baryonic matter would

<sup>1</sup> GEM = 1 Galactic Event per Megayear.

accrete into it, spinning it up. This mechanism is often discussed under the title of hypernovae [4].

More recently Bethe and Brown [5] found that in binary neutron star evolutions, an order of magnitude more low-mass black-hole, neutron-star binaries were formed than binary neutron stars. The low-mass black-hole mass of  $\sim 2.4M_{\odot}$  looks favorable for the Blandford-Znajek mechanism.

In some calculations which begin with a neutron star binary, one of the neutron stars evolves into a black hole in the process of accretion, and the resulting binary might also be a good candidate for GRBs. In any case, there are various possibilities furnished by black-hole, neutron-star binaries.

In this paper we will discuss the Blandford-Znajek mechanism in quantitative detail. In Section 2, an overview of the proposed central engine for gamma-ray bursts using Blandford-Znajek process is given. The Blandford-Znajek process and the evolution of the black hole will be discussed in detail in Sections 3 and 4, respectively. The structure of ambient magnetic field surrounding black hole and accretion disk is discussed in Section 5. In Section 6, we will give a rough estimation of the possible angular momentum of the black hole which might emerge as a final compact object during the merging or collapsing processes. The possible constraint from the surrounding accretion disk is discussed in Section 7 and the results are summarized and discussed in Section 8.

## 2. Overview of the proposed central engine for GRB

Two decades ago Blandford and Znajek [2,6] proposed a process (BZ) in which rotational energy of black hole can be efficiently extracted. If there are sufficient charge distributions around the black hole to provide the force-free condition, then the magnetic field lines exert no force and corotate rigidly with the rotating black hole. The induced current loops which pass along the black hole's stretched horizon feel the forces by the magnetic field supported by the environment. Hence, these forces give magnetic braking of the black-hole rotation. The maximum amount of energy which can be extracted out of the black hole without violating the second law of thermodynamics is the rotational energy,  $E_{\text{rot}}$ , which is defined as

$$E_{\text{rot}} = Mc^2 - M_{\text{irr}}c^2, \quad (1)$$

where

$$M_{\text{irr}} = \sqrt{A_H c^4 / 16\pi G^2} \quad (2)$$

$$= \sqrt{(S_H / 4\pi k_B)} M_{\text{planck}}, \quad (3)$$

where  $A_H$  and  $S_H$  are surface area and entropy of a black hole, respectively,<sup>2</sup> and  $k_B$  is Boltzman's constant. The rotational energy of a black hole with angular momentum  $J$  is a fraction of the

---

<sup>2</sup> For a solar mass black hole,  $A_H \sim 10^{42} \text{ cm}^2$ ,  $S_H \sim 10^{77} k_B$ . The Planck mass is  $M_{\text{planck}} = \sqrt{c\hbar/G} = 2.18 \times 10^{-5} \text{ g}$ .

black-hole mass  $M$ ,

$$E_{\text{rot}} = f(\tilde{a})Mc^2, \quad (4)$$

$$f(\tilde{a}) = 1 - \sqrt{\frac{1}{2}[1 + \sqrt{1 - \tilde{a}^2}]}, \quad (5)$$

where  $\tilde{a} = Jc/M^2G$  is the rotation parameter. For a maximally rotating black hole ( $\tilde{a} = 1$ ),  $f = 0.29$ . In BZ, the efficiency of extracting the rotational energy is determined by the ratio between the angular velocities of black hole,  $\Omega_H$ , and magnetic field angular velocity  $\Omega_F$ ,

$$\varepsilon_\Omega = \Omega_F/\Omega_H. \quad (6)$$

The rest of the rotational energy is dissipated into the black hole increasing the entropy or equivalently irreducible mass.<sup>3</sup> The total BZ energy available is

$$E_{\text{BZ}} = 1.8 \times 10^{54} \varepsilon_\Omega f(\tilde{a}) \left( \frac{M}{M_\odot} \right) \text{erg}. \quad (7)$$

For the optimal processes  $\varepsilon_\Omega \sim 0.5$  [6].

Since the energy transport is in the form of Poynting flow in BZ the outgoing energy flux from the black hole is basically  $B^2c$ . An order of magnitude calculation is now in order. There are basically three parameters: Mass of the black hole  $M$  and magnetic field on the horizon  $B$ , which are dimensionful, and angular momentum parameter of the black hole  $\tilde{a}$ . The time scale for the BZ process can be calculated by the ratio of the black-hole mass to the output power from the black-hole surface  $B^2R^2c$ ,

$$\begin{aligned} \tau_{\text{BZ}} &\sim Mc^2/B^2R^2c \sim Mc^5/B^2M^2G^2 \\ &= c^5/B^2MG^2 = 2.7 \times 10^3 (10^{15} \text{ G}/B)^2 (M_\odot/M) \text{ s}, \end{aligned} \quad (8)$$

where  $M_\odot^{-1}(10^{15} \text{ G})^{-2}c^5/G^2 = 2.7 \times 10^3 \text{ s}$  and the radius of horizon,  $R$ , is taken to be  $\sim GM/c^2$ . Also the outgoing Poynting power is

$$P_{\text{BZ}} \sim B^2R^2c = 6.7 \times 10^{50} (B/10^{15} \text{ G})^2 (M/M_\odot)^2 \text{ erg/s}. \quad (9)$$

The fluence of the recently observed GRB971214 [1] corresponds to  $E_\gamma = 10^{53.5} (\Omega_\gamma/4\pi) \text{ erg}$  which is consistent with  $E_{\text{BZ}}$ . That of GRB990123 may be as large as  $E_\gamma = 3.4 \times 10^{54} (\Omega_\gamma/4\pi) \text{ [7]}$ . This suggests that if a strong enough magnetic field ( $\sim 10^{15}$ ) on the black hole can be supported by the surrounding material (accretion disk) the BZ process is a good candidate to provide the powerful energy of the GRB in the observed time interval up to 1000 s, which is comparable to the BZ time scale  $\tau_{\text{BZ}}$ .

In recent years a black hole plus debris torus system (or accretion disk) has been considered to be a plausible structure for the GRB central engine [8]. The presence of the accretion disk is important for the BZ process because it is the supporting system of the strong magnetic field on the black hole, which would disperse without the pressure from the fields anchored in the accretion disk. Recent numerical calculations [9] show that accretion disks formed by various merging processes are found to have large enough pressure such that they can support  $\sim 10^{15} \text{ G}$  assuming

<sup>3</sup> The mass of a nonrotating black hole itself is its irreducible mass.

a value of the disk viscosity parameter  $\alpha \sim 0.1$ , where  $\alpha$  is the usual parameter in scaling the viscosity. This gives a relevant order of magnitude magnetic field on the black hole, which is, however, not considered to be much larger than magnetic field of the inner accretion disk [10,11]. The discovery that soft gamma-ray bursts are magnetars [12] also supports the presence of the strong magnetic fields of  $\sim 10^{15}$  G in nature. The identification by now of three soft gamma repeaters as strong-field pulsars indicates that there may be a large population of such objects: since the pulsar spindown times scale as  $B^{-1}$ , we would expect to observe only 1 magnetar for every 1000 normal pulsars if they were formed at the same rate, and if selection effects were the same for the two populations. We see three magnetars and about 700 normal pulsars, but since they are found in very different ways the selection effects are hard to quantify. It is, nonetheless, clear that magnetars may be formed in our Galaxy at a rate not very different from that of normal pulsars.

The life time of the accretion disk is also very important for the GRB time scale because it supports the magnetic field on the black hole. According to numerical simulations of merging systems which evolve eventually into black hole – accretion disk configuration [9] the viscous life times are 0.1–150 s, which are not inconsistent with the GRB time scale. Also it has been pointed out [8] that a residual cold disk of  $\sim 10^{-3} M_{\odot}$  can support  $10^{15}$  G, even after the major part of the accretion disk has been drained into the black hole or dispersed away.

Recent hydrodynamic simulations of merging neutron stars and black holes [13] show that along the rotation axis of the black hole an almost baryon-free funnel is possible. This can be easily understood since the material above the hole axis has not much angular momentum so that it can be drained quickly, leaving a baryon-free funnel. Hence relativistically expanding jets along the funnel, fueled by Poynting outflow which is collimated along the rotation axis [14], can give rise to gamma-ray bursts effectively. It has been observed that the BZ process is also possible from the disk since the magnetic field on the disk is not much less than that on the black hole [10]. However the energy outflow from the disk is mostly directed vertically from the disk where the baryon loading is supposed to be relatively high enough to keep the baryons from being highly relativistic. Therefore the BZ process from the disk can be considered to have not much to do with gamma-ray burst phenomena. However the BZ from the disk could power an outflow with lower  $\Gamma$ , but nonetheless high energy, which could cause an afterglow at large angles. That would lead to more afterglows being visible than GRBs, because the afterglows are less beamed. The BZ mechanism can also play a very important role in the disk accretion because it carries out angular momentum from the disk. We will discuss this in more detail later.

The structure we are proposing as a central engine of the GRB is a system of black hole – accretion disk (debris of the torus):

The rotating black hole is threaded by a strong magnetic field. Along the baryon-free funnel relativistic jets fueled by Poynting outflow give rise to the GRB. The interaction between disk and black hole is characterized by accretion and magnetic coupling. We consider the BZ process only after the main accretion process is completed, leaving an accretion disk of cold residual material, which can support a strong enough magnetic field.

### 3. Blandford-Znajek process

Consider a half hemisphere (radius  $R$ ) rotating with angular velocity  $\Omega$  and a circle on the surface at fixed  $\theta$  (in the spherical polar coordinate system) across which a surface current  $I$  flows down

from the pole. When the external magnetic field  $B$  is imposed to thread the surface outward normally, the surface current feels a force and the torque due to the Lorentz force exerted by the annular ring of width  $R d\theta$  is

$$\Delta T = - R \sin \theta \times I R \Delta B d\theta \quad (10)$$

$$= -(I/2\pi) B \Delta A_{\text{ann.}} \quad (11)$$

$$= -(I/2\pi) \Delta \Psi , \quad (12)$$

where  $\Delta \Psi$  is the magnetic flux through annular ring extended by  $d\theta$  with surface area  $A_{\text{ann.}}$ . We consider an axially symmetric situation. From this magnetic braking, we can calculate the rotational energy loss rate

$$\Delta P_{\text{rot}} = - \Omega \times \Delta T \quad (13)$$

$$= \Omega (I/2\pi) \Delta \Psi . \quad (14)$$

Blandford and Znajek [2] demonstrated that such a magnetic braking is possible, provided that the external charge distribution can support an electric current with the magnetic field threading the horizon. The original formulation of Blandford and Znajek [2] is summarized in Appendix C. Macdonald and Thorne [15] reformulated the Blandford-Zanek process using a (3 + 1)-dimensional formalism, in which the complicated physics beyond the horizon can be expressed in terms of physical quantities defined on the stretched horizon [6]. It can be shown that the rotational energy loss due to magnetic braking can be obtained from Eq. (14) by simply replacing  $\Omega$  and  $B$  with

$$\Omega \rightarrow \Omega_H, \quad B \rightarrow B_H , \quad (15)$$

where  $H$  denotes the quantity on the stretched horizon [6]:

$$\Delta P_{\text{rot}} = - \Omega_H \times \Delta T \quad (16)$$

$$= \Omega_H (I/2\pi) \Delta \Psi . \quad (17)$$

The current  $I$  induced by the black-hole rotation and the angular velocity  $\Omega_R(\theta)$  of the rigidly rotating magnetic field which is dragged by the rotating black hole can be determined together with the magnetic field by solving Maxwell's equations:

$$F^{\mu\nu}_{;\nu} = 4\pi J^\mu \quad (18)$$

with the force-free condition

$$F^{\mu\nu} J_\nu = 0 , \quad (19)$$

where  $J^\mu$  is a current density vector. The detailed structure of the magnetic field will be discussed in Section 5. In the BZ process, however, the power which can be carried out as Poynting outflow along the field lines is

$$\Delta P_{\text{mag}} = \Omega_F (I/2\pi) \Delta \Psi . \quad (20)$$

Then the rest of the rotational energy is used to increase the entropy (or equivalently irreducible mass) of the black hole. Therefore, the efficiency of the BZ process which is defined as the ratio of

$P_{\text{mag}}$  to  $P_{\text{rot}}$  is

$$\varepsilon_{\Omega} = P_{\text{mag}}/P_{\text{rot}} = \Omega_F/\Omega_H . \quad (21)$$

The ideal efficiency,  $\varepsilon_{\Omega} = 1$ , is meaningless because for  $\Omega_F = \Omega_H$  the Poynting outflow itself is zero as can be seen below in Eq. (25). The optimal power can be obtained at  $\varepsilon_{\Omega} = 1/2$  [15]. Then the rest of the rotational energy is used to increase the entropy(or equivalently irreducible mass) of the black hole. Now consider the loading region far from the black hole, onto which magnetic fields out of the black-hole anchor. In most of the cases we are interested in, the inertia of the loading region can be considered to be so large that the transported angular momentum cannot give rise to any substantial increase of the angular velocity of the loading region. Therefore, the angular velocity of the loading region can be assumed to be zero and the power delivered by the torque, Eq. (12), along the field line is given

$$\Delta P_L = -\Omega_F \Delta T \quad (22)$$

$$= \Omega_F (I/2\pi) \Delta \Psi \quad (23)$$

which is identified as the BZ power for GRB:

$$\Delta P_{\text{BZ}} = \Delta P_{\text{mag}} = \Delta P_L . \quad (24)$$

Since the current  $I$  induced by black-hole rotation is given by [6]

$$I(\theta) = (1/2c)(\Omega_H - \Omega_F(\theta))\tilde{\omega}^2 B_H , \quad (25)$$

the rotational energy loss and the BZ power can be given by

$$\Delta P_{\text{rot}} = [\Omega_H(\Omega_H - \Omega_F)/4\pi]\tilde{\omega}^2 B_H \Delta \Psi , \quad (26)$$

$$\Delta P_{\text{BZ}} = [\Omega_F(\Omega_H - \Omega_F)/4\pi c]\tilde{\omega}^2 B_H \Delta \Psi , \quad (27)$$

where  $\tilde{\omega}$  is a kind of cylindrical radius defined in Boyer-Lindquist coordinates (Appendix A).

To get the total power, Eq. (27) should be integrated from  $\theta \sim 0$  to  $\theta_{\text{BZ}}$ , up to which the magnetic field lines from the black hole anchored on to the loading region for GRB. As a first approximation, we put

$$\theta_{\text{BZ}} \sim \pi/2 , \quad (28)$$

$$B_H(\theta) \sim \langle B_H \rangle \quad (29)$$

and we get an optimal power

$$P_{\text{BZ}} = 1.7 \times 10^{50} \tilde{a}^2 (M/M_{\odot})^2 (\langle B_H \rangle / 10^{15} \text{ G})^2 f(h) \text{ erg/s} . \quad (30)$$

The details are given in Appendix B. The rest of the rotational energy given by

$$P_H = P_{\text{rot}} - P_{\text{BZ}} \quad (31)$$

$$= [(\Omega_H - \Omega_F)/\Omega_F] P_{\text{BZ}} \quad (32)$$

$$= P_{\text{BZ}} \quad \text{for optimal case} \quad (33)$$

is dissipated into the black hole increasing the irreducible part of the black-hole mass (equivalently increasing the entropy of the black hole).

For  $\theta_{BZ} < \theta \leq \pi/2$ , the magnetic field lines from the black hole can be anchored onto elsewhere than the loading region, for example, onto the inner accretion disk. Then the angular velocity of the field line is determined by the angular velocity of the disk  $\Omega_D$ ,  $\Omega_F = \Omega_D$ . The innermost radius of the disk can be considered to be the marginally stable orbit,  $r_{ms}$  [16]. For an orbit rotating in the same direction as the black-hole rotation, we have

$$r_{ms} = (G/c^2)M[3 + Z_2 - \sqrt{(3 - Z_1)(3 + Z_1 + 2Z_2)}] , \quad (34)$$

$$Z_1 = 1 + (1 - \tilde{a}^2)^{1/3}[(1 + \tilde{a})^{1/3} + (1 - \tilde{a})^{1/3}] , \quad (35)$$

$$Z_2 = \sqrt{3\tilde{a}^2 + Z_1^2} . \quad (36)$$

Here  $r_{ms}$  becomes  $GM/c^2$  as  $\tilde{a} \rightarrow 1$  (extreme rotation). In this limit, the angular velocity of disk  $\Omega_D \sim \Omega_H$ . Then one can expect  $\Omega_F \sim \Omega_H$  so that there is no Poynting flow to the disk. However in this case there is no BZ Poynting outflow. For finite  $\tilde{a} < 1$ , we can have either  $\Omega_H > \Omega_D$  or  $\Omega_H < \Omega_D$ , which can be determined by solving Eq. (63). However if we can assume that the power from/to the disk can be much suppressed compared to that of the loading region, we can ignore the magnetic coupling between the black hole and the accretion disk. For example, suppose the portion of black hole-threading magnetic fields anchored onto the disk is somehow suppressed,  $\theta_{BZ} \sim \pi/2$ , then we can assume there is no significant energy and angular momentum feedback into the disk due to magnetic coupling. If it is not so we can discuss only the limiting case.

#### 4. Evolution of a black hole via the Blandford–Znajek process

While the black hole is slowed down and a part of the rotational energy is carried out as Poynting outflow, the rest of the rotational energy increases the entropy of the black hole or its irreducible mass. The increasing rate of the irreducible mass is given by using Eqs. (26) and (27)

$$\frac{dM_{irr}}{dt} = P_{rot} - P_L \quad (37)$$

$$= \int [(\Omega_H - \Omega_F)^2/4\pi c] \tilde{\omega}^2 B_H \Delta \Psi . \quad (38)$$

The irreducible mass eventually becomes the mass of Schwarzschild black hole when it stops rotating. The difference between the initial Kerr black-hole mass,  $M_0$ , and the final Schwarzschild mass,  $M$ , is the energy output from black hole via the Blandford–Znajek process. The evolution of a Kerr black hole is determined by the evolution of its mass and angular momentum given by

$$dMc^2/dt = -P_L , \quad (39)$$

$$dJ/dt = T . \quad (40)$$

Using Eqs. (22) and (40) in the Blandford–Znajek process, the evolution of the mass and the angular momentum are related by

$$dM/dt = \Omega_F dJ/dt \quad (41)$$

For the optimal case,  $\varepsilon_\Omega = 1/2$  ( $\Omega_F = \Omega_H/2$ ), we can obtain an analytic expression for the mass in terms of the angular momentum. With the angular velocity of a black hole expressed in the angular momentum given by in the natural units  $G = c = 1$

$$\Omega_H = J/2M^3(1 + \sqrt{1 - J^2/M^4}) , \quad (42)$$

we get

$$dM/dt = [J/4M^3(1 + \sqrt{1 - J^2/M^4})] dJ/dt , \quad (43)$$

$$2 dM^4/dt = [1/(1 + \sqrt{1 - J^2/M^4})] dJ^2/dt . \quad (44)$$

Introducing a new variable  $H$  defined as

$$H = r_H/M = 1 + \sqrt{1 - (J^2/M^4)} , \quad (45)$$

Eq. (44) can be written

$$2 dM^4/dt = (1/H) dJ^2/dt . \quad (46)$$

Using the identities

$$(d/dt)(J^2/M^4) = 2(1 - H) dH/dt , \quad (47)$$

$$dJ^2/dt = (J^2/M^4) dM^4/dt + M^4 2(1 - H) dH/dt , \quad (48)$$

$$J^2/M^4 = 2H - H^2 , \quad (49)$$

Eq. (46) can be written

$$(1/M) dM/dt = [(1 - H)/2H^2] dH/dt , \quad (50)$$

which can be integrated analytically to give [17]

$$M = M_0 e^{i(H - H_0)/2HH_0} \sqrt{H_0/H} , \quad (51)$$

where  $H_0$  and  $M_0$  represent the initial angular momentum and mass of the black hole. From Eq. (45) one can see that  $H = 1$  for the maximally rotating ( $\tilde{\alpha} = 1$ ) black hole and  $H = 2$  for the non-rotating one.

Consider the black hole initially maximally rotating, which is slowed down by the Blandford-Znajek mechanism in the optimal mode ( $\Omega_F = \Omega_H/2$ ).<sup>4</sup> The final black-hole mass is given by

$$M = (e^{1/4}/\sqrt{2})M_0 \quad (52)$$

$$= 0.91M_0 . \quad (53)$$

We see that 9% of the initial mass or 31% of the rotational energy can be taken out to power the gamma-ray burst from the maximally rotating black hole. The extracted energy is therefore less than a half of the initial rotational energy. This can be easily understood by noticing that the fraction of the rotational energy drops faster because of the decreasing total mass and at the same

<sup>4</sup>The more general discussion has been given by Okamoto [17] where  $\zeta = 1/\varepsilon_D - 1$ .

time increasing irreducible mass in Eq. (1). For  $\tilde{a} = 0.5$ ,  $M = 0.98M_0$  or 2% of the initial mass can be used to power the gamma-ray burst. The time dependence of the power can be obtained from Eq. (D.25) using Eq. (49):

$$P = -dMc^2/dt = (f(h)/4)(2H - H^2)M^2B_H^2G^2/c^3, \quad (54)$$

where  $f(h)$  is defined in Appendix B and  $H = 2/(1 + h^2)$ . The initial rate for the maximally rotating black hole can be obtained by taking  $H = 1$  and  $f = \pi - 2$ . Eq. (54) can be written as

$$(1/M^2)dM/dt = -[(\pi - 2)/4]B_H^2G^2/c^5. \quad (55)$$

Assuming there is no change in magnetic field which is supported by the environment, we get

$$M = M_0/(1 + (\pi - 2)M_0B_H^2G^2t/4c^5). \quad (56)$$

If we extrapolate Eq. (56) until the black hole stops rotating,  $M \rightarrow 0.91M_0$ , we get the time scale of Blandford-Znajek process as

$$\tau = 0.35c^5/M_0B_H^2G^2 \quad (57)$$

$$\sim 10^3(10^{15}\text{ G}/B_H)^2(M_\odot/M)\text{s}. \quad (58)$$

Since there is no considerable change of the weighting factor  $f(h)$  in Eq. (54) from the maximally rotating black hole to the nonrotating black hole ( $f = 2/3$ ), Eq. (58) can be considered as a reasonable estimate of a time scale, which is consistent with the rough estimate given in Eq. (8).

## 5. Magnetic field and force-free plasma

The rotating black hole immersed in the magnetic field induces an electric field around the black hole [6]. The electromagnetic field in the vicinity of rotating black hole immersed in the uniform magnetic field in free space has been obtained [6,18,19] as a solution of the source-free Maxwell equation

$$F_{;\nu}^{\mu\nu} = 0. \quad (59)$$

From the analytic expression (see Appendix B) which gives an asymptotically uniform magnetic field at infinity ( $r \rightarrow \infty$ )

$$\mathbf{B} = B\hat{z}, \quad r \rightarrow \infty, \quad (60)$$

one can see that the radial component of the magnetic field on the horizon,  $B_r^H$ ,

$$\begin{aligned} B_r^H = & [B \cos \theta/(r_H^2 + (a/c)^2 \cos^2 \theta)^2][(r_H^2 - (a/c)^2)(r_H^2 - (a/c)^2 \cos^2 \theta) \\ & + 2(a/c)^2 r_H(r_H - M)(1 + \cos^2 \theta)] \end{aligned} \quad (61)$$

vanishes as the rotation of black hole approaches extreme rotation

$$a \rightarrow GM/c, \quad r_H \rightarrow GM/c^2. \quad (62)$$

This is what has been observed [19,20] as the absence of the magnetic flux across the maximally rotating black hole. One can also see that there is no outward Poynting flow, which means no outflow of energy to infinity. Essentially, it is because of the absence of the toroidal component of the magnetic field due to the vacuum environment which is charge and current free space. In other words, there is no current on the stretched horizon on which the external magnetic field exerts torques to slow the black hole down so as to extract its rotational energy. Therefore, it is necessary to have a magnetosphere with charges and currents to extract the rotational energy of the black hole. The force-free magnetosphere around a rotating black hole has been proposed [2]:

$$F^{\mu\nu}_{;\nu} = 4\pi J^\mu, \quad (63)$$

where  $J^\mu$  is a current density vector and

$$F^{\mu\nu}J_\nu = 0 \quad (64)$$

which is the force-free condition. From Eq. (64), we get the degenerate condition

$$\mathbf{E} \cdot \mathbf{B} = 0. \quad (65)$$

In the case of a rotating black hole in charge-free space one can see using equations in Appendix B

$$\mathbf{E} \cdot \mathbf{B} \neq 0. \quad (66)$$

It should be remarked that since the force-free condition is essential for the BZ process the expulsion of magnetic field on the rotating black hole demonstrated for the charge-free space cannot be directly addressed in the BZ process [20].

To maintain the current flows charged particles (electrons and positrons) should be supplied by a source outside the horizon since the black hole itself cannot provide the outgoing particle from inside the horizon. Blanford and Znajek [2] proposed that the strong electric field induced by the rotating black hole can give rise to a spark gap in which sufficient charged particles are created to supply the currents [21]. Another mechanism provides charged particles around black hole: electron–positron pair creation by neutrino annihilation [22]. Recent numerical studies of merging binary systems [9,23] which result in black-hole-accretion disk configurations show that the power of electron–positron pairs by the annihilation of neutrinos and antineutrinos which are radiated out of the disk is

$$\dot{E}_v \sim 10^{50} \text{ erg/s}. \quad (67)$$

This power is being poured into the space above the black hole for 0.01–1 s. If we divide it by the average neutrino energy  $\langle \epsilon_v \rangle \sim 10 \text{ MeV}$  [23], we get a rough estimate of the numbers of  $e^+e^-$  pairs

$$\dot{N}_{\text{pair}} \sim 10^{56}/\text{s} \quad (68)$$

or the charge producing rate for  $e^+$  (or  $e^-$ ; the sum is always zero) is

$$\dot{Q}_e = e\dot{N}_{\text{pair}} \sim 10^{37} \text{ C/s}. \quad (69)$$

The magnitude of the currents involved in the optimal BZ process can be obtained from Eq. (D.28)

$$I \sim \sqrt{(10^{50} \text{ erg/s})/R_H} \sim 10^{20} \text{ C/s} \quad (70)$$

for a black hole with solar mass threaded by a  $10^{15}$  G magnetic field.  $R_H$  is the surface resistance of the horizon,  $377 \Omega$ . From the comparison of Eq. (69) with (70), the neutrino annihilation process produces orders of magnitude more than enough pairs to keep the necessary currents for the optimal BZ process. The possible effects of neutrino annihilation during the active neutrino cooling of the accretion disk is that the magnetosphere for the BZ process might be disturbed so much that the BZ process is suspended until the burst of  $e^+e^-$  pairs clears out. However, since the pairs are produced with strong directionality along the rotation axis, most of the produced pairs expand along the axis in less than a second [23]. This is reasonable because the electric field along the magnetic field lines is almost negligible and therefore it will take too long for the  $e^+$  to reverse its velocity to make the same current as the  $e^-$ . Also the strong magnetic field keeps them from moving perpendicular to the magnetic field lines so that the separations between them cannot be effective. Therefore the pair contribution to the net current flow can be negligible and it might not disturb the magnetosphere for the BZ process so violently.<sup>5</sup> However, the effects of the neutrino cooling process of the accretion disk can be considered as an additional disturbing burst by the  $\bar{v}v$  driven  $e^+e^-$  burst, which lasts less than few seconds of the BZ burst.

The structure of a force-free magnetosphere can be described by a stream function  $\psi(r, \theta)$  and two functions of  $\psi$ ;  $\Omega_F(\psi)$  and  $B_\phi(\psi)$  (or equivalently  $I = -\frac{1}{2}\alpha\tilde{\omega}B_\phi$ ) [2,24]. Hence  $\psi$  at a point of  $(r, \theta)$  is equal to the total magnetic flux upward through the azimuthal loop at  $(r, \theta)$ . On the other hand from the equation of motion, the stream function is proportional to the toroidal component of the vector potential,  $\psi = 2\pi A_\phi$ . On the horizon, it determines the total magnetic flux,  $\Psi(\theta)$ , through the horizon up to  $\theta$ ,  $\Psi(\theta) = \psi(r_H, \theta)$ . The poloidal and toroidal components of the electromagnetic field are

$$\mathbf{E}^P = -[(\Omega_F - \omega)/2\pi\alpha]\nabla\psi, \quad \mathbf{E}^T = 0, \quad (71)$$

$$\mathbf{B}^P = (\nabla\psi \times \hat{\phi})/2\pi\tilde{\omega}, \quad \mathbf{B}^T = -(2I/\alpha\tilde{\omega})\hat{\phi}, \quad (72)$$

where  $I(r, \theta)$  is the total current downward through the loop at  $(r, \theta)$ .  $\Omega_F(\psi)$  is the angular velocity of the magnetic field relative to absolute space. Blandford and Znajek [2] derived the differential equation (stream equation) for the stream function, which takes the form [24]

$$\nabla \left\{ \frac{\alpha}{\tilde{\omega}^2} \left[ 1 - \frac{(\Omega_F^2 - \omega^2)\tilde{\omega}^2}{\alpha^2} \right] \nabla \psi \right\} + \frac{(\Omega_F - \omega)}{\alpha} \frac{d\Omega_F}{d\psi} (\nabla\psi)^2 + \frac{16\pi^2 I}{\alpha\tilde{\omega}^2} \frac{dI}{d\psi} = 0. \quad (73)$$

It includes two functions of  $\psi$ , which are not known a priori. In solving Eq. (73) the boundaries of the force-free region and the boundary conditions on  $\psi$ ,  $\Omega_F$ , and  $I$  should be specified. There is no known analytical solution of Eq. (73). Blandford and Znajek obtained perturbative solutions for small  $a/M$  using the analytic solutions in charge free space around the non-rotating Schwarzschild black hole. With  $\Omega_F = 0$ ,  $\omega = 0$  and  $I = 0$ , the stream equation reduces to

$$\nabla \{(\alpha/\tilde{\omega}^2)\nabla\psi\} = 0. \quad (74)$$

<sup>5</sup> Although very hard to estimate, a very small fraction of pairs which has very small momentum component along the axis can contribute to BZ currents along the magnetic field lines, which might be sufficient for the force-free configuration.

MacDonald [24] developed a numerical method to obtain solutions with finite  $a/M$ , in which the solutions of Eq. (74) with appropriate boundary conditions are spun up numerically. It has been found that the poloidal field structure does not change greatly as the holes and fields are spun up. This result implies that the flux of the poloidal magnetic field threading black hole does not decrease greatly as  $a \rightarrow M$  in contrast to the rotating black hole in free space.

One of the interesting solutions is the poloidal field structure generated by the paraboloidal magnetic field solution of Eq. (74):

$$\psi = (\psi_0/4 \ln 2) \{ (r - 2)(1 - \cos \theta) + 2[2 \ln 2 - (1 + \cos \theta) \ln(1 + \cos \theta)] \}, \quad (75)$$

where  $\psi_0$  is the total flux threading the hole. Since this solution extends not only onto the horizon but also onto the equatorial plane where the accretion disk is supposed to be placed, the boundary conditions on the disk should be satisfied by the spun up solutions generated by Eq. (75). The boundary conditions depend strongly on the accretion disk model, which will be discussed in the next section. The presence of the accretion disk may be considered to be the main source of the difficulty in obtaining a solution with these complicated boundary conditions. However, the solutions with proper boundary conditions provide us a way as to how we can infer the magnetic field from those developed in the disk. It has been demonstrated [10,11] that the strength of the magnetic field on the black hole is not much stronger than those on the inner disk.

## 6. Rotation of the black hole

It has been suggested that merging compact binary systems [8] or hypernovae [4] may result in a disk with rapidly rotating black hole at the center. In this section we will demonstrate how this is possible using semiquantitative arguments. The basic idea is that a substantial part of the orbital angular momentum of the binary system or the spin angular momentum of the progenitor of hypernova can be imparted onto the black hole.

Consider first the BH-NS merger. The typical distance for the merging system in this case is the tidal radius,  $R_t$ , at which the neutron star (radius  $R_{\text{NS}}$ ) fills the Roche lobe [25]:

$$R_t = (R_{\text{NS}}/0.46)((1+q)/q)^{1/3}, \quad (76)$$

where

$$q = M_{\text{NS}}/M_{\text{BH}}. \quad (77)$$

If the tidal radius is greater than the last stable orbit radius,<sup>6</sup>  $R_{ls}$ , we can calculate the Keplerian orbital angular velocity

$$\Omega_K = \sqrt{GM/R_t^3}, \quad (78)$$

where

$$M = M_{\text{BH}} + M_{\text{NS}}. \quad (79)$$

---

<sup>6</sup> This leads to the condition for the black-hole mass [26]:  $M_{\text{BH}} < M_{\text{NS}}[(0.4R_{\text{NS}}c^2/GM_{\text{NS}})^{3/2} - 1] \sim 2.3M_\odot$ .

Then the orbital angular momentum of the binary system can be written

$$J_{\text{binary}} = \mu R_t^2 \Omega_K \quad (80)$$

$$= M_{\text{BH}} M_{\text{NS}} \sqrt{G R_t / M}, \quad (81)$$

where  $\mu$  is the reduced mass,  $\mu = M_{\text{BH}} M_{\text{NS}} / M$ . Using Eq. (76)

$$J_{\text{binary}} = 1.47 M_{\text{BH}} M_{\text{NS}} \sqrt{(G R_{\text{NS}} / M)(M / M_{\text{NS}})^{1/6}}. \quad (82)$$

During the collapse, a part of angular momentum is carried off by gravitational waves (or possibly in the later stage by neutrino cooling) and a part of the total mass explodes away or remains in the torus around the black hole. Assuming that a fraction ( $x$ ) of the orbital angular momentum goes into the black hole, which keeps a fraction ( $y$ ) of the total mass  $M$ ,<sup>7</sup> we can calculate the angular momentum parameter of the black hole,  $\tilde{a}$ :

$$J_{\text{BH}} = \tilde{a} (y M)^2 G/c = x y J_{\text{binary}}, \quad (83)$$

$$\tilde{a} = 1.47 \frac{x}{y} \frac{M_{\text{BH}} M_{\text{NS}}}{M^2} \sqrt{\frac{R_{\text{NS}} c^2}{GM}} \left( \frac{M}{M_{\text{NS}}} \right)^{1/6}. \quad (84)$$

For  $M_{\text{BH}} = 2.5 M_{\odot}$ ,  $M_{\text{NS}} = 1.5 M_{\odot}$ ,  $R_t = 10^6 \text{ cm}$ ,

$$\tilde{a} = 0.53 x/y \quad (85)$$

which is a quite reasonable value for an efficient Blandford-Znajek process.

For the NS-NS merger, the radius of physical contact ( $2R_{\text{NS}}$ ) is smaller than the tidal radius,

$$R_t = 2.46 R_{\text{NS}}^{2/3} = 3 R_{\text{NS}}. \quad (86)$$

Hence it is reasonable to consider the orbital angular momentum at the tidal radius. Following the same procedure as in the BH-NS merger, we get

$$\begin{aligned} \tilde{a} &= 0.31(x/y) \sqrt{\frac{R_{\text{NS}} c^2}{GM_{\text{NS}}}} \\ &= 0.67 x/y, \end{aligned} \quad (87)$$

where we replaced  $M_{\text{BH}}$  by  $M_{\text{NS}}$  in Eq. (84).

In the hypernova model [4,27], a massive rapidly spinning progenitor ( $M_0 \sim 40 M_{\odot}$ ) is supposed to collapse into a rapidly rotating black hole. If we assume a critically rotating progenitor ( $\Omega = \Omega_K$  at the surface,  $R_0$ ) and solid body rotation throughout, the angular momentum of the core which eventually collapses into the black hole is

$$J_{\text{core}} = \frac{2}{3} M_{\text{core}} R_{\text{core}}^2 \Omega_K (R_0), \quad (88)$$

where the moment of inertia of the core is assumed to be that of a uniformly distributed spherical object,<sup>8</sup>  $I = \frac{2}{5} M_{\text{core}} R_{\text{core}}^2$ . Using

$$\Omega_K = \sqrt{GM_0/R_0^3}, \quad (89)$$

<sup>7</sup>  $x$  is a fraction of the specific angular momentum.

<sup>8</sup> Since in general  $I = k^2 M_{\text{core}} R_{\text{core}}^2$ , and  $k^2 \ll 1$  for radiative stars [28], this is an upper limit to the true angular momentum.

we get

$$J_{\text{core}} = \frac{2}{3} M_{\text{core}} R_{\text{core}}^2 \sqrt{GM_0/R_0^3}. \quad (90)$$

As in the previous merger case, a part of the core angular momentum ( $x$ ) and the core mass ( $y$ ) collapses into the black hole ( $M_{\text{BH}} = yM_{\text{core}}$ ). Then we get

$$J_{\text{BH}} = xy J_{\text{core}} \quad (91)$$

$$= xy \frac{2}{3} M_{\text{core}} R_{\text{core}}^2 \sqrt{GM_0/R_0^3} \quad (92)$$

$$= \tilde{a}(yM_{\text{core}})^2 G/c. \quad (93)$$

Then the black-hole angular momentum parameter is given by

$$\tilde{a} = \frac{2}{3}(x/y)(R_{\text{core}}^2 c/GM_{\text{core}}) \sqrt{GM_0/R_0^3} \quad (94)$$

$$= \frac{2}{3}(x/y)(R_{\text{core}}/R_0)^{3/2} \sqrt{R_{\text{core}} c^2/GM_{\text{core}}} \sqrt{M_0/M_{\text{core}}}. \quad (95)$$

With

$$\begin{aligned} R_0 &\sim 10^5 \text{ km}, \quad R_{\text{core}} \sim 10^3 \text{ km}, \\ M_0 &\sim 15M_\odot, \quad M_{\text{core}} \sim 2M_\odot (\sim 3 \text{ km}) \end{aligned} \quad (96)$$

we get

$$\tilde{a} \sim 0.1 x/y. \quad (97)$$

This depends strongly on the numbers taken in Eq. (96) and on how the core angular momentum is determined. The specific angular momentum of the core can be calculated using Eqs. (90) and (96) as

$$a_{\text{core}} = J_{\text{core}}/M_{\text{core}} = \frac{2}{3} R_{\text{core}}^2 \sqrt{GM_0/R_0^3} \quad (98)$$

$$= 1.8 \times 10^{14} \text{ cm}^2/\text{s} \quad (99)$$

which is much smaller than that from the numerical simulation of collapsars [9]:  $a_{\text{core}} \sim 10^{16} \text{ cm}^2/\text{s}$ . Of course, if the core angular momentum is not completely redistributed during the precollapse evolution the core is likely to be spinning faster than  $\Omega_K$ . The maximum possible value comes from replacing  $\Omega_K$  determined at the progenitor radius in Eq. (90) by the one determined at the core radius

$$\Omega_K^{\text{core}} = \sqrt{GM_{\text{core}}/R_{\text{core}}^3} \quad (100)$$

$$= (1.3 \times 10^2) \Omega_K \quad (101)$$

and we get

$$a_{\text{core}} = 2.34 \times 10^{16} \text{ cm}^2/\text{s}. \quad (102)$$

Then the angular momentum parameter can be given by

$$\tilde{a} = (x/y) a_{\text{core}}/a_{\text{max}} = 2.3 x/y. \quad (103)$$

In summary, it is very plausible to have a rapidly rotating black hole ( $\tilde{a} > 0.1$ ) as a resulting object in the center in the merging systems and also in hypernovae of large angular momentum progenitors, but a precise value of  $\tilde{a}$  will be difficult to calculate.

## 7. Magnetized accretion disks

A black hole by itself cannot keep magnetic fields on it for a long time. Magnetic fields diffuse away in a short time  $\sim R_H/c$  [6,10]. The most plausible environments which can support a magnetic field threading the black hole are accretion disks surrounding the black hole. There are two issues about accretion disks that we need to consider in order to decide whether the Blandford-Znajek process is a viable power source for gamma-ray bursts: The life time should be long enough to extract the bulk of the black-hole spin energy and also the magnetic field on the disk should be strong enough to power the gamma-ray bursts from the spinning black hole. A strong magnetic field on the inner part of the accretion disk is necessary to keep the magnetic pressure comparable to that of magnetic fields on the black hole [6]. It has also been demonstrated [11] from the axisymmetric solutions discussed in Section 4. Since the magnetic field on the disk affects the angular momentum transfer of the accretion disk via magnetic braking [29] and/or magnetic viscosity, the accretion process is not independent of the magnetic field on the disk. The magnetic field on the disk should not be larger than the value from the equipartition argument:

$$B_{\text{eq}}^2/8\pi \sim P_{\text{disk}} \quad (104)$$

and  $B_{\text{eq}}$  can be considered as an upper limit to the magnetic field which can be supported by the accretion disk. Recent numerical calculations on the hyper-accreting black hole by Popham et al. [9] show

$$P_{\text{disk}} \sim 10^{30} \text{ erg/cm}^3, \quad (105)$$

$$B_{\text{eq}} \sim 5 \times 10^{15} \text{ G}, \quad (106)$$

which implies the accretion disk of the hyper-accreting black hole may support a magnetic field strong enough for the gamma-ray burst. However, it depends on the detailed mechanism how strong a magnetic field can be built up on the disk. One of the approaches is that the magnetic field is evolved magnetohydrodynamically during the accreting process, which depends on the magnetic viscosity.

The magnetic viscosity,  $\nu^{\text{mag}}$ , is defined [30] as

$$B_r B_\phi / 4\pi = - \nu^{\text{mag}} (r d\Omega_{\text{disk}}/dr) \rho, \quad (107)$$

which can be parametrized by the viscosity parameter,  $\alpha$ , in terms of dimensionful quantities of the disk as [32,33],

$$\nu^{\text{mag}} = \alpha^{\text{mag}} c_s H, \quad (108)$$

where  $c_s$  is the sound velocity of the disk ( $c_s = \sqrt{\gamma P_{\text{disk}}/\rho}$ ) and  $H$  is a half of the disk thickness. Using hydrostatic equilibrium perpendicular to the disk plane,

$$H = \sqrt{P_{\text{disk}}/\rho} / \Omega_{\text{disk}} \quad (109)$$

$$= c_s / \Omega_{\text{disk}} \sqrt{\gamma}, \quad (110)$$

we get

$$v^{\text{mag}} = \alpha^{\text{mag}} c_s^2 / \Omega_{\text{disk}} \sqrt{\gamma}. \quad (111)$$

For a Keplerian orbit

$$r d\Omega_{\text{disk}}/dr = -\frac{3}{2}\Omega_{\text{disk}}, \quad (112)$$

Eq. (107) can be written as

$$B_r^{\text{dyn}} B_\phi^{\text{dyn}} / 4\pi = \frac{3}{2} \alpha^{\text{mag}} P_{\text{disk}} \sqrt{\gamma}. \quad (113)$$

We can see that the Maxwell stress of the magnetic field which has been built up by the accretion dynamo is also proportional to the disk pressure but for a different reason from that of the equipartition argument, (104). Numerical estimations of the viscosity parameter,  $\alpha^{\text{mag}}$ , obtained for various boundary conditions range from 0.001 to 0.005 [30]. This means that the magnetic pressure is only a small fraction of the disk pressure [10]. Using the estimation of disk pressure by Popham et al., we can estimate the dynamically generated magnetic field by the accretion:

$$B^{\text{dyn}} \sim 10^{13} \text{ G} \quad (114)$$

which may not be strong enough for a gamma-ray bursts powered by the Blandford–Znajek process.

However, an accretion dynamo might not be the only process responsible for the magnetic fields on the disk. We know from the recent observations that there are a number of pulsars, magnetars, which are believed to have a strong magnetic field of  $\sim 10^{15}$  G. Although the origin [31] of such strong magnetic fields is not well known at the moment, one may consider the case that a debris torus or disk around the black hole which was formed by the disruption of a neutron star retain the high-ordered field of that neutron star.

Now consider the axisymmetric solution around a disk with a force-free magnetosphere which has been discussed by Blandford [29]. Here we adopt cylindrical coordinates,  $(r, \phi, z)$ , where the  $z$ -direction is perpendicular to the disk. The sum of current flows into the disk up to radius  $r$  defines the surface current density,  $J_r$ , which is proportional to the poloidal component of the magnetic field on the disk,  $B_\phi$ :

$$4\pi J_r / 2 = -B_\phi, \quad (115)$$

where 2 in the denominator comes from the fact that the radial current density  $J_r$  includes the currents into the disk both from above and below. The torque exerted by the annular ring with width  $\Delta r$  of the disk due to the Lorentz force is given by

$$\Delta T = -r 2\pi r J_r B_z \Delta r \quad (116)$$

$$= 2r(B_\phi B_z / 4\pi) \Delta S, \quad \Delta S = 2\pi r \Delta r. \quad (117)$$

For the steady-state accretion disk with surface density,  $\Sigma$ , the angular momentum conservation can be written by

$$\Sigma v_r [\partial(r^2 \Omega) / \partial r] 2\pi r \Delta r = \Delta T + (\partial G / \partial r) \Delta r, \quad (118)$$

$$\dot{M} [\partial(r^2 \Omega) / \partial r] \Delta r = \Delta T + (\partial G / \partial r) \Delta r, \quad (119)$$

where the torque due to the shear force of differential rotation,  $G$  [33], is given by

$$G = 2\pi r\nu\Sigma(r(\partial\Omega/\partial r))r. \quad (120)$$

To see the effect of magnetic field, we consider only the magnetic viscosity for the moment. Using Eq. (107), we get

$$G^{\text{mag}} = 2\pi r^2(\Sigma/\rho)B_\phi B_r/4\pi \quad (121)$$

$$= 4\pi r^2 H B_\phi B_r/4\pi \quad (122)$$

$$= 4\pi r^2(c_s/\Omega_{\text{disk}})B_\phi B_r/4\pi, \quad (123)$$

where  $2H\rho = \Sigma$ . Assuming a simple power dependence of  $H B_\phi B_r/4\pi \propto r^n$ ,

$$\partial G^{\text{mag}}/\partial r = 4(2+n)\pi r H B_\phi B_r/4\pi \quad (124)$$

and we get from Eq. (119)

$$\dot{M} \partial(r^2\Omega_{\text{disk}})/\partial r = r^2 B_\phi B_z [1 + (2+n)(H/r)B_r/B_z]. \quad (125)$$

Since  $H \ll r$ , the accretion rate is determined by the magnetic braking as far as  $B$ , is not much larger than  $B_z$  and the second term on the right-hand side of the above equation can be neglected. For  $r \gg GM_{\text{BH}}/c^2$  where the disk angular velocity can be approximated by a Keplerian velocity,  $\partial(r^2\Omega)/\partial r = r\Omega_{\text{disk}}/2$ ,

$$\dot{M} = 2rB_\phi B_z/\Omega_{\text{disk}}. \quad (126)$$

Using the axisymmetric solution [29],

$$B_\phi = 2r\Omega_{\text{disk}}B_z/c, \quad (127)$$

we get

$$\dot{M} = 4r^2 B_z^2/c. \quad (128)$$

Since for a steady accretion  $\dot{M}$  is independent of  $r$ ,  $B_z \sim 1/r$  at large distance. For a numerical estimation, we take  $r = 10^6$  cm and  $B_z = 10^{14}$  G. Then we get

$$\dot{M} = 6 \times 10^{-4} (B_z/10^{14} \text{ G})^2 M_\odot \text{ s}^{-1} \quad (129)$$

which is much larger than Eddington luminosity,  $\dot{M}_{\text{Ed}} \sim 10^{-16} (M_{\text{BH}}/M_\odot) M_\odot \text{ s}^{-1}$ . The total accretion during the gamma-ray burst period is at most  $10^{-1} M_\odot$ , which may not change the discussions on Blandford-Znajek process in the previous sections. Hence black holes surrounded by the hyper-accretion disks might be a good candidate for the central engine of the gamma-ray bursts. The magnetic field on the disk extracts not only the angular momentum as described above but also a substantial energy out of disk as it does on the black hole. The power of the disk magnetic braking can be calculated using Eq. (117)

$$\Delta P_{\text{disk}} = \Omega_{\text{disk}} \Delta T \quad (130)$$

$$= 2(B_\phi B_z/4\pi)(r\Omega_{\text{disk}})\Delta S. \quad (131)$$

Using Eq. (127) with Keplerian angular velocity of the disk, we get

$$\Delta P_{\text{disk}} = 4c(B_z^2/4\pi)(GM_{\text{BH}}/rc^2)\Delta S, \quad (132)$$

$$dP_{\text{disk}} = (1/\pi c)B_z^2 r^2 (GM_{\text{BH}}/r^3) dS, \quad (133)$$

then setting  $dS = 2\pi r dr$  and using the steady-state relation  $B_z^2 r^2 = B_z(r_{\text{in}})^2 r_{\text{in}}^2$ , we can obtain the total power by integrating Eq. (133) from  $r_{\text{in}}$  to infinity,

$$P_{\text{disk}} = 2B_z(r_{\text{in}})^2 r_{\text{in}} GM_{\text{BH}}/c, \quad (134)$$

where  $r_{\text{in}}$  is the distance from the black hole to the inner edge of the accretion disk. Compared to the BZ power from the spinning black hole, Eq. (D.24), the ratio can be given by

$$\frac{P_{\text{disk}}}{P_{\text{BZ}}} = 8[r_{\text{in}}c^2/GMf(h)\tilde{a}^2](B_z(r_{\text{in}})/B_H)^2. \quad (135)$$

The current conservation condition, namely that the total current flows onto the black hole should go into the inner edge ( $r_{\text{in}}$ ) of the accretion disk, implies

$$2MB_{\phi}^H(\theta = \pi/2) = \tilde{\omega}(r_{\text{in}})B_{\phi}^{\text{disk}}(r_{\text{in}}). \quad (136)$$

Since the cylindrical radius in Kerr geometry  $\tilde{\omega}(r_{\text{in}}) > 2M$ , we can see that  $B_{\phi}^H$  is larger than  $B_{\phi}^{\text{disk}}$ . From the boundary conditions on the horizon [6] in the optimal case,

$$B_{\phi}^H = -\Omega_H MB_H \quad (137)$$

and on the accretion disk [29] with angular velocity  $\Omega_D$ ,

$$B_{\phi}^{\text{disk}} = -2\Omega_D r B_z^{\text{disk}}, \quad (138)$$

we get

$$B_z(r_{\text{in}})/B_H = \sqrt{(GM/r_{\text{in}}c^2)(\tilde{a}/2)GM/r_Hc^2} < 1, \quad (139)$$

and the power of the disk magnetic braking  $P_{\text{disk}} = (2/f(\tilde{a}))(GM/r_Hc^2)^2 P_{\text{BZ}}$ . It shows that the magnetic field on the horizon cannot be smaller than that on the inner edge of the accretion disk [34] and the disk power need not be substantially larger than that from the black hole [10,11].

## 8. Conclusion

We have evaluated the power and energy that can be extracted from a rotating black hole immersed in a magnetic field, the Blandford-Znajek effect. We improve on earlier calculations to find that the power from a black hole of given mass immersed in an external field is 10 times greater than previously thought. The amount of energy that can be extracted from a black hole in this way is limited by the fact that only 29% of the rest mass of a black hole can be in rotational energy, and that the maximum efficiency with which energy can be extracted from the hole via the Blandford-Znajek effect is 31%. The net amount of energy that can be extracted is therefore 9% of the rest energy of the black hole. We consider various scenarios for the formation of

rotating black holes in gamma-ray burst engines, and while the resulting angular momenta are quite uncertain in some cases, it seems that the required values of the rotation parameter,  $\tilde{a} \gtrsim 0.5$ , are achievable.

The rate at which angular momentum is extracted depends on the magnetic field applied to the hole. A field of  $10^{15}$  G will extract the energy in less than 1000 s, so time scales typical of gamma-ray bursts can be obtained. Since the black hole cannot carry a field, there must be an ambient gas in which the field is anchored that drives the Poynting flux from the black hole. The most obvious place for it is the accretion disk or debris torus surrounding the black hole just after it formed.

It has been argued that a field turbulently generated in the disk would not give a strong Blandford-Znajek flux [10], because the disk would dominate the total Poynting output, and no more than the disk's binding energy could be extracted. We show explicitly that a field in the disk could be much greater, for example if it is derived from the large, ordered field of a neutron star that was disrupted. The field distribution proposed by Blandford [29] for such a case would allow the field on the hole to be much greater than on the disk, such that the Poynting flow would not be dominated by the disk and not subject to any obvious limits imposed by the disk.

We also note that a Poynting flow may provide an alternative way of providing a very large magnetic field for the shocked material that radiates the afterglow and the gamma-ray burst itself: the standard assumption is that the required high fields grow turbulently in the shocked gas, up to near-equipartition values. But the field in the Poynting flow only decreases as  $1/r$ , so if it is  $10^{15}$  G at  $r = 10^5$  cm, it could be as high as  $10^4$  G at the deceleration radius ( $10^{16}$  cm), ample to cause an energetic gamma-ray burst.

## Acknowledgements

We would like to thank Roger Blandford, Sterl Phinney and Kip Thorne for guidance and useful discussions. HKL is very grateful to the Nuclear Theory Group, SUNY at Stony Brook for the hospitality during his one year stay, where the most of collaborated works on this paper has been done. This work is supported by the U.S. Department of Energy under Grant No. DE-FG02-88ER40388. HKL is also supported in part by KOSEF-985-0200-001-2 and BSRI-98-2441.

## Appendix A. Rotating black hole

A rotating black hole is defined as its angular momentum ( $J$ ) and mass ( $M$ ). The angular momentum is measured by the non-Newtonian gravitational effect (gravitomagnetic effect) far from the black hole and the mass is measured by the Newtonian gravitational field far from the black hole. The solution of Einstein's equation [35] defines the metric around the rotating black hole with specific angular momentum  $a = J/M$ . Using the Boyer-Lindquist coordinates [36], in the natural unit  $G = c = 1$  the metric can be written as

$$ds^2 = -\alpha^2 dt^2 + g_{ij}(dx^i + \beta^i dt)(dx^j + \beta^j dt). \quad (\text{A.1})$$

$\alpha$  and  $\beta$  are lapse and shift functions [6] defined, respectively, as

$$\alpha = \rho \sqrt{\Delta}/\Sigma , \quad (\text{A.2})$$

$$\beta^\phi = -2Mr/\Sigma^2 , \quad (\text{A.3})$$

$$\beta^t = \beta^r = 0 , \quad (\text{A.4})$$

where

$$\Delta = r^2 + a^2 - 2Mr , \quad (\text{A.5})$$

$$\rho^2 = r^2 + a^2 \cos^2 \theta , \quad (\text{A.6})$$

$$\Sigma^2 = (r^2 + a^2)^2 - a^2 \Delta \sin^2 \theta . \quad (\text{A.7})$$

The metric tensor  $g_{ij}$  is given by

$$g_{rr} = \rho^2/\Delta, \quad g_{\theta\theta} = \rho^2, \quad g_{\phi\phi} = \tilde{\omega}^2 , \quad (\text{A.8})$$

where

$$\tilde{\omega} = (\Sigma/\rho)\sin \theta . \quad (\text{A.9})$$

The volume and area elements  $dV$  and  $dS$  are defined in the standard way [6] by

$$dV = \sqrt{\det(g_{ij})} dr d\theta d\phi \quad (\text{A.10})$$

$$= (\rho^2 \tilde{\omega} / \sqrt{\Delta}) dr d\theta d\phi , \quad (\text{A.11})$$

$$dS_i = \sqrt{\det(g_{ij})} \epsilon_{ijk} (\partial x^j / \partial b) (\partial x^k / \partial c) db dc . \quad (\text{A.12})$$

The circumference of a circle around the axis of symmetry is  $2\pi\sqrt{g_{\phi\phi}} = 2\pi\tilde{\omega}$ . The horizon,  $r_H$ , is defined as a larger root of  $\Delta_{r=r_H} = 0$ ,

$$r_H = M + (M^2 - a^2)^{1/2} , \quad (\text{A.13})$$

$$2Mr_H = r_H^2 + a^2 . \quad (\text{A.14})$$

The surface area of the horizon,  $A_H$ , is given by

$$A_H = 4\pi(r_H^2 + a^2) \quad (\text{A.15})$$

and the entropy and the irreducible mass of the black hole are given, respectively, by

$$S_H = (k_B/4\hbar)A_H \quad (\text{A.16})$$

$$= (\pi k_B/\hbar)(r_H^2 + a^2) = (2\pi k_B/\hbar)Mr_H , \quad (\text{A.17})$$

$$M_{\text{irr}} = \sqrt{A_H/16\pi} = \frac{1}{2}\sqrt{r_H^2 + a^2} = \sqrt{S_H/4\pi} . \quad (\text{A.18})$$

The mass of the black hole can be rewritten in terms of  $J$ ,  $M_{\text{irr}}$ , and  $S_H$ ,

$$M = \sqrt{S_H/4\pi + J^2/S_H} = \sqrt{M_{\text{irr}}^2 + J^2/4M_{\text{irr}}^2}. \quad (\text{A.19})$$

The angular velocity of the black hole is given by

$$\Omega_H = -\beta_H^\phi = a/2Mr_H = J/2M^2r_H \quad (\text{A.20})$$

$$= J/2M^3(1 + \sqrt{1 - J^2/M^4}) = (1/2M)\tilde{a}/(1 + \sqrt{1 - \tilde{a}^2}), \quad (\text{A.21})$$

where  $\tilde{a}$  is the angular momentum parameter defined as  $\tilde{a} = J/M^2$ .

## Appendix B. Electromagnetic fields in vacuum around rotating black hole

For a rotating black hole immersed in an asymptotically uniform magnetic field

$$\mathbf{B} = B\hat{z}, \quad r \rightarrow \infty, \quad (\text{B.1})$$

the poloidal components of electric and magnetic fields can be written analytically [6]

$$B_r = (B/2\Sigma \sin \theta) \partial X / \partial \theta, \quad (\text{B.2})$$

$$B_\theta = -(B\sqrt{4/2\Sigma \sin \theta}) \partial X / \partial r, \quad (\text{B.3})$$

$$E_r = \frac{-Ba\Sigma}{\rho^2} \left[ \frac{\partial \alpha^2}{\partial r} + \frac{M \sin^2 \theta}{\rho^2} (\Sigma^2 - 4a^2Mr) \frac{\partial}{\partial r} \left( \frac{r}{\Sigma^2} \right) \right], \quad (\text{B.4})$$

$$E_\theta = \frac{-Ba\Sigma}{\rho^2 \sqrt{4}} \left[ \frac{\partial \alpha^2}{\partial \theta} + \frac{\sin^2 \theta}{\rho^2} (\Sigma^2 - 4a^2Mr) \frac{\partial}{\partial \theta} \left( \frac{1}{\Sigma^2} \right) \right], \quad (\text{B.5})$$

where  $X = (\sin^2 \theta/\rho^2)(\Sigma^2 - 4a^2Mr)$  and  $a$  is the specific angular momentum of the black hole,  $J = aM$ . Here we adopt the natural units  $G = c = 1$ .

The radial components of magnetic field on the horizon,  $B_r^H$ , can be written explicitly by

$$B_r^H = (B \cos \theta / (r_H^2 + (a/c)^2 \cos^2 \theta)^2), \\ \times [(r_H^2 - (a/c)^2)(r_H^2 - (a/c)^2 \cos^2 \theta) + 2(a/c)^2 r_H(r_H - M)(1 + \cos^2 \theta)], \quad (\text{B.6})$$

$$E_r^H = -(aB/2r_H c(r_H^2 + (a/c)^2 \cos^2 \theta)^2) \quad (\text{B.7})$$

$$\times (r_H^2 - (a/c)^2)[(3r_H^2 + (a/c)^2(1 + \cos^2 \theta))\cos^2 \theta - r_H^2]. \quad (\text{B.8})$$

It is clear that  $B_r^H$  vanishes as the rotation of black hole approaches extreme rotation

$$a \rightarrow GM/c, \quad r_H \rightarrow GM/c^2. \quad (\text{B.9})$$

This is what has been observed [19,20] as the absence of the magnetic flux across the maximally rotating black hole. There are no  $\phi$  components of electric and magnetic fields in Eq. (B.5), and the  $\theta$  components vanish as  $\sqrt{4}$  as the horizon is approached. The main reason for these vanishing tangential components is that there are no currents outside the horizon in charge-free space. In the

absence of in/out currents there is no toroidal magnetic field on the horizon. This also requires a vanishing tangential component of electric field on the horizon, which induces a charge separation on the horizon [6]. Therefore, there is no current on which the external magnetic field exerts torques to slow the black hole down so as to extract its rotational energy. One can verify from Eqs. (B.2) to (B.5) that there is no outward component of the Poynting vector (essentially it is due to the absence of the  $\phi$  component of the magnetic field), which means no outflow of energy to infinity.

### Appendix C. Axial-symmetric force-free magnetosphere

In this appendix, the original formulation [2] of the Blandford–Znajek process is summarized. Consider the stationary and axial symmetric situation where all the partial derivatives of the physical quantities with respect to time and azimuthal angle  $\phi$  are vanishing. The electromagnetic field tensor is given by the vector potential  $A_\mu$  as

$$F_{\mu\nu} = A_{\nu,\mu} - A_{\mu,\nu} \quad (\text{C.1})$$

and the magnetic field is given by

$$B_r = F_{\theta\phi} = A_{\phi,\theta}, \quad (\text{C.2})$$

$$B_\theta = F_{\phi r} = -A_{\phi,r}, \quad (\text{C.3})$$

$$B_\phi = F_{r\theta} = A_{\theta,r} - A_{r,\theta}. \quad (\text{C.4})$$

The magnetic flux,  $\Psi$ , through a circuit encircling  $\phi = 0 \rightarrow 2\pi$ ,

$$\Psi = \oint A_\phi d\phi = 2\pi A_\phi. \quad (\text{C.5})$$

It defines a magnetic surface on which  $A_\phi(r, \theta)$  is constant and therefore is characterized by the magnetic flux  $\Psi$  contained inside it. Magnetic field lines are spiraling on the surface and no magnetic field lines can cross the magnetic surface.

#### C.1. Force-free magnetosphere

The force-free condition for the magnetosphere with the current density  $J^\mu$

$$F_{\mu\nu} J^\mu = 0 \quad (\text{C.6})$$

can be written by

$$A_{0,r} J^r + A_{0,\theta} J^\theta = 0, \quad (\text{C.7})$$

$$A_{\phi,r} J^r + A_{\phi,\theta} J^\theta = 0, \quad (\text{C.8})$$

$$-A_{0,r} J^0 - A_{\phi,r} J^\phi = B_\phi J^\theta, \quad (\text{C.9})$$

$$A_{0,\theta} J^0 + A_{\phi,\theta} J^\phi = B_\phi J^r. \quad (\text{C.10})$$

From Eqs. (C.7) and (C.8) we get

$$A_{0,r}/A_{0,\theta} = A_{\phi,r}/A_{\phi,\theta} = - J^\theta/J^r . \quad (\text{C.11})$$

We can see that  $A_0$  is also constant along the magnetic field lines and the electric field is always perpendicular to the magnetic surface. We can define a function  $\omega(r, \theta)$ ,

$$A_{0,r}/A_{\phi,r} = A_{0,\theta}/A_{\phi,\theta} = - \omega , \quad (\text{C.12})$$

$$dA_0 = - \omega dA_\phi , \quad (\text{C.13})$$

which is also constant along the magnetic surface.  $\omega$  can be identified as an electromagnetic angular velocity (an angular velocity of magnetic field line on a magnetic surface). We can also suppose a function  $\mu(r, \theta)$  satisfying

$$\begin{aligned} 4\pi J^r &= (\mu/\sqrt{g})A_{\phi,\theta} , \\ 4\pi J^\theta &= (\mu/\sqrt{g})A_{\phi,r} , \end{aligned} \quad (\text{C.14})$$

where  $\sqrt{g} = \sqrt{-\det(g_{\mu\nu})} = \rho^2 \sin \theta$ . The current conservation

$$(1/\sqrt{g})(\sqrt{g}J^\mu)_{,\mu} = 0 \quad (\text{C.15})$$

leads to

$$(\mu A_{\phi,\theta})_r = (\mu A_{\phi,r})_\theta \quad (\text{C.16})$$

which implies that  $\mu$  is also constant on the magnetic surface.

From Eq. (C.9) (or equivalently from Eq. (C.10)), we get

$$J^\phi = \omega J^0 + (\mu/4\pi\sqrt{g})B_\phi . \quad (\text{C.17})$$

Together with Eq. (C.14), we now have an expression of the current in terms of the magnetic field and charge density with yet undetermined function  $\omega$  and  $\mu$ . The outward current  $I$  can be calculated from the  $r$  and  $\theta$  components of the current. Since the outward current is proportional to the magnetic field, the current between the magnetic surface,  $dI$ , can be written in terms of the magnetic flux as

$$dI = (1/4\pi)\mu d\Psi \quad (\text{C.18})$$

$$= \frac{1}{2}\mu dA_\phi . \quad (\text{C.19})$$

## C.2. Energy and angular momentum outflow

For a stationary and axial symmetric system, the conserved fluxes can be defined about the axis of symmetry. The force-free condition ensures the conserved electromagnetic energy flux

$$\mathcal{E}^\mu = T_b^\mu \quad (\text{C.20})$$

and angular momentum flux

$$\mathcal{L}^\mu = - T_\phi^\mu , \quad (\text{C.21})$$

where the electromagnetic energy momentum tensor is given by

$$T^{\mu\nu} = (1/4\pi)(F_\rho^\mu F^{\nu\rho} - \frac{1}{4}g^{\mu\nu}F_{\rho\sigma}F^{\rho\sigma}) . \quad (\text{C.22})$$

The energy flux and angular momentum flux have a simple relation,

$$\mathcal{E}^\mu = \omega \mathcal{L}^\mu \quad (\text{C.23})$$

and the poloidal components can be written as

$$\mathcal{E}^r = -(1/4\pi)\omega(\Delta/\rho^4)A_{\phi,\theta}B_\phi , \quad (\text{C.24})$$

$$\mathcal{E}^\theta = (1/4\pi)\omega(\Delta/\rho^4)A_{\phi,r}B_\phi . \quad (\text{C.25})$$

Using the physical component,  $\mathcal{E}_r$ ,

$$\mathcal{E}_r = \sqrt{g_{00}/g_{rr}}\mathcal{E}_r = \sqrt{g_{00}g_{rr}}\mathcal{E}^r , \quad (\text{C.26})$$

we get the power at infinity given by

$$P = \int \mathcal{E}_r \cdot dS_r \quad (\text{C.27})$$

$$= -\frac{1}{4\pi} \int \omega \sqrt{\frac{g_{00}}{g_{\theta\theta}g_{rr}}} B_\phi A_{\phi,\theta} \sqrt{g_{\theta\theta}g_{\phi\phi}} d\theta d\phi \quad (\text{C.28})$$

$$= -\frac{1}{4\pi} \int \omega \sqrt{\frac{g_{00}g_{\phi\phi}}{g_{rr}g_{\theta\theta}}} B_\phi A_{\phi,\theta} d\theta d\phi \quad (\text{C.29})$$

$$= -\frac{1}{4\pi} \int \omega \left( \frac{\Delta \sin \theta}{\rho^2} B_\phi \right) A_{\phi,\theta} d\theta d\phi . \quad (\text{C.30})$$

Defining  $B_T$  [2] as

$$B_T = (\Delta \sin \theta / \rho^2) B_\phi , \quad (\text{C.31})$$

the power can be written as

$$P = -\frac{1}{2} \int \omega B_T A_{\phi,\theta} d\theta \quad (\text{C.32})$$

$$= -\frac{1}{2} \int \omega B_T dA_\phi . \quad (\text{C.33})$$

From the inhomogeneous Maxwell equations

$$F_{;\nu}^{\mu\nu} = 4\pi J^\mu , \quad (\text{C.34})$$

we get

$$B_{T,\theta} = 4\pi\sqrt{gJ^r}, \quad (C.35)$$

$$B_{T,r} = -4\pi\sqrt{gJ^\theta}. \quad (C.36)$$

Using Eq. (C.14),

$$\mu A_{\phi,\theta} = B_{T,\theta}, \quad \mu A_{\phi,r} = B_{T,r}, \quad (C.37)$$

which shows that  $B_T$  is also constant along the magnetic surface. Since

$$\mu dA_\phi = dB_T, \quad (C.38)$$

the outward current can be calculated as

$$dI = \frac{1}{2}\mu dA_\phi \quad (C.39)$$

$$= \frac{1}{2}dB_T \quad (C.40)$$

and we get

$$I(\theta) = \frac{1}{2} \int dB_T = \frac{1}{2}B_T(\theta), \quad (C.41)$$

where  $B_T(\theta = 0) = 0$  has been used. Using the physical component,  $B_\phi = B_\phi/\sqrt{g_{rr}g_{\theta\theta}}$ , Eq. (C.41) can be written

$$I = \frac{1}{2}\tilde{\omega}\alpha B_\phi \quad (C.42)$$

which is nothing but the Ampère's law on the stretched horizon

$$2\pi\tilde{\omega}B_\phi^H = 4\pi I \quad (C.43)$$

using the tangential field on the stretched horizon [6]

$$B_\phi^H = \alpha B_\phi. \quad (C.44)$$

Using Eqs. (C.33), (C.39) and (C.41) we get

$$dP = -\omega I dA_\phi \quad (C.45)$$

$$= -(1/2\pi)\omega I d\Psi, \quad (C.46)$$

$$P = -\frac{1}{2\pi} \int_{\text{horizon}} \omega I d\Psi. \quad (C.47)$$

Since  $\omega, I$  and  $\Psi$  are constant along the magnetic surface we can evaluate the integral on the horizon of the black hole:

$$P = -\frac{1}{2\pi} \int_{\text{horizon}} \omega I d\Psi. \quad (C.48)$$

Hence we can see that it is a proof of the simple-minded derivation of Eq. (20) in the text with  $\omega$  identified with  $\Omega_r$ .

## Appendix D. Power from rotating black hole

Using the boundary condition [37] that  $B_r$  is finite and

$$B_T = [\sin \theta [\omega(r_H^2 + a^2) - a]/(r_H^2 + a^2 \cos \theta)] B_r, \quad (\text{D.1})$$

$$= [\sin \theta (\omega - \Omega_H) \Sigma_H / \rho_H^2] B_r, \quad (\text{D.2})$$

which can be written for  $B_\phi$  as

$$B_\phi = (\rho^2/\Delta)[(\omega - \Omega_H) \Sigma_H / \rho^2] B_r, \quad (\text{D.3})$$

$$= [(\omega - \Omega_H) \Sigma_H / \Delta] B_r, \quad (\text{D.4})$$

$$= [(\omega - \Omega_H) \Sigma_H / \Delta] \rho \tilde{\omega} B_r, \quad (\text{D.5})$$

and the toroidal component on the stretched horizon is

$$B_\phi^H = (\omega - \Omega_H) \tilde{\omega} B_H, \quad (\text{D.6})$$

where  $B_H = B_r$ . Then the current up to  $\theta = \theta_H$  can be written using Eq. (C.41)

$$I(\theta_H) = \frac{1}{2}(\Delta/\rho^2) \sin \theta B_\phi = \frac{1}{2}(\omega - \Omega_H) \tilde{\omega}^2 B_H \quad (\text{D.7})$$

which is exactly the same current derived in the membrane paradigm [6].

The total power out of the black hole can be calculated using Eq. (C.32), (C.41), and (D.7),

$$P = \frac{1}{2} \int \omega B_T B_r d\theta \quad (\text{D.8})$$

$$= - \int \omega I \rho \tilde{\omega} B_H d\theta \quad (\text{D.9})$$

$$= - \frac{1}{2} \int \omega (\omega - \Omega_H) B_H^2 \rho \tilde{\omega}^3 d\theta. \quad (\text{D.10})$$

Adopting the optimal condition [6],  $\omega \sim \Omega_H/2$ , the power can be written as

$$P = - \frac{1}{2} \int \omega (\omega - \Omega_H) B_H^2 \rho \tilde{\omega}^3 d\theta \quad (\text{D.11})$$

$$= \frac{1}{8} \left( \frac{a}{2Mr_H} \right)^2 \int \frac{(r_H^2 + a^2)^3 \sin^3 \theta}{r_H^2 + a^2 \cos^2 \theta} B_H^2 d\theta \quad (\text{D.12})$$

$$= \frac{1}{32} \left( \frac{a}{M} \right)^2 \frac{(2Mr_H)^2}{r_H^2} \left( 1 + \left( \frac{r_H}{a} \right)^2 \right) \int \frac{\sin^3 \theta}{(r_H/a)^2 + \cos^2 \theta} B_H^2 d\theta. \quad (\text{D.13})$$

Using the integral identity,

$$\int_0^{\pi/2} \frac{\sin^3 \theta}{(r_H/a)^2 + \cos^2 \theta} d\theta = [(h + 1/h) \arctan h - 1], \quad (\text{D.14})$$

where  $h$  is defined in Ref. [17] as

$$h = a/r_H \quad (\text{D.15})$$

$$= \frac{J/M^2}{1 + \sqrt{1 - J^2/M^4}} = \tilde{a}/(1 + \sqrt{1 - \tilde{a}^2}), \quad (\text{D.16})$$

$$M/r_H = (M/a)a/r_H = h/\tilde{a}, \quad (\text{D.17})$$

we integrate the angular integral assuming the average magnetic field,  $\langle B_H^2 \rangle$ ,

$$P = \frac{1}{32}(a/M)^2 [(2Mr_H)^2/r_H^2][(1+h^2)/h^2]\langle B_H^2 \rangle 2[(h+1/h)\arctan h - 1] \quad (\text{D.18})$$

$$= \frac{1}{4}(a/M)^2 M^2 \langle B_H^2 \rangle f(h), \quad (\text{D.19})$$

where

$$f(h) = [(1+h^2)/h^2][(h+1/h)\arctan h - 1] \quad (\text{D.20})$$

$$\rightarrow 2/3 \quad \text{for } h \rightarrow 0, \quad (\text{D.21})$$

$$= \pi - 2 \quad \text{for } h = 1 \text{ (extreme rotation, } a = M\text{).} \quad (\text{D.22})$$

Now inserting  $G$  and  $c$  properly, we get

$$P = \frac{1}{4}(ac/GM)^2 M^2 \langle B_H^2 \rangle f(h) G^2/c^3 \quad (\text{D.23})$$

$$= \frac{1}{4}\tilde{a}^2 M^2 \langle B_H^2 \rangle f(h) G^2/c^3 \quad (\text{D.24})$$

$$= (1/4c)(J^2/M^4)M^2 \langle B_H^2 \rangle f(h). \quad (\text{D.25})$$

Numerically,

$$(G^2/c^3)M_\odot^2(10^{15} \text{ G})^2 = (2 \times 10^{33} \times 10^{15})^2 \times (6.7 \times 10^{-8})^2/(3 \times 10^{10})^3 \quad (\text{D.26})$$

$$= 6.7 \times 10^{50} \text{ erg/s}. \quad (\text{D.27})$$

Then the total power of Poynting flux is

$$P = 1.7 \times 10^{50}(ac/GM)^2(M/M_\odot)^2(\langle B_H \rangle/10^{15} \text{ G})^2 f(h) \text{ erg/s}, \quad (\text{D.28})$$

One of the frequently quoted numbers for the total power of the BZ process from Ref. [6] is

$$P \sim 10^{49}(ac/GM)^2(M/M_\odot)^2(\langle B_H \rangle/10^{15} \text{ G})^2 \text{ erg/s} \quad (\text{D.29})$$

which is one order of magnitude lower than Eq. (D.23).<sup>9</sup> In Ref. [6], the angular integration of Eq. (C.48) is approximated by replacing

$$\tilde{\omega}^2 \rightarrow r_H^2/2, \quad (\text{D.30})$$

$$\Delta\Psi \rightarrow \pi r_H^2 B_H, \quad (\text{D.31})$$

<sup>9</sup> Recently Popham et al. [9] use

$$P \sim 10^{50}(ac/GM)^2(\langle B_H \rangle/10^{15} \text{ G})^2 \text{ erg/s} \quad (\text{D.39})$$

for  $M = 3M_\odot$  hence equivalent to Eq. (D.29).

$$P \sim \frac{1}{128} (ac/GM)^2 M^2 \langle B_H^2 \rangle g^2 (a/M) G^2 / c^3, \quad (\text{D.32})$$

$$g^2 (ac/GM) = 1 + (1 - (ac/GM)^2)^{1/2} \quad (\text{D.33})$$

$$\rightarrow 1 \quad \text{for } a \rightarrow 0, \quad (\text{D.34})$$

$$= 2 \quad \text{for } a = M \text{ (extreme rotation).} \quad (\text{D.35})$$

Then

$$P(\text{Eq. (D.32)})/P(\text{Eq. (D.23)}) = g^2 (ac/GM)/32f(h) \quad (\text{D.36})$$

$$\rightarrow 3/64 \quad \text{for } a \rightarrow 0 \quad (\text{D.37})$$

$$\sim 5/56 \quad \text{for } a = M \text{ (extreme rotation)} \quad (\text{D.38})$$

which explains the difference of an order-of-magnitude.

### Note added in proof

The effect of plasma fluids, a part of which comprise the current flows of the global structure proposed by Blandford and Znajek, on the Blandford-Znajek process has been discussed by Park and Vishniac [38], Takahashi et al. [39] and Hirotani et al. [40] (cf. Punsly and Coroniti [41] for a different point of view). It has been demonstrated [39] that the fluid flows along the magnetic field line can cross the critical points with causal connection (see, also Phinney, [42]), from the originating injection points through the Alfvén points and the magnetosonic points to the horizon (inflow) and infinity (outflow). The magnetically dominated MHD as originally proposed by Blandford and Znajek can be obtained if the fast magnetosonic point is very close to the horizon [40].

It is found that the Blandford-Znajek process is possible when the injection point is inside the ergosphere. Hence accreting material interacting with the magnetic field can also extract the rotational energy of the black hole [39,40]. This is an interesting observation in connection with the proposed model for GRB in this paper, since it is likely that the magnetic energy may be comparable to that of the accreting material from the accretion disk in the transition region between the horizon and the accretion disk near the equatorial plane. Close to the rotation axis of the black hole there is no substantial accretion and the magnetically dominated Blandford-Znajek process can be an effective process in extracting the rotational energy of the black hole. Recent discussions on the effect of magnetic field in the transition region can be found in [43–45].

Earlier suggestions that the central engine of GRB (for a recent review of the gamma ray bursts, see [48]) may be powered by the Blandford-Znajek process can be found in [46,47,4]. Recently an attempt to explain the temporal structure of gamma ray bursts using the Blandford-Znajek process has been made by Portegies-Zwart et al. [49].

### References

- [1] S.R. Kulkarni et al., Nature 393 (1998) 35.
- [2] R.D. Blandford, R.L. Znajek, Mon. Not. R. Astron. Soc. 179 (1977) 433.

- [3] S.E. Woosley, *Astrophys. J.* 405 (1993) 405.
- [4] B. Paczyński, *Astrophys. J.* 494 (1998) L45.
- [5] H. Bethe, G.E. Brown, *Astrophys. J.* 506 (1998) 780.
- [6] K.S. Thorne, R.H. Price, D.A. MacDonald, *Black Holes; The Membrane Paradigm*, Yale University Press, New Haven and London, 1986.
- [7] S.R. Kulkarni et al., astro-ph/9902272 v2.
- [8] P. Mészáros, M.J. Rees, R.A.M.J. Wijers, *New Ustron.* 4 (1999) 303.
- [9] R. Popham et al., *Astrophys. J.* 518 (1999) 356.
- [10] M. Livio, G.I. Ogilvie, J.E. Pringle, *Astrophys. J.* 512 (1999) 100.
- [11] P. Ghosh, M.A. Abramowicz, *Mon. Not. R. Astron. Soc.* 292 (1997) 887.
- [12] C. Kouveliotou et al., *Nature* 393 (1998) 235.
- [13] H. Janka, M. Ruffert, T. Eberl, astro-ph/9810057.
- [14] C. Fendt, *Astron. Astrophys.* 319 (1997) 1025.
- [15] D. Macdonald, K.S. Thorne, *Mon. Not. R. Astron. Soc.* 198 (1982) 345.
- [16] S. Shapiro, S.A. Teukolsky, *Black Holes, White Dwarfs, and Neutron Stars*, Wiley, New York, 1983.
- [17] I. Okamoto, *Mon. Not. R. Astron. Soc.* 294 (1992) 192.
- [18] R.M. Wald, *Phys. Rev. D* 10 (1974) 1680.
- [19] A.R. King, J.P. Laosta, W. Kundt, *Phys. Rev. D* 12 (1975) 3037.
- [20] J. Bičák, V. Janis, *Mon. Not. R. Astron. Soc.* 212 (1985) 899.
- [21] See, for example K. Hirotani, I. Okamoto, *Astrophys. J.* 497 (1998) 563.
- [22] S.E. Woosley, E. Baron, *Astrophys. J.* 391 (1992) 228.
- [23] H. Janka, M. Ruffert, H.-Th. Janka, astro-ph/9804132.
- [24] D.A. MacDonald, *Mon. Not. R. Astron. Soc.* 211 (1984) 313.
- [25] Z. Kopal, *Close Binary Systems*, Chapman & Hall, London.
- [26] L. Bildsten, C. Cutler, *Astrophys. J.* 400 (1992) 175.
- [27] K. Iwamoto et al., *Nature* 395 (1998) 672.
- [28] L. Motz, *Astrophys. J.* 115 (1952) 562.
- [29] R. Blandford, *Mon. Not. R. Astron. Soc.* 176 (1976) 465.
- [30] See, for example, U. Torkelsson, A. Brandenburg, A. Nordlund, R.F. Stein, *Astro. Lett. Comm.* 34 (1996) 383.
- [31] W. Kluzniak, M. Ruderman.
- [32] N.I. Shakura, R.A. Sunyaev, *Astron. Astrophys.* 24 (1973) 337.
- [33] J.E. Pringle, *Ann. Rev. Astron. Astrophys.* 19 (1981) 137.
- [34] L.-X. Li, astro-ph/9902352.
- [35] R.P. Kerr, *Phys. Rev. Lett.* 11 (1963) 237.
- [36] R.H. Boyer, R.W. Lindquist, *J. Math. Phys.* 8 (1967) 265.
- [37] R.L. Znajek, *Mon. Not. R. Astron. Soc.* 179 (1977) 457.
- [38] S.J. Park, E.T. Vishniac, *Astrophys. J.* 332 (1988) 135.
- [39] M. Takahashi, S. Nitta, Y. Tatematsu, A. Tomimatsu, *Astrophys. J.* 363 (1990) 206.
- [40] K. Hirotani, M. Takahashi, S.-Y. Nitta, A. Tomimatsu, *Astrophys. J.* 386 (1992) 455.
- [41] B. Punsly, F.V. Coroniti, *Astrophys. J.* 350 (1990) 518.
- [42] S. Phinney, 1983, Ph.D. thesis, Univ. Cambridge.
- [43] B. Punsly, *Astrophys. J.* 506 (1998) 790.
- [44] J.H. Krolik, *Astrophys. J.* 515 (1999) L73.
- [45] C.F. Gammie, *Astrophys. J.* 522 (1999) L57.
- [46] P. Mészáros, M. Rees, *Astrophys. J.* 482 (1997) L29.
- [47] J. Katz, *Astrophys. J.* 490 (1997) 633.
- [48] T. Piran, *Phys. Rep.* 314 (1999) 576.
- [49] S.F. Portegies-Zwart, C.-H. Lee, H.K. Lee, *Astrophys. J.* 520 (1999) 666.

---

## Chapter 17

# A Theory of Gamma-Ray Bursts

G.E. Brown, C.-H. Lee, R.A.M.J. Wijers, H.K. Lee, G. Israelian and H.A. Bethe

Reprinted with permission from *New Astronomy*, 5 (2000) 191–210

Copyright (2000) by Elsevier Science B.V.

### **Commentary**

One of the earliest and most popular models for gamma-ray bursters (GRB) was the Woosley (1993) Collapsar model, updated by MacFadyen & Woosley (1999), in which the center of a rapidly rotating massive Wolf-Rayet falls into a black hole. In the Woosley & MacFadyen work, magnetohydrodynamic mechanisms, such as the Blandford-Znajek one, which we favor, “jump-start” the GRB in case not enough energy is furnished through the jet.

It seemed clear to us that the Woosley mechanism, supplemented by Blandford-Znajek, would work, but the problems were: (i) a naked Wolf-Rayet would blow away, down to a low-mass CO core, as we showed in several papers, culminating in Paper 21, and would go into a low-mass black hole; (ii) especially with the very high magnetic fields in the He star necessary to run the later magnetohydrodynamics, the viscosity would be close to its upper limit (i.e.,  $\alpha \sim 1$  in the Shakura & Sunyaev (1973) parameterization of viscosity), so the massive progenitor of the black hole would have been brought into common rotation (Spruit & Phinney 1998). Thus, there would not be the angular momentum to bring the black hole into the rapid rotation necessary to drive the Blandford-Znajek mechanism.

In Paper 21 we show how the difficulty (i) can be overcome; i.e., by keeping a hydrogen envelope on the Wolf-Rayet for at least some part of its lifetime — i.e., by starting with a WNL star. This cuts down on the wind loss.

The more difficult point (ii) suggested taking as progenitor for the GRB the same progenitor as for transient sources; namely, a massive progenitor of the high-mass black hole, in a binary with a lower-mass star. The function of the latter is to remove the common envelope of the more massive star, but from our earlier work catalogued in Papers 9, 13, 14, 15 and 20, it should do this only after the more massive star has completed He core burning; i.e., after the supergiant stage of that star. In removing the envelope, the companion spins up the He core of the massive star. It continues spinning it up by tidal interaction.

It has become increasingly clear that gamma-ray bursts are accompanied by extremely energetic supernova explosions, now termed hypernovae. We learned from Roger Blandford during our annual visit to Caltech in January 2000 that these follow naturally from the Blandford-Znajek mechanism. The gamma-ray burst is run by plasma currents up open field lines, forming a hot fireball in the loading region. The black hole couples to the accretion disk through closed field lines and since the black hole initially rotates about twice as fast as the accretion disk, it will torque it up, delivering energy into it. The amount of this energy is expected to be about equal to that in the GRB, as we develop in our paper. After publishing our paper we found that Maurice van Putten (1999) had earlier foreseen this deposition of energy in the accretion disk ("torus"). See also van Putten & Wilson (1999).

We were very excited to hear of the Israeli *et al.* work (on public radio one day in September 2000) where the F-star companion in the soft X-ray transient source Nova Scorpii had been bombarded (about one billion years ago) by  $\alpha$ -particle nuclei, the relative abundances of which could only have come from a hypernova. This story is told in Section 7.4 of our paper.

Indeed, Brown, Lee, Lee & Bethe had already submitted a presentation, given by C.-H. Lee, "Is Nova Sco 1994 (GRO 1655-40) a Relic of a GRB?", to the 5th Huntsville Symposium on Gamma-Ray Bursts, Huntsville, Alabama, 18-22 October 1999, and it is published in the Proceedings. In our present paper we find that only about half of the rotational energy of the black hole is used up before the hypernova disrupts the magnetic fields from the accretion disk which results in the collapse of the He star once the rapidly rotating black hole is formed in its center. The black hole surface is still rotating with  $\sim 70\%$  of the speed of light.

The referee of our paper emphasized the difficulty of depositing the necessary  $0.25M_{\odot}$  of oxygen on the companion star in the explosion leading to the black hole. Indeed, assuming spherical symmetry we deposited only about 0.1 of this amount, Eq. (34). We explained how the nonspherical geometry enhanced the amount of oxygen that would be sent out along the accretion disk, and this material being much slower would be more easily captured. In fact, a detailed numerical calculation of an aspherical hypernova explosion with very nearly the geometry we proposed (Maeda *et al.* 2002) has been carried out. From Fig. 2 of this paper one can see that at least an order of magnitude more goes out sideways, through what would be our disk than vertically, up the direction of the jet. Thus, our discussion is confirmed.

We tell the story of Nova Scorpii in detail in Paper 17. An even more spectacular story may be in the offing, "Black Hole Binary May Be a Relic of a Gamma-Ray Burst in Our Galaxy" (<http://saturn.sron.nl/~erikk/v4641sgr/v4641sgr.html>), which is described in "A Black Hole in the Superluminal Source SAX J1819.3-2525 (V4641 Sgr)" by Orosz *et al.* This is a transient source of a black hole of mass  $9.61^{+2.08}_{-0.88} M_{\odot}$  with late B-star companion of  $6.53^{+1.60}_{-1.03} M_{\odot}$  which has evolved off the main sequence.

---

## References

- (1) MacFadyen, A.I., and Woosley, S.E. (1999), *Astrophysical Journal* **524**, 262.
- (2) Maeda, K., Nakamura, T., Nomoto, K., Mazzali, P.A., Patat, F., and Hachisu, I. (2002), *Astrophysical Journal* **565**, 405.
- (3) Orosz, J.A., Kuulkers, E., van der Klis, M., McClintoch, J.E., Garcia, M.R., Collanan, P.J., Bailyn, C.D., Jain, R.K., and Remillard, R.A. (2000), *Astrophysical Journal* **555**, 489.
- (4) Shakura, N.I., and Sunyaev, R.A. (1973), *Astronomy & Astrophysics* **24**, 337.
- (5) Spruit, H., and Phinney, E.S. (1998), *Nature* **393**, 139.
- (6) van Putten, M.H.P.M. (1999), *Science* **284**, 115.
- (7) van Putten, M.H.P.M., and Wilson, A. (1999), in Proc. ITP Conf. Black Holes, *Theory Confronts Reality*, Feb. 2-5 (R. Antonucci, N. Gehrels, S.M. Kahn, R. Narayan, E.S. Phinney, D.O. Richstone, J.H. Swank and A.S. Wilson, Eds.); [http://www.itp.ucsb.edu/online/bhole\\_c99](http://www.itp.ucsb.edu/online/bhole_c99)
- (8) Woosley, S.E. (1993), *Astrophysical Journal* **405**, 273.

This page is intentionally left blank



ELSEVIER

## A theory of gamma-ray bursts

G.E. Brown<sup>a</sup>, C.-H. Lee<sup>a</sup>, R.A.M.J. Wijers<sup>a</sup>, H.K. Lee<sup>b</sup>, G. Israelian<sup>c</sup>, H.A. Bethe<sup>d</sup>

<sup>a</sup>Department of Physics & Astronomy, State University of New York, Stony Brook, New York 11794, USA

<sup>b</sup>Department of Physics, Hanyang University, Seoul 133-791, South Korea

<sup>c</sup>Instituto de Astrofísica de Canarias, E-38200 La Laguna, Tenerife, Spain

<sup>d</sup>Floyd R. Newman Laboratory of Nuclear Studies, Cornell University, Ithaca, New York 14853, USA

Received 20 March 2000; received in revised form 18 May 2000; accepted 25 May 2000

Communicated by S.E. Woosley

### Abstract

Recent observations and theoretical considerations have linked gamma-ray bursts with ultra-bright type Ib/c supernovae ('hypernovae'). We here work out a specific scenario for this connection. Based on earlier work, we argue that especially the longest bursts must be powered by the Blandford-Znajek mechanism of electromagnetic extraction of spin energy from a black hole. Such a mechanism requires a high angular momentum in the progenitor object. The observed association of gamma-ray bursts with type Ib/c supernovae leads us to consider massive helium stars that form black holes at the end of their lives as progenitors. In our analysis we combine the numerical work of MacFadyen & Woosley with analytic calculations in Kerr geometry, to show that about  $10^{53}$  erg each are available to drive the fast GRB ejecta and the supernova. The GRB ejecta are driven by the power output through the open field lines threading the black hole, whereas the supernova can be powered both by the shocks driven into the envelope by the jet, and by the power delivered into the disk via field lines connecting the disk with the black hole. We also present a much simplified approximate derivation of these energetics.

Helium stars that leave massive black-hole remnants can only be made in fairly specific binary evolution scenarios, namely the kind that also leads to the formation of soft X-ray transients with black-hole primaries, or in very massive WNL stars. Since the binary progenitors will inevitably possess the high angular momentum we need, we propose a natural link between black-hole transients and gamma-ray bursts. Recent observations of one such transient, GRO J1655-40/Nova Scorpii 1994, explicitly support this connection: its high space velocity indicates that substantial mass was ejected in the formation of the black hole, and the overabundance of  $\alpha$ -nuclei, especially sulphur, indicates that the explosion energy was extreme, as in SN 1998bw/GRB 980425. Furthermore, X-ray studies of this object indicate that the black hole may still be spinning quite rapidly, as expected in our model. We also show that the presence of a disk during the powering of the GRB and the explosion is required to deposit enough of the  $\alpha$  nuclei on the companion. © 2000 Elsevier Science B.V. All rights reserved.

PACS: 98.70.Rz; 97.10.Cv; 97.10.Tk; 97.60.Lf; 95.30.Sf

Keywords: Accretion, accretion disks; Black hole physics; MHD; Supernovae: general; Gamma-rays: bursts; Gamma-rays: theory

E-mail addresses: epopenoe@nuclear.physics.sunysb.edu (G.E. Brown), chlee@silver.physics.sunysb.edu (C.-H. Lee), rwijers@mail.astro.sunysb.edu (R.A.M.J. Wijers), hklee@hepth.hanyang.ac.kr (H.K. Lee), gil@ll.iac.es (G. Israelian).

## 1. Introduction

The discovery of afterglows to gamma-ray bursts has greatly increased the possibility of studying their physics. Since these afterglows have thus far only been seen for long gamma-ray bursts (duration  $\gtrsim 2$  s), we shall concentrate on the mechanism for this subclass. The shorter bursts (duration  $\leq 2$  s) may have a different origin; specifically, it has been suggested that they are the result of compact-object mergers and therefore offer the intriguing possibility of associated outbursts of gravity waves. (Traditionally, binary neutron stars have been considered in this category (Eichler et al., 1989; Janka et al., 1999). More recently, Bethe & Brown (1998) have shown that low-mass black-hole, neutron-star binaries, which have a ten times greater formation rate and are stronger gravity-wave emitters, may be the more promising source of this kind.)

An important recent clue to the origin of long bursts is the probable association of some of them with ultra-bright type Ibc supernovae (Bloom et al., 1999; Galama et al., 1998, 2000). The very large explosion energy<sup>1</sup> implied by fitting the light curve of SN 1998bw, which was associated with GRB 980425 (Galama et al., 1998), indicates that a black hole was formed in this event (Iwamoto et al., 1998). This provides two good pieces of astrophysical information: it implicates black holes in the origin of gamma-ray bursts, and it demonstrates that a massive star can explode as a supernova even if its core collapses into a black hole.

In this paper, we start from the viewpoint that the gamma-ray burst is powered by electromagnetic energy extraction from a spinning black hole, the so-called Blandford–Znajek mechanism (Blandford & Znajek, 1977). This was worked out in detail by Lee et al. (1999), and further details and comments were discussed by Lee et al. (2000), who built on work by Thorne et al. (1986) and Li (2000). They have shown that with the circuitry in a 3+1 dimensional description using the Boyer-Lindquist metric, one can have a simple pictorial model for the BZ mechanism.

The simple circuitry which involves steady state current flow is, however, inadequate for describing dissipation of the black hole rotational energy into the accretion disk formed from the original helium envelope. In this case the more rapidly rotating black hole tries to spin up the inner accretion disk through the closed field lines coupling the black hole and disk. Electric and magnetic fields vary wildly with time. Using the work of Blandford & Spruit (2000) we show that this dissipation occurs in an oscillatory fashion, giving a fine structure to the GRB, and that the total dissipation should furnish an energy comparable to that of the GRB to the accretion disk. We use this energy to drive the hypernova explosion.

Not any black-hole system will be suitable for making GRB: the black hole must spin rapidly enough and be embedded in a strong magnetic field. Moreover, the formation rate must be high enough to get the right rate of GRB even after accounting for substantial collimation of GRB outflows. We explore a variety of models, and give arguments why some will have sufficient energy and extraction efficiency to power a GRB and a hypernova. We argue that the systems known as black-hole transients are the relics of GRBs, and discuss the recent evidence from high space velocities and chemical abundance anomalies that these objects are relics of hypernovae and GRBs; we especially highlight the case of Nova Scorpii 1994 (GRO J1655-40).

The plan of this paper is as follows. We first show that it is reasonable to expect similar energy depositions into the GRB outflow and the accretion disk (Section 2) and discuss the amount of available energy to be extracted (Section 3). Then we show the agreement of those results with the detailed numerical simulations by MacFadyen & Woosley, and use those simulations to firm up our numbers (Section 4). We continue by presenting a simple derivation of the energetics that approximates the full results well (Section 5). Finally, we discuss some previously suggested progenitors (Section 6) and present our preferred progenitors: soft X-ray transients (Section 7).

## 2. Simple circuitry

Although our numbers are based on the detailed review of Lee et al. (1999), which confirms the

<sup>1</sup>Höflich et al. (1999) have proposed that the explosion energy was not much larger than usual, but that the explosion was very asymmetric; this model also provides a reasonable fit to the light curve of SN 1998bw.

original Blandford–Znajek paper (Blandford & Znajek, 1977), we illustrate our arguments with the pictorial treatment of Thorne et al. (1986) in ‘The Membrane Paradigm’. Considering the time as universal in the Boyer-Lindquist metric, essential electromagnetic and statistical mechanics relations apply in their 3+1 dimensional manifold. We summarize their picture in our Fig. 1.

The surface of the black hole can be considered as a conductor with surface resistance  $R_{BH} = 4\pi/c = 377 \Omega$ . A circuit that rotates rigidly with the black hole can be drawn from the loading region, the low-field region up the axis of rotation of the black hole in which the power to run the GRB is delivered, down a magnetic field line, then from the North pole of the black hole along the (stretched) horizon to its equator. From the equator we continue the circuit through part of the disk and then connect it upwards with the loading region. We can also draw circuits starting from the loading region which pass along only the black hole or go through only the disk, but adding these would not change the results of our schematic model.

Using Faraday’s law, the voltage  $V$  can be found by integrating the vector product of charge velocity,  $v$ , and magnetic field,  $B$ , along the circuit:

$$V = \int [v \times B] \cdot dl, \quad (1)$$

( $dl$  is the line element along the circuit). Because this law involves  $v \times B$  the integrals along the field lines make no contribution. We do get a contribution  $V$  from the integral from North pole to equator along the black hole surface. Further contributions to  $V$  will come from cutting the field lines from the disk. We assume the field to be weak enough in the loading region to be neglected.

The GRB power,  $E_{GRB}$ , will be

$$\dot{E}_{GRB} = I_{BH+D}^2 R_L, \quad (2)$$

where  $R_L$  is the resistance of the loading region, and the current is given by

$$I_{BH+D}^2 = \left( \frac{V_D + V_{BH}}{R_D + R_{BH} + R_L} \right)^2. \quad (3)$$

(The index BH refers to the black hole, L to the load region, and D to the disk.)

The load resistance has been estimated in various ways and for various assumptions by Lovelace et al. (1979) and by MacDonald & Thorne (1982), and by Phinney (1983). All estimates agree that to within a factor of order unity  $R_L$  is equal to  $R_{BH}$ .

In a similar fashion, some power will be deposited into the disk

$$\dot{E}_{disk} = I_{BH+D}^2 R_D \quad (4)$$

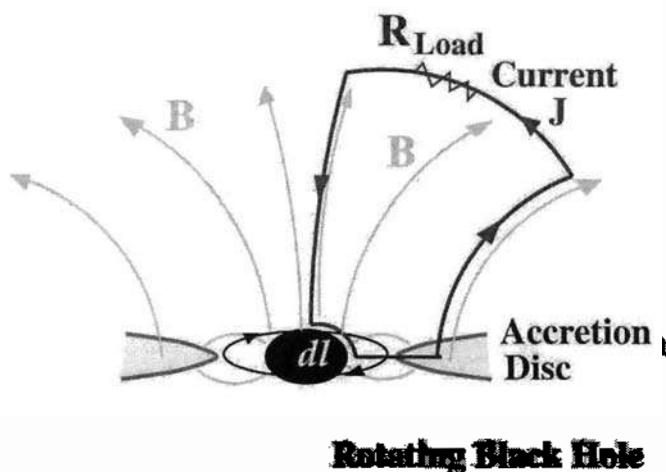


Fig. 1. The black hole in rotation about the accretion disk. A circuit, in rigid rotation with the black hole is shown. This circuit cuts the field lines from the disk as the black hole rotates, and by Faraday’s law, produces an electromotive force. This force drives a current. More detailed discussion is given in the text.

but this equilibrium contribution will be small because of the low disk resistance  $R_D$ .

Blandford & Spruit (2000) have shown that important dissipation into the disk comes through magnetic field lines coupling the disk to the black hole rotation. As shown in Fig. 2 these lines, anchored in the inner disk, thread the black hole.

The more rapidly rotating black hole will provide torques, along its rotation axis, which spin up the inner accretion disk, in which the closed magnetic field lines are anchored. With increasing centrifugal force the material in the inner disk will move outwards, cutting down the accretion. Angular momentum is then advected outwards, so that the matter can drift back inwards. It then delivers more matter to the black hole and is flung outwards again. The situation is like that of a ball in a roulette wheel (R.D. Blandford, private communication). First of all it is flung outwards and then drifts slowly inwards. When it hits the hub it is again thrown outwards. The viscous inflow time for the fluctuations is easily estimated to be

$$\tau_d \sim \Omega_{\text{disk}}^{-1} \left( \frac{r}{H} \right)^2 \alpha_{\text{vis}}^{-1}, \quad (5)$$

where  $H$  is the height of the disk at radius  $r$ ,  $\Omega_{\text{disk}}$  its angular velocity, and  $\alpha_{\text{vis}}$  is the usual  $\alpha$ -parameterization of the viscosity. We choose  $\alpha_{\text{vis}} \sim 0.1$ ,  $r/H \sim 10$  for a thin disk and then arrive at  $\tau_d \sim 0.1$  s. We therefore expect variability on all time scales between the Kepler time (sub-millisecond) and the viscous time, which may explain the very erratic light curves of many GRBs.

We suggest that the GRB can be powered by  $\dot{E}_{\text{GRB}}$  and a Type Ibc supernova explosion by  $\dot{E}_{\text{SN}}$  where  $\dot{E}_{\text{SN}}$  is the power delivered through dissipation into the disk. To the extent that the number of closed field lines coupling disk and black hole is equal to the number of open field lines threading the latter,

the two energies will be equal. In the spectacular case of GRB 980326 (Bloom et al., 1999), the GRB lasts about 5 s, which we take to be the time that the central engine operates. We shall show that up to  $\sim 10^{53}$  erg is available to be delivered into the GRB and into the accretion disk, the latter helping to power the supernova (SN) explosion. This is more energy than needed and we suggest that injection of energy into the disk shuts off the central engine by blowing up the disk and thus removing the magnetic field needed for the energy extraction from the black hole. If the magnetic field is high enough the energy will be delivered in a short time, and the quick removal of the disk will leave the black hole still spinning quite rapidly.

### 3. Energetics of GRBs

The maximum energy that can be extracted from the BZ mechanism (Lee et al., 1999) is

$$(E_{\text{BZ}})_{\text{max}} = 0.09 M_{\text{BH}} c^2. \quad (6)$$

This is 31% of the black hole rotational energy, the remainder going toward increasing the entropy of the black hole. This maximum energy is obtained if the extraction efficiency is

$$\epsilon_n = \frac{\Omega_{\text{disk}}}{\Omega_H} = 0.5. \quad (7)$$

In Appendix A we give numerical estimates for this ratio for various  $\omega = \Omega_{\text{disk}}/\Omega_K$  and various radii in the region of parameter space we consider. As explained in Section 2 we expect the material in the inner disk to swing in and out around the marginally stable radius,  $r_m$ . It can be seen from the Table 2 and Appendix A that the relevant values of  $\epsilon_n$  are close to that of Eq. (7).



Fig. 2. Magnetic field lines, anchored in the disk, which thread the black hole, coupling the disk rotation to that of the black hole.

For a  $7 M_{\odot}$  black hole, such as that found in Nova Sco 1994 (GRO J1655-40),

$$E_{\max} = 1.1 \times 10^{54} \text{ erg.} \quad (8)$$

We estimate below that the energy available in a typical case will be an order of magnitude less than this. Without collimation, the estimated gamma-ray energy in GRB 990123 is about  $4.5 \times 10^{54}$  erg (Andersen et al., 1999). The BZ scenario entails substantial beaming, so this energy should be multiplied by  $d\Omega/4\pi$ , which may be a small factor (perhaps  $10^{-2}$ ).

The BZ power can be delivered at a maximum rate of

$$P_{\text{BZ}} = 6.7 \times 10^{50} \left( \frac{B}{10^{15} \text{ G}} \right)^2 \left( \frac{M_{\text{BH}}}{M_{\odot}} \right)^2 \text{ erg s}^{-1}, \quad (9)$$

(Lee et al., 1999) so that high magnetic fields are necessary for rapid delivery.

The above concerns the maximum energy output into the jet and the disk. The real energy available in black-hole spin in any given case, and the efficiency with which it can be extracted, depend on the rotation frequency of the newly formed black hole and the disk or torus around it. The state of the accretion disk around the newly formed black hole, and the angular momentum of the black hole, are somewhat uncertain. However, the conditions should be bracketed between a purely Keplerian, thin disk (if neutrino cooling is efficient) and a thick, non-cooling hypercritical advection-dominated accretion disk (HADAF), of which we have a model (Brown et al., 2000b). Let us examine the result for the Keplerian case. In terms of

$$\bar{a} = \frac{Jc}{M^2 G}, \quad (10)$$

where  $J$  is the angular momentum of the black hole, we find the rotational energy of a black hole to be

$$E_{\text{rot}} = f(\bar{a})Mc^2, \quad (11)$$

where

$$f(\bar{a}) = 1 - \sqrt{\frac{1}{2}(1 + \sqrt{1 - \bar{a}^2})}. \quad (12)$$

For a maximally rotating black hole one has  $\bar{a} = 1^2$ .

We begin with a neutron star in the middle of a Keplerian accretion disk, and let it accrete enough matter to send it into a black hole. In matter free regions the last stable orbit of a particle around a black hole in Schwarzschild geometry is

$$r_{\text{iso}} = 3R_{\text{Sch}} = 6 \frac{GM}{c^2}. \quad (13)$$

This is the marginally stable orbit  $r_{\text{ms}}$ . However, under conditions of hypercritical accretion, the pressure and energy profiles are changed and it is better to use (Abramowicz et al., 1988)

$$r_{\text{iso}} \geq 2R_{\text{Sch}}. \quad (14)$$

With the equal sign we have the marginally bound orbit  $r_{\text{mb}}$ . With high rates of accretion we expect this to be a good approximation to  $r_{\text{iso}}$ . The accretion disk can be taken to extend down to the last stable orbit (refer to Appendix B for the details).

We take the angular velocity to be Keplerian, so that the disk velocity  $v$  at radius  $2R_{\text{Sch}}$  is given by

$$v^2 = \frac{GM}{2R_{\text{Sch}}} = \frac{c^2}{4}, \quad (15)$$

or  $v = c/2$ . The specific angular momentum,  $l$ , is then

$$l \geq 2R_{\text{Sch}}v = 2 \frac{GM}{c}. \quad (16)$$

which in Kerr geometry indicates  $\bar{a} \sim 1$ . Had we taken one of the slowest-rotating disk flows that are possible, the advection-dominated or HADAF case (Brown et al., 2000b; Narayan & Yi, 1994), which has  $\Omega^2 = 2\Omega_K^2/7$ , we would have arrived at  $\bar{a} \sim 0.54$ , so the Kerr parameter will always be high.

Further accretion will add angular momentum to the black hole at a rate determined by the angular

<sup>2</sup>As an aside, we note a nice mnemonic: if we define a velocity  $v$  from the black-hole angular momentum by  $J = MR_{\text{Sch}}v$ , so that  $v$  carries the quasi-interpretation of a rotation velocity at the horizon, then  $\bar{a} = 2v/c$ . A maximal Kerr hole, which has  $R_{\text{event}} = R_{\text{sch}}/2$ , thus has  $v = c$ . For  $\bar{a} \leq 0.5$ , the rotation energy is well approximated by the easy-to-remember expression  $E_{\text{rot}} = \frac{1}{2}Mc^2$ .

velocity of the inner disk. The material accreting into the black hole is released by the disk at  $r_{\text{iso}}$ , where the angular momentum delivered to the black hole is determined. This angular momentum is, however, delivered into the black hole at the event horizon  $R_{\text{Sch}}$ , with velocity at least double that at which it is released by the disk, since the lever arm at the event horizon is only half of that at  $R_{\text{Sch}}$ , and angular momentum is conserved. With more rapid rotation involving movement towards a Kerr geometry where the event horizon and last stable orbit coincide at

$$r_{\text{iso}} = R_{\text{event}} = \frac{GM}{c^2}. \quad (17)$$

Although we must switch over to a Kerr geometry for quantitative results, we see that  $\tilde{a}$  will not be far from its maximum value of unity. Again, for the

lower angular-momentum case of a HADAF, the expected black-hole spin is not much less.

#### 4. Comparison with numerical calculation

Our schematic model has the advantage over numerical calculations that one can see analytically how the scenario changes with change in parameters or assumptions. However, our model is useful only if it reproduces faithfully the results of more complete calculations which involve other effects and much more detail than we include. We here make comparison with Fig. 19 of MacFadyen & Woosley (1999). Accretion rates, etc., can be read off from their figure which we reproduce as our Fig. 3. MacFadyen & Woosley prefer  $\tilde{a}_{\text{initial}} = 0.5$  (We have

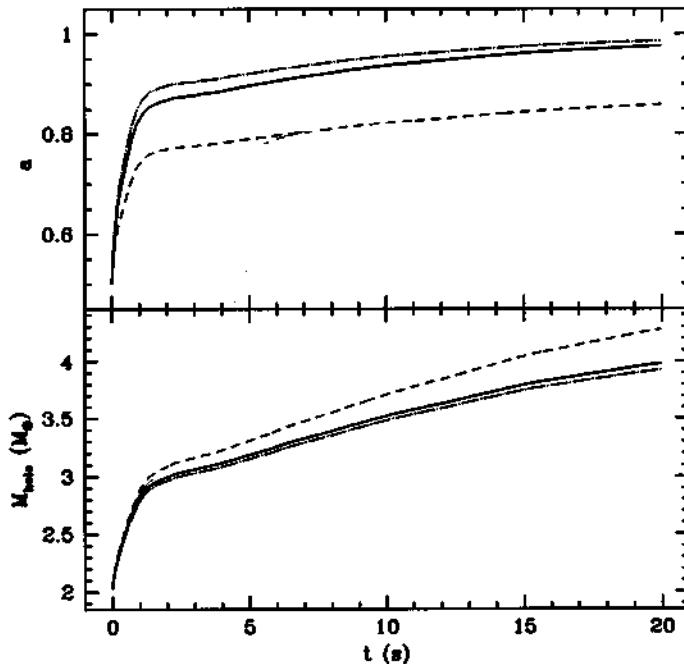


Fig. 3. Time evolution of BH mass and angular momentum taken from Fig. 19 of MacFadyen & Woosley (1999). The upper panel shows the increase in the Kerr parameter for various models for the disk interior to the inner boundary at 50 km. "Thin" (dash-dot), neutrino-dominated (thick solid) and advection dominated (short dash) models are shown for initial Kerr parameter  $\tilde{a}_{\text{init}} = 0.5$ . The lower panel shows the growth of the gravitational mass of the black hole. The short-dashed line shows the growth in baryonic mass of the black hole since for a pure advective model no energy escapes the inner disk.

removed their curve for  $\tilde{a}_{\text{initial}} = 0$ ). This is a reasonable value if the black hole forms from a contracting proto-neutron star near breakup. MacFadyen & Woosley find that  $\tilde{a}_{\text{initial}} = 0.5$  is more consistent with the angular momentum assumed for the mantle than  $\tilde{a}_{\text{initial}} = 0$ . (They take the initial black hole to have mass  $2 M_{\odot}$ ; we choose the Brown & Bethe (1994) mass of  $1.5 M_{\odot}$ .) We confirm this in the next section.

After 5 seconds (the duration of GRB 980326) the MacFadyen & Woosley black hole mass is  $\sim 3.2 M_{\odot}$  and their Kerr parameter  $\tilde{a} \sim 0.8$ , which gives  $f(\tilde{a})$  of our Eq. (12) of 0.11. With these parameters we find  $E = 2 \times 10^{53}$  erg, available for the GRB and SN explosion.

One can imagine that continuation of the MacFadyen & Woosley curve for  $M_{\text{BH}}(M_{\odot})$  would ultimately give something like our  $\sim 7 M_{\odot}$ , but the final black hole mass may not be relevant for our considerations. This is because more than enough energy is available to power the supernova in the first 5 seconds; as the disk is disrupted, the magnetic fields supported by it will also disappear, which turns off the Blandford–Znajek mechanism.

Power is delivered at the rate given by Eq. (9). Taking a black hole mass relevant here,  $\sim 3.2 M_{\odot}$ , we require a field strength of  $\sim 5.8 \times 10^{15}$  G in order for our estimated energy ( $4 \times 10^{52}$  erg) to be delivered in 5 s (the duration of GRB 980326). For such a relatively short burst, we see that the required field is quite large, but it is still not excessive if we bear in mind that magnetic fields of  $\sim 10^{15}$  G have already been observed in magnetars (Kouveliotou et al., 1998, 1999). Since in our scenario we have many more progenitors than there are GRBs, we suggest that the necessary fields are obtained only in a fraction of all potential progenitors.

Thus we have an extremely simple scenario for powering a GRB and the concomitant SN explosion in the black hole transients, which we will discuss in Section 7.2. After the first second the newly evolved black hole has  $\sim 10^{53}$  erg of rotational energy available to power these. The time scale for delivery of this energy depends (inversely quadratically) on the magnitude of the magnetic field in the neighborhood of the black hole, essentially that on the inner accretion disk. The developing supernova

explosion disrupts the accretion disk; this removes the magnetic fields anchored in the disk, and self-limits the energy the B-Z mechanism can deliver.

## 5. An even more schematic model

Here we calculate the energy available in a rotating black hole just after its birth (before accretion adds more). Our model is to take a  $1.5 M_{\odot}$  neutron star which co-rotates with the inner edge of the accretion disk in which it is embedded. The neutron star then collapses to a black hole, conserving its angular momentum. Since the accretion disk is neutrino cooled, but perhaps not fully thin, its angular velocity will be somewhere between the HADAF value and the Keplerian value. We parameterize it as  $\Omega = \omega \Omega_K$ , where  $\omega = 1$  for Keplerian and  $\omega = \sqrt{2/7} \sim 0.53$  for the HADAF.

The moment of inertia,  $I$ , of a neutron star is well fitted for many different equations of state with the simple expression

$$I = \frac{0.21 MR^2}{1 - 2GM/Rc^2} \quad (18)$$

(Lattimer & Prakash, 2000). With  $J = \omega I \Omega_K$  and a neutron star of  $1.5 M_{\odot}$ , with a radius of 10 km, we find

$$\tilde{a}^2 = \left( \frac{Jc}{GM^2} \right)^2 = 0.64\omega^2. \quad (19)$$

We choose  $\omega = 1.0$  to roughly reproduce the MacFadyen & Woosley value of  $\tilde{a}$ , see our Fig. 3. We do not really believe the disk to be so efficiently neutrino cooled that its angular velocity is Keplerian; i.e.  $\omega = 1$ , but it may be not far from it. Our  $\omega$  should be more properly viewed as a fudge factor which allows us to match the more complete MacFadyen & Woosley calculation. MacFadyen & Woosley find that, while the accretion disk onto the black hole is forming, an additional solar mass of material is added to it “as the dense stellar core collapses through the inner boundary at all polar angles”. We shall add this to our  $1.5 M_{\odot}$  and take the black hole mass to be  $2.5 M_{\odot}$ . We neglect the increase in spin of the black hole by the newly accreted matter; this is already included in the

MacFadyen & Woosley results. For  $\tilde{a}^2 = 0.64$  we find  $f(\tilde{a}^2) = 0.11$ , so that the black hole rotation energy becomes

$$E_{BZ} = 1.5 \times 10^{53} \text{ erg} \quad (20)$$

in rough agreement with the estimates of MacFadyen & Woosley in the last section.

## 6. Previous models

### 6.1. Collapsar

We have not discussed the Collapsar model of Woosley (1993), and MacFadyen & Woosley (1999). In this model the center of a rotating Wolf-Rayet star evolves into a black hole, the outer part being held out by centrifugal force. The latter evolves into an accretion disk and then by hypercritical accretion spins the black hole up. MacFadyen & Woosley point out that "If the helium core is braked by a magnetic field prior to the supernova explosion to the extent described by Spruit & Phinney (1998) then our model will not work for single stars." Spruit & Phinney argue that magnetic fields maintained by differential rotation between the core and envelope of the star will keep the whole star in a state of approximately uniform rotation until 10 years before its collapse. As noted in the last section, with the extremely high magnetic fields we need the viscosity would be expected to be exceptionally high, making the Spruit & Phinney scenario probable. Livio & Pringle (1998) have commented that one finds evidence in novae that the coupling between layers of the star by magnetic fields may be greatly suppressed relative to what Spruit & Phinney assumed. However, we note that even with this suppressed coupling, they find pulsar periods from core collapse supernovae no shorter than 0.1 s. Independent evidence for the fact that stellar cores mostly rotate no faster than this comes from the study of supernova remnants: Bhattacharya (1990, 1991) concludes that the absence of bright, pulsar-powered plerions in most SNRs indicates that typically pulsar spin periods at birth are no shorter than 0.03–0.05 s. Translated to our black holes, such spin periods would imply  $\tilde{a} \leq 0.01$ , quite insufficient to power a

GRB. As a cautionary note, we might add that without magnetic coupling the cores of evolved stars can spin quite rapidly (Heger et al., 2000). This rapid initial spin may be reconciled with Bhattacharya's limit if r-mode instabilities cause very rapid spin-down in the first few years of the life of a neutron star (e.g., Heger et al., 2000; Lindblom & Owen, 1999).

### 6.2. Coalescing low-mass black holes and helium stars

Fryer & Woosley (1998) suggested the scenario of a black hole spiraling into a helium star. This is an efficient way to spin up the black hole.

Bethe & Brown (1998) evolved low-mass black holes with helium star companion, as well as binaries of compact objects. In a total available range of binary separation  $0.04 < a_{13} < 4$ , low-mass black-hole, neutron-star binaries were formed when  $0.5 < a_{13} < 1.4$  where  $a_{13}$  is the initial binary separation in units of  $10^{13}$  cm. The low-mass black hole coalesces with the helium star in the range  $0.04 < a_{13} < 0.5$ . Binaries were distributed logarithmically in  $a$ . Thus, coalescences are more common than low-mass black-hole, neutron-star binaries by a factor of  $\ln(0.5/0.04)/\ln(1.9/0.5) = 1.9$ .

In Bethe & Brown (1998), the He-star, compact-object binary was disrupted ~50% of the time by the He-star explosion. This does not apply to the coalescence. Thus, the rate of low-mass black-hole, He-star mergers is 3.8 times the formation rate of low-mass black-hole, neutron-star binaries, or

$$R = 3.8 \times 10^{-4} \text{ yr}^{-1} \quad (21)$$

in the Galaxy. The estimated empirical rate of GRBs, with a factor of 100 for beaming, is  $10^{-5} \text{ yr}^{-1}$  in the Galaxy (Appendix C of Brown et al. (1999a)). Thus, the number of progenitors is more than adequate.

In Bethe & Brown (1998) the typical black hole mass was  $\sim 2.4 M_\odot$ , somewhat more massive than their maximum assumed neutron star mass of  $1.5 M_\odot$ . As it enters the helium star companion an accretion disk is soon set up and the accretion scenario will follow that described above, with rotating black holes of various masses formed. Brown et al. (2000b) find that the black hole will be spun up quickly. We have not pursued this scenario

beyond the point that it was developed by Fryer & Woosley (1998).

## 7. Soft X-ray transients as relics of hypernovae and GRB

### 7.1. Our model: angular momentum

We favor a model of hypernovae similar to MacFadyen & Woosley (1999) in that it involves a failed supernova as a centerpiece. But, in distinction to MacFadyen & Woosley, our initial system is a binary, consisting of a massive star A (which will later become the failed SN) and a lighter companion B, which serves to provide ample angular momentum.

Failed supernovae require a ZAMS mass of 20–35 M<sub>⊙</sub>, according to the calculations of Woosley & Weaver (1995) as interpreted by Brown et al. (1999b). The limits 20 and 35 M<sub>⊙</sub> are not accurately known, but it is a fairly narrow range, so we shall in many of our calculations assume a “typical” ZAMS mass of 25 M<sub>⊙</sub>. The heavy star A must not be in a close binary because then its hydrogen envelope would be removed early in its evolution and therefore the star would lose mass by wind at a very early stage and become a low-mass compact object (Brown et al., 1996). Instead, we assume a wide binary, with a separation,  $a$  in the range

$$a = 500 - 1000 R_{\odot}, \quad (22)$$

so star A evolves essentially as a single star through its first few burning stages. It is essential that most of the He core burning is completed before its hydrogen envelope is removed (Heger & Wellstein, 2000; Wellstein & Langer, 1999). We assume the initial distance  $a$  between the two stars to be in this range. When star A fills its Roche lobe, the companion, star B, will spiral inwards.

The initiation and early development of the common envelope has been best treated by Rasio & Livio (1996). This is the only phase that can at present be modeled in a realistic way. They find a short viscous time in the envelope, but emphasize that numerical viscosity may play an important role in their results. However, we believe the viscosity to

be large. Torkelsson et al. (1996) showed the Shakura–Sunyaev viscosity parameter (Shakura & Sunyaev, 1973),  $\alpha_{SS}$ , to range from 0.001 to 0.7, with the higher values following from the presence of vertical magnetic fields. Since in our Blandford–Znajek model extremely high magnetic fields  $\sim 10^{15}$  G are needed in the He envelope to deliver the energy rapidly, we believe  $\alpha_{SS}$  to be not much less than unity. Given such high viscosities, it seems reasonable to follow the Rasio–Livio extrapolation, based on a short viscous transport time, to later times. The most significant new result of Rasio & Livio “is that, during the dynamical phase of common envelope evolution, a corotating region of gas is established near the central binary. The corotating region has the shape of an oblate spheroid encasing the binary (i.e., the corotating gas is concentrated in the orbital plane).”

A helium core, which we deal with, is not included in their calculations, because they do not resolve the inner part of the star numerically. However, since the physics of the spiral-in does not really change as it proceeds past the end of their calculations, it seems most likely that during further spiraling, the spin-up of material inside the orbit of the companion will continue to be significant.

Star B will stop spiraling in when it has ejected the H envelope of A. Since we assume that all stars A have about the same mass, and that  $a_f$  is very large, we expect

$$\frac{M_B}{a_f} \approx \text{const.} \quad (23)$$

From Section 7.2 we conclude that  $a_f$  is a few R<sub>⊙</sub> for  $M_B = (0.4 - 1) M_{\odot}$ . Now the He cores of stars of ZAMS mass  $M = 20 - 35 M_{\odot}$  have a radius about equal to R<sub>⊙</sub>. Therefore small  $M_B$  stars will spiral into the He core of A. There they cannot be stopped but will coalesce with star A. However, they will have transmitted their angular momentum to star A.

Star B of larger mass will stop at larger  $a_f \gg R_{\odot}$ . It is then not clear whether they will transfer all of their angular momentum to star A. In any case, they must generally wait until they evolve off the main sequence into the subgiant or possibly even the giant stage before they can fill their Roche Lobes and later accrete onto the black hole resulting from star A.

The Kepler velocity of star B at  $a_f$  is

$$V_K^2 = G \frac{M_{A,f}}{a_f}. \quad (24)$$

We estimate the final mass of A, after removal of its hydrogen envelope, to be about  $10 M_\odot$ ; then

$$V_K \approx 1.2 \times 10^8 a_{f,11}^{-1/2} \text{ cm s}^{-1}, \quad (25)$$

where  $a_{f,11}$  is  $a_f$  in units of  $10^{11}$  cm. The specific angular momentum of B is then

$$j(B) = a_f V_K = 1.2 \times 10^{19} a_{f,11}^{1/2} \text{ cm}^2 \text{s}^{-1}. \quad (26)$$

If B and A share their angular momentum, the specific angular momentum is reduced by a factor  $M_B/(M_{A,f} + M_B)$  which we estimate to be  $\sim 0.1$ . Since  $a_f$  should be  $\gtrsim 3 R_\odot$  (See Table 1), the specific angular momentum of A should be

$$j(A) \gtrsim 10^{18} \text{ cm}^2 \text{s}^{-1}. \quad (27)$$

Star B has now done its job and can be disregarded.

## 7.2. Supernova and collapse

Star A now goes through its normal evolution, ending up as a supernova. But since we have chosen its mass to be between 20 and  $35 M_\odot$ , the SN shock cannot penetrate the heavy envelope but is stopped at some radius

$$R_{SN} \approx 10^{10} \text{ cm}, \quad (28)$$

well inside the outer edge of the He envelope. We estimate  $R_{SN}$  by scaling from SN 1987A; in that supernova, with progenitor mass  $\sim 18 M_\odot$ , most of the He envelope was returned to the galaxy. The separation between compact object and ejecta was estimated to occur at  $R \sim 5 \times 10^8$  cm (Bethe, 1990; Woosley, 1988) at mass point  $1.5 M_\odot$  (gravitational). Woosley & Weaver (1995) find remnant masses of  $\sim 2 M_\odot$ , although with large fluctuations, for ZAMS masses in the range  $20\text{--}35 M_\odot$ , which go into high-mass black holes. From Table 3 of Brown et al. (1996) we see that fallback between  $R = 3.5$  and  $4.5 \times 10^8$  cm is  $0.03 M_\odot$ . Using this we can extrapolate to  $R = 10^{10}$  cm as the distance within which matter has to begin falling in immediately in our heavier stars, to make up a compact object of  $2 M_\odot$ .

Unlike in 1987A the shock energy in the more massive star does not suffice to eject the envelope beyond this point, and the remaining outer envelope will also eventually fall back.

At  $R_{SN}$ , the specific angular momentum of Kepler motion around a central star of mass  $10 M_\odot$  is, cf. Eq. (26)

$$\begin{aligned} j_K(10 M_\odot) &= 1.2 \times 10^{19} R_{f,11}^{1/2} \text{ cm}^2 \text{s}^{-1} \\ &= 4 \times 10^{18} \text{ cm}^2 \text{s}^{-1}. \end{aligned} \quad (29)$$

In reality, at this time the central object has a mass  $M \sim 1.5 M_\odot$  (being a neutron star) and since  $j_K \sim V_K \sim M^{1/2}$

$$j_K(1.5 M_\odot) = 1.5 \times 10^{18} \text{ cm}^2 \text{s}^{-1}. \quad (30)$$

The angular momentum inherent in star A, Eq. (27), is therefore greater than the Kepler angular momentum. This would not be the case had our initial object been a single star, a collapsar. (The collapsar may work none the less, but our binary model is more certain to work.)

The supernova material is supported by pressure inside the cavity, probably mostly due to electromagnetic radiation. The cavity inside  $R_{SN}$  is rather free of matter. After a while, the pressure in the cavity will reduce. This may happen by opening toward the poles, in which case the outflowing pressure will drive out the matter near the poles and create the vacuum required for the gamma ray burst. Reduction of pressure will also happen by neutrino emission. As the pressure gets reduced, the SN material will fall in toward the neutron star in the center. But because the angular momentum of the SN material is large (Eq. (27)) the material must move more or less in Kepler orbits; i.e., it must spiral in. This is an essential point in the theory.

If  $j(A)$  is less than  $j_K$  at  $R_{SN}$ , the initial motion will have a substantial radial component in addition to the tangential one. But as the Kepler one decreases, cf. Eq. (29), there will come a point of  $r$  at which  $j_K = j(A)$ . At this point an accretion disk will form, consisting of SN material spiraling in toward the neutron star. The primary motion is circular, but viscosity will provide a radial component inward

$$v_r \sim \alpha v_K, \quad (31)$$

where  $\alpha$  is the viscosity parameter. It has been

argued by Brandenburg et al. (1996) that  $\alpha \sim 0.1$  in the presence of equipartition magnetic fields perpendicular to the disk, and it may be even larger with the high magnetic fields required for GRBs. Narayan & Yi (1994) have given analytical solutions for such accretion disks. The material will arrive at the neutron star essentially tangentially, and therefore its high angular momentum will spin up the neutron star substantially. Accretion will soon make the neutron star collapse into a black hole. The angular momentum will be conserved, so the angular velocity is increased since the black hole has smaller radius than the neutron star. Thus the black hole is born with considerable spin.

A large fraction of the material of the failed supernova will accrete onto the black hole, giving it a mass of order  $7 M_{\odot}$ . All this material adds to the angular momentum of the black hole since all of it has the Kepler velocity at the black hole radius. Our estimates show that the black hole would be close to an extreme Kerr hole (Section 5), were it to accrete all of this material. It may, however, be so energetic that it drives off part of the envelope in the explosion before it can all accrete (see Section 5).

### 7.3. Soft X-ray transients with black-hole primaries

Nine binaries have been observed which are black-hole X-ray transients. All contain a high-mass black hole, of mass  $\sim 7 M_{\odot}$ . In seven cases the lower-mass companion (star B) has a mass  $\leq M_{\odot}$ . The two stars are close together, their distance being of order  $5 R_{\odot}$ . Star B fills its Roche Lobe, so it spills over some material onto the black hole. The accretion disk near the black hole emits soft X rays. Two of the companions are subgiants, filling their Roche lobes at a few times larger separations from the black hole.

In fact, however, the accretion onto the central object is not constant, so there is usually no X-ray emission. Instead, the material forms an accretion disk around the black hole, and only when enough material has been assembled, it falls onto the black hole to give observable X rays. Hence, the X-ray source is transient. Recent observation of a large space velocity of Cygnus X-1 (Nelemans et al., 1999) suggests that it has evolved similarly to the transient sources, with the difference that the companion to the black hole is an  $\sim 18 M_{\odot}$  O star. The latter pours enough matter onto the accretion disk so that Cyg X-1 shines continuously. We plan to describe the evolution of Cyg X-1 in a future paper (Brown et al., 2000a).

Table 1 is an abbreviated list of data on transient sources. A more complete table is given in Brown et al. (1999b). Two of the steady X-ray sources, in the LMC, have been omitted, because we believe the LMC to be somewhat special because of its low metallicity; also masses, etc., of these two are not as well measured. Of the others, 6 are main-sequence K stars, one is main-sequence M, and the other two have masses greater than the Sun. The masses given are geometric means of the maximum and minimum masses given by the observers. The distance  $a$  between the black hole and the optical (visible) star is greater for the heavier stars than for the K- and M stars (except the more evolved one of them) as was expected in Section 7.1 for the spiraling in of star B. The table also gives the radius of the Roche Lobe and the specific orbital angular momentum of star B.

Five K stars have almost identical distance  $a \sim 5 R_{\odot}$ , and also Roche Lobe sizes,  $\sim 1.0 R_{\odot}$ . These Roche Lobes can be filled by K stars on the main sequence. The same is true for the M star. Together, K and M stars cover the mass range from 0.3 to  $1 M_{\odot}$ . The two heavier stars have Roche Lobes of 3

**Table 1**  
Properties of transient X-ray sources

Spectral type	$M_B [M_{\odot}]$	$a [R_{\odot}]$	$R_L [R_{\odot}]$	$j_{B,\text{orb}} [10^{19} \text{ cm}^2 \text{ s}^{-1}]$
5 K-type	0.4–0.9	$4.5 \pm 0.8$	$0.9 \pm 0.2$	$1.5 \pm 0.6$
1 K-type	0.8	34	6.1	5.7
1 M-type	0.5	3.1	0.6	1.5
F (Nova Scorpii 1994)	2.2	16	4.7	1.9
A2 (1543-47)	2.0	9.0	2.6	1.4

and  $5 R_{\odot}$  which cannot possibly be filled by main-sequence stars of mass  $\sim 2 M_{\odot}$ . We must therefore assume that these stars are subgiants, in the Herzsprung gap. These stars spend only about 1% of their life as subgiants, so we must expect that there are many “silent” binaries in which the  $2 M_{\odot}$  companion has not yet evolved off the main sequence and sits well within its Roche lobe, roughly 100 times more. The time as subgiants is even shorter for more massive stars; this explains their absence among the transient sources.

Therefore we expect a large number of “silent partners”: stars of more than  $1 M_{\odot}$ , still on their main sequence, which are far from filling their Roche Lobe and therefore do not transfer mass to their black hole partners. In fact, we do not see any reason why the companion of the black hole could not have any mass, up to the ZAMS mass of the progenitor of the black hole; it must only evolve following the formation of the black hole. It then crosses the Herzsprung gap in such a short time, less than the thermal time scale, that star A cannot accept the mass from the companion, so that common envelope evolution must ensue. If we include these ‘silent partners’ in the birth rate, assuming a flat mass ratio distribution, we enhance the total birth rate of black-hole binaries by a factor 25 over the calculations by Brown et al. (1999b).

On the lower mass end of the companions, there is only one M star. This is explained in terms of the model of Section 7.1 by the fact that stars of low mass will generally spiral into the He core of star A, and will coalesce with A, see below Eq. (23), so no relic is left. (Since the core is left spinning rapidly, these complete merger cases could also be suitable GRB progenitors.) As the outcome of the spiral-in depends also on other factors, such as the initial orbital separation and the primary mass, one may still have an occasional survival of an M star binary (note that the one M star companion is M0, very nearly in the K star range).

The appearance of the black hole transient X-ray binaries is much like our expectation of the relic of the binary which has made a hypernova: a black hole of substantial mass, and an ordinary star, possibly somewhat evolved, of smaller mass. We expect that star B would stop at a distance  $a_f$  from star A which is greater if the mass of B is greater (see Section 7.1). This is just what we see in the black-hole

binaries: the more massive companion stars ( $\sim 2 M_{\odot}$ ) are further from the black hole than the K stars. We also note that the estimated birth rate of these binaries is high enough for them to be the progenitors of GRB, even if only in a modest fraction of them the conditions for GRB powering are achieved.

#### 7.4. Nova Scorpii 1994 (GRO J1655-40)

Nova Sco 1994 is a black hole transient X-ray source. It consists of a black hole of  $\sim 7 M_{\odot}$  and a subgiant of about  $2 M_{\odot}$ . Their separation is  $17 R_{\odot}$ . Israeli et al. (1999) have analyzed the spectrum of the subgiant and have found that the  $\alpha$ -particle nuclei O, Mg, Si and S have abundances 6 to 10 times the solar value. This indicates that the subgiant has been enriched by the ejecta from a supernova explosion; specifically, that some of the ejecta of the supernova which preceded the present Nova Sco (a long time ago) were intercepted by star B, the present subgiant. Israeli et al. (1999) estimate an age since accretion started from the assumption that enrichment has only affected the outer layers of the star. We here reconsider this: the time that passed since the explosion of the progenitor of the black hole is roughly the main-sequence lifetime of the present subgiant companion, which given its mass of  $\sim 2 M_{\odot}$  will be about 1 Gyr. This is so much longer than any plausible mixing time in the companion that the captured supernova ejecta must by now be uniformly mixed into the bulk of the companion. This rather increases the amount of ejecta that we require the companion to have captured. (Note that the accretion rate in this binary is rather less than expected from a subgiant donor, though the orbital period leaves no doubt that the donor is more extended than a main-sequence star (Regos et al., 1998). It is conceivable that the high metal abundance has resulted in a highly non-standard evolution of this star, in which case one might have to reconsider its age.)

The presence of large amounts of S is particularly significant. Nomoto et al. (2000) have calculated the composition of a hypernova from an  $11 M_{\odot}$  CO core, see Fig. 4. This shows substantial abundance of S in the ejecta. Ordinary supernovae produce little of this element, as shown by the results of Nomoto et al. (2000) in Fig. 4. The large amount of S, as well

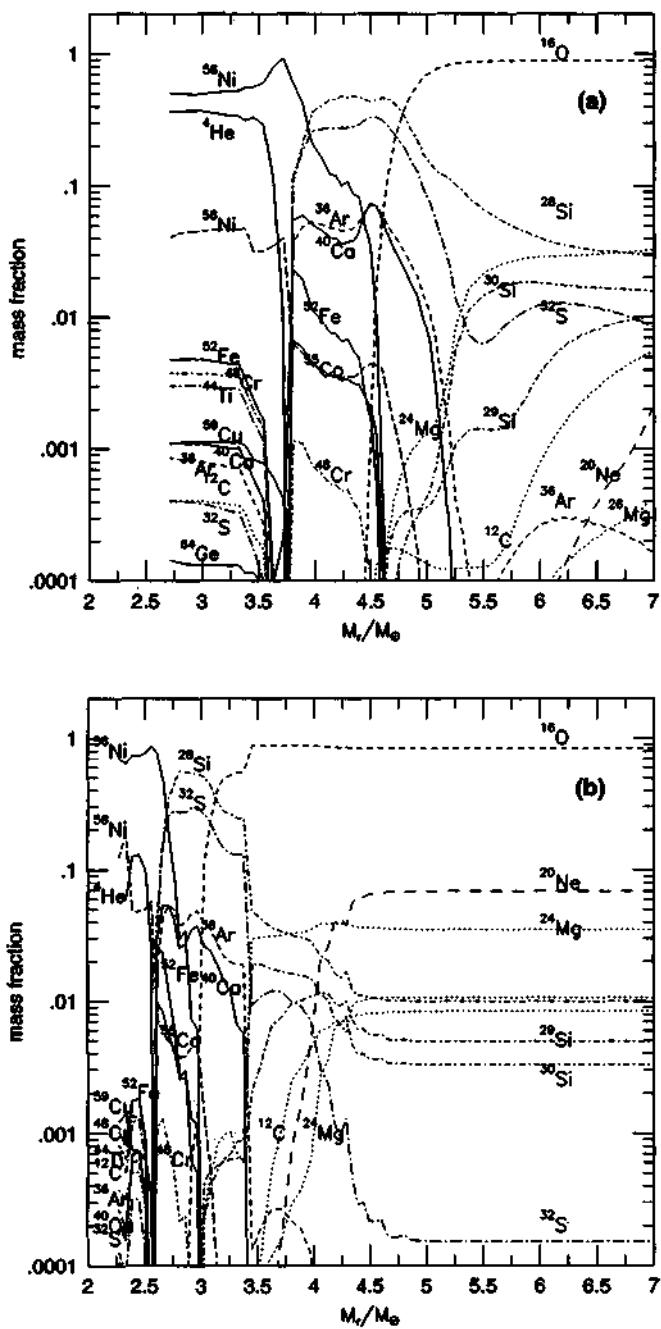


Fig. 4. The isotopic composition of ejecta of the hypernova ( $E_k = 3 \times 10^{52}$  erg; upper) and the normal supernova ( $E_k = 1 \times 10^{51}$  erg; lower) for a  $16 M_\odot$  He star, from Nomoto et al. (2000). Note the much higher sulphur abundance in the hypernova.

as O, Mg and Si we consider the strongest argument for considering Nova Sco 1994 as a relic of a hypernova, and for our model, generally.

Fig. 4 also shows that  $^{56}\text{Ni}$  and  $^{52}\text{Fe}$  are confined to the inner part of the hypernova, and if the cut between black hole and ejecta is at about  $5 M_{\odot}$ , there will be no Fe-type elements in the ejecta, as observed in Nova Scorpii 1994. By contrast hypernova 1998bw shows a large amount of Ni, indicating that in this case the cut was at a lower included mass.

The massive star A in Nova Sco will have gone through a hypernova explosion when the F-star B was still on the main sequence, its radius about  $1.5 R_{\odot}$ . Since the explosion caused an expansion of the orbit, the orbital separation  $a$  was smaller at the time of the supernova than it is now, roughly by a factor

$$a_{\text{then}} = a_{\text{now}} / (1 + \Delta M / M_{\text{now}}). \quad (32)$$

( $\Delta M$  is the mass lost in the explosion; see, e.g., Verbunt et al. (1990)). With  $\Delta M \sim 0.8 M_{\text{now}}$ , as required by the high space velocity, this means  $a_{\text{then}} = 10 R_{\odot}$ . Therefore the fraction of solid angle subtended by the companion at the time of explosion was

$$\frac{\Omega}{4\pi} = \frac{\pi(1.5 R_{\odot})^2}{4\pi(10 R_{\odot})^2} \approx 6 \times 10^{-3}. \quad (33)$$

Assuming the ejecta of the hypernova to have been at least  $5 M_{\odot}$  (Nelemans et al., 1999), the amount deposited on star B was

$$M_B \gtrsim 0.03 M_{\odot}. \quad (34)$$

The solar abundance of oxygen is about 0.01 by mass, so with the abundance in the F star being 10 times solar, and oxygen uniformly mixed, we expect  $0.1 \times 2.5 = 0.25 M_{\odot}$  of oxygen to have been deposited on the companion, much more than the total mass it could have captured from a spherically symmetric supernova. [Si/O] is 0.09 by mass in the Sun, and [S/O] is 0.05, so since the over-abundances of all three elements are similar we expect those ratios to hold here, giving about  $0.02 M_{\odot}$  of captured Si and  $0.01 M_{\odot}$  of captured S. We therefore need a layer of stellar ejecta to have been captured which has twice as much Si as S, at the same time as having about 10 times more O. From Fig. 4, we see that this occurs nowhere in a normal supernova, but

does happen in the hypernova model of Nomoto et al. (2000) at mass cuts of  $6 M_{\odot}$  or more. This agrees very nicely with the notion that a hypernova took place in this system, and that the inner  $7 M_{\odot}$  or so went into a black hole.

What remains is to explain how the companion acquired ten times more mass than the spherical supernova model allows, and once again we believe that the answer is given in recent hypernova calculations (MacFadyen & Woosley, 1999; Wheeler et al., 2000): hypernovae are powered by jet flows, which means they are very asymmetric, with mass outflow along the poles being much faster and more energetic than along the equator. The disk provides a source for easily captured material in two ways: First, it concentrates mass in the equatorial plane, which will later be ejected mostly in that plane. Second, the velocity acquired by the ejecta is of the order of the propagation speed of the shock through it. This propagation speed is proportional to  $\sqrt{P_2/\rho_1}$ , where  $P_2$  is the pressure behind the shock and  $\rho_1$  the density ahead of it. The driving pressure will be similar in all directions (or larger, due to the jet injection, in the polar regions), whereas the disk density is much higher than the polar density. Hence, the equatorial ejecta will be considerably slower than even normal supernova ejecta, greatly increasing the possibility of their capture by the companion. Other significant effects of the disk/jet geometry are (1) that the companion is shielded from ablation of its outer layers by fast ejecta, which is thought to occur in spherical supernovae with companion stars (Marietta et al., 2000) and (2) that there is no iron enrichment of the companion, because the iron—originating closest to the center—is either all captured by the black hole or ejected mainly in the jet, thus not getting near the companion (Wheeler et al. (2000); note that indeed no overabundance of Fe is seen in the companion of GRO J1655-40).

For the companion to capture the required  $0.2-0.3 M_{\odot}$  of ejecta it is sufficient that the ejecta be slow enough to become gravitationally bound to it. However, the material may not stay on: when the companion has so much mass added on a dynamical time scale it will be pushed out of thermal equilibrium, and respond by expanding, as do main-sequence stars that accrete mass more gradually on a time scale faster than their thermal time scale (e.g., Kippenhahn & Meyer-Hofmeister, 1977). During

this expansion, which happens on a time scale much longer than the explosion, the star may expand beyond its Roche lobe and transfer some of its mass to the newly formed black hole. However, because the dense ejecta mix into the envelope on a time scale between dynamical and thermal, i.e., faster than the expansion time, this back transfer will not result in the bulk of the ejecta being fed back, though probably the material lost is still richer in heavy elements than the companion is now. Since the outer layers of the star are not very dense, and the mass transfer is not unstable because the black hole is much more massive than the companion, the total amount of mass transferred back is probably not dramatic. However, the expansion does imply that the pre-explosion mass of the companion was somewhat higher than its present mass, and that the amount of ejecta that needs to be captured in order to explain the abundances observed today is also somewhat higher than the present mass of heavy elements in the companion.

A further piece of evidence that may link Nova Sco 1994 to our GRB/hypernova scenario are the indications that the black hole in this binary is spinning rapidly. Zhang et al. (1997) argue from the strength of the ultra-soft X-ray component that the black hole is spinning near the maximum rate for a Kerr black hole. However, studies by Sobczak et al. (1999) show that it must be spinning with less than 70% maximum. Gruzinov (1999) finds the inferred black hole spin to be about 60% of maximal from the 300 Hz QPO. Our estimates of the last section indicate that enough rotational energy will be left in the black hole so that it will still be rapidly spinning.

We have already mentioned the unusually high space velocity of  $-150 \pm 19 \text{ km s}^{-1}$ . Its origin was first discussed by Brandt et al. (1995), who concluded that significant mass must have been lost in the formation of the black hole in order to explain this high space velocity: it is not likely to acquire a substantial velocity in its own original frame of reference, partly because of the large mass of the black hole. But the mass lost in the supernova explosion is ejected from a moving object and thus carries net momentum. Therefore, momentum conservation demands that the center of mass of the binary acquire a velocity; this is the Blaauw–Boersma kick (Blaauw, 1961; Boersma, 1961). Note that the F-star companion mass is the largest among the

black-hole transient sources, so the center of mass is furthest from the black hole and one would expect the greatest kick. Nelemans et al. (1999) estimate the mass loss in this kick to be  $5\text{--}10 M_{\odot}$ .

In view of the above, we consider it well established that Nova Sco 1994 is the relic of a hypernova. We believe it highly likely that the other black-hole transient X-ray sources are also hypernova remnants. We believe it likely that the hypernova explosion was accompanied by a GRB if, as in GRB 980326, the energy was delivered in a few seconds. It is not clear what will happen if the magnetic fields are so low that the power is delivered only over a much longer time. There could then still be intense power input for a few seconds due to neutrino annihilation deposition near the black hole (Janka et al., 1999), but that may not be enough for the jet to pierce through the He star and cause a proper GRB (MacFadyen & Woosley, 1999). At this point, we recall that the GRB associated with SN 1998bw was very sub-luminous,  $10^5$  times lower than most other GRB. While it has been suggested that this is due to us seeing the jet sideways, it is in our view more likely that the event was more or less spherical (Kulkarni et al., 1998) and we see a truly lower-power event. A good candidate would be the original suggestion by Colgate (1968, 1974) of supernova shock break-out producing some gamma rays. Indications are that the expansion in SN 1998bw was mildly relativistic (Kulkarni et al., 1998) or just sub-relativistic (Waxman & Loeb, 1999). In either case, what we may have witnessed is a natural intermediate event in our scenario: we posit that there is a continuum of events varying from normal supernovae, delivering 1 or more or less spherically in ten seconds, to extreme hypernovae/GRB that deliver 100 or more in a highly directed beam. In the middle, there will be cases where the beam cannot pierce through the star, but the total energy delivered is well above a supernova, with as net result a hypernova accompanied by a very weak GRB.

### 7.5. Numbers

Nearly all observed black hole transient X-ray sources are within 5 kpc of the Sun. Extrapolating to the entire Galaxy, a total of 8,800 black-hole trans-

ients with main-sequence K companions has been suggested (Brown et al., 1999b).

The lifetime of a K star in a black hole transient X-ray source is estimated to be  $\sim 10^{10}$  yr (Van Paradijs, 1996) but we shall employ  $10^9$  yr for the average of the K-stars and the more massive stars, chiefly those in the “silent partners”. In this case the birth rate of the observed transient sources would be

$$\lambda_K = 10^4 / 10^9 = 10^{-5} \text{ per galaxy yr}^{-1}. \quad (35)$$

We see no reason why low-mass companions should be preferred, so we assume that the formation rate of binaries should be independent of the ratio

$$q = M_{B,i}/M_{A,i}. \quad (36)$$

In other discussions of binaries, e.g., in Portegies Zwart & Yungelson (1998), it has often been assumed that the distribution is uniform in  $q$ . This is plausible but there is no proof. Since all primary masses  $M_A$  are in a narrow interval, 20 to  $35 M_\odot$ , this means that  $M_B$  is uniformly distributed between zero and some average  $M_A$ , let us say  $25 M_\odot$ . Then the total rate of creation of binaries of our type is

$$\lambda = \frac{25}{0.7} \lambda_K = 3 \times 10^{-4} \text{ galaxy}^{-1} \text{ yr}^{-1}. \quad (37)$$

This is close to the rate of mergers of low mass black holes with neutron stars which Bethe & Brown (1998) have estimated to be

$$\lambda_m = 2 \times 10^{-4} \text{ galaxy}^{-1} \text{ yr}^{-1}. \quad (38)$$

These mergers have been associated speculatively with short GRBs, while formation of our binaries is supposed to lead to “long” GRBs (Fryer et al., 1999). We conclude that the two types of GRB should be equally frequent, which is not inconsistent with observations. In absolute number both of our estimates Eqs. (37) and (38) are substantially larger than the observed rate of  $10^{-7}$  galaxy $^{-1}$  yr $^{-1}$  (Wijers et al., 1998); this is natural, since substantial beaming is expected in GRBs produced by the Blandford-Znajek mechanism. Although we feel our mechanism to be fairly general, it may be that the magnetic field required to deliver the BZ energy within a suitable time occurs in only a fraction of the He cores.

## 8. Discussion and conclusion

Our work here has been based on the Blandford-Znajek mechanism of extracting rotational energies of black holes spun up by accreting matter from a helium star. We present it using the simple circuitry of ‘The Membrane Paradigm’ (Thorne et al., 1986). Energy delivered into the loading region up the rotational axis of the black hole is used to power a GRB. The energy delivered into the accretion disk powers a SN Ib explosion.

We also discussed black-hole transient sources, high-mass black holes with low-mass companions, as possible relics for both GRBs and Type Ib supernova explosions, since there are indications that they underwent mass loss in a supernova explosion. In Nova Sco 1994 there is evidence from the atmosphere of the companion star that a very powerful supernova explosion (‘hypernova’) occurred.

We estimate the progenitors of transient sources to be formed at a rate of 300 GEM (Galactic Events per Megayear). Since this is much greater than the observed rate of GRBs, there must be strong collimation and possible selection of high magnetic fields in order to explain the discrepancy.

We believe that there are strong reasons that a GRB must be associated with a black hole, at least those of duration several seconds or more discussed here. Firstly, neutrinos can deliver energy from a stellar collapse for at most a few seconds, and sufficient power for at most a second or two. Our quantitative estimates show that the rotating black hole can easily supply the energy as it is braked, provided the ambient magnetic field is sufficiently strong. The black hole also solves the baryon pollution problem: we need the ejecta that give rise to the GRB to be accelerated to a Lorentz factor of 100 or more, whereas the natural scale for any particle near a black hole is less than its mass. Consequently, we have a distillation problem of taking all the energy released and putting it into a small fraction of the total mass. The use of a Poynting flux from a black hole in a magnetic field (Blandford & Znajek, 1977) does not require the presence of much mass, and uses the rotation energy of the black hole, so it provides naturally clean power.

Of course, nature is extremely inventive, and we do not claim that all GRBs will fit into the frame-

work outlined here. We would not expect to see all of the highly beamed jets following from the BZ mechanism head on, the jets may encounter some remaining hydrogen envelope in some cases, jets from lower magnetic fields than we have considered here may be much weaker and delivered over longer times, etc., so we speculate that a continuum of phenomena may exist between normal supernovae and extreme hypernovae/GRBs. This is why we call our effort "A Theory of Gamma Ray Bursts" and hope that it will be a preliminary attempt towards systematizing the main features of the energetic bursts.

### Acknowledgements

We would like to thank Stan Woosley for much useful information. Several conversations with Roger Blandford made it possible for us to greatly improve our paper, as did valuable comments from Norbert Langer. This work is partially supported by the U.S. Department of Energy Grant No. DE-FG02-88ER40388. HKL is supported also in part by KOSEF Grant No. 1999-2-112-003-5 and by the BK21 program of the Korean Ministry of Education.

### Appendix A. Estimates of $\epsilon_\Omega = \Omega_{\text{disk}}/\Omega_H$

We collect here useful formulas needed to calculate  $\epsilon_\Omega = \Omega_{\text{disk}}/\Omega_H$ . First of all

$$\begin{aligned}\Omega_H &= \frac{\tilde{a}}{1 + \sqrt{1 - \tilde{a}^2}} \left( \frac{c^3}{2MG} \right) \\ &= \frac{\sqrt{2} \tilde{a}}{1 + \sqrt{1 - \tilde{a}^2}} \left( \frac{r}{R_{\text{Sch}}} \right)^{3/2} \Omega_K,\end{aligned}\quad (\text{A.1})$$

$$\begin{aligned}\Omega_{\text{disk}} &= \omega \Omega_K \left[ 1 + \tilde{a} \frac{GM}{c^2} \sqrt{\frac{GM}{c^2 r^3}} \right]^{-1} \\ &= \omega \Omega_K \left[ 1 + \tilde{a} \left( \frac{R_{\text{Sch}}}{2r} \right)^{3/2} \right]^{-1},\end{aligned}\quad (\text{A.2})$$

where  $\Omega_K = \sqrt{GM/R^3}$  and  $\omega$  is dimensionless parameter ( $0 < \omega < 1$ ). Thus

$$\begin{aligned}\frac{\Omega_{\text{disk}}}{\Omega_H} &= \omega \frac{1 + \sqrt{1 - \tilde{a}^2}}{\sqrt{2} \tilde{a}} \\ &\times \left( \frac{R_{\text{Sch}}}{r} \right)^{3/2} \left[ 1 + \tilde{a} \left( \frac{R_{\text{Sch}}}{2r} \right)^{3/2} \right]^{-1}.\end{aligned}\quad (\text{A.3})$$

The numerical estimates are summarized in Table 2 for various  $\omega$  and radii.

### Appendix B. Spin-up of black holes by accretion

The specific angular momentum and energy of test particles in Keplerian circular motion, with rest mass  $\delta m$ , are

$$\begin{aligned}\tilde{E} &= \frac{E}{\delta m} = c^2 \left[ \frac{r^2 - R_{\text{Sch}}r + a\sqrt{R_{\text{Sch}}r/2}}{r(r^2 - \frac{3}{2}R_{\text{Sch}}r + a\sqrt{2R_{\text{Sch}}r})^{1/2}} \right], \\ \tilde{l} &= \frac{l}{\delta m} \\ &= c \sqrt{\frac{R_{\text{Sch}}r}{2}} \left[ \frac{(r^2 - a\sqrt{2R_{\text{Sch}}r} + a^2)}{r(r^2 - \frac{3}{2}R_{\text{Sch}}r + a\sqrt{2R_{\text{Sch}}r})^{1/2}} \right],\end{aligned}\quad (\text{B.1})$$

where  $R_{\text{Sch}} = 2GM/c^2$  and BH spin  $a = J/Mc = \tilde{a}(GM/c^2)$ . The accretion of  $\delta m$  changes the BH's total mass and angular momentum by  $\Delta M = \tilde{E}\delta m$  and  $\Delta J = \tilde{l}\delta m$ . The radii of marginally bound ( $r_{\text{mb}}$ ) and stable ( $r_{\text{ms}}$ ) orbits are given as

Table 2  
Estimates of  $\epsilon_\Omega = \Omega_{\text{disk}}/\Omega_H$  as a function of spin parameter and radius, where  $r_{\text{mb}}$  is the marginally bound radius and  $r = r_{\text{ms}}$  the marginally stable radius

$\omega$	$\tilde{a}$	$\Omega_{\text{disk}}/\Omega_H$			
		$r = r_{\text{mb}}(\tilde{a})$	$r = r_{\text{ms}}(\tilde{a})$	$r = 2R_{\text{Sch}}$	$r = 3R_{\text{Sch}}$
1.0	0.80	1.00	0.69	0.45	0.26
0.9	0.72	0.99	0.63	0.49	0.27
0.8	0.64	0.93	0.58	0.51	0.29
0.7	0.56	0.89	0.54	0.53	0.30
0.6	0.48	0.84	0.50	0.55	0.31
0.5	0.40	0.80	0.46	0.57	0.32

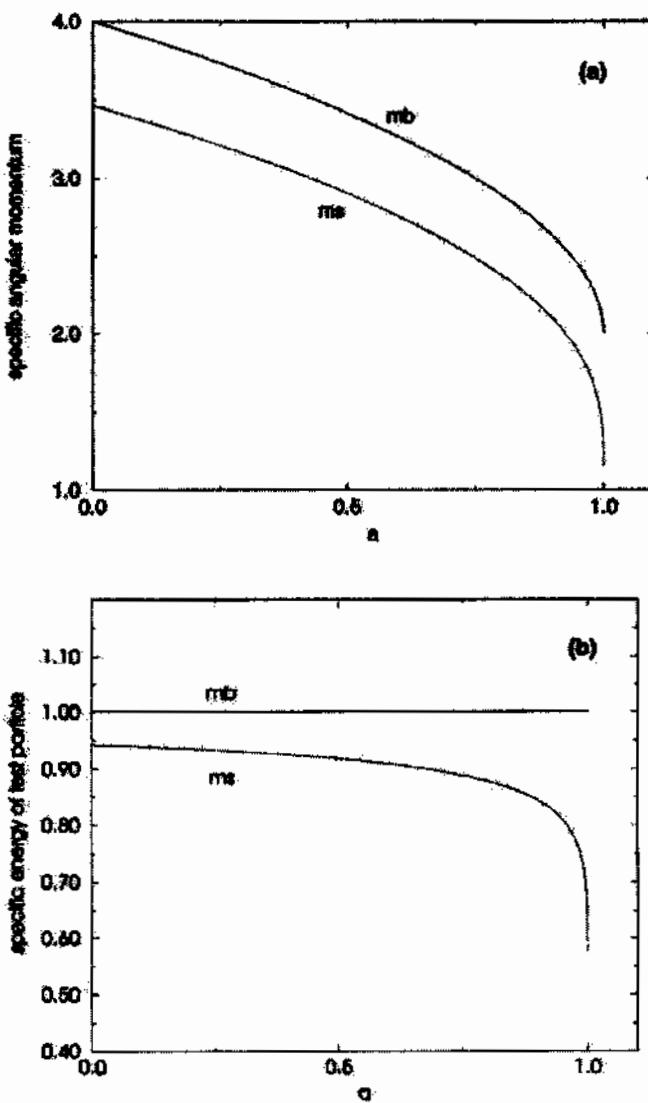


Fig. 5. Specific angular momentum and energy of test particle in units of  $[GM/c]$  and  $[c^2]$ . BH spin  $a$  is given in unit of  $[GM/c^2]$ . For the limiting values at  $a = 0$  and  $GM/c^2$ , refer Table 3.

$$\begin{aligned}
 r_{mb} &= R_{Sch} - a + \sqrt{R_{Sch}(R_{Sch} - 2a)}, & & \\
 r_{ms} &= \frac{R_{Sch}}{2} \{3 + Z_2 - [(3 - Z_1)(3 + Z_1 + 2Z_2)]^{1/2}\}, & & \times \left[ \left(1 + \frac{2a}{R_{Sch}}\right)^{1/3} + \left(1 - \frac{2a}{R_{Sch}}\right)^{1/3} \right], \\
 Z_1 &= 1 + \left(1 - \frac{4a^2}{R_{Sch}^2}\right)^{1/3} & & Z_2 = \left(3 \frac{4a^2}{R_{Sch}^2} + Z_1^2\right)^{1/2}. \tag{B.2}
 \end{aligned}$$

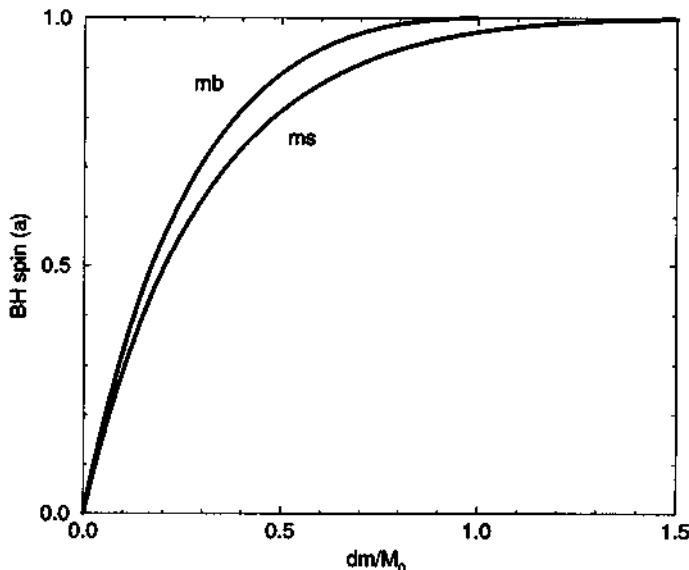


Fig. 6. Spinning up of Black Holes. The BH spin  $a$  is given in units of  $[GM/c^2]$  and  $\delta m$  is the total rest mass of the accreted material.

Table 3

Properties of Schwarzschild & Kerr BH. (a)  $r_{iso} = r_{ms}$  case: 6% (42%) of energy can be released during the spiral-in for Schwarzschild (maximally-rotating Kerr) BHs. (b)  $r_{iso} = r_{mb}$  case: The released energy during the spiral-in is almost zero

$a [GM/c^2]$	$\tilde{r} [GM/c]$	$r [R_{Sch}]$	$\tilde{E} [c^2]$
$r_{ms}$	$2\sqrt{3} \approx 3.46$	3	$\sqrt{8/9} \approx 0.943$
	$2/\sqrt{3} \approx 1.15$	1	$\sqrt{1/3} \approx 0.577$
$r_{mb}$	4	2	1
	2	1	1

The numerical values of the specific angular momentum and energy of test particles are summarized in Table 3 and Fig. 5. In Fig. 6, we test how much mass we need in order to spin up the non-rotating black hole up to given  $\tilde{a}$ . Note that the last stable orbit is almost Keplerian even with the accretion disk, and we assume 100% efficiency of angular momentum transfer from the last stable Keplerian orbit to BH. In order to spin-up the BH up to  $\tilde{a} = 0.9$ , we need ~68% (52%) of original non-rotating BH mass in case of  $r_{iso} = r_{ms}$  ( $r_{mb}$ ). For a very rapidly rotating BH with  $\tilde{a} = 0.99$ , we need 122% and 82%, respectively. For  $r_{iso} = r_{ms}$ , there is an upper limit,  $\tilde{a} =$

0.998, which can be obtained by accretion (Thorne, 1974). In the limit where  $r_{iso} = r_{mb}$ , however, spin-up beyond this limit is possible because the photons can be captured inside thick accretion disk, finally into BH (Abramowicz et al., 1988).

## References

- Abramowicz, M.A., Czerny, B., Lasota, J.P., & Szuszkiewicz, E., 1988, ApJ, 332, 646.
- Andersen, M.I., Castro-Tirado, A.J., Hjorth, J., Moller, P., Pedersen, H., Caon, N., Marina Cairos, L., Korhonen, H., Zapatero Osorio, M.R., Perez, E., & Frontera, F., 1999, Sci, 283, 2075.
- Bethe, H.A., 1990, Rev. Mod. Phys., 62, 801.
- Bethe, H.A. & Brown, G.E., 1998, ApJ, 506, 780.
- Bhattacharya, D., 1990, JApA, 11, 125.
- Bhattacharya, D., 1991, in: J. Ventura and D. Pines, (Eds.), Neutron Stars: Theory and Observation, NATO ASI Ser. C, vol. 344, p. 103.
- Blaauw, A., 1961, Bull. Astron. Inst. Netherlands, 15, 265.
- Blandford, R.D. & Spruit, H., 2000, in preparation.
- Blandford, R.D. & Znajek, R.L., 1977, MNRAS, 179, 433.
- Bloom, J.S., et al., 1999, Natur, 401, 453.
- Boersma, J., 1961, Bull. Astron. Inst. Netherlands, 15, 291.
- Brandenburg, A., Nordlund, Å., Stein, R.F., & Torkelsson, U., 1996, ApJ, 458, L45.

- Brandt, W.N., Podsiadlowski, P., & Sigurdsson, S., 1995, MNRAS, 277, L35.
- Brown, G.E. & Bethe, H.A., 1994, ApJ, 423, 659.
- Brown, G.E., Weingartner, J.C., & Wijers, R.A.M.J., 1996, ApJ, 463, 297.
- Brown, G.E., Lee, C.-H., Wijers, R.A.M.J., & Bethe, H.A., 1999a, Phys. Reports, in press, astro-ph/9910088.
- Brown, G.E., Lee, C.-H., & Bethe, H.A., 1999b, NewA, 4, 313.
- Brown, G.E., Heger, A., Langer, N., Lee, C.-H., Wellstein, S., & Bethe, H.A., 2000a, in preparation.
- Brown, G.E., Lee, C.-H., & Bethe, H.A., 2000b, ApJ, to be published, astro-ph/9909132.
- Colgate, S.A., 1968, Can. J. Phys., 46, 5471.
- Colgate, S.A., 1974, ApJ, 187, 333.
- Eichler, D., Livio, M., Piran, T., & Schramm, D., 1989, Natur, 340, 126.
- Fryer, C.L. & Woosley, S.E., 1998, ApJ, 502, L9.
- Fryer, C.L., Woosley, S.E., & Hartmann, D.H., 1999, ApJ, 526, 152.
- Galama, T., et al., 1998, Natur, 395, 670.
- Galama, T., et al., 2000, ApJ, 536, 185.
- Gruzinov, A., 1999, ApJ, 517, L105.
- Heger, A. & Wellstein, S., 2000, in website of A. Heger.
- Heger, A., Langer, N., & Woosley, S., 2000, ApJ, 528, 368.
- Höflich, P., Wheeler, J.C., & Wang, L., 1999, ApJ, 521, 179.
- Israeli, G., Rebolo, R., Basri, G., Casares, J., & Martin, E.L., 1999, Natur, 401, 142.
- Iwamoto, K., et al., 1998, Natur, 395, 672.
- Janka, H.T., Eberl, T., Ruffert, M., & Fryer, C.L., 1999, ApJ, 527, L39.
- Kippenhahn, R. & Meyer-Hofmeister, E., 1977, A&A, 54, 539.
- Kouveliotou, C., et al., 1998, Natur, 393, 235.
- Kouveliotou, C., et al., 1999, ApJ, 510, L115.
- Kulkarni, S.R., Frail, D.A., Wieringa, M.H., Ekers, R.D., Sadler, E.M., Wark, R.M., Higdon, J.L., Phinney, E.S., & Bloom, J.S., 1998, Natur, 395, 663.
- Lattimer, J.M. & Prakash, M., 2000, ApJ, submitted, astro-ph/0002232.
- Lee, H.-K., Wijers, R.A.M.J., & Brown, G.E., 1999, Phys. Rep., 325, 83.
- Lee, H.-K., Brown, G.E., & Wijers, R.A.M.J., 2000, ApJ, 536, 416.
- Li, L.-X., 2000, PhRvD, 61, 084016.
- Lindblom, L. & Owen, B.J., 1999, PhRvL, 80, 4843.
- Livio, M. & Pringle, J.E., 1998, ApJ, 505, 339.
- Lovelace, R.V.E., MacAuslan, J., & Burns, M., 1979, in: AIP Conf. Ser. 56: Particle Acceleration Mechanisms in Astrophysics, p. 399.
- MacDonald, D. & Thorne, K.S., 1982, MNRAS, 198, 345.
- MacFadyen, A.I. & Woosley, S.E., 1999, ApJ, 524, 262.
- Marietta, E., Burrows, A., & Fryxell, B., 2000, ApJ, in press, astro-ph/9908116.
- Narayan, R. & Yi, I., 1994, ApJ, 428, L13.
- Nelemans, G., Tauris, T.M., & van den Heuvel, E.P.J., 1999, A&A, 352, L87.
- Nomoto, K., Mazzali, P.A., Nakamura, T., Iwamoto, K., Maeda, K., Suzuki, T., Turatto, M., Danziger, I.J., & Patat, F., 2000, in: M. Livio, N. Panagia, & K. Sahu, (Eds.), The Greatest Explosions since the Big Bang: Supernovae and Gamma-Ray Bursts. CUP:Cambridge, astro-ph/0003077.
- Phinney, E.S., 1983, Ph.D. Thesis, University of Cambridge, Cambridge, UK.
- Portegies Zwart, S.F. & Yungelson, L.R., 1998, A&A, 332, 173.
- Rasio, F.A. & Livio, M., 1996, ApJ, 471, 366.
- Regos, E., Tout, C.A., & Wickramasinghe, D., 1998, ApJ, 509, 362.
- Shakura, N.I. & Sunyaev, R.A., 1973, A&A, 24, 337.
- Sobczak, G.J., McClintock, J.E., Remillard, R.A., Bailyn, C.D., & Orosz, J.A., 1999, ApJ, 520, 776.
- Spruit, H. & Phinney, E.S., 1998, Natur, 393, 139.
- Thorne, K.S., 1974, ApJ, 191, 507.
- Thorne, K.S., Price, R.H., & MacDonald, D.A., 1986, The Membrane Paradigm, Yale University Press, New Haven.
- Torkelsson, U., Brandenburg, A., Nordlund, A., & Stein, R.F., 1996, Astron. Lett. & Comm., 34, 383.
- Van Paradijs, J., 1996, ApJ, 494, L45.
- Verbunt, F.W.M., Wijers, R.A.M.J., & Burm, H.M.G., 1990, A&A, 234, 195.
- Waxman, E. & Loeb, A., 1999, ApJ, 515, 721.
- Wellstein, S. & Langer, N., 1999, A&A, 350, 148.
- Wheeler, J.C., Höflich, P., Wang, L., & Yi, I., 2000, in: Proc. Fifth Huntsville Conf. on Gamma-Ray Bursts, astro-ph/9912080.
- Wijers, R.A.M.J., Bloom, J.S., Bagla, J.S., & Natarajan, P., 1998, MNRAS, 294, L13.
- Woosley, S.E., 1988, ApJ, 350, 218.
- Woosley, S.E., 1993, ApJ, 405, 273.
- Woosley, S.E. & Weaver, T.A., 1995, ApJS, 101, 181.
- Zhang, S.N., Cui, W., & Chen, W., 1997, ApJ, 482, L155.

---

## Chapter 18

# Hypercritical Advection-Dominated Accretion Flow

G.E. Brown, C.-H. Lee and H.A. Bethe

Reprinted with permission from *Astrophysical Journal*, 541 (2000) 918-923

Copyright (2000) by American Astronomical Society

### Commentary

We wanted to understand accretion disks. Brown and Lee began with the standard disk equation (A1)-(A5) plus the equation of state (A8), and then found that a reasonable solution was the Narayan & Yi (1994) self-similar one for advection-dominated accretion flow. This applied particularly simply because the hypercritical accretion was nearly adiabatic, with very small energy losses.

Section 2 contains Hans Bethe's equivalent solution, essentially using Newton's laws.

We were ready to dump the accreted matter onto the neutron star, and simply say that the latter accepted it and grew in mass. However, luckily we had a knowledgeable referee, Roger Chevalier, who persisted in telling us how to get things right. The low temperature of the accreting material means that it cannot get rid of its energy rapidly by neutrino emission, so it piles up, pushing its way up through the accretion disk. Zel'dovitch *et al.* had found this already in 1972. At this stage we had to go over to two-dimensional hydro calculations by Armitage & Livio (2000). These showed that an accretion disk reformed inside of the accretion shock, allowing matter to accrete onto the neutron star. Armitage & Livio, however, suggested that jets might drive the hypercritical accreting matter off, saving the neutron star from going into a black hole. We do not believe, given our maximum neutron star mass of  $1.5M_{\odot}$ , that much matter has to accrete onto a neutron star in order to send it into a black hole, but none the less it would be much more comfortable for us if a low-mass black-hole, neutron-star binary were observed. See our estimate of the probability of seeing one in the Commentary on the next chapter.

### References

- (1) Armitage, P.J., and Livio, L. (2000), *Astrophysical Journal* 532, 540.

- (2) Narayan, R., and Yi, I. (1994), *Astrophysical Journal* **428**, L13.
- (3) Zel'dovich, Ya.B., Ivanova, L.N., and Nadezhin, D.K. (1972), *Soviet Astronomy* **16**, 209.

## HYPERCritical ADVECTION-DOMINATED ACCRETION FLOW

G. E. BROWN AND C.-H. LEE

Department of Physics and Astronomy, State University of New York, Stony Brook, New York 11794

AND

H. A. BETHE

Floyd R. Newman Laboratory of Nuclear Studies, Cornell University, Ithaca, New York 14853

Received 1999 September 15; accepted 2000 May 5

### ABSTRACT

In this paper we study the accretion disk that arises in hypercritical accretion of  $\dot{M} \sim 10^8 M_{\text{Edd}}$  onto a neutron star while it is in common envelope evolution with a massive companion. Such a study was carried out by Chevalier, who had earlier suggested that the neutron star would go into a black hole in common envelope evolution. In his later study he found that the accretion could possibly be held up by angular momentum.

In order to raise the temperature high enough that the disk might cool by neutrino emission, Chevalier found a small value of the  $\alpha$ -parameter, where the kinematic coefficient of shear viscosity is  $\nu = \alpha c_s H$ , with  $c_s$  the velocity of sound and  $H$  the disk height; namely,  $\alpha \sim 10^{-6}$  was necessary for gas pressure to dominate. He also considered results with higher values of  $\alpha$ , pointing out that radiation pressure would then predominate. With these larger  $\alpha$ -values, the temperatures of the accreting material are much lower,  $\gtrsim 0.35$  MeV. The result is that neutrino cooling during the flow is negligible, satisfying very well the advection-dominating conditions.

The low temperature of the accreting material means that it cannot get rid of its energy rapidly by neutrino emission, so it piles up, pushing its way through the accretion disk. An accretion shock is formed, far beyond the neutron star, at a radius  $\lesssim 10^8$  cm, much as in the earlier spherically symmetric calculation, but in rotation. Two-dimensional numerical simulation shows that an accretion disk is reformed inside of the accretion shock, allowing matter to accrete onto the neutron star with pressure high enough so that neutrinos can carry off the energy.

*Subject headings:* accretion, accretion disks — binaries: close — black hole physics — stars: evolution — stars: neutron

### 1. INTRODUCTION

In one stage of the theory of evolution of compact objects the neutron star from the explosion of the primary progenitor enters into common envelope evolution of the evolving secondary. Recently, developing ideas suggested by Chevalier (1993) and later by Brown (1995), Bethe & Brown (1998) introduced hypercritical accretion into a population synthesis of binary compact objects. They found that in the standard scenario of compact binary evolution the neutron star essentially always went into a low-mass black hole.

In scrutiny of hypercritical accretion during common envelope evolution, Chevalier (1996) discussed the effects of rotation. He studied accretion disks and, in particular, neutrino-cooled disks and advection-dominated disks. He chiefly considered disks with viscosity parameter  $\alpha \sim 10^{-6}$ . Such small  $\alpha$ -values were needed in order to reach high enough temperatures, in situations where gas pressure predominates, so that the temperature was sufficient for neutrino emission.

Calculations (Brandenburg et al. 1996; Torrelles et al. 1996) suggest that the higher values of  $\alpha \sim 0.1$ , which Shakura & Sunyaev (1973) suggested as arising from magnetic turbulence, are more appropriate. Chevalier (1996) notes that “Radiation pressure tends to dominate at higher values of  $\alpha$ ; for  $M = 1.4 M_{\odot}$ ,  $\dot{M} = 1 M_{\odot} \text{ yr}^{-1}$ , and  $r = 10^8$  cm, the crossover point is  $\alpha \approx 10^{-6}$ ”

In this paper we discuss the situation for  $\alpha \sim 0.05$ , finding that radiation pressure predominates to the extent that we can take the adiabatic index to be  $\gamma = 4/3$ . We show the

resulting hypercritical accretion to be essentially that of Houck & Chevalier (1991) and Brown (1995), although angular momentum was not discussed in these earlier papers.

### 2. ADVECTION

We consider super-Eddington accretion, so electromagnetic radiation is negligible compared to the energy carried by the infalling matter. We shall show that, for reasonable parameters, the temperature remains sufficiently low that neutrino emission is also small. Therefore the accreting matter keeps nearly constant energy per unit mass.

The accreting matter orbits around the neutron star. Because of collisions between the accreting particles, we may consider their orbits circular. From Narayan & Yi (1994), the orbital velocity is

$$v^2 = (r\Omega)^2 \approx \frac{2}{3} v_K^2, \quad (1)$$

with  $v_K = (GM/r)^{1/2}$  is the Keplerian velocity and  $M$  the mass of the neutron star.

Because of viscosity, the velocity has a radial (inward) component that is conventionally set equal to<sup>1</sup>

$$v_r = \frac{dr}{dt} = - \frac{3}{7} \alpha v_K. \quad (2)$$

<sup>1</sup> Our  $\alpha$  here is (2/3) of the  $\alpha_m$  of Shakura & Sunyaev (1973), who took the kinematic coefficient of shear viscosity to be  $\nu = \frac{2}{3} \alpha_m c_s H$ . Thus eq. (2) is that of Narayan & Yi (1994),  $v = \alpha c_s H$ . Our treatment is equivalent to their advection-dominated solution with their  $f = \epsilon' - \epsilon = 1$ .

## HYPERCRITICAL ADAF

We shall assume  $\alpha = 0.05$ .

The potential energy per unit mass is  $-GM/r$ , and from equation (1) the kinetic energy from the bulk motion of disk is  $\sim \frac{1}{2}v^2 = GM/7r$ . Since we assume negligible heat loss, the heat energy in the local frame rotating with the disk is

$$\epsilon = \frac{3}{4} \times \frac{6}{7} \frac{GM}{r}, \quad (3)$$

where the factor 6/7 comes from the sum of the local potential and kinetic energies in the disk, and 1/3 of  $\epsilon$  having gone into  $PV$  work (enthalpy rather than energy is conserved). The density is related to the rate of energy accretion by

$$\dot{M} = 2\pi r^2 H \rho (-\dot{r}), \quad (4)$$

where  $H$  is the height of the accretion disk. An explicit solution of the accretion problem, by Narayan & Yi (1994) finds  $H = \sqrt{2/7}r \approx 1/2r$ , which we adopt. Then, with equation (2),

$$\dot{M} = 2\pi r^2 (\frac{3}{7}\alpha v_K) \rho. \quad (5)$$

Using equation (1), and assuming  $\alpha$  to be constant,

$$\rho \propto r^{-3/2}. \quad (6)$$

The thermal energy per unit volume is

$$\rho\epsilon = \frac{3\dot{M}v_K}{4\pi\alpha r^2} \propto r^{-5/2}, \quad (7)$$

which also gives the  $r$ -dependence of the pressure,  $p = \rho\epsilon/3$ . Let us assume the mass of a neutron star to be  $1.5 M_\odot$  and its radius to be 10 km,<sup>2</sup> then

$$\frac{v_K}{4\pi r^2} = 1.2 \times 10^{-3} \text{ cm}^{-1} \text{ s}^{-1}. \quad (8)$$

Let us further assume  $\dot{M} = 1 M_\odot \text{ yr}^{-1} = 6.3 \times 10^{25} \text{ g s}^{-1}$ , and  $\alpha = 0.05$ ,

$$\rho\epsilon = 4.5 \times 10^{24} \text{ ergs cm}^{-3}. \quad (9)$$

The energy density of electromagnetic radiation plus electrons and positrons is

$$\rho\epsilon = \frac{11}{3} aT^4 = \frac{11}{4} (1.37) \times 10^{26} T_{\text{MeV}}^4 \text{ ergs cm}^{-3}. \quad (10)$$

Equating the two expressions,

$$T = 0.33 \text{ MeV}. \quad (11)$$

This is only a moderate temperature. The emission of neutrino pairs, according to Dicus (1972) and Brown & Weingartner (1995) is

$$\begin{aligned} \dot{\epsilon}_n &= 1.0 \times 10^{25} \left( \frac{T}{\text{MeV}} \right)^9 \text{ ergs s}^{-1} \text{ cm}^{-3} \\ &= 4.6 \times 10^{20} \text{ ergs s}^{-1} \text{ cm}^{-3}. \end{aligned} \quad (12)$$

This is the energy emission at the surface of the neutron star. It can easily be seen that this quantity goes like  $r^{-45/8}$ .

<sup>2</sup> For the neutron star with  $M = 1.5 M_\odot$ , the Schwarzschild radius is  $R_{\text{Sch}} = 4.4 \text{ km}$ , and the marginally stable orbit is  $R_{\text{MS}} = 3R_{\text{Sch}}$ , which is larger than the radius of neutron star. From Abramowicz et al. (1988), however, the inner disk can be extended to the marginally bound orbit  $R_{\text{MS}} = 2R_{\text{Sch}}$ , with hypercritical accretion. In this paper, therefore, we assumed that the accretion disk extends to the surface of the neutron star.

Multiplying by the volume element  $4\pi r^2 dr$  and integrating over  $r$ , we get for the total energy loss by neutrinos.

$$\dot{W}_n = \int \dot{\epsilon}_n 4\pi r^2 dr = 2.2 \times 10^{39} \text{ ergs s}^{-1}. \quad (13)$$

On the other hand, the energy advected by infalling matter is

$$\begin{aligned} \dot{M}\epsilon &= 6.3 \times 10^{25} \text{ g s}^{-1} \times 1.3 \times 10^{20} \text{ ergs g}^{-1} \\ &= 8.2 \times 10^{45} \text{ ergs s}^{-1}. \end{aligned} \quad (14)$$

So the fraction of energy radiated in neutrinos is

$$F = \frac{\dot{W}_n}{\dot{M}\epsilon} = 2.7 \times 10^{-7}. \quad (15)$$

This proves our contention at the beginning of this section that the energy lost during infall in neutrino emission is negligible.

If the rate of infall is changed or  $\alpha$  is altered, the fraction of neutrino radiation changes as

$$F \sim \dot{M}^{5/4} \alpha^{-9/4}. \quad (16)$$

Only if  $\alpha$  were reduced by a factor  $10^{3.5}$ , to about  $2 \times 10^{-5}$ , would neutrino radiation become appreciable.

Whereas the above might appear to solve the accretion problem, it will not give a steady state solution because the accreted matter, which is at too low a temperature to cool sufficiently rapidly by neutrino radiation, cannot be fitted into the narrow gap between neutron star and inner disk. Thus, it pushes outward through the accretion disk (R. A. Chevalier 1999, private communication). This was first seen in the numerical calculation of Zeldovich et al. 1972. In the spherically symmetric case an accretion shock must be built into the flow. Freely falling gas outside of this shock goes into subsonic flow inside, with formation of an envelope at small radii hot enough so that neutrinos can be emitted to carry off the energy. Detailed calculation (Houck & Chevalier 1991; Brown 1995) gives

$$r_{\text{sh}} \simeq 2.6 \times 10^8 \text{ cm} \left( \frac{\dot{M}}{M_\odot \text{ yr}^{-1}} \right)^{-0.37} \quad (17)$$

R. A. Chevalier (1999, private communication) illustrates the necessity for the accretion shock simply. Starting from hydrostatic equilibrium in the gap region

$$\frac{dP}{dr} = -\rho \frac{GMT}{r^2} \quad (18)$$

with  $r$  initially set to  $R$  because the gap between neutron star and the innermost part of the accretion disk is narrow, and  $\Gamma = (1 - R_{\text{sh}}/r)^{-1}$ , one can integrate from point A, the inner edge of the hypercritical advection-dominated accretion flow, to point B, basically at the outer edge of the neutron star. Taking

$$P = K\rho^\gamma \quad (19)$$

since the flow is adiabatic, and integrating from point A to point B near the neutron star surface, one finds that

$$\epsilon_B - \epsilon_A \approx \left( \frac{\gamma - 1}{\gamma} \right) \frac{GMT}{R^2} (R_A - R_B), \quad (20)$$

Now from equation (3)  $\epsilon_A \sim GM/R$ . For  $\epsilon$  to increase by a factor  $\lesssim 2$  so that neutrinos could carry off the energy in going from point A to B would require  $(R_A - R_B) \lesssim R$ , violating the assumption that the gap is narrow. This argument is general, not depending on the dimensionality of the description.

The above discussion is at least roughly consistent with the results of Armitage & Livio (2000) who carry out a two-dimensional numerical simulation of the hypercritical accretion. They find that an accretion disk forms inside the accretion shock and with the usually considered values of the viscosity the disk is probably able to transfer mass inward at the Bondi-Hoyle (hypercritical) rate. Pressures at the neutron star are high enough so that the accreted energy can be carried off by neutrinos. They warn, however, that the pressure of a disk also makes the formation of outflows or jets probable, which could reduce the accretion rates onto the neutron star.

Our results up through equation (16) are consistent with those of Minishige et al. (1997) for black hole disk accretion, once our accretion rate is scaled down to the  $\sim \dot{M}/\dot{M}_{\text{Edd}} \sim 10^6$  that they use. For a black hole accretion energy can be lost from the system because the energy can be advected across the horizon. However, even in the case of a black hole, an accretion shock is developed in the numerical calculations of MacFadyen & Woosley (1999). Their rate of accretion is  $\sim 0.1 M_{\odot} \text{ s}^{-1}$ ,  $\sim 10^6$  times ours. In their evolution of the accretion first the low angular momentum material in the equator and material along the axes falls through the inner boundary onto the black hole. An accretion shock at  $\sim 350$  km with an interior centrifugally supported disk then forms. The MacFadyen & Woosley (1999) calculation deals with a different situation from ours; namely, advection of material that is originally He onto a black hole. Here nuclear dissociation energy as well as neutrino emission plays an important role in lowering the pressure. From the Minishige et al. (1997) calculations at an accretion rate of  $\dot{M} \sim 10^6 \dot{M}_{\text{Edd}}$  into a black hole one can see the formation of the accretion disk, but cannot discern an accretion shock.

From the calculation most relevant to our work, that of Armitage & Livio (2000) one sees that the case the flow has initial angular momentum, the main effect of the angular momentum in their two-dimensional simulation is to set both accretion shock and accretion disk into rotation. The matter then being advected smoothly onto the neutron star. Thus, modulo the possibility of jets, the resulting accretion is very much like that in the one-dimensional calculations of Houck & Chevalier (1991) and Brown (1995).

It should be noted, however, that in the one-dimensional case the pressure drops off as  $p \propto r^{-4}$ , appropriate for a  $\Gamma = 4/3$  accretion envelope (Chevalier 1989; Houck & Chevalier 1991; Brown & Weingartner 1994), whereas in the advection-dominated accretion flow solution  $p \propto r^{-5/2}$ . The flatter profile in the latter case makes it more difficult to reach the energy densities necessary for neutrino emission.

Perhaps the most important result is that the orbital motion of the accreting material is preserved. So the neutron star, in the natural course of events, acquires an angular momentum,

$$J = \Delta M \times r_N v_N, \quad (21)$$

where  $\Delta M$  is the total mass accreted,  $r_N$  is the radius of the neutron star, and  $v_N$  the corresponding orbital velocity,

$$r_N v_N = (\frac{3}{2} G M r_N)^{1/2} = 0.76 \times 10^{16} \text{ cm}^2 \text{ s}^{-1} \quad (22)$$

an appreciable angular momentum.

These considerations apply equally to accretion to a black hole. The thermal energy, in this case, is not radiated away, but  $\epsilon_{\text{th}} c^{-2}$  is added to the mass of the black hole. If we assume that the black hole has a mass  $M_B = 2 M_{\odot}$ , its Schwarzschild radius is

$$r_B = 6 \text{ km}. \quad (23)$$

From Abramowicz et al. (1988), for the hypercritical accretion, the inner disk can be extended to the marginally bound orbit  $R_{MB} = 2R_{\text{Sch}} = 12 \text{ km}$ .<sup>3</sup> So the orbital velocity at the inner disk boundary is

$$v_B = 0.80 \times 10^{10} \text{ cm s}^{-1}, \quad (24)$$

and the specific angular momentum is

$$r_B v_B = 0.96 \times 10^{16} \text{ cm}^2 \text{ s}^{-1}, \quad (25)$$

about the same as for a neutron star. Going over to a Kerr Geometry with decreased event horizon this will be lowered somewhat. Bethe & Brown (1998) have considered a neutron star of  $1.4 M_{\odot}$  in common envelope with a giant, and have found that the neutron star will accrete about  $\Delta M = 1 M_{\odot}$  of the giant's mass and thereby become a black hole. This accreted material will have an angular momentum

$$J_B = \Delta M r_B v_B. \quad (26)$$

Assuming that the core of the giant does not merge with the black hole, this core will later become a supernova, yielding a second neutron star. Gravitational waves will be emitted, and consequently later the two compact objects will merge. It is important that the black hole has a big angular momentum  $J_B$  that may be extracted (Blandford & Znajek 1977; Lee, Wijers, & Brown 2000) and may possibly give a gamma ray pulse: As we have seen, no special mechanism is needed to give that angular momentum to the black hole, but it arises naturally as the black hole is formed.

### 3. DISCUSSION

Although there is considerable uncertainty in the  $\alpha$ -parameter in the viscosity, we believe our value to be in a reasonable range. Torkelsson et al. (1996) have analyzed simulations of the magnetic turbulence different groups used to estimate the strength of the turbulent viscosity. Estimates of the  $\alpha$ -parameter, which range from 0.001 to 0.7 in the Shakura & Sunyaev (1973) definition, were verified using the same code for all simulations. The higher values of  $\alpha$  are the result of an applied vertical (roughly equipartition) magnetic field. Without this field the typical value is  $\alpha \sim 0.005$ .

Armitage & Livio (2000) have raised the possibility of jets removing the infalling material, so that it will not accrete onto the neutron star. Jets are prevalent in situations with accretion disks, but the mechanisms for driving them are poorly understood (Livio 1999). Suggestions have been made that accretion disks could generate a magnetically

<sup>3</sup> Near the black hole, there should be relativistic corrections to the advection-dominated accretion flow solutions. In this paper, however, we assumed that the ADAF solution can be extended to the marginally stable orbit.

driven outflow, first by Blandford (1976) and by Lovelace (1976).

If the magnetic energy above the disk is larger than the thermal and kinetic energy densities, the ionized outflowing material is forced to follow field lines, some of which form an angle with the perpendicular to the disk surface. The material may then be accelerated by the centrifugal force like a bead on a wire, as described by Livio (1999).

It would seem reasonable for jets to be driven in the range of accretion rates  $\dot{M}_{\text{Edd}}$  to  $10^4 \dot{M}_{\text{Edd}}$ , the latter being the lower limit at which neutrinos can carry off the energy sufficiently rapidly that hypercritical accretion can take place. In this accretion interval, only  $\dot{M}_{\text{Edd}}$  can be accreted, and since the surplus material must leave, it seems reasonable that magnetic forces can collimate it in, at least, some cases. We believe that the jets in Cyg X-3 may be an example.

It is more difficult to envisage such jet propulsion with hypercritical accretion, especially at rates close to  $\dot{M} \sim 10^9 \dot{M}_{\text{Edd}}$  where the Alfvén radius comes right down to the surface of the neutron star, i.e., where the ram pressure always exceeds the magnetic pressure, unless the latter is increased by differential rotation, etc. Of course, as in Eggum, Coroniti, & Katz (1988), who calculated only slightly super Eddington accretion ( $\dot{M} \sim 4 \dot{M}_{\text{Edd}}$ ), a jet may be formed up the center of the disk axis, where there is little infalling matter, but this would not drastically lower the accretion.

In fact, in a similar case of hypercritical accretion we have some observational support for the main part of the matter being accreted. In the evolution of the transient sources (Brown, Lee, & Bethe 1999) the H envelope of an  $\sim 25 M_{\odot}$  zero-age main-sequence star is removed in common envelope evolution with an  $\gtrsim 1 M_{\odot}$  main-sequence star. Following this, the Fe core of the massive star goes into a black hole of mass  $\sim 1.5 M_{\odot}$  and most of the He envelope must be accreted in order to obtain the final  $\sim 7 M_{\odot}$  black hole. In general the He envelope will be rotating. In fact, in the case of GRO J1655-40 arguments have been made (Zhang, Cui, & Chen 1997) that the black hole is presently rotating close to the Kerr limit although later studies

(Sobczak et al. 1999; Gruzinov 1999) lower the velocity to  $\sim 0.6$ – $0.7 c$ . Because the  $\sim R_{\odot}$  radius of the He envelope is  $\sim 10^5$  times larger than the neutron star radius, even with slow rotation of the envelope an accretion disk will be formed before the matter can be advected into the black hole. The central He density is  $\rho \sim 10^5 \text{ g cm}^{-3}$ , about the same as the H density in our work above in the inner disk, for  $\dot{M} \sim 10^8 \dot{M}_{\text{Edd}}$ . Thus, the difference is that the matter is He, rather than H, in the transient sources. Clearly jet formation does not strongly affect the amount of He accreted onto the black hole.

It should also be realized that observed masses of neutron stars are not very far below the Brown & Bethe (1994) proposed upper mass limit of  $\sim 1.5 M_{\odot}$ . Thus, only a small fraction of the Bethe & Brown (1998) accretion of  $\sim 1 M_{\odot}$  need be realized for the neutron star to go into a black hole.

Since entropy gradients become negative, advection-dominated flows are unstable to convection, as found by Eggum et al. (1988) and as worked out in considerable detail by Narayan & Yi (1995). Advection-dominated accretion is found to dominate in a thick belt about the equator, with somewhat reduced velocity, depending upon the polar angle. Narayan & Yi do find that jets can be driven in fairly wide angles about the poles, but in agreement with Eggum et al. (1988), they find that very little mass is driven off in these jets.

Equation (7) shows the energy density to go inversely with  $\alpha$ , as  $\alpha^{-1}$ . Thus a much higher temperature is reached for small  $\alpha$ -values,  $\sim 10$  times greater for  $\alpha \sim 10^{-5}$  than for  $\alpha \sim 0.1$  (Chevalier 1996), the energy density going as  $T^4$ . The rate of cooling goes as  $T^9 \propto \alpha^{-9/4}$  and increases a factor  $\sim 10^9$  for the small  $\alpha$ -values. In this regime of temperatures the disk is cooled very rapidly by neutrino emission.

We would like to thank Roger A. Chevalier and Insu Yi for useful suggestions and advice. This work was partially supported by the US Department of Energy under grant DE-FG02-88ER40388.

## APPENDIX

### ADVECTION-DOMINATED DISKS

In this appendix we give a more conventional derivation of our results.

In the case that the disk is thin to neutrino emission, the steady state disk equation can be written as (Shakura & Sunyaev 1993; Shapiro & Teukolsky 1983, § 14.5)

$$\Sigma = 2H\rho, \quad (\text{A1})$$

$$\dot{M} = 2\pi r \Sigma v_r = \text{const.}, \quad (\text{A2})$$

$$6\pi r^2 H \dot{\epsilon}_p = \dot{M}(GMr)^{1/2}, \quad (\text{A3})$$

$$H\dot{\epsilon}_p = \frac{3\dot{M}GM}{8\pi r^3}, \quad (\text{A4})$$

$$H = \left(\frac{p}{\rho}\right)^{1/2} \left(\frac{r^3}{GM}\right)^{1/2}, \quad (\text{A5})$$

where we use the notation of Chevalier (1996).<sup>4</sup> Here  $\Sigma$  is the surface density of the disk. Our equation of state is

$$p = \frac{11}{4} \frac{\alpha T^4}{3}, \quad (\text{A6})$$

where we neglect the gas pressure. We assume that the neutrino pair process (Dicus 1972) is the dominant source of neutrino cooling, using the Brown & Weingartner (1999) expression

$$\dot{\epsilon}_n = 1.0 \times 10^{25} \left( \frac{T}{\text{MeV}} \right)^9 \text{ergs s}^{-1} \text{cm}^{-3}. \quad (\text{A7})$$

Equation (A4) is however, not correct for the main part of our accretion disk with radiation-dominated pressure, because the temperature of the material is not high enough for it to be cooled appreciably by neutrinos. In the absence of appreciable cooling, we can use the Narayan & Yi (1994) advection-dominated accretion, with a self-similar solution:

$$v_r \approx -\frac{3\alpha}{5+2\epsilon} \left( \frac{GM}{r} \right)^{1/2}, \quad (\text{A8})$$

$$\Omega \approx \left( \frac{2\epsilon}{5+2\epsilon} \right)^{1/2} \left( \frac{GM}{r^3} \right)^{1/2}, \quad (\text{A9})$$

$$\rho \approx \frac{(5+2\epsilon)^{3/2} \dot{M}}{2^{5/2} 3 \pi \alpha r^{3/2} (GM)^{1/2}}, \quad (\text{A10})$$

$$p \approx \frac{2^{1/2} (5+2\epsilon)^{1/2} \dot{M} (GM)^{1/2}}{12 \pi \alpha r^{5/2}}, \quad (\text{A11})$$

where we take  $\epsilon = 1$  for the case of  $\gamma = 4/3$ .

Energy is produced by viscous heating at a rate

$$\dot{\epsilon}_{\text{prod}} = \frac{3\dot{M}GM}{8\pi r^3 H}, \quad (\text{A12})$$

where we use  $H = \sqrt{2/7}r$  from equations (A5), (A10) and, (A11). In order to use equation (A7) we must obtain a temperature. For  $\alpha = 0.05$ ,  $M = 1.5 M_{\odot}$ , and  $\dot{M} = 1 M_{\odot} \text{ yr}^{-1} = 0.63 \times 10^{26} \text{ g s}^{-1}$  we find

$$\dot{\epsilon}_{\text{prod}} = 0.28 \times 10^{20} r_s^{-4} \text{ ergs cm}^{-3} \text{ s}^{-1}. \quad (\text{A13})$$

Using

$$v_r = \frac{dr}{dt} = -\frac{3}{7} \alpha \sqrt{\frac{GM}{r}}, \quad (\text{A14})$$

we can integrate this expression:

$$\begin{aligned} \epsilon_{\text{prod}} &= \int_0^r \dot{\epsilon}_{\text{prod}} dt = \int_{\infty}^r \frac{\dot{\epsilon}_{\text{prod}}}{v_r} dr \\ &= - \int_{\infty}^r dr \frac{3\dot{M}GM}{3/7\alpha\sqrt{GM8\pi(H/r)}} r^{-7/2} \\ &= \frac{2}{5} \times \frac{r}{3/7\alpha\sqrt{GM/r}} \times \frac{3\dot{M}GM}{8\pi r^3 H} \\ &= \frac{2}{5} \frac{r}{|v_r|} \dot{\epsilon}_{\text{prod}}, \end{aligned} \quad (\text{A15})$$

where the last expression shows  $\epsilon_{\text{prod}}$  to be  $(2/5)$  of the value that would be obtained from dimensional analysis. The quantity  $(2/5)r/|v_r|$  can be interpreted as a dynamical time  $\tau_d$ ,

$$\tau_d = \frac{2}{5} \frac{r}{|v_r|} = 1.3 r_s^{3/2} \text{ s}, \quad (\text{A16})$$

<sup>4</sup> In this paper, we take the kinematic coefficient of shear viscosity  $\nu = \alpha cH$ , while  $\nu = \frac{1}{2}\alpha cH$  is used in eq. (17) of Chevalier (1996).

and

$$\epsilon_{\text{prod}} = \tau_d \dot{\epsilon}_{\text{prod}} = 3.6 \times 10^{19} r_8^{-5/2} \text{ ergs cm}^{-3}. \quad (\text{A17})$$

We obtain a temperature of the advected material by relating  $T$  to the thermal energy density

$$\begin{aligned} \epsilon_{\text{th}} &= aT^4 = \frac{11}{4} (1.37) \times 10^{26} T_{\text{MeV}}^4 \\ &= 3.8 \times 10^{26} T_{\text{MeV}}^4 \text{ ergs cm}^{-3}. \end{aligned} \quad (\text{A18})$$

From equation (A17) we find<sup>5</sup>

$$T_{\text{MeV}}^4 = 0.96 \times 10^{-7} r_8^{-5/2}, \quad (\text{A19})$$

showing that  $T^4$  goes as  $r_8^{-5/2}$ . The total cooling by equation (A7) will be

$$\delta\epsilon = \tau_d \dot{\epsilon}_n, \quad (\text{A20})$$

so that

$$\frac{\delta\epsilon}{\epsilon_{\text{prod}}} = \frac{\dot{\epsilon}_n}{\dot{\epsilon}_{\text{prod}}} = \frac{0.29 \times 10^{-9}}{r_8^{13/8}}. \quad (\text{A21})$$

Thus  $\delta\epsilon/\epsilon_{\text{prod}} \sim 1.6 \times 10^{-7}$  for  $r = 10^6$  cm at the inner disk just outside the neutron star, in rough agreement with equation (15). In other words, temperatures of the advected material are so low that the neutrino cooling is negligible. Thus, the material will be deposited essentially adiabatically onto the neutron star, heating it.

<sup>5</sup> Eq. (A19) shows that  $T$  reaches  $\sim 0.31$  MeV at the inner disk just outside the neutron star surface  $10^6$  cm. There the dynamical time is only  $\sim 1.3 \times 10^{-3}$  s, so the neutrino cooling by eq. (A7) is negligible.

#### REFERENCES

- Abramowicz, M. A., Czerny, B., Lasota, J. P., & Szuszkiewicz, E. 1988, *ApJ*, 332, 646
- Armitage, P. J., & Livio, M. 2000, *ApJ*, 532, 540
- Bethe, H. A., & Brown, G. E. 1998, *ApJ*, 506, 780
- Blandford, R. 1976, *MNRAS*, 176, 465
- Blandford, R. D., & Znajek, R. L. 1977, *MNRAS*, 179, 433
- Brandenburg, A., Nordlund, Å., Stein, R. F., & Torkelsson U. 1996, *ApJ*, 458, L45
- Brown, G. E. 1995, *ApJ*, 440, 270
- Brown, G. E., & Bethe, H. A. 1994, *ApJ*, 423, 659
- Brown, G. E., Lee, C.-H., & Bethe, H. A. 1999, *NewA*, 4, 313
- Brown, G. E., & Weingartner, J. C. 1994, *ApJ*, 423, 843
- Chevalier, R. A. 1989, *ApJ*, 346, 847
- , 1993, *ApJ*, 411, L33
- , 1996, *ApJ*, 459, 322
- Dicus, D. A. 1972, *Phys. Rev. D*, 6, 941
- Eggum, G. E., Coroniti, F. V., & Katz, J. I. 1988, *ApJ*, 330, 142
- Gruzinov, A. 1999, *ApJ*, 517, L105
- Houck, J. C., & Chevalier, R. A. 1991, *ApJ*, 376, 234
- Lee, H. K., Wijers, R. A. M. J., & Brown, G. E. 2000, *Phys. Rep.*, 325, 83
- Livio, M. 1999, *Phys. Rep.*, 311, 225
- Lovelace, R. V. E. 1976, *Nature*, 262, 649
- MacFadyen, A. I., & Woosley, S. E. 1999, *ApJ*, 524, 262
- Mineahige, S., Nomura, H., Hirose, M., Nomoto, K., & Suzuki, T. 1997, *ApJ*, 489, 227
- Narayan, R., & Yi, I. 1994, *ApJ*, 428, L13
- , 1995, *ApJ*, 444, 231
- Shakura, N. I., & Sunyaev, R. A. 1973, *A&A*, 24, 337
- Shapiro, S. L., & Teukolsky, S. A. 1983, *Black Holes, White Dwarfs, and Neutron Stars* (New York: Wiley)
- Sobczak, G. J., McClintock, J. E., Remillard, R. A., Bailyn, C. D., & Orosz, J. A. 1999, *ApJ*, 520, 776
- Torkelsson, U., Brandenburg, A., Nordlund, Å., & Stein, R. F. 1996, *Astrophys. Lett. Commun.*, 34, 383
- Zeldovich, Ya. B., Ivanova, L. N., & Nadezhin, D. K. 1972, *Soviet Astron.*, 16, 209
- Zhang, S. N., Cui, W., & Chen, W. 1997, *ApJ*, 482, L155

---

## Chapter 19

# Evolution of Neutron Star, Carbon-Oxygen White Dwarf Binaries

G.E. Brown, C.-H. Lee, S. Portegies Zwart and H.A. Bethe

Reprinted with permission from *Astrophysical Journal*, 547 (2001) 345–354

Copyright (2001) by American Astronomical Society

### Commentary

We teamed up here with Simon Portegies Zwart, who had written a very complete computer-driven population synthesis of the various kinds of binaries involving at least one compact object. In his paper on the evolution of binary neutron stars with Lev Yungelson (1998) they had a Case H which included hypercritical accretion and we remarked in our Paper 12 that once we normalized to the same number of supernova explosions in the Galaxy, we had very good agreement between their results and ours. This comparison is carried out in more detail in this paper, showing the agreement to be very good indeed.

Extending our previous work, we found that with hypercritical accretion in the common envelope evolution in the red giant, neutron star stage the neutron star should accrete sufficient matter to send it into a black hole. If such binaries survived they would have had their orbits circularized and had their magnetic fields brought down by a factor  $\gtrsim 100$  during the He-star, neutron star stage, as described in Paper 10. On the other hand, eccentric binaries in which the neutron star was formed last would be unrecycled and still retain their magnetic fields of  $\sim 10^{12}$  G. Therefore, the circular binaries should have an observability premium of  $\Pi \sim 100$  (see Paper 10). Since we see at least one (B2303+46) and probably a second one (PSR J1141–6545) of the eccentric binaries, we should see  $\sim 100$  or 200 of the circular ones that have been recycled, given roughly equal production rates of the two. (More precise estimates are made in Paper 19.)

At the time of writing Paper 19 it appeared as if no neutron star, carbon-oxygen white dwarf binaries that had gone through common envelope evolution had been observed. Those that had earlier been evolved in this way were found to be preferably evolved with avoidance of common envelope evolution. In a sense, we wrote our paper at the most propitious time for our theory, when it appeared that the ratio of unrecycled to binaries which had undergone common envelope evolution was infinite.

To be fair we must say that well before our paper was accepted for publication, the preprint of Tauris, van den Heuvel and Savonije (2000) had proposed five binaries, four of them new pulsars from the Parkers Multibeam Survey, which they proposed should be evolved through common envelope. Two of these recycled pulsars in relativistic orbits, PSR 1157–5112 and J1757–5322, are discussed by Edwards and Bailes (2001). Two others, J1435–6100 and J1454–5846, are discussed by Camilo *et al.* (2001).

The fifth of the systems favored for common envelope evolution by Tauris, van den Heuvel and Savonije (2000) is J1022+1001. This closely resembles PSR J2145–0750, aside from a more massive white dwarf companion, as remarked by van den Heuvel (1994), who evolved the latter through common envelope when the white dwarf progenitor was on the AGB (Case C mass transfer). Van den Heuvel suggested for J2145–0750 that there was considerable mass loss because of possible instabilities on the AGB caused by the presence of the neutron star. This is one possibility of saving our general theme; i.e., that most of the neutron star, carbon-oxygen white dwarf binaries would end up as low-mass black hole carbon-oxygen white dwarf binaries, although some might be saved with neutron stars because of the possible instabilities caused by the neutron star while the white dwarf progenitor is on the AGB. In fact, as noted in our Appendix C, after Eq. (C4) van den Heuvel chose  $\lambda = 1/2$  for the parameter that characterizes the structure of the hydrogen envelope of the massive star that is removed in common envelope evolution. Dewi & Tauris (2001) have since carried out detailed calculations that in some cases, “particularly on the asymptotic giant branch of lower-mass stars, it is possible that  $\lambda > 4$ ”. This lowers the binding energy of the envelope by a large factor, so that it can be removed in common envelope evolution and still leave a reasonably wide orbit, as remarked by Dewi & Tauris. We believe that this may be the reason that some binaries have survived common envelope evolution.

A possible result less good for our theory is that an order of magnitude more recycled binaries, in which the pulsar field is low,  $B \sim 10^8\text{--}10^{10}$  G, will turn up in the substantially more sensitive surveys now being conducted, destroying the large discrepancy predicted in our Paper 19. However, after publishing Paper 19 we found another unrecycled pulsar, PSR B0820+02, in Hansen & Phinney (1998). It has a magnetic field of  $B = 3 \times 10^{11}$  G and is most consistently described with carbon core white dwarf of mass  $0.5\text{--}1.0 M_\odot$ . Thus, even with the new discoveries, the number of unrecycled and recycled pulsars is comparable. Since the inferred magnetic fields of the recycled pulsars are all  $\gtrsim 100$  times lower than of the unrecycled ones, we believe the discrepancy to be large.

In the discussion following Eq. (A2) we find that population synthesis brings the ratio of  $(ns, ns)$  binaries to the eccentric neutron star, carbon-oxygen white dwarf binaries down to  $\sim 1$  or 2, close to observation. We might well ask what the ratio of low-mass black hole binaries to  $(ns, ns)$  binaries would be predicted to be. In Paper 12 we find that 10 times more of the former are produced than  $(ns, ns)$  binaries. Periods would be about the same, so the Ramachandran & Portegies Zwart (1998) observation penalty would apply to both. The pulsar in the  $(ns, ns)$  binary would be recycled with fields  $B \sim 10^{10}$  G, whereas that in the low-mass black hole, neutron star binary would be fresh, with field  $B \sim 10^{12}$  G. Thus,

the observational premium would favor the ( $ns, ns$ ) binaries by a factor of 100, with a net factor of 10 in their favor, given that they are produced only 10% as often as the binaries with low-mass black holes. Thus, our probability for having seen a low-mass black hole, neutron star binary is only  $\sim 0.1\text{--}0.2$ , and it is reasonable that none has been seen.

## References

- (1) Camilo, F., *et al.* (2001), *Astrophysical Journal* **548**, L187.
- (2) Dewi, J.D.M., and Tauris, T.M. (2000), *Astronomy & Astrophysics* **360**, 1043.
- (3) Edwards, R.T., and Bailes, M. (2001), *Astrophysical Journal* **547**, L37.
- (4) Hansen, B.M.S., and Phinney, E.S. (1998), *Monthly Notices of the Royal Astronomical Society* **294**, 569.
- (5) Tauris, T.M., Van den Heuvel, E.P.J., and Savonije, G.J. (2000), *Astrophysical Journal* **530**, L93.
- (6) Portegies Zwart, S.F., and Yungelson, L.R. (1998), *Astronomy & Astrophysics* **332**, 173.
- (7) Ramachandran, R., and Portegies Zwart, S.F. (1998), Abstrs. 19th Texas Symp. on Relativistic Astrophysics and Cosmology, ed. J. Paul, T. Montmerle and E. Aubourg.
- (8) Van den Heuvel, E.P.J. (1994), *Astronomy & Astrophysics* **291**, L39.

This page is intentionally left blank

## EVOLUTION OF NEUTRON STAR, CARBON-OXYGEN WHITE DWARF BINARIES

G. E. BROWN

Department of Physics and Astronomy, State University of New York, Stony Brook, NY 11794

C.-H. LEE

Korea Institute for Advanced Study, Seoul 130-012, Korea

S. F. PORTEGIES ZWART<sup>1</sup>

Department of Astronomy, Boston University, 725 Commonwealth Avenue, Boston, MA 02581

AND

H. A. BETHE

Floyd R. Newman Laboratory of Nuclear Studies, Cornell University, Ithaca, NY 14853

Received 1999 December 21; accepted 2000 September 11

### ABSTRACT

At least one, but more likely two or more, eccentric neutron star, carbon-oxygen white dwarf binaries with an unrecycled pulsar have been observed. According to the standard scenario for evolving neutron stars that are recycled in common envelope evolution, we expect to observe  $\gtrsim 50$  such circular neutron star, carbon-oxygen white dwarf binaries, since their formation rate is roughly equal to that of the eccentric binaries and the time over which they can be observed is 2 orders of magnitude longer, as we shall outline. We observe at most one or two such circular binaries, and from that we conclude that the standard scenario must be revised. Introducing hypercritical accretion into common envelope evolution removes the discrepancy by converting the neutron star into a black hole which does not emit radio waves, and therefore would not be observed.

*Subject headings:* binaries: close — stars: evolution — stars: neutron — stars: statistics — white dwarfs

### 1. INTRODUCTION

We consider the evolution of neutron star, carbon-oxygen white dwarf binaries using both the Bethe & Brown (1998) schematic analytic evolutions and the Portegies Zwart & Yungelson (1998) numerical population syntheses.

The scenario in which the circular neutron star, carbon-oxygen white dwarf binaries [which we denote hereafter as  $(ns, co)_c$ ] have gone through common envelope evolution is considered. In conventional common envelope evolution for the circular binaries it is easy to see that the observed ratio of these to eccentric binaries [hereafter  $(ns, co)_e$ ] should be  $\sim 50$  because (1) the formation rate of the two types of binaries is, within a factor of 2, the same; (2) the magnetic fields in the circular binaries will be brought down by a factor of  $\sim 100$  by He accretion in the neutron star He star phase following common envelope evolution just as the inferred pulsar magnetic field strengths in the double neutron star binaries are brought down (Brown 1995). In the eccentric binaries the neutron star is formed last, after the white dwarf, so there is nothing to circularize its orbit. More important, its magnetic field will behave like that of a single star and will not be brought down from the  $B \sim 10^{12}$  G with which it is born. (At least empirically, neutron star magnetic fields are brought down only in binaries, by accreting matter from the companion star [Taam & van den Heuvel 1986], although Wijers 1997 shows the situation to be more complex.) Neutron stars with higher magnetic fields can be observed only for shorter times, because of more rapid spin-down from magnetic dipole radiation. The time of possible observation goes inversely with the magnetic field  $B$ . We use the observability premium

$$\Pi = 10^{12} G/B \quad (1)$$

(Wettig & Brown 1996), which gives the relative time a neutron star can be observed. Given our point 2 above, the circular binaries should have an observability premium  $\Pi \sim 100$  as compared with  $\Pi \sim 1$  for the higher magnetic field neutron star in an eccentric orbit. Correcting for the factor of 2 higher formation rate of the eccentric binaries (point 1 above), this predicts the factor of  $\sim 50$  ratio of circular to eccentric binaries.

In our paper we cite one firm eccentric neutron star, carbon-oxygen white dwarf binary  $(ns, co)_e$ , B2303+46, and argue for a recently observed second one, J1141-65. Portegies Zwart & Yungelson (1999) suggest that PSR 1820-11 may also be in this class but cannot exclude the possibility that the neutron star companion is a main-sequence star (Phinney & Verbunt 1991). This would imply that  $\gtrsim 100$  such binaries with circular orbits should be observed. But, in fact, only one,<sup>2</sup> B0655+64, is observed if we accept the developing consensus (§ 2.3) that those observed  $(ns, co)_e$ 's are evolved with avoidance of common envelope evolution. We are thus confronted by a big discrepancy, for which we suggest a solution.

In order to understand our solution, we need to review three past works. In the earlier literature the observed circular  $(ns, co)_c$ 's were evolved through common envelope; e.g. see van den Heuvel (1994) and Phinney & Kulkarni (1994). Accretion from the evolving giant progenitor of the white dwarf was neglected, since it was thought that the accretion would be held to the Eddington rate of  $\dot{M}_{\text{Edd}} \sim 1.5 \times 10^{-8} M_{\odot} \text{ yr}^{-1}$ , and in the  $\sim 1$  year long common envelope evolution a negligible amount of matter would be accreted. We term this the standard scenario. However, Chevalier (1993) showed that once  $\dot{M}$  exceeded  $\sim 10^4 \dot{M}_{\text{Edd}}$ , it was no longer held up by the radiative pressure due to the X-rays from the neutron star but swept them inward in an adiabatic inflow.

<sup>1</sup> Hubble Fellow.

<sup>2</sup> We have a special scenario for evolving it; see § 3.2.

Bethe & Brown (1998) employed this hypercritical accretion in their evolution of double neutron star [(ns, ns)] and neutron star, low-mass black hole binaries [(ns, lmbh)], and we shall use the same techniques in binary evolution here. In particular, these authors found that, including hypercritical accretion in the standard scenario for double neutron star binary evolution, the first-born neutron star evolved into a low-mass black hole. To avoid the neutron star going through the companion's envelope, a new scenario was introduced beginning with a double He star binary. It gives about the right number of double neutron star binaries.

A new development has been that most of the circular (ns, co)'s are currently evolved with avoidance of common envelope evolution. In § 2.3, we shall summarize this work, carried out independently by King & Ritter (1999) and Tauris, van den Heuvel, & Savonije (2000). If we accept the new scenario, at most one or two circular (ns, co)'s that went through common envelope evolution have been observed. Yet, in the standard scenario at least  $\sim 50$  of them should be seen.

In this paper we find that in (ns, co)'s which do go through common envelope evolution, the neutron star evolves into a black hole. The (ns, co) binaries observed to date have been identified through radio emission from the neutron star. Thus, binaries containing a low-mass black hole would not have been seen. We discuss masses for which neutron stars evolve into black holes.

Although the main point of our paper relies only on relative formation rates, we shall show in Appendix A that the Bethe & Brown (1998) schematic, analytic analysis agrees well with the detailed numerical population synthesis of Portegies Zwart, once both are normalized to the same supernova rate.

## 2. THE PROBLEM

### 2.1. Standard Scenario versus Observations

Portegies Zwart & Yungelson (1998), in a very careful population synthesis, have calculated the expected number of newly created binaries of compact stars (neutron stars or black holes) and white dwarfs. Among the latter, they distinguish between those consisting of helium and those consisting of carbon-oxygen (denoted by "co"). To make an eccentric binary containing a neutron star, the supernova must occur after the carbon-oxygen star has formed (Portegies Zwart & Yungelson 1999). To make a circular binary containing a neutron star, it is necessary that its companion be close so that at some earlier stage in evolution (but after formation of the neutron star) there was mass transfer or strong tidal interaction, which requires the companion to (nearly) fill its Roche lobe.

Since the Bethe & Brown (1998) schematic calculations did not include mass exchange, which is very important in evolving (ns, co)<sub>e</sub> binaries, we need the more complete calculations of Portegies Zwart & Yungelson, which are listed in Table 1. We discuss this in more detail later. These do not include hypercritical accretion, i.e., they follow the standard scenario. In this case the formation ratio of (ns, co)<sub>e</sub> to (ns, co)<sub>c</sub> is  $17.7/32.1 = 0.55$ . We now make the case that if the (ns, co)<sub>e</sub> were to be formed through common envelope evolution (Phinney & Kulkarni 1994; van den Heuvel 1994) in the standard scenario, their pulsar magnetic fields would be brought down to  $B \sim 10^{10}$  G because of the similarity to binary neutron star systems in which this occurs. In detail,

TABLE 1  
SIMULATIONS NORMALIZED TO  
SUPERNOVA RATE OF  
 $0.025 \text{ yr}^{-1}$

Binary	Model B ( $10^{-5} \text{ yr}^{-1}$ )
(ns, ns) .....	10.6
(bh, ns) .....	1.9
(ns, co) <sub>e</sub> .....	17.7
(ns, co) <sub>c</sub> .....	32.1

\* Assuming 100% binarity following case B of Portegies Zwart & Yungelson 1998. These simulations do not include hypercritical accretion.

this results from helium accretion during the neutron star, He star stage which precedes the final binary (Brown 1995; Wettig & Brown 1996).

Detailed calculation of Iben & Tutukov (1993) for original donor masses 4–6  $M_{\odot}$  of the white dwarf progenitor show that following common envelope evolution the remnant stars fill their Roche lobes and continue to transfer mass to their companion neutron star. These remnants consist of a degenerate carbon-oxygen core and an evolving envelope undergoing helium shell burning. The mass transfer to the neutron star is at a rate  $\dot{M} < 10^4 M_{\odot} \text{ day}^{-1}$ , the lower limit for hypercritical accretion, so it is limited by Eddington. Van den Heuvel (1994) estimates that the neutron star accretes about  $0.045$  and  $0.024 M_{\odot}$  in the case of the ZAMS 5 and  $6 M_{\odot}$  stars, and  $0.014 M_{\odot}$  for a  $4 M_{\odot}$  star, where these ZAMS masses refer to the progenitors of the white dwarfs. The accretion here is of the same order, roughly double,<sup>3</sup> the wind accretion used by Wettig & Brown (1996) in the evolution of the relativistic binary pulsars B1534+12 and B1913+16. There the magnetic fields were brought down by a factor of  $\sim 100$  from  $B \sim 10^{12}$  to  $B \sim 10^{10}$  G, increasing the observability premium  $\Pi$  by a factor of  $\sim 100$ . Thus, the scenario in which the (ns, co)<sub>e</sub>'s are produced through common envelope evolution without hypercritical accretion should furnish them with  $\Pi \sim 100$ , by helium accretion following the common envelope. Although the detailed description may not be correct, the similarity of the evolution of (ns, co)<sub>e</sub> to that of binary neutron stars in the older works (Phinney & Kulkarni 1994; van den Heuvel 1994) should furnish these with about the same  $\Pi$ .

There is one confirmed (ns, co)<sub>e</sub>, namely, B2303+46 (see Table 2), so there should be about 50 circular ones which went through common envelope evolution. Indeed, several circular ones have been observed (see Table 2), and one or two of these may have gone through common envelope evolution. Thus we have a big discrepancy between the standard scenario and the observations. In the next section, we discuss the possibility of PSR J1141–6545 being (ns, co)<sub>e</sub>, which enhances the discrepancy.

### 2.2. Is PSR J1141–6545 (ns, co)<sub>e</sub>?

Not only is the eccentric B2303+46 quite certain, but a relativistic counterpart, PSR J1141–6545 has recently been

<sup>3</sup> The He-burning time to be used for the progenitor of the white dwarf is  $\sim 10^6$  yr, whereas for the relativistic binary pulsars the average time of  $5 \times 10^5$  yr is more appropriate, so one would expect a factor of  $\sim 2$  greater accretion.

TABLE 2  
BINARY RADIO PULSAR SYSTEMS: (ns, ns) AND (ns, co) BINARIES<sup>a</sup>

Pulsar	$P_{\text{orb}}$ (days)	$P_{\text{spin}}$ (ms)	$f$ ( $M_{\odot}$ )	$M_p$ ( $M_{\odot}$ )	$M_c$ ( $M_{\odot}$ )	$\epsilon$	$d$ (kpc)	$B$ (G)	$\Pi$	References
(ns, ns)										
J1518+4904 .....	8.634	40.9	0.116	<1.75	>0.93	0.249	0.70	<1.3 × 10 <sup>9</sup>	>769	
B1534+12 .....	0.421	37.9	...	1.339	1.339	0.274	1.1	10 <sup>10</sup>	100	
B1913+16 .....	0.323	59.0	...	1.441	1.387	0.617	7.13	2.3 × 10 <sup>10</sup>	43	
B2127+11C <sup>b</sup> .....	0.335	30.5	...	1.349	1.363	0.681	10	1.2 × 10 <sup>10</sup>	83	
(ns, co) <sub>c</sub>										
B2303+46 .....	12.34	1066	0.246	<1.44	>1.20	0.658	4.35	7.9 × 10 <sup>11</sup>	1.26	1, 2
J1141-6545 <sup>c</sup> .....	0.198	394	0.177	<1.348	>0.97	0.172	3.2	1.3 × 10 <sup>12</sup>	0.77	3
(ns, co) <sub>w</sub> <sup>d</sup>										
J2145-0750 .....	6.839	16.1	0.024	...	0.515	2.1 × 10 <sup>-5</sup>	0.5	6 × 10 <sup>8</sup>	1667	4
J1022+1001 .....	7.805	16.5	0.083	...	0.872	9.8 × 10 <sup>-5</sup>	0.6	8.4 × 10 <sup>8</sup>	1190	5
J1603-7202 .....	6.309	14.8	0.009	...	0.346	<2 × 10 <sup>-5</sup>	1.6	4.6 × 10 <sup>8</sup>	2173	6
J0621+1002 .....	8.319	28.9	0.027	...	0.540	0.00245	1.9	1.6 × 10 <sup>9</sup>	625	7
B0655+64 .....	1.029	195.7	0.071	...	0.814	0.75 × 10 <sup>-5</sup>	0.48 <sup>e</sup>	1.26 × 10 <sup>10</sup>	79	8
J1810-2005 .....	15.01	32.8	0.0085	...	0.34	...	...	2.1 × 10 <sup>9</sup>	476	
J1157-5112 .....	3.507	43.6	0.2546	...	>1.20	...	...	<6.3 × 10 <sup>8</sup>	159	
J1232-6501 .....	1.863	88.3	0.0014	...	0.175	...	...	9.5 × 10 <sup>8</sup>	105	
J1453-58 .....	12.42	45.3	0.13	...	1.07	0.0019	...	6.1 × 10 <sup>9</sup>	164	
J1453-60 .....	1.355	9.35	0.14	...	1.10	1 × 10 <sup>-5</sup>	...	4.7 × 10 <sup>8</sup>	2127	
J1756-5322 .....	0.453	8.87	0.0475	...	0.683	...	...	...	...	

<sup>a</sup> The observability premium  $\Pi = (10^{12} \text{ G})/B$ .  $M_p$  ( $M_{\odot}$ ) means the pulsar (companion) mass, and  $f$  the mass function.

<sup>b</sup> Binary in globular cluster M15.

<sup>c</sup> Not confirmed yet.

<sup>d</sup> The white dwarf mass  $M_c$  is calculated assuming  $M_p = 1.4 M_{\odot}$  and  $i = 60^{\circ}$ .

<sup>e</sup> Assumed distance.

REFERENCES.—(1) Thorsett & Chakrabarty 1999; (2) van Kerkwijk & Kulkarni 1999; (3) Kaspi et al. 2000; (4) Bailes et al. 1994; (5) Camilo 1995; (6) Lorimer et al. 1996; (7) Camilo et al. 1996; (8) van Kerkwijk & Kulkarni 1995.

observed (Kaspi et al. 2000), in an eccentric orbit. The inferred magnetic dipole strength is  $1.3 \times 10^{12} \text{ G}$ , and the total mass is  $2.300 \pm 0.012 M_{\odot}$ . Kaspi et al. argue that the companion of the neutron star can only be a white dwarf, or neutron star. With a total mass of  $2.3 M_{\odot}$ , if J1141-65 were to contain two neutron stars, each would have to have a mass of  $\sim 1.15 M_{\odot}$ , well below the 19 accurately measured neutron star masses; see Figure 1 (Thorsett & Chakrabarty 1999).

We can understand the absence of binary neutron stars with masses below  $\sim 1.3 M_{\odot}$ , although neutron stars of this mass are expected to result from the relatively copious main-sequence stars of ZAMS mass  $\sim 10-13 M_{\odot}$  from the argument of Brown (1997). The He stars in the progenitor He star, pulsar binary of mass  $\lesssim 4 M_{\odot}$  (Habets 1986a, 1986b) expand substantially during He-shell burning. Accretion onto the nearby pulsar sends it into a black hole. Indeed, with inclusion of mass loss by helium wind, He stars of masses up to 6 or  $7 M_{\odot}$  expand in this stage (Woosley, Langer, & Weaver 1995). Fryer & Kalogera (1997) find that special kick velocities need to be selected in order to avoid the evolution of PSR 1913+16 and PSR 1534+12 from evolving into a black hole by reverse case C mass transfer (mass transfer from the evolving He star companion onto the pulsar in the He star, neutron star stage which precedes that of the binary of compact objects).

Our argument above says that the first neutron star formed in these binaries would be sent into a black hole when its companion He star evolved and poured mass on it. Therefore, we believe the companion in J1141-65

must be a white dwarf. Earlier Tauris & Sennels (2000) developed the case that J1141-65 was an eccentric neutron star, white dwarf binary. Given the high magnetic field of J1141-65 ( $1.3 \times 10^{12} \text{ G}$ ) with low observability premium of 0.77, this would increase the predicted observed number of circular (ns, co)<sub>w</sub>'s which had gone through common envelope evolution to  $\sim 130$  in the standard scenario.

### 2.3. Evolution of Neutron Star, Carbon-Oxygen White Dwarf Binaries with Avoidance of Common Envelope Evolution

Our discussion of the common envelope evolution in the preceding section applied to convective donors. In case the donor is radiative or semiconvective, common envelope evolution can be avoided. Starting from the work of Savonije (1983), van den Heuvel (1995) proposed that most low-mass X-ray binaries would evolve through a Hercules X-1 type scenario, where the radiative donor, more massive than the neutron star, poured matter onto its accretion disk at a super-Eddington rate, during which time almost all of the matter was flung off. This involved Roche lobe overflow. Although van den Heuvel limited the ZAMS mass of the radiative donor to  $2.25 M_{\odot}$  in order to evolve the helium white dwarf, neutron star binaries, his scenario has been extended to higher ZAMS mass donors in order to evolve the carbon-oxygen white dwarf neutron star binaries. The advection-dominated inflow-outflow solutions (ADIOSs) of Blandford & Begelman (1999) suggest that the binding energy released at the neutron star can carry away mass, angular momentum and energy from the gas accreting onto

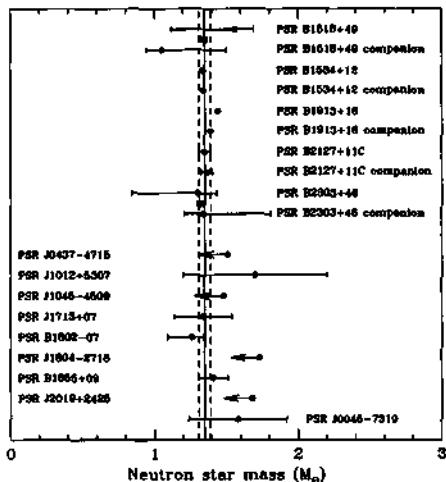


FIG. 1.—Neutron star masses from observations of radio pulsar system (Thorsett & Chakrabarty 1999). All error bars indicate central 68% confidence limits, except that upper limits are one-sided 95% confidence limits. Five double neutron star systems are shown at the top of the diagram. In two cases, the average neutron star mass in a system is known with much better accuracy than the individual masses; these average masses are indicated with open circles. Eight neutron star-white dwarf binaries are shown in the center of the diagram, and one neutron star-main-sequence star binary is shown at bottom. Vertical lines are drawn at  $m = 1.35 \pm 0.04 M_{\odot}$ . As noted in our paper, PSR B2303+46 has since been shown to have a white dwarf companion.

the accretion disk, provided the latter does not cool too much. In this way the binding energy of gas at the neutron star can carry off  $\sim 10^3$  g of gas at the accretion disk for each gram accreting onto the neutron star. King & Begelman (1999) suggest that such radiatively driven outflows allow the binary to avoid common envelope evolution.

As noted above, for helium white dwarf companions, van den Heuvel (1995) had suggested Cygnus X-2 as an example following the Her X-1 scenario. King & Ritter (1999) calculated the evolution of Cyg X-2 in the ADIOS scenario in detail. These authors also evolved the  $(ns, co)_s$  binaries in this way, using donor stars of ZAMS masses 4–7  $M_{\odot}$ .

TABLE 3  
FLUX DENSITIES AT 400 MHz\*

Pulsar	$S_{400}$ (mJy)	$S_{400} d^2$ (mJy kpc $^2$ )
$(ns, ns)$		
B1534+12.....	36.00	43.56
B1913+16.....	4.00	203.35
B2127+11C.....	0.60	56.45
$(ns, co)_s$		
B2303+46.....	1.90	35.95
$(ns, co)_e$		
J2145-0750.....	50.00	12.50
J1022+10.....	23.00	8.28
B0655+64.....	5.00	1.15

\* From Princeton Pulsar Group, The Pulsar Catalog (<http://pulsar.princeton.edu>).

Tauris et al. (1999) have carried out similar calculations, with stable mass transfer. These authors find that even for extremely high mass-transfer rates, up to  $\dot{M} \sim 10^4 M_{\odot} \text{ yr}^{-1}$ , the system will be able to avoid a common envelope and spiral-in evolution.

Tauris et al. (2000) evolve J1453–58, J1435–60, and J1756–5322, the three lowest entries in our Table 2, through common envelope. We obtained the eccentricities and values of  $\dot{P}$  for the first two of these (F. Fernando Camilo 2000, private communication). The binary J1453–58, quite similar to J0621+1002, has a substantial eccentricity and clearly should be evolved with a convective donor as Tauris et al. did for J0621+1002. The spin periods of J1435–60 and J1756–5322 are short, indicating greater recycling than the other listed pulsars. It would seem difficult to get the inferred magnetic field down to the  $4.7 \times 10^8$  G of J1435–60 by the Iben & Tutukov or Wettig & Brown accretion scenarios following common envelope evolution as discussed in § 2.1. If, however, one does believe that J1435–60 and J1756–5322 have gone through common envelope, the discrepancy between predicted and observed circular binaries in the standard scenario is only slightly relieved.

#### 2.4. Are There Observational Selection Effects?

In Table 3 we have tabulated  $S_{400} d^2$  in order to see whether the normalized intensity gives strong selection effects. Note that 35.95 for B2303+46 is not so different from 43.56 and 203.35 for B1534+12 and B1913+16, respectively. For the circular binaries  $(ns, co)_s$ , the intensities are less, but their empirical observability premium  $\Pi$  is much larger. There may be other observational selection effects, but, we believe that there are no observational selection effects strong enough to compensate for the factor of 100 discrepancy between the observed population and the one expected from the standard model. So the problem remains the same.

### 3. THE ANSWER

#### 3.1. Black Hole Formation in Common Envelope Evolution

We believe the explanation of the missing binaries is that the neutron star goes into a black hole in common envelope evolution, as we now describe. We label the mass of the neutron star as  $M_A$  and that of the giant progenitor of the white dwarf as  $M_B$ . Following Bethe & Brown (1998), we choose as variables the neutron star mass  $M_A$  and  $Y \equiv M_B/a$ , where  $a$  is the orbital radius. From their equation (5.12) we find

$$\frac{M_{A,f}}{M_{A,i}} = \left( \frac{Y_f}{Y_i} \right)^{c_d - 1}, \quad (2)$$

where  $c_d$  is the drag coefficient. From Shima et al. (1985) we take

$$c_d = 6. \quad (3)$$

We furnish the energy to remove the hydrogen envelope of the giant B (multiplied by  $\alpha_{ce}^{-1}$ , where  $\alpha_{ce}$  is the efficiency of coupling of the orbital motion of the neutron star to the envelope of B) by the drop in orbital energy of the neutron star; i.e.,

$$\frac{0.6GM_{B,i}Y_i}{\alpha_{ce}} = \frac{1}{2} GM_{A,i}Y_i \left( \frac{Y_f}{Y_i} \right)^{6/5}. \quad (4)$$

Here the  $0.6GM_{B,i}Y_i$  is just the binding energy of the initial giant envelope, found by Applegate (1997) to be  $0.6GM_{B,i}^2a_i^{-1}$ , and the right-hand side of the equation is the final gravitational binding energy  $\frac{1}{2}GM_{A,f}M_{B,f}a_f^{-1}$  in our variables. Using equations (2) and (3) in equation (4), one finds

$$\frac{M_{A,f}}{M_{A,i}} = \left( \frac{1.2M_{B,i}}{\alpha_{ee} M_{A,i}} \right)^{1/c_4}. \quad (5)$$

For the sake of argument, we take the possible range of initial neutron star mass to be  $1.2\text{--}1.5 M_\odot$  (the upper bound is the Brown & Bethe 1994 mass at which a neutron star goes into a low-mass black hole), and the main-sequence progenitor masses of the carbon-oxygen white dwarf to be  $M_{B,i} = 2.25\text{--}10 M_\odot$ . As we show in Appendix C, in the Bethe & Brown (1998) schematic model, mass transfer was assumed to take place when the evolving giant reached the neutron star, whereas more correctly it begins when the envelope of the giant comes to its Roche lobe. For the masses we employ, main-sequence progenitors of the carbon-oxygen white dwarf of  $2.25\text{--}10 M_\odot$ , the fractional Roche lobe radius is

$$r_L \sim 0.5. \quad (6)$$

The binding energy of the progenitor giant at its Roche lobe is, thus, double what it would be at  $a_i$ , the separation of giant and neutron star. Therefore, a Bethe & Brown  $\alpha_{ee} = 0.5$  corresponds to a true efficiency  $\alpha_{ee} \sim 1$ , if the latter is defined as the value for which the envelope removal energy, at its Roche lobe, is equal to the drop in neutron star orbital energy as it moves from  $a_i$  to  $a_f$ . If we take  $\alpha_{ee} = 0.5$  in equation (5), we find, given our assumed possible intervals,

$$1.54 M_\odot \lesssim M_{A,f} \lesssim 2.38 M_\odot. \quad (7)$$

These are above the neutron star mass limit  $1.5 M_\odot$  (Brown & Bethe 1994) beyond which a neutron star goes into a low-mass black hole. Thus, all neutron stars with common envelope evolution in our scenario evolve into black holes. This solves the big discrepancy between the standard scenario and observation. The only remaining problem is the evolution of B0655+64, which survived the common envelope evolution, and we suggest a special scenario for it in the next section.

### 3.2. Is B0655+64 a Problem?

Van den Heuvel & Taam (1984) were the first to notice that the (*ns, co*)<sub>+</sub> system B0655+64 might have been formed in a similar way to the double neutron stars. The short period of 1.03 days, magnetic field  $\sim 10^{10}$  G, and the high companion mass of  $\sim 1 M_\odot$  make this binary most similar to a binary neutron star, but with a carbon-oxygen white dwarf companion, resulting from probable ZAMS masses  $\sim 5\text{--}8 M_\odot$ . For a  $1.4 M_\odot$  neutron star with a  $1 M_\odot$  white dwarf companion,  $a_f = 5.7 R_\odot$ .

The similarity of B0655+64 to the close neutron star binaries suggests the double helium star scenario (Brown 1995) to calculate the evolution. The ZAMS mass of the primary is chosen to be just above the limit for evolving into a neutron star, that of the secondary just below. For the double He star scenario the ZAMS masses of primary and secondary cannot be more than  $\sim 5\%$  different. However, in this case the ratio  $q$  of masses is so close to unity that the secondary will not be rejuvenated (Braun & Langer 1995): If

the core burning of hydrogen to helium in the companion star is nearly complete, the accreted matter would have to cross a molecular weight barrier in order to burn, and if  $q$  is near unity there is not time enough to do so. Thus He cores of both stars will evolve as if the progenitors never had more than their initial ZAMS mass.)

What we have learned recently about effects of mass loss (Wellstein & Langer 1999) will change the Brown (1995) scenario in detail, but not in general concept. An  $\sim 10 M_\odot$  ZAMS star that loses mass in RLOF to a lower mass companion will burn helium as a lower mass star due to subsequent mass loss by helium winds, roughly as an  $8 M_\odot$  star (S. Wellstein & N. Langer 2000, in preparation). Thus, the primary must have ZAMS mass  $\gtrsim 10 M_\odot$  in this case in order to evolve into a neutron star following mass loss. Although the secondary will not be rejuvenated as mass is transferred to it, it will burn helium without helium wind loss because it is clothed with a hydrogen envelope. Thus, a secondary of ZAMS  $8 M_\odot$  will burn He roughly as the primary of  $10 M_\odot$  in the situation considered. Given these estimates, a primary of ZAMS mass  $M \lesssim 10 M_\odot$  will evolve into a white dwarf, whereas a secondary of mass  $\gtrsim 8 M_\odot$  will end up as a neutron star. Of course, the former must be more massive than the latter, but stars in this mass range are copious because this is the lowest mass range from which neutron stars can be evolved, so there will be many such cases.

This scenario might not be as special as outlined because the fate of stars of ZAMS mass  $8\text{--}10 M_\odot$ , which do not form iron cores but do burn in quite different ways from more massive stars, is somewhat uncertain in the literature. Whereas it is generally thought that single stars in this range end up as neutron stars, it has also been suggested that some of them evolve as AGB stars ending as white dwarfs. In terms of these discussions it does not seem unlikely that with two stars in the binary of roughly the same mass, the first to evolve will end up as a neutron star and the second as a white dwarf, especially if the matter transferred in RLOF cannot rejuvenate the companion.

Van den Heuvel & Taam (1984) evolved B0655+64 by common envelope evolution. In taking up the problem again, Tauris et al. (2000) in agreement with King & Ritter (1999) find that B0655+64 cannot be satisfactorily evolved with their convective donor scenario. Tauris et al. suggest that a spiral-in phase is the most plausible scenario for the formation of this system, but we find that the neutron star would evolve into a black hole in this scenario, unless the two progenitors burn He at the same time.

### 3.3. Neutron Star Masses

There is by no means agreement about maximum and minimum neutron star masses in the literature. The mass determinations of Vela X-1 have been consistently higher than the Brown & Bethe  $1.5 M_\odot$ , which is consistent with well-measured neutron star masses in Figure 1. In a recent careful study at ESO, Barziv et al. (2000, in preparation), as reported by van Kerkwijk (2000), obtain

$$M_{\text{NS}} = 1.87^{+0.23}_{-0.17} M_\odot. \quad (8)$$

Even at the 99% confidence level,  $M_{\text{NS}} > 1.6 M_\odot$ . Taking the maximum mass to be  $1.87 M_\odot$  and  $\alpha_{ee} = 0.5$ ,  $(M_{\text{NS}})_{\text{min}} = 1.2 M_\odot$  one finds from equation (5) that the maximum carbon-oxygen white dwarf progenitor mass of

$(ns, co)_*$  is

$$(M_{B,i})_{\max} = \alpha_{\infty} \left( \frac{M_{A,i}}{M_{A,i}} \right)^6 \frac{M_{A,i}}{1.2} \approx 7.2 M_{\odot}. \quad (9)$$

Although there is some uncertainty in the efficiency  $\alpha_{\infty}$ , the ratio  $M_{B,i}/M_{\odot}$  is much more sensitive to  $M_{A,i}$  because of the 6th power of the ratio in equation (9).<sup>4</sup> But then one cannot explain why no  $(ns, co)_*$  (except B0655+64) that survived the common envelope evolution are seen, since this mass is high enough to give those of the white dwarf companions.

Distortion of the  $\sim 20 M_{\odot}$  B star companion by the neutron star in Vela X-1 brings in large corrections (Zuiderwijk et al. 1977; van Paradijs, Takens, & Zuiderwijk 1977a; van Paradijs et al. 1997b), making measurement of neutron star masses in high-mass X-ray binaries much more difficult than those with degenerate companions.

Given that  $Y_e \approx 0.43$  at the collapse of the core of a large star (Aufderheide et al. 1990), one finds the cold Chandrasekhar mass to be

$$M_{CS} = 5.76 Y_e^2 M_{\odot} \approx 1.06 M_{\odot}, \quad (10)$$

where  $Y_e$  is the ratio of the number of electrons to the number of nucleons. Thermal corrections increase this a bit, whereas lattice corrections on the electrons decrease it, so that when all is said and done,  $M_{CS} \gtrsim 1.1 M_{\odot}$  (Shapiro & Teukolsky 1983). The major dynamical correction to this is from fallback following the supernova explosion. We believe that fallback in supernova explosions will add at least  $\sim 0.1 M_{\odot}$  to the neutron star, since bifurcation of the matter going out and in happens at about 4000 km (Bethe & Brown 1995). Thus our lower limit of  $\sim 1.2 M_{\odot}$  is reasonable.

#### 4. DISCUSSION AND CONCLUSIONS

At least one, but more likely two or more,  $(ns, co)_*$  binaries with an unrecycled pulsar have been observed. According to the standard scenario for evolving neutron stars that are recycled in a common envelope evolution, we then expect to observe  $\gtrsim 50$   $(ns, co)_*$ 's. We only observe B0655+64 (which we evolve in our double He star way) and possibly one or two binaries that went through common envelope evolution, and from that we conclude that the standard scenario must be revised. Introducing

<sup>4</sup> With  $M_{NS} = 1.5 M_{\odot}$  we get  $(M_{B,i})_{\max} = 1.9 M_{\odot}$ , which is below the minimum  $M_B$  ( $\sim 2.25 M_{\odot}$ ) for forming a carbon-oxygen white dwarf, so no  $(ns, co)_*$ 's survive the common envelope evolution.

hypercritical accretion into common envelope evolution (Brown 1995; Bethe & Brown 1998) removes the discrepancy.

We believe that the evolution of the other  $(ns, co)_*$  binaries may originate from systems with a neutron star with a radiative or semiconvective companion. The accretion rate in these systems can be as high as  $10^4 M_{\odot} \text{ yr}^{-1}$ , but common envelope evolution is avoided. This possibility, however, does not affect our conclusion concerning hypercritical accretion.

It is difficult to see "fresh" (unrecycled) neutron stars in binaries because they do not shine for long. B2303+46 (Table 2) is the firmest example of a  $(ns, co)_*$  binary with a fresh neutron star. Although binaries where a "fresh" neutron star is accompanied by a black hole have similar birthrates ( $\sim 10^{-4} \text{ yr}^{-1}$  for both types; Bethe & Brown 1998; Portegies Zwart & Yungelson 1999) and lifetime, none are observed. In Appendix A we quote results of Ramachandran & Portegies Zwart (1998) that show that there is an observational penalty which disfavors the observation of neutron stars with black holes as companions, because of the difficulty in identifying the pulsar, owing to the Doppler shift which smears out the signal in these short-period objects. Because of the longer orbital period and lower companion mass of  $(ns, co)_*$ 's, such binaries are less severely plagued by this effect, although the recently discovered J1141-6545 is a relativistic binary with 5 hr period. We therefore argue that it is not unreasonable that no  $(imbh, ns)$  binaries have yet been observed, but that they should be actively searched for, since the probability of seeing them is not far lower than that of seeing the  $(ns, co)_*$ 's.

We are grateful to Fernando Camilo for information in Table 2. We would like to thank Brad Hansen, Marten van Kerkwijk, Ralph Wijers, Thomas Tauris, and Lev Yungelson for useful discussions and advice, and thank Justin Holmer for providing us with the convective envelopes for stars in the giant phase. G. E. B. would like to thank Brad Hansen also for correspondence which started him on this problem. G. E. B. and C. H. L. were supported by the US Department of Energy under grant DE-FG02-88ER40388. S. P. Z. was supported by NASA through Hubble Fellowship grant HF-01112.01-98A by the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., for NASA under contract NAS 5-26555. S. P. Z. is grateful for the hospitality of the State University of New York at Stony Brook.

#### APPENDIX A

#### COMPARISON OF POPULATION SYNTHESES

In this appendix we first compare results that the Bethe & Brown (1998) schematic analytic evolution would have given without hypercritical accretion with the Portegies Zwart & Yungelson (1998, 1999) results of Table 1, which do not include hypercritical accretion. We can then illustrate how hypercritical accretion changes the results.

Without hypercritical accretion the  $(imbh, ns)$  binaries of Bethe & Brown would end up rather as neutron star binaries,  $(ns, ns)$ , giving a summed formation rate of  $1.1 \times 10^{-4} \text{ yr}^{-1}$ , to compare with  $1.1 \times 10^{-4} \text{ yr}^{-1}$  from the Portegies Zwart numerical driven population synthesis results presented in Table 1. This good agreement indicates that kicks that the neutron star receives in formation were implemented in the same way in the two syntheses. Introduction of hypercritical accretion leaves only those neutron stars that do not go through a common envelope, i.e., those in the double He star scenario of Brown (1995), with formation rate  $10^{-5} \text{ yr}^{-1}$ . This is much closer to the estimated empirical rate of  $8 \times 10^{-6} \text{ yr}^{-1}$  of van den Heuvel

& Lorimer (1996), which equals the rate derived independently by Narayan, Piran, & Shemi (1991) and Phinney (1991). Large poorly known factors are introduced in arriving at these "empirical" figures, so it is useful that our theoretical estimates end up close to them. In our theoretical estimates the possibility described earlier that the pulsar in the lower mass binary pulsars goes into a black hole in the He shell burning stage of the progenitor He star, neutron star binary (Brown 1997) was not taken into account, and this process may change roughly half of the remaining neutron star binaries in our evolution into  $(mbh, ns)$  binaries.

Bethe & Brown (1998) had a numerical symmetry between high-mass binaries in which both massive stars become supernovae and those in which the more massive one becomes a supernova and the other, below the assumed dividing mass of  $10 M_{\odot}$ , does not; i.e., the number of binaries was equal in the two cases. Taking ZAMS mass progenitors of  $2.3-10 M_{\odot}$  for carbon-oxygen white dwarfs, we then find a rate of

$$R = 2 \left( \frac{10-2.3}{10} \right) \times 1.1 \times 10^{-4} \text{ yr}^{-1} = 16.9 \times 10^{-5} \text{ yr}^{-1} \quad (\text{A1})$$

for the formation rate of  $(ns, co)_s$  binaries. The  $1.1 \times 10^{-4} \text{ yr}^{-1}$  is taken from the last paragraph and applies here because of the numerical symmetry mentioned above. The factor of 2 results because there is no final explosion of the white dwarf to disrupt the binary, as there was above in the formation of the neutron star. This rate  $R$  is to be compared with the  $(ns, co)_s$  rate of  $17.7 \times 10^{-5} \text{ yr}^{-1}$  in Table 1. These  $(ns, co)_s$  binaries are just the ones in which the neutron star becomes a black hole in common envelope evolution, unless the masses of the two initial progenitors are so close that they burn He at the same time. Then a binary such as B0655 + 64 can result, since the two helium stars then go through a common envelope rather than the neutron star and main-sequence star.

It is of interest to compare the populations of  $(ns, ns)$  binaries with the  $(ns, co)_s$  binaries. We must rely on the Portegies Zwart result for the latter, which cannot be evolved without mass transfer, which is not included in the Bethe & Brown evolution. The  $(ns, ns)$  binaries involve common envelope evolution, whereas the  $(ns, co)_s$  do not. Thus, results for the rates should differ substantially in the standard scenario, which does not include hypercritical accretion, and our scenario which does. The ratio for  $(ns, ns)$  and  $(ns, co)_s$  binaries from Table 1 are  $10.6 \times 10^{-5} \text{ yr}^{-1}$  and  $32.1 \times 10^{-5} \text{ yr}^{-1}$ . The  $(ns, ns)$ 's are recycled in the He star pulsar stage by the He wind, giving an observability premium of  $\Pi \sim 100$  (Brown 1995). The pulsar in the  $(ns, co)_s$  is not recycled. Thus, the expected observational ratio is

$$\frac{(ns, ns)}{(ns, co)_s} \sim \frac{100 \times 10.6}{32.1} \sim 33. \quad (\text{A2})$$

Now B2303 + 46, and possibly B1820 - 11 and J1141 - 6545 lie in the  $(ns, co)_s$  class, whereas B1534 + 12 and B1913 + 16 are relativistic binary neutron stars with recycled pulsars. We do not include the neutron star binary 2127 + 11C, although it has the same  $B$  as the other two. It is naturally explained as resulting from an exchange reaction between a neutron star and a binary which took place less than  $10^8$  years ago in the cluster core of M15 (Phinney & Sigurdsson 1991). Thus, the empirical ratio in equation (A2) is not much different from unity. In Bethe & Brown (1998) common envelope evolution cuts the  $(ns, ns)$  rate down by a factor of 11, only the 1/11 of the binaries which burn He at the same time surviving. The remaining factor of 3 is much closer to observation. Furthermore, Ramachandran & Portegies Zwart (1998) point out that there is an observational penalty of a factor of several disfavoring the relativistic binary neutron stars because of the difficulty in identifying them caused by the Doppler shift which smears out the signal in these short-period objects. Some observational penalty should, however, also be applied to J1141 - 6545, which is a relativistic binary. We estimate that the combination of neutron stars evolving into black holes in common envelope evolution and the greater difficulty in seeing them will bring the ratio of 33 in equation (A2) down to  $\sim 1$  or 2, close to observation.

We have shown that there is remarkable agreement between the Bethe & Brown (1998) schematic analytic population synthesis and the computer-driven numerical synthesis of Portegies Zwart & Yungelson (1998). This agreement can be understood by the scale invariance in the assumed logarithmic distribution of binary separations. In general we are interested in the fraction of binaries that end up in a given interval of  $a$ . E.g., in Bethe & Brown (1998) that fraction was

$$d\phi = \frac{d(\ln a)}{7}, \quad (\text{A3})$$

where  $d(\ln a)$  was the logarithmic interval between the  $a_i$  below which the star in the binary would merge in common envelope evolution and  $a_f$ , the largest radius for which they would merge in a Hubble time. Here 7 was the assumed initial logarithmic interval over which the binaries were distributed. Thus, the desired fraction

$$d(\ln a) = \Delta a/a \quad (\text{A4})$$

is scale invariant. Mass exchange in the evolution of the binary will change the values of  $a_i$  and  $a_f$ , delineating the favorable logarithmic interval, but will not change the favorable  $d(\ln a)$ . Of course, when He stars become neutron stars, the probability of the binary surviving the kick velocity does depend on the actual value of  $a$ , violating the scale invariance. But this does not seem to be a large effect in the calculations. In the case of the formation of  $(ns, co)_s$  binaries, the neutron star is formed last, out of the more massive progenitor. Mass transfer is required for this, because otherwise the more massive progenitor would explode first. The mass not only must be transferred but must be accepted, so that the companion star is rejuvenated (unless  $q \sim 1$  as discussed). We need the Portegies Zwart & Yungelson detailed numerical program for this. In fact, in calculations with this program (see Table 1) the formation of  $(ns, co)_s$  binaries is nearly double that of  $(ns, co)_e$  binaries. However, for

$q \gtrsim 0.75$ , where  $q$  is the mass ratio of original progenitors, of the ZAMS progenitors Braun & Langer (1995) showed that the transferred hydrogen has trouble passing the molecular weight barrier in the companion, so that the latter would not be rejuvenated. We have not included this effect here, but roughly estimate that it will lower the predicted numbers of  $(ns, co)_s$ 's by a factor greater than 2, bringing it down below the number of  $(ns, co)_s$ 's, exacerbating the problems of the standard model of binary evolution.

In the literature one sees statements such as "Population syntheses are plagued by uncertainties." It is, therefore, important to show that when the same assumptions about binary evolution are made and when the syntheses are normalized to the same supernova rates, similar results are obtained. The evolution in the Bethe & Brown (1998) schematic way is simple, so that effects of changes in assumptions are easily followed.

## APPENDIX B

### HYPERCRITICAL ACCRETION

We develop here a simple criterion for the presence of hypercritical accretion. We further show that if it holds in common envelope evolution for one separation  $a$  of the compact object met by the expanding red giant or supergiant, it will also hold for other separations and for other times during the spiral-in. We assume the envelope of the giant to be convective.

In the rest frame of the compact object, Bondi-Hoyle-Lyttleton accretion of the envelope matter (hydrogen) of density  $\rho_\infty$  and velocity  $V$  is (for  $\Gamma = 5/3$  matter)

$$\dot{M} = 2.23 \times 10^{29} (M_\infty/M_\odot)^2 V_s^{-3} \rho_\infty \text{ g s}^{-1}, \quad (B1)$$

where  $M_\infty$  is the mass of the compact object, and  $V_s$  is the velocity in units of  $1000 \text{ km s}^{-1}$ ,  $\rho_\infty$  is given in  $\text{g cm}^{-3}$ . From Brown (1995) the minimum rate for hypercritical accretion is

$$\frac{\dot{M}_\text{cr}}{\dot{M}_\text{Edd}} = 1.09 \times 10^4. \quad (B2)$$

For hydrogen,

$$\dot{M}_\text{cr} = 0.99 \times 10^{22} \text{ g s}^{-1}. \quad (B3)$$

Using equations (B1) and (B2), we obtain

$$(\rho_\infty)_\text{cr} = 0.44 \times 10^{-7} (M_\odot/M_\infty)^2 V_s^3 \text{ g cm}^{-3}. \quad (B4)$$

Using Kepler for circular orbits

$$V^2 = \frac{GM_\text{tot}}{a}, \quad (B5)$$

where  $M_\text{tot}$  is the mass of the compact object plus the mass of the helium core of the companion plus the envelope mass interior to the orbit of the compact object. One finds

$$(\rho_\infty)_\text{cr} = 2.1 \times 10^{-9} \left( \frac{M_\odot}{M_\infty} \right)^2 \left( \frac{M_\text{tot}/10 M_\odot}{a_{12}} \right)^{3/2} \text{ g cm}^{-3}. \quad (B6)$$

The  $a$ -dependence of  $(\rho_\infty)_\text{cr}$  is the same as the asymptotic density for the  $n = 3/2$  polytrope which describes the convective envelope. Thus, if the criterion for hypercritical accretion is satisfied at one time and at one radius, it will tend to be satisfied for other times and for other radii. The change of  $M_\text{tot}$  with  $a$  is unimportant because from Table 4 it can be seen that  $\rho > (\rho_\infty)_\text{cr}$  already near the surface of the star.

TABLE 4  
DENSITIES FOR THE HYDROGEN  
ENVELOPE OF A  $4 M_\odot$  STAR  
OF RADIUS  $100 R_\odot$ \*

$r$ ( $R_\odot$ )	$\rho$ ( $\text{g cm}^{-3}$ )	$(\rho_\infty)_\text{cr}$
1 .....	0	...
0.95 .....	5.0(-11)	7.8(-13)
0.90 .....	3.5(-9)	8.5(-13)
0.85 .....	5.8(-8)	9.3(-13)
0.80 .....	3.9(-7)	10.1(-13)

\* From Holmer 1998.

In order to check the applicability of hypercritical accretion to the compact object in the relatively low-mass stars we consider in this paper, we make application to a  $4 M_{\odot}$  red giant of radius  $R = 100 R_{\odot}$ , evolved as pure hydrogen but with inclusion of dissociation by Holmer (1998). In Table 4 we compare the densities in the outer part of the hydrogen envelope with those needed for hypercritical accretion. From the table it can be seen that hypercritical accretion sets in quickly, once the compact object enters the envelope of the evolving giant.

Note that the accretion through most of the envelope will be greater than  $1 M_{\odot} \text{ yr}^{-1}$ . Since the total mass accreted by the neutron star is  $\sim 1 M_{\odot}$ , this gives a dynamical time of  $\lesssim 1$  year, although the major part of the accretion takes place in less time. This is in agreement with the dynamical time found, without inclusion of accretion, by Terman, Taam, & Hernquist (1995).

## APPENDIX C EFFICIENCY

We discuss the definition of the efficiency of the hydrodynamical coupling of the orbital motion of the neutron star to the envelope of the main-sequence star.

Van den Heuvel (1994) starts from the Webbink (1984) energetics in which the gravitational binding energy of the hydrogen envelope of the giant is taken to be

$$E_{\text{env}} = -\frac{G(M_{\text{core}} + M_{\text{env}})M_{\text{env}}}{R}, \quad (\text{C1})$$

which results in the envelope gravitational energy

$$E_{\text{env}} = -\frac{0.7GM^2}{R}, \quad (\text{C2})$$

where  $M = M_{\text{core}} + M_{\text{env}}$  is the total stellar mass and the Bethe & Brown (1998) approximation  $M_{\text{core}} \simeq 0.3M$  has been used.

Applegate (1997) has calculated the binding energy of a convective giant envelope, obtaining

$$E_B = -\frac{0.6GM^2}{R} = \frac{1}{2} E_{\text{env}}. \quad (\text{C3})$$

Note that  $M$  is the total stellar mass, also that  $E_B$  is just half of the gravitational potential energy, the kinetic energy being included in  $E_B$ . Equation (C3) was checked independently by Holmer (1998). Unfortunately, this work was never published. Van den Heuvel and others have introduced an additional parameter  $\lambda$  that takes into account both the kinetic energy and the density distribution of the star,  $R$  in equations (C1) and (C2) being replaced by  $R\lambda$ . They use

$$E_B = -\frac{0.7GM^2}{\lambda R}. \quad (\text{C4})$$

With  $\lambda = 7/6$  this is the same as  $E_B$  in equation (C3). Van den Heuvel (1994) chooses  $\lambda = 1/2$ . The result is that his efficiency  $\eta$  is a factor of  $7/3$  too high. Thus, his suggested efficiency  $\eta = 4$  is more like  $\eta \sim 12/7 = \hat{\alpha}_{\text{ee}}$ .

Bethe & Brown (1998) used the Applegate result but incorrectly took the energy necessary to expel the giant envelope as

$$E_g = -0.6GM^2/a_1, \quad (\text{C5})$$

rather than

$$E_g = -0.6GM^2/a_1 r_L, \quad (\text{C6})$$

the latter being the correct energy needed to remove the giant envelope at its Roche lobe. The correct efficiency is

$$\hat{\alpha}_{\text{ee}} = (\alpha_{\text{ee}})_{\text{BB}}/r_L, \quad (\text{C7})$$

and since for the binaries considered here with  $q \sim 4$  the fractional Roche lobe is  $r_L \sim 0.5$ ,

$$\hat{\alpha}_{\text{ee}} \simeq 2(\alpha_{\text{ee}})_{\text{BB}} = 1, \quad (\text{C8})$$

with the Bethe & Brown (1998)  $\alpha_{\text{ee}} = 0.5$ .

Of course,  $\hat{\alpha}_{\text{ee}}$  should not vary with the Roche lobe, the Bethe & Brown (1998) usage of  $\alpha_{\text{ee}}$  being in error.

For  $\hat{\alpha}_{\text{ee}} = 1$ , the envelope would be removed from the giant but would end up with zero kinetic energy, which is unreasonable. Thus, without additional energy sources, as discussed earlier in this paper, one would expect  $\hat{\alpha}_{\text{ee}} \sim 0.5$ , in which case the kinetic energy of the envelope would remain unchanged in its expulsion. The Bethe & Brown (1998) results were insensitive to changes in  $\hat{\alpha}_{\text{ee}}$ , which changed the location but not the magnitude of the favored logarithmic intervals, as noted by those authors.

In the present case van den Heuvel's  $\hat{\alpha}_{\text{ee}} = 1.2$  definitely indicated the presence of energy sources additional to the drop in orbital energy, although they are not as large as he indicated. We have checked that with  $\hat{\alpha}_{\text{ee}} = 1.2$  and  $c_s \gg 1$  in the Bethe & Brown (1998) formation we obtain the numerical results of Table 1 of van den Heuvel (1994).

## REFERENCES

- Applegate, J. H. 1997, Columbia Univ. preprint
- Aukierheide, M. B., Brown, G. E., Kuo, T. T. S., Stout, D. B., & Vogel, P. 1990, *ApJ*, 362, 241
- Bailes, M., et al. 1994, *ApJ*, 425, L41
- Bethe, H. A., & Brown, G. E. 1995, *ApJ*, 445, L129
- , 1998, *ApJ*, 506, 780
- Blandford, R. D., & Begelman, M. C. 1999, *MNRAS*, 303, L1
- Braun, H., & Langer, N. 1995, *A&A*, 297, 483
- Brown, G. E. 1995, *ApJ*, 440, 270
- , 1997, *Phys. Blätter*, 53, 671
- Brown, G. E., & Bethe, H. A. 1994, *ApJ*, 423, 659
- Camilo, F. 1995, Ph.D. thesis, Princeton Univ.
- Camillo, F., Nice, D. J., Shrauner, J. A., & Taylor, J. H. 1996, *ApJ*, 469, 819
- Chevalier, R. A. 1993, *ApJ*, 411, L33
- Fryer, C., & Kalogera, V. 1997, *ApJ*, 489, 244
- Habets, G. M. B. H. 1986a, *A&A*, 165, 95
- , 1986b, *A&A*, 167, 61
- Holmer, J. 1998, Stony Brook Senior thesis
- Iben, I., & Tutukov, A. V. 1993, *ApJ*, 418, 343
- Kaspi, V. M., et al. 2000, *ApJ*, 543, 321
- King, A. R., & Begelman, M. C. 1999, *ApJ*, 519, L169
- King, A. R., & Ritter, H. 1999, *MNRAS*, 309, 253
- Lorimer, D. R., Lyne, A. G., Bailes, M., Manchester, R. N., D'Amico, N., Stappers, B. W., Johnston, S., & Camilo, F. 1996, *MNRAS*, 283, 1383
- Narayan, R., Piran, T., & Shemi, A. 1991, *ApJ*, 379, L17
- Phinney, E. S. 1991, *ApJ*, 380, L17
- Phinney, E. S., & Kulkarni, S. R. 1994, *ARA&A*, 32, 591
- Phinney, E. S., & Sigurdsson, S. 1991, *Nature*, 349, 220
- Phinney, E. S., & Verbunt, F. 1991, *MNRAS*, 248, 21P
- Portegies Zwart, S. F., & Yungelson, L. R. 1998, *A&A*, 332, 173
- Portegies Zwart, S. F., & Yungelson, L. R. 1999, *MNRAS*, 309, 26
- Ramachandran, R., & Portegies Zwart, S. F. 1998, Abstrs. 19th Texas Symp. on Relativistic Astrophysics and Cosmology, ed. J. Paul, T. Montmerle, & E. Aubourg (Amsterdam: North-Holland)
- Savonije, G. J. 1983, *Nature*, 304, 422
- Shapiro, S. L., & Teukolsky, S. 1983, *Black Holes, White Dwarfs, and Neutron Stars: The Physics of Compact Objects* (New York: Wiley).
- Shima, E., Matsuda, T., Takeda, H., & Sawada, K. 1985, *MNRAS*, 217, 367
- Taam, R. E., & van den Heuvel, E. P. J. 1986, *ApJ*, 305, 235
- Tauris, T. M., & Sennels, T. 2000, *A&A*, 355, 236
- Tauris, T. M., van den Heuvel, E. P. J., & Savonije, G. J. 2000, *ApJ*, 530, L93
- Terman, J. L., Taam, R. E., & Hernquist, L. 1995, *ApJ*, 445, 367
- Thorsett, S. E., & Chakrabarty, D. 1999, *ApJ*, 512, 288
- van den Heuvel, E. P. J. 1994, *A&A*, 291, L39
- , 1995, *J. Astrophys. Astron.*, 16, 255
- van den Heuvel, E. P. J., & Lorimer, D. R. 1996, *MNRAS*, 283, L37
- van den Heuvel, E. P. J., & Taam, R. E. 1984, *Nature*, 309, 235
- van Kerkwijk, M. H. 2000, in Proc. ESO Workshop on Black Holes in Binaries and Galactic Nuclei, ed. I. Kaper, E. P. J. van den Heuvel, & P. A. Woudt (Berlin: Springer), in press (astro-ph/0001077)
- van Kerkwijk, M. H., & Kulkarni, S. R. 1995, *ApJ*, 454, L141
- , 1999, *ApJ*, 516, L25
- van Paradijs, J., Takens, R., & Zuiderwijk, E. 1977a, *A&A*, 57, 221
- van Paradijs, J., Zuiderwijk, E., Takens, R. J., & Hammerschlag-Hensberge, G. 1977b, *A&AS*, 30, 195
- Webbink, R. F. 1984, *ApJ*, 277, 355
- Weinstein, S., & Langer, N. 1999, *A&A*, 350, 148
- Wettig, T., & Brown, G. E. 1996, *NewA*, 1, 17
- Wijers, R. A. M. J. 1997, *MNRAS*, 287, 607
- Woosley, S. E., Langer, N., & Weaver, T. A. 1995, *ApJ*, 448, 315
- Zuiderwijk, E., Hammerschlag-Hensberge, G., van Paradijs, J., Sterken, C., & Hensberge, H. 1977, *A&A*, 54, 167

---

## Chapter 20

# Formation and Evolution of Black Hole X-Ray Transient Systems

G.E. Brown, C.-H. Lee and T. Tauris

Reprinted with permission from *New Astronomy*, 6 (2001) 331–338

Copyright (2001) by Elsevier Science B.V.

### **Commentary**

Imagine our surprise when Gijs Nelemans from the Astronomical Institute Anton Pannekoek, Amsterdam, visited us and told us that our formation of high-mass black holes in low-mass X-ray binaries, Paper 14, mostly didn't work. He had read the paper of Portegies Zwart, Verbunt & Ergma (1997) more carefully than we.

The point is that massive stars lose a lot of mass during He core burning. A binary composed of a massive star and low-mass companion, either main sequence or evolved, will move apart during this mass loss, as

$$\frac{\dot{a}}{a} = \frac{M}{\dot{M}}.$$

Equivalently,

$$\frac{a_f}{a_i} = \frac{M_i}{M_f},$$

where *i* and *f* denote *initial* and *final*. This means that early on in the binary evolution, the initial separation  $a_0$  must be quite small, if it is to widen to the separation where Roche Lobe overflow from the massive star takes place only following He core burning. In Fig. 1 of the paper we see that this is possible in the case of the ZAMS  $20M_\odot$  star with  $1M_\odot$  main sequence companion. However, in the case of the ZAMS  $25M_\odot$  star the massive star will undergo Roche Lobe overflow towards the end of H shell burning. Then it will burn He as a naked star and evolve into either a low-mass black hole or neutron star.

There is, given the winds used in present stellar evolution, only a narrow interval of ZAMS masses around  $20M_\odot$  in which high-mass black holes can be made. Paper 21, following this paper, will show that again high-mass black holes can be made in binaries above ZAMS masses  $\gtrsim 60M_\odot$ . This latter number is corrected downwards from the  $\sim 80M_\odot$

arrived at in Paper 12, due to lower mass loss rates from the He wind during the Wolf-Rayet stage (see Paper 21). Thus, we have a bimodal distribution for evolving high-mass black holes in binaries: a small interval around ZAMS  $20M_{\odot}$  and, then again, in WNL's for ZAMS  $\gtrsim 60M_{\odot}$ .

We find that high-mass black hole evolution around ZAMS  $20M_{\odot}$  depends sensitively on the  $^{12}\text{C}(\alpha, \gamma)^{16}\text{O}$  rate and that the Woosley & Weaver rate of 170 keV barns at  $E = 300$  keV is exactly right for our scenario.

Although our results are at variance with all of the high-mass black hole evolution to date, they are tightly structured and hang together.

Although the window for forming high-mass black holes in binaries is very narrow, it is sufficient to produce the observed number of black hole X-ray transient systems because of their long lifetimes. The narrowness of this window allows us to set the limits of  $0.2 < \lambda\alpha_{ce} < 0.5$  on the common envelope efficiency.

After this manuscript was completed, we ran into difficulty in evolving the spectacular new transient source V4641 Sgr discussed at the end of the Commentary to Paper 17. The black hole mass of  $9.61_{-0.88}^{+2.08} M_{\odot}$  requires that we begin from an  $\sim 25M_{\odot}$  progenitor. Given mass loss in the explosion, the progenitor would have to be even more massive. We are able to extend the narrow mass limit of the present paper from  $22M_{\odot}$  up to  $\sim 30$  if we assume the Schaller *et al.* mass loss rates to be a factor of  $2\frac{1}{2}$ –3 too high, as we found for the Wolf-Rayet star mass loss rates used until recently (see Paper 21). Determinations of winds from resonance lines of metals, etc. are fraught with uncertainties.

## Reference

- (1) Portegies Zwart, S.F., Verbunt, F., and Ergma, E. (1997), *Astronomy & Astrophysics* **321**, 207.



## Formation and evolution of black hole X-ray transient systems

G.E. Brown<sup>a</sup>, C.-H. Lee<sup>a,b,\*</sup>, T.M. Tauris<sup>c</sup>

<sup>a</sup>Department of Physics & Astronomy, State University of New York, Stony Brook, New York, NY 11794, USA

<sup>b</sup>School of Physics, Korea Institute for Advanced Study, Seoul 130-012, South Korea

<sup>c</sup>Nordic Institute for Theoretical Physics (NORDITA), Blegdamsvej 17, DK-2100 Copenhagen Ø, Denmark

Received 1 February 2001; received in revised form 5 May 2001; accepted 5 May 2001

Communicated by S.E. Woosley

### Abstract

We study the formation of low-mass black hole X-ray binaries with main sequence companions that have formed through case C mass transfer (mass transfer following the helium core burning phase of the black hole progenitor). We identify these objects with the observed soft X-ray transients. Although this scenario requires a set of fine tuned conditions, we are able to produce a current Galactic population of ~2000 objects, in agreement with estimates based on observations. The narrow interval in initial separations leading to case C mass transfer, combined with the allowed narrow range of separations after the common envelope evolution, constrains the common envelope efficiency in this scenario:  $\lambda\alpha_{ce} \approx 0.2 - 0.5$ . © 2001 Elsevier Science B.V. All rights reserved.

PACS: 97.80.Jp; 97.60.Lf; 97.80.Fk

Keywords: Black hole physics; Binaries: close; Accretion, accretion disks

### 1. Introduction

In the literature there have been severe difficulties in evolving a sufficient number of black hole soft X-ray transient sources (SXTs) (Ergma and van den Heuvel, 1998; Portegies Zwart et al., 1997). For example, Portegies Zwart et al. obtain a birth rate of  $9.6 \times 10^{-9} \text{ yr}^{-1}$  for the SXTs, to be compared with one of  $2.2 \times 10^{-6} \text{ yr}^{-1}$  for the binaries with a neutron star, whereas they remark that observations indicate equal formation rates for these two types of

binaries. Their large discrepancy arises from their limit ZAMS mass  $40 M_\odot$  for black hole formation.

In fact, however, for the evolution of the SXTs, a ZAMS mass more like  $\sim 20 M_\odot$  should be chosen as limit for high-mass black hole formation. This is the limit for single stars (Brown et al., 2001). If the massive black hole progenitor in the SXTs can complete He core burning before it is removed in common envelope evolution (Case C mass transfer) then it will evolve like a single star since the remaining  $\sim 10^4$  yrs of its lifetime is too short for He wind losses to effect its evolution. Indeed, Ergma and van den Heuvel (1998) realized that “ $M_{BH}$  should not be larger than 20 to  $25 M_\odot$ ” in order to have sufficient SXTs, but they did not realize that the more massive stars in a binary evolve in a different way from a single massive star.

\*Corresponding author.

E-mail addresses: [popenoe@nuclear.physics.sunysb.edu](mailto:popenoe@nuclear.physics.sunysb.edu) (G.E. Brown), [chlee@kias.re.kr](mailto:chlee@kias.re.kr) (C.-H. Lee), [tauris@nordita.dk](mailto:tauris@nordita.dk) (T.M. Tauris).

If the H envelope of the more massive star in a binary is removed by early Roche Lobe overflow, Case A or Case B, before He core burning is completed, the resulting “naked” He core will blow away, leaving an Fe core too low in mass to evolve into a high-mass black hole (Brown et al., 2001).

The common envelope evolution of Portegies Zwart et al. (1997) for a  $20 M_{\odot}$  turns out to be useful for us, because this will be the same, regardless of whether the resulting He star goes into a neutron star as they believed, or into a high-mass black hole. It turns out that this evolution can only be carried out in Case C, so the He star is clothed during He core burning, appropriate for our high mass black hole evolution.

Given the Schaller et al. (1992) evolution, we find that Case C evolution does not work for stars more massive than  $\sim 22 M_{\odot}$ , in agreement with Portegies Zwart et al. (1997). However, we need the more massive black hole progenitors to evolve the transient sources with subgiant companions, as we shall discuss.

In this note we consider the evolution of those low-mass black hole X-ray binaries with main sequence companions which are formed through case C mass transfer followed by common envelope and spiral-in evolution. We estimate their expected number currently present in our Galaxy. We therefore discuss the following.

1. The initial separation interval of binaries that go through case C mass transfer and the ZAMS mass interval of stars being able to produce high-mass black holes (Section 2).
2. The initial masses of the black hole companions (the donors in the present-day X-ray binaries) and the allowed orbital periods after the common envelope required for the systems to become observable X-ray binaries within the age of our Galaxy before the donors leave the main-sequence (Section 3).
3. The resulting common envelope efficiency that can be estimated combining points 1 and 2 (Section 4).
4. The evolution of the mass-transfer rate in the X-ray transient phase and the lifetimes of the systems (Section 5).

5. The total number of systems expected in our Galaxy at present (Section 6).

Following the discussion of the above five points, we briefly discuss the black hole systems with (sub)giant companions Nova Scorpii (GRO J1655 – 40) with a  $2.4 M_{\odot}$  F-star companion and a  $6.3 M_{\odot}$  black hole and V4641 Sgr (XTE J1819 – 254) with a  $6.5 M_{\odot}$  B-star companion and a  $9.6 M_{\odot}$  black hole. At least in the former case there has been substantial mass loss in the black hole formation,  $\geq 5 M_{\odot}$ . Thus, a He core of at least  $11 M_{\odot}$  is needed for the black hole progenitor, or a ZAMS  $35 M_{\odot}$  star. From this we conclude that the Schaller et al. (1992) wind losses are too large and that they must be substantially decreased in order to evolve these two binaries.

## 2. Case C mass transfer: limits on initial separations and black hole progenitor masses

In order to have case C mass transfer the radius of the star has to expand after core helium burning has stopped. In the calculation of Schaller et al. (1992), stars with ZAMS masses  $20 M_{\odot}$  and  $25 M_{\odot}$  satisfy the requirement. However, during the core helium burning stage, the orbit expands due to the mass loss:

$$\frac{a'_{\Delta M}}{a} = \frac{M + M_d}{M - \Delta M + M_d} \quad (1)$$

where  $a$  is the orbital separations,  $M$  is the mass of the black hole progenitor,  $\Delta M$  is the mass lost from the black hole progenitor during the core helium burning, and  $M_d$  is the donor star mass. In order to initiate mass transfer after core helium burning the star has to expand sufficiently that this widening of the orbit is compensated for.

In the upper curve of Fig. 1 we draw the orbital separation of the binary required to initiate Roche lobe overflow calculated at each stage from the radii given by Schaller et al. (1992). For the masses considered the radius of giant star is at  $\sim 2/3$  of the radial separation between the two stars. The solid curve gives the corresponding *initial* separation between the massive star and the low-mass main sequence companion after strong wind mass loss is

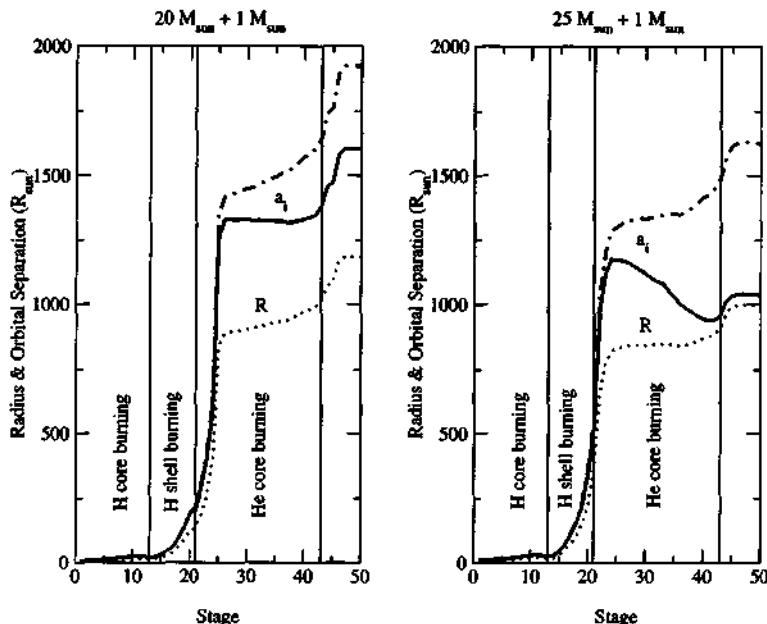


Fig. 1. Radius of black hole progenitors ( $R$ ) and the initial orbital separations ( $a_i$ ) of the progenitors of X-ray transient binaries with a  $1 M_{\odot}$  companion. Upper dot dashed curves correspond to the orbital separations required to initiate Roche lobe overflow at the radius taken for each stage from Schaller et al. (1992). The solid curves correspond to the required initial separations after corrections of the orbit widening due to the wind mass loss. The lower dotted curves correspond to the radius of the black hole progenitors taken from Schaller et al. (1992).

switched on. Here we assume that the wind mass loss does not affect the stellar radius.<sup>1</sup> We see that in the case of a  $20 M_{\odot}$  star the mass transfer can start following He core burning for binaries with an initial separation in the interval of  $1370 R_{\odot} < a_i < 1605 R_{\odot}$ . In the case of a  $25 M_{\odot}$  star, however, mass transfer is only possible up to just after the beginning of core helium burning (with initial separations around  $1170 R_{\odot}$ ). Binaries with larger initial separations will become too wide during core helium burning so even though the star expands slightly after core helium burning, it will not be able to fill its Roche lobe (see also Fig. 2 of Portegies Zwart et al., 1997).

Brown et al. (1996) found that stars with ZAMS masses above  $19 M_{\odot}$ , evolved by Woosley and

Weaver (1995), evolve into high-mass ( $\sim 7 M_{\odot}$ ) black holes. As outlined in Brown et al. (2001), this requires the special rate of 170 keV barns at energy  $E = 300$  keV for the  $^{12}\text{C}(\alpha, \gamma)^{16}\text{O}$  reaction used by Woosley and Weaver. Recent experiments including both E1 and E2 components, obtain  $S_{\text{tot}}^{300} = (165 \pm 50)$  keV barns (Kunz et al., 2001). Brown et al. (2001) discuss quantitatively how changes in this rate change the ZAMS mass at which the calculated Fe core mass increases rapidly with mass. The fact that the progenitor mass of SN1987A lay in the narrow interval of  $18 - 20 M_{\odot}$  appropriate for forming black holes with the Woosley and Weaver 170 keV barns gives further support to this value. A ZAMS  $18 M_{\odot}$  star, which is usually taken as progenitor for SN 1987A, will form a low-mass ( $\sim 1.5 - 1.8 M_{\odot}$ ) black hole or a neutron star.

Hence, the mass interval of stars that can go through case C mass transfer and collapse into a high-mass black hole is limited between 19 and

<sup>1</sup>One of us (T.M.T.) has carried out stellar evolution calculations, using Eggleton's code, with/without the wind mass losses, and find that the radius behaves almost the same. The minor difference will give rise to adjusting our numbers slightly, but the outline will remain the same.

$20\text{--}25 M_{\odot}$ . We shall (arbitrarily) take  $22 M_{\odot}$  as an upper limit. Stellar evolution calculations with lower wind mass losses might give a higher mass limit.

Using the above derived limits on the mass of the black hole progenitors and the initial separation, the rate of formation of black hole transient sources with main sequence companions we can evolve (see Section 6) is nearly an order of magnitude less than in Brown et al. (1999) where wind mass losses were ignored. We realize that our evolution requires a fine-tuned parameter space, but we shall argue that it all hangs together. The exact numbers may alter with changes in stellar evolution models.

### 3. Limits on donor masses and orbital periods

During the common envelope evolution the low-mass main sequence star of mass  $\sim 1 M_{\odot}$  will spiral into the envelope of the massive giant. The orbit will shrink dramatically and the outcome is a close binary consisting of the core of the massive giant and the low-mass star (Paczynski, 1976). Shortly after the spiral-in and ejection of the envelope of the giant, the remaining core will collapse into a black hole. Whether or not this collapse is associated with mass loss is still an open question. The orbit will then shrink further due to the loss of orbital angular momentum via magnetic braking and gravitational wave radiation (e.g. Verbunt, 1990) until the low-mass (donor) star begins to transfer mass to the black hole — forming an X-ray binary. Since the wind mass-loss rate of the low-mass donor star is very low, and its accretion during the short common envelope phase is negligible, it will have a mass at the onset of the X-ray phase which is about equal to its initial mass.

The time evolution of the orbital separation is described as

$$\frac{\dot{a}}{a} = \frac{2j_{\text{gw}}}{J_{\text{orb}}} + \frac{2j_{\text{mb}}}{J_{\text{orb}}} - 2\dot{M}_d \left( \frac{M_{\text{BH}} - M_d}{M_{\text{BH}} M_d} \right). \quad (2)$$

Here we assume that the mass lost from the companion star is all accreted onto the black hole. The orbital angular momentum loss by gravitational wave radiation (gw) and magnetic braking (mb) is given by:

$$\begin{aligned} \frac{j_{\text{gw}}}{J_{\text{orb}}} &= -\frac{32G^3}{5c^5} \frac{M_{\text{BH}} M_d (M_{\text{BH}} + M_d)}{a^4} s^{-1} \\ \frac{j_{\text{mb}}}{J_{\text{orb}}} &\approx -0.5 \times 10^{-28} b_{\text{mb}} \frac{IR_d^2}{a^5} \frac{G(M_{\text{BH}} + M_d)^2}{M_{\text{BH}} M_d} s^{-1} \end{aligned} \quad (3)$$

where  $R_d$  is the radius of the donor star,  $b_{\text{mb}}$  is an efficiency parameter given below, and  $I \approx 0.1 M_d R_d^2$  is the moment of inertia of the donor star. Here we have assumed a magnetic braking law based on observations of slowly rotating single stars by Skumanich (1972). (However, this law may be too strong, i.e. overestimating the dependence on the angular velocity, see e.g. Stepień, 1995.) We further assumed the the magnetic braking efficiency suggested by Kalogera and Webbink (1998),

$$b_{\text{mb}}(M_d) = \begin{cases} 0, & M_d \leq 0.37 M_{\odot} \\ 1, & 0.37 M_{\odot} < M_d < 1.03 M_{\odot} \\ \exp(-4.15(M_d - 1.03)), & M_d > 1.03 M_{\odot} \end{cases} \quad (4)$$

Here the effect of magnetic braking is assumed to be strongest for stars with a mass  $\approx 1.0 M_{\odot}$ . Donor stars with mass  $> 1.5 M_{\odot}$  do not have a convective envelope on the main sequence and the magnetic braking effects are therefore negligible, so the binary evolution is dominated by gravitational radiation until the mass transfer starts.

In Fig. 2, we show the possible ranges of orbital periods and donor masses immediately after the formation of the black hole, such that an X-ray binary with a main-sequence donor is formed within the Hubble time (Kalogera, 1999). If no mass is lost in the formation of the black hole, these periods are equal to the periods immediately after the common envelope as a result of the short time interval before the collapse of the core.

The upper limit is set by two conditions: (i) the binary should start mass transfer within the Hubble time (line denoted by A), (ii) the mass transfer should start before the donor evolves off the main-sequence (line denoted by B) which also determines the lower limit on  $\lambda\alpha_{\text{ce}} = 0.2$  (see Section 4). The kink in line B around  $\sim 1 M_{\odot}$  comes from the fact that the effect of magnetic braking, Eq. (4), is

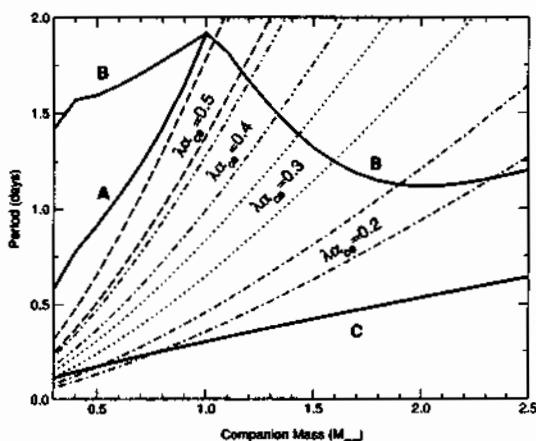


Fig. 2. Limits on the orbital period after the formation of the black hole as function of the mass of the low-mass companion. The upper and lower boundaries (thick solid lines) of mass and period are set by the conditions (see Section 3). (A) Mass transfer begins within the Hubble time, (B) mass transfer begins before the companion evolves off the main sequence, (C) mass transfer from the companion begins right after finishing common envelope evolution. The expected ranges of periods and companion masses, after the common envelope phase, are plotted for various given efficiencies (see Section 4). The width of each band for an assumed  $\lambda\alpha_{ce}$  is determined by the initial band of possible Roche Lobe overflows with Case C mass transfer. We took the ZAMS mass of the black hole progenitor to be  $20 M_\odot$ , which corresponds to  $M_p \sim 16.8 M_\odot$  in the beginning of case C mass transfer, and  $M_{He} = M_{BH} = 7 M_\odot$ .

strongest for stars with a mass  $\sim 1 M_\odot$ . For this plot, we took the simple interpolation of Schaller's results for the evolution of a  $1 M_\odot$  star with two extrapolations for the radius  $R$  and the lifetimes of the main-sequence star  $t_{end\ ms}$ :

$$R(M_d, t) = R_{M_\odot}(t) \times M_d^{0.88} \\ t_{end\ ms}(M_d) = \begin{cases} t_{M_\odot}/M_d^2 & \text{for } M_d < 1 M_\odot \\ t_{M_\odot}/M_d^{2.5} & \text{for } M_d > 1 M_\odot \end{cases} \quad (5)$$

where  $t_{M_\odot}$  is the time at the stage 13 of the  $1 M_\odot$  model of Schaller et al. (1992) (see Fig. 1). The applied formulae may not be very accurate, but the qualitative behaviour after more realistic numerical calculations will remain the same. The lower boundary of the period immediately after the common envelope (line denoted by C) is set by the condition that the main sequence star should not overflow its Roche-lobe at the end of the common envelope.

#### 4. Common envelope evolution: Binding energy and efficiency parameter

Since our initial separation  $a_i$  is sharply defined by the condition of case C mass transfer, and our final separations are constrained by the distance that can be traversed during magnetic braking and gravitational wave radiation, we can determine the range of allowed common envelope efficiencies in order to form low-mass black hole X-ray transients.

During the common envelope phase the energy needed to expel the hydrogen envelope of the black hole progenitor is tapped from the drop in binary orbital potential energy with efficiency  $\alpha_{ce}$  (also denoted  $\eta$  in the literature):

$$\frac{GM_p M_e}{\lambda R} = \frac{GM_p M_e}{\lambda r_L a_i} \\ = \alpha_{ce} \left( \frac{GM_{He} M_d}{2a_i} - \frac{GM_p M_d}{2a_f} \right) \quad (6)$$

where  $M_p$  is the total mass of the BH progenitor star just before the common envelope forms,  $M_e$  is the mass of its hydrogen envelope,  $M_{He}$  is the mass of its core,  $a_i$  and  $a_f$  is the initial and final separation, before and after the common envelope, respectively.  $r_L$  is the dimensionless Roche-lobe radius. Given the parameters of the system at the start of the common envelope, the final separation is determined by  $\lambda$ , describing the structure of the giant and  $\alpha_{ce}$ , the efficiency of the energy conversion. In our case the final separation is limited as shown in Fig. 2 so the product  $\lambda\alpha_{ce}$  can be constrained. In the literature  $\lambda = 0.5$  had often been used, and a high efficiency  $\alpha_{ce} > 1$  was often required in order to explain the observations. However, recent detailed stellar evolution calculations by Dewi and Tauris (2000) show that  $\lambda$  can be substantially larger. These high values of  $\lambda$  solve the problem of unrealistically high efficiencies  $\alpha_{ce}$ . For a  $20 M_\odot$  star, corresponding to  $16.2 M_\odot$  at the tip of the AGB, Tauris and Dewi (2001) find  $\lambda$ -values in a large interval:  $0.1 < \lambda < 3$  depending on the exact location of the core mass boundary and amount of internal thermodynamic energy included.

In Fig. 2 we plotted the possible ranges of periods after the common envelope. We took the ZAMS mass of the black hole progenitor to be  $20 M_\odot$ ,

which corresponds to  $M_p \sim 16.8 M_\odot$  in the beginning of case C mass transfer, and assumed  $M_{He} = M_{BH} = 7 M_\odot$ . For each value of  $\lambda\alpha_{ce}$  the two lines are for the limiting initial separations  $a_i$  (see Section 2). Donor stars with masses above  $1.5 M_\odot$  can only be formed if the common envelope efficiency  $\lambda\alpha_{ce}$  is around 0.2.

## 5. Life time of X-ray transient sources

Brown et al. (1999), assuming a mass-loss rate of  $10^{-9} M_\odot \text{ yr}^{-1}$ , obtained a lifetime of the X-ray transients of  $10^9$  yrs. However, as Ergma and Fedorova (1998) discussed, the mass-loss rate changes as a function of time.

Once the Roche lobe overflow starts, at any given time, the orbital separation is self determined in order for the companion star to fill its Roche lobe,

$$a = R_d / r_L(M_d, M_{BH}) \quad (7)$$

Hence, the mass loss rate is automatically determined by the feedback effects between the orbital widening due to the mass loss and the orbital

contraction due to the gravitational wave radiation and the magnetic braking, Eq. (2). In the numerical simulation of mass transfer due to Roche lobe overflow, for a given  $\delta M_d$ , we can get the  $\delta t$ , equivalently  $\dot{M}_d$ , by requiring the donor star fill its Roche lobe.

In Fig. 3 we plot the evolution of the donor star mass and the mass loss rate as a function of time for  $M_{d,i} = 1.25 M_\odot$  with a  $7 M_\odot$  black hole. There are two tracks after Roche lobe overflow.

1. Case TO: we assumed that the stripped star after Roche lobe overflow continues to follow exactly the same evolution time scale as the original star.
2. Case TS: we assumed that the stripped donor star loses all its initial information. When the star become smaller in mass, it follows the time scale of the main sequence star of that reduced mass.

We believe the realistic situation to lie between case TO and TS. The radius of the donor star, which essentially determines the orbital separation, is smaller in case TS because the donor star is less evolved than that in TO case. Even though, for any given donor star mass  $M_d$ , the lifetime of the donor

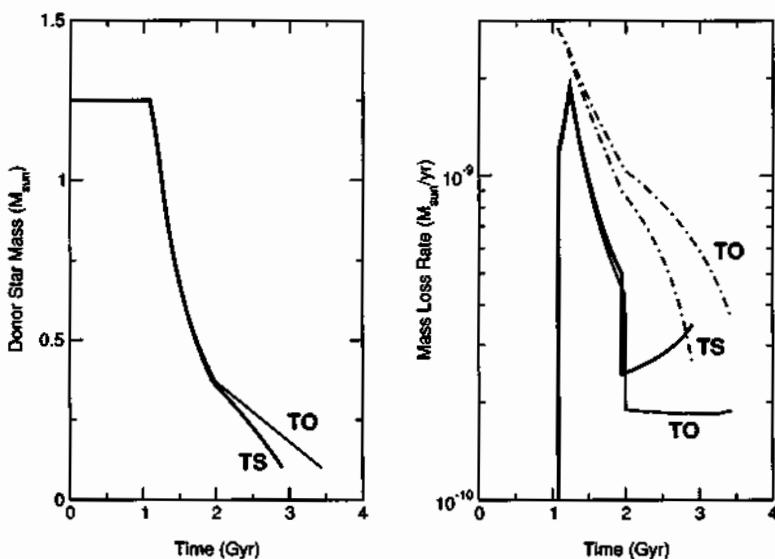


Fig. 3. Time evolution of the donor star mass and mass loss rate for  $M_{d,i} = 1.25 M_\odot$  with a  $7 M_\odot$  black hole and an initial period of 20 h. The dot-dashed lines on the right panel correspond to the critical mass loss rate for the steady X-ray sources, Eq. (8). The two tracks correspond to the different time scales of stripped donor stars (see the discussion of Section 5).

star itself is longer in TS case, the smaller radius of the donor star gives the shorter orbital separation with larger mass loss rates, finally decreasing the lifetime of X-ray transients in case TS.

In Fig. 3, the time scales of X-ray transients are  $\sim 2 \times 10^9$  yrs for a  $1.25 M_{\odot}$  donor star. The mass of the donor star (initially a  $1.25 M_{\odot}$  star) drops quickly in the beginning with a high mass loss rate  $> 10^{-9} M_{\odot} \text{ yr}^{-1}$ , which slows down afterwards. Any donor stars  $> 0.5 M_{\odot}$  show the similar behaviour. Hence, independently of initial donor masses, most of the time X-ray transients will appear as low-mass X-ray transients (i.e. with a donor star mass  $< 0.5 M_{\odot}$ ) having a mass-loss rate of  $\sim 10^{-10} M_{\odot} \text{ yr}^{-1}$ , which is consistent with the empirical rate of  $1.3 \times 10^{-10} M_{\odot} \text{ yr}^{-1}$  found by Van Paradijs (1996). The dot-dashed lines on the right panel of Fig. 3 correspond to the critical mass loss rate for the steady X-ray sources (King et al., 1997)

$$\dot{M}_{\text{crit}} = 2.86 \times 10^{-11} M_{\text{BH}}^{5/6} M_d^{-1/6} P_h^{4/3} M_{\odot} \text{ yr}^{-1}. \quad (8)$$

We interpret the different masses of the main sequence companion as resulting firstly from varying initial masses, and secondly from the star being observed at different times in its evolution.

## 6. Population synthesis

We calculated the total number of expected X-ray transients evolving through case C mass transfer in the same way as Brown et al. (1999). From the discussions in Section 2, for a ZAMS star of mass  $20 M_{\odot}$ , we see that only the interval of  $235 R_{\odot}$  between  $1370 R_{\odot}$  and  $1605 R_{\odot}$  of  $a(t=0)$  is available for case C mass transfer. It is this small fractional interval that allows us to obtain a narrow interval of values for  $\lambda\alpha_{\text{ee}}$ . This is a logarithmic interval of only  $\ln(1605/1370) = 0.16$  compared with our total logarithmic interval of  $\ln(4 \times 10^9 \text{ km}/4 \times 10^6 \text{ km}) = 7$  (Brown et al., 1999) so that we have a fraction of only 0.023 (as compared with 0.11 in Brown et al., 1999).

Since the possible ZAMS range for case C mass transfer is very narrow near  $20 M_{\odot}$ , we somewhat arbitrarily choose a ZAMS mass interval of  $19 M_{\odot} < M < 22 M_{\odot}$ . This means (Bethe and Brown, 1998)

that the fraction of binaries with primaries in this range (assuming an IMF,  $P(m) \propto M^{-2.5}$ ) is:  $(1.9)^{-3/2} - (2.2)^{-3/2} = 0.08$  which is smaller by a factor of 2.5 from that for the interval of  $20 - 35 M_{\odot}$  chosen by Brown et al. (1999).

From the discussion in Sections 3 and 4, the upper limit of the mass of the donor star is sensitive to the efficiency parameter  $\lambda\alpha_{\text{ee}}$ . For  $\lambda\alpha_{\text{ee}} = 0.4$  this upper limit is  $1.6 M_{\odot}$ . Assuming a flat  $q$  distribution, we get  $\Delta q \approx 1/20$ . From Section 5, we take the average lifetime of X-ray transients as  $2 \times 10^9$  yrs, a factor two higher than the assumed lifetime in Brown et al. (1999).

By taking the supernova rate as  $2 \times 10^{-2} \text{ yr}^{-1}$  per Galaxy, we have the number of X-ray transients in our Galaxy as

$$2 \times 10^{-2} \text{ yr}^{-1} \times \frac{1}{20} \times \frac{1}{2} \times 0.08 \times 0.023 \times 2 \times 10^9 \text{ yr} \\ \approx 1840, \quad (9)$$

where a binarity of a  $1/2$  is considered as in Brown et al. (1999). The total numbers in the Galaxy of such systems is estimated to be between a few hundred and a few thousand (Ergma and Fedorova, 1998).<sup>2</sup> Our estimate is thus consistent with theirs.

Our birth rate for the black hole binaries obtained from Eq. (9) is  $10^{-6} \text{ yr}^{-1}$ , roughly half the Portegies Zwart et al. rate for binaries with a neutron star.

## 7. Black hole X-ray binaries with (sub)giant donors

Transient black hole binaries with (sub)giant donors might be expected to have followed the same scenario, but with initially higher companion masses and larger separations after the common envelope, so that they start mass transfer only when the donor has evolved off the main sequence. However, the large space velocity of Nova Scorpii 1994, best explained with a large amount of mass loss ( $5-10 M_{\odot}$ ) during the explosion in which the black hole was formed (Nelemans et al., 1999), and the high mass of the black hole in V404 Cyg ( $\sim 10 M_{\odot}$ ) suggest that the

<sup>2</sup>Of the Wijers (1996) lower limit of 3000 transient black hole sources, 6 out of 9 sources had a main sequence companion with a short period, so he would have had  $\sim 2000$  of the latter.

helium cores of stars around  $20 M_{\odot}$  may not be massive enough to explain these systems. Hence, the progenitor masses of the black hole may have been larger.

Brown and Lee (2001) require a He core mass of  $\sim 11 M_{\odot}$  corresponding to ZAMS  $\sim 35 M_{\odot}$  for the black hole progenitor of Nova Scorpii and a somewhat higher He core mass for the black hole progenitor of V4641 Sgr. The wind losses employed by Schaller et al. (1992) must be substantially reduced if these SXTs are to be evolved in Case C mass transfer. This would take us back to the interval of ZAMS masses  $20\text{--}35 M_{\odot}$  suggested by Brown et al. (1999), possibly even up to  $\sim 40 M_{\odot}$ .

## 8. Conclusion

We evolve the low-mass black-hole X-ray binaries, which are identified as the observed soft X-ray transients, and show that these systems are formed via case C mass transfer following helium core burning phase of the black hole progenitor. Although this scenario requires a set of fine tuned conditions, we are able to produce a current Galactic population of  $\sim 2000$  objects, in agreement with estimates based on observations. Combining the narrow interval in initial separations leading to case C mass transfer with the allowed narrow range of separations after the common envelope evolution, we put constraints on the common envelope efficiency as  $\lambda \alpha_{ce} \approx 0.2\text{--}0.5$ .

Since our analysis requires fine tuned parameter space, more detailed calculations of the stellar evolution of stars with ZAMS masses around  $20 M_{\odot}$  are required. More uncertain is the common envelope efficiency, which is essential for the formation of final short orbital period of the binaries.

## Acknowledgements

This work was initiated by Gijs Nelemans' visit to Stony Brook. We acknowledge his helpful comments and discussions. GEB and CHL were supported by the U.S. Department of Energy under grant DE-FG02-88ER40388.

## References

- Brown, G.E., Lee, C.-H., 2001. In preparation.
- Brown, G.E., Lee, C.-H., Bethe, H.A., 1999. *NewA* 4, 313.
- Brown, G.E., Heger, A., Langer, N., Lee, C.-H., Wellstein, S., Bethe, H.A., 2001, *astro-ph/0102379*.
- Brown, G.E., Weingartner, J.C., Wijers, R.A.M.J., 1996. *ApJ* 463, 297.
- Dewi, J.D.M., Tauris, T.M., 2000. *A&A* 360, 1043.
- Ergma, E., Fedorova, A., 1998. *A&A* 338, 69.
- Ergma, E., van den Heuvel, E.P.J., 1998. 331, L29.
- Kalogera, V., 1999. *ApJ* 521, 723.
- Kalogera, V., Webbink, R.F., 1998. *ApJ* 493, 351.
- King, A.R., Kolb, U., Szuszkiewicz, E., 1997. *ApJ* 488, 89.
- Kunz, R., Jaeger, M., Mayer, A., Hammer, J.W., Staudt, G., Harissopoulos, S., Paradellis, T., 2001. *PhRvL* 86, 3244.
- Nelemans, G., Tauris, T.M., van den Heuvel, E.P.J., 1999. *A&A* 352, L87.
- Paczynski, B., 1976. In: Eggleton, P.P., Mitton, S., Whelan, J. (Eds.), *Structure and Evolution of Close Binary Systems*. Kluwer, Dordrecht, p. 75.
- Portegies Zwart, S., Verbunt, F., Ergma, E., 1997. *A&A* 321, 207.
- Schaller, G., Schaerer, D., Meynet, G., Maeder, A., 1992. *A&AS* 96, 269.
- Skumanich, A., 1972. *ApJ* 171, 565.
- Stepien, K., 1995. *MNRAS* 274, 1019.
- Tauris, T.M., Dewi, J.D.M., 2001. *A&A* 369, 170.
- Van Paradijs, J., 1996. *ApJ* 464, L139.
- Verbunt, F., 1990. In: Kundt, W. (Ed.), *Neutron Stars and Their Birth Events*. Kluwer, Dordrecht, p. 179.
- Wijers, R.A.M.J., 1996. In: Wijers, R.A.M.J. et al. (Ed.), *Evolutionary Processes in Binary Stars*, Vol. 327. Kluwer Acad. Publ, Dordrecht.
- Woosley, S.E., Weaver, T.A., 1995. *ApJS* 101, 181.

## Chapter 21

# Formation of High-Mass X-Ray Black Hole Binaries

G.E. Brown, A. Heger, N. Langer, C.-H. Lee, W. Wellstein and H.A. Bethe

Reprinted with permission from *New Astronomy*, 6 (2001) 457–470

Copyright (2001) by Elsevier Science B.V.

### **Commentary**

In Sec. 3 of this article we repeat Woosley's argument that at some ZAMS mass stars will skip convective carbon burning, with extremely important consequences that have been detailed in Papers 13, 14, 15, 16, 17, 20 and 21. Until Paper 21, some of our arguments were based on conjecture, especially in the case of "naked" helium stars, those that have lost their hydrogen envelopes, because evolutions of these in the previous literature had used 2–3 times too large He wind losses. The history of how Stephen Wellstein with Norbert Langer burned naked He stars down to CO cores, and then Alexander Heger evolved these, using the Woosley & Weaver KEPLER computer program, down to Fe cores is outlined in the present paper with appropriate references. These more or less definite results show the earlier conjectured ones to be correct. Appreciable corrections would only occur for wind losses lower by a factor of 4 or more from the ones previously used, and such a factor seems excessive.

It is essential for the evolution of a star whether the reaction C+C occurs or is "skipped". This reaction is referred to in the literature as "convective carbon burning". We don't think this is a good designation, because its important consequences do not depend on its convectivity, nor is it clear that the alternative destruction of carbon, by the reaction  $^{12}\text{C}(\alpha, \gamma)^{16}\text{O}$ , is radiative. The important difference is that the C +  $\alpha$  reaction takes place at  $T \simeq 20$  keV, while the C+C reaction requires a temperature of about 80 keV.

There is a striking difference between naked and clothed He cores. Clothed He cores have a protective layer of hydrogen around them. This diminishes the loss of mass by wind from the surface. Moreover, in the surface layer the H+H reaction proceeds, creating additional He. In naked He stars, the amount of He is limited and can only diminish by wind. Once it is used up, the remaining C can only be used up by the reaction  $^{12}\text{C} + ^{12}\text{C} \rightarrow ^{24}\text{Mg}$  (or  $^{20}\text{Ne} + \alpha$ ), at the much higher temperature.

By contrast, in clothed He cores there is essentially unlimited He available. Once the He concentration falls below a certain range, the chief reaction of He is with C and the  $3\alpha \rightarrow C$  reaction becomes relatively small. If enough He is available, all the C can be used up by the C+ $\alpha$  reaction. Convection of matter (rather than energy) is needed, namely He from the shell into the core which still has C. In this way all of the C may be used up at about 20 keV rather than 80 keV. Consequences of this are detailed below.

Heger used immensely improved physics in the present calculations. In particular, the beta decays, which Woosley & Weaver (1995) had not turned on, were included and the presently best weak interaction rates of Martinez-Pinedo and Langanke were inserted into the program. This work is summarized by Heger, Woosley, Martinez-Pinedo & Langanke (2001).

We first make crude estimates of how much entropy is carried away for stars in which C+C burning takes place, i.e., ZAMS  $\lesssim 19M_{\odot}$ . Because of the high temperature of the C+C reaction of 80 keV, neutrinos can be emitted which carry away a lot of energy and entropy. At the  $\sim 20$  keV at which the  $^{12}C(\alpha, \gamma)^{16}O$  reaction takes place, neutrino emission is negligible. Burning times given in the table are only approximate. Relativistically the neutrino pair emission goes as  $T^{11}$ , and nonrelativistically as  $T^9$ , so we use an average  $T^{10}$ .

Element	$T$ (keV)	Burning Time ( $\tau$ )	$T_9^{10}$	$T_9^9 \times \tau$ (yrs)
Carbon	80	1000 yrs	0.11	134
Oxygen	200	0.5 yr	1024	256
Silicon	350	1 day	$2.76 \times 10^5$	216

\* $T_9 = 100$  keV assumed.

In the number in the final column we used  $T_9^9$  instead of  $T_9^{10}$  because the entropy removed involves dividing the energy removed by  $T$ .

In fact, we have detailed calculations from Alexander Heger giving the entropy. From his figure we see that in the range of ZAMS  $20M_{\odot}$ , roughly equal amounts of entropy are lost in the carbon burning and silicon burning stages (between O depletion and core collapse in the figure), and not much in oxygen burning. However, the point of carbon depletion was taken to be that where the carbon abundance had dropped to  $10^{-4}$ , very small. A better choice may be the time when the amount of O consumed in the O-O reaction equals the amount of C left over. Then, we expect, the decrease of entropy in O-O reaction will be about the same as in C-C reaction.

It makes sense to combine the entropy change during carbon and oxygen burning, because most of the energy, 1.18 MeV per nucleon, is furnished in going from carbon to oxygen and only 0.16 MeV is going from oxygen to silicon. From silicon to iron 2.16 MeV per nucleon is provided. Thus, the large entropy loss in our table for oxygen burning is

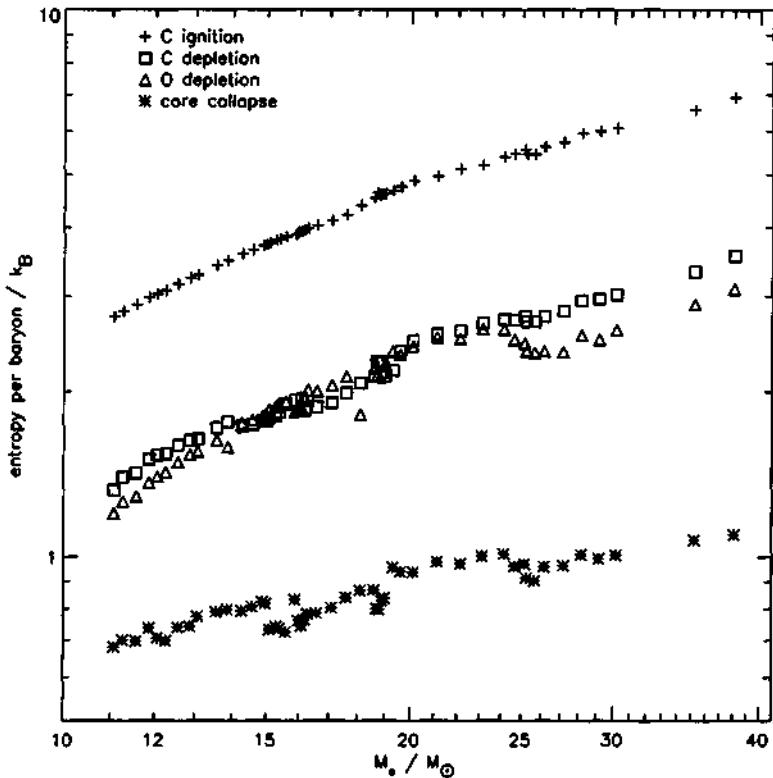


Figure. Entropy per baryon at different stages of the evolution. Calculations are done by Alexander Heger (private communication).

misleading, although some of the entropy emitted during this time would have come from energy produced during C+C burning. For an  $\sim 18M_\odot$  star in which C+C burning takes place the figure shows that roughly 2.3 units of entropy are emitted during carbon and oxygen burning, and about 1.3 units during silicon burning. The ratio is not far out of line from the  $(134 + 256)/216$  from our very rough schematic model in the table.

Whereas the curves of entropy per particle are nearly flat through the region about  $20M_\odot$ , it can be seen from Fig. 2 of Paper 21 that the Fe core mass increases rapidly just in this region. This tells us that the total entropy of the Fe core increases rapidly, the total entropy being roughly proportional to the number of nucleons since the entropy per nucleon is nearly constant. We can turn this around and say that since the C+C burning is skipped, entropy cannot be carried off during this phase, as it is in lower mass stars. Therefore, the skipping of the C+C burning is responsible for the rapid increase in Fe core mass.

In Paper 9, Table 3 gives estimates of the mass that falls back after a supernova explosion (assumed to be successful), and in Paper 8, Eq. (15) gives the gravitational binding energy of the final compact core. These corrections to the final compact core mass roughly cancel, so we equate it with the Fe core mass. Until successful supernova calculations are carried out, this is the best we can do.

In Papers 5 and 13 we laid out our estimates of the maximum neutron star mass of  $1.5M_{\odot}$ , lower than generally accepted. This low mass resulted from kaon condensation. At a density  $\sim 3\rho_0$ , where  $\rho_0$  is nuclear matter density, the  $K^-$  in medium mass is brought down to  $\sim 20$ – $250$  MeV, just in the range of the electron chemical potential. At this density electrons turn into  $K^-$  mesons. The latter are bosons, so they can go into a Bose condensate. This softens the equation of state and results in a rather low maximum neutron star mass. Alternative schemes introduce strangeness through hyperons, especially  $\Sigma^-$ -mesons, since a  $\Sigma^-$ -meson can replace both a neutron and an electron. However, with movement toward chiral restoration (Paper 25 in the Appendix) the vector meson field which favors the introduction of  $\Sigma^-$ 's decreases and kaon condensation takes over. Having both of these possibilities can only decrease the pressure, since the system goes from one minimum to the other deeper one in the free energy. Thus, we are confident in our conclusion. Compact cores in the range of  $1.5$ – $1.7$  to  $1.8M_{\odot}$  can be supported for a few seconds as neutron stars by the trapped neutrinos. As these leave, the pressure is lowered and the core goes into a low-mass black hole. It is low in mass because the supernova explosion, which takes only 2–3 seconds, has already taken place and blown off the He envelope. Thus, only the Fe core plus a small amount of external material collapse. We label the upper limit for low-mass black hole formation as  $M_{PC}$ , where PC stands for protocompact (Paper 13).

If the compact core mass is higher than  $M_{PC}$ , then the compact core collapses immediately (He envelope included) into a high-mass black hole, typically  $7M_{\odot}$ . Note that this is for a single star. The theme of Paper 21 is that in binaries if mass is transferred from the primary to the companion, then because of the large wind loss from the naked He star, we must go to much higher ZAMS masses,  $\gtrsim 60M_{\odot}$ , before we can make high-mass black holes.

Our maximum neutron star and protocompact mass are shown in Fig. 2 of Paper 21.

### Miscellaneous Remarks

There are reports, most recently by Middeditch *et al.* (2000), of observation of periodicities in radiation from 1987A, which would imply a neutron star. In Paper 3 we have made our case that the remnant must be a black hole.

The Fe core masses go down again after  $\sim$  ZAMS  $25M_{\odot}$  stars. However, by this time the envelopes are so massive with such great binding energies that they cannot be blown off by the shock and they will go into high-mass black holes.

In the earlier work of Wellstein & Langer 1999 (referenced in Paper 21), naked He stars evolved into either low-mass black holes or neutron stars. Even the more reasonable reduced (by a factor of 2–4) He winds employed by these authors did not leave a sufficiently massive He star to go into a high-mass black hole.

Naked He stars do not skip convective C+C reaction, whatever their mass. Helium is continuously being blown off their surfaces, rather than added to them by hydrogen burning in clothed He stars. In clothed He stars this newly formed He is convected

throughout the star. As noted in Sec. 3 of Paper 21, this helps get rid of the last  $^{12}\text{C}$  in clothed stars.

Silicon burning is so much more energetic than the earlier burnings that it proceeds roughly independently of the initial conditions, especially of core mass. Thus, in the ashes from the clothed He star from a ZAMS  $20M_{\odot}$  star, the central entropy following Si burning is 0.76, taking the Boltzmann constant to be unity. The He core burned as a naked one ended up at central entropy 0.775 per nucleon, only  $\sim 2\%$  different than in the clothed case. In both cases, the central entropy is not far from that of BBAL, Paper 1, where  $\sigma/N = 0.93$ .

Finally, we note the sensitivity to the ZAMS mass at which the C+C core burning is skipped to the  $^{12}\text{C}(\alpha, \gamma)^{16}\text{O}$  reaction. Schaller *et al.* (1992) used a value of 100 keV barns at energy 300 keV, smaller than the Woosley & Weaver 175 keV barns. From their table one finds that the central carbon abundance following He core burning goes below 15% for ZAMS mass  $25M_{\odot}$ , six solar masses higher than the corresponding mass in Woosley & Weaver. Thus, each lowering of the 175 keV by about 12 keV decreases the ZAMS mass at which the C+C reaction is skipped by about  $1M_{\odot}$ .

We want the narrow window in which the Fe core mass goes from  $M_{\text{NS}}$  to  $M_{\text{PC}}$  in Fig. 2 of Paper 21 to occur around  $18\text{--}19M_{\odot}$ , however, because this is the only region in which a low-mass black hole can be formed, so we favor the Woosley & Weaver value of  $^{12}\text{C}(\alpha, \gamma)^{16}\text{O}$ .

## References

- (1) Heger, A., Woosley, S.E., Martinez-Pinedo, G., and Langanke, K. (2001), *Astrophysical Journal* **560**, 307.
- (2) Middleditch, J., Kristian, J.A., Kunkel, W.E., Hill, K.M., Watson, R.D., Lucinio, R., Imamura, J.N., Steiman-Cameron, T.Y., Shearer, A., Butler, R., Redfern, M., and Danks, A.C. (2000), *New Astronomy* **5**, 243.
- (3) Schaller, G., Schaerer, D., Meynet, G., and Maeder, A. (1992), *Astronomy & Astrophysics Supplement Series* **96**, 269.
- (4) Woosley, S.E., and Weaver, T.A. (1995), *Astrophysical Journal Supplement Series* **101**, 181.

This page is intentionally left blank



ELSEVIER

## Formation of high mass X-ray black hole binaries

G.E. Brown<sup>a</sup>, A. Heger<sup>b</sup>, N. Langer<sup>c</sup>, C.-H. Lee<sup>a,d,\*</sup>, S. Wellstein<sup>e</sup>, H.A. Bethe<sup>f</sup>

<sup>a</sup>Department of Physics & Astronomy, State University of New York, Stony Brook NY 11794, USA

<sup>b</sup>Department of Astronomy and Astrophysics, University of California, Santa Cruz, CA 95064, USA

<sup>c</sup>Astronomical Institute, P.O. Box 80000, NL-3508 TA Utrecht, The Netherlands

<sup>d</sup>School of Physics, Korea Institute for Advanced Study, Seoul 130-012, South Korea

<sup>e</sup>Institut für Physik, Universität Potsdam, Am Neuen Palais 10, D-14415 Potsdam, Germany

<sup>f</sup>Floyd R. Newman Laboratory of Nuclear Studies, Cornell University, Ithaca, New York, NY 14853, USA

Received 22 February 2001; received in revised form 20 July 2001; accepted 23 July 2001

Communicated by S.E. Woosley

### Abstract

The discrepancy in the past years of many more black-hole soft X-ray transients (SXTs), of which a dozen have now been identified, had challenged accepted wisdom in black hole evolution. Reconstruction in the literature of high-mass X-ray binaries has required stars of up to  $\sim 40 M_{\odot}$  to evolve into low-mass compact objects, setting this mass as the limit often used for black hole formation in population syntheses. On the other hand, the sheer number of inferred SXTs requires that many, if not most, stars of ZAMS masses  $20-35 M_{\odot}$  end up as black holes (Portegies Zwart et al., 1997; Ergma and van den Heuvel, 1998).

In this paper we show that this can be understood by challenging the accepted wisdom that the result of helium core burning in a massive star is independent of whether the core is covered by a hydrogen envelope, or ‘naked’ while it burns. The latter case occurs in binaries when the envelope of the more massive star is transferred to the companion by Roche Lobe overflow while in either main sequence or red giant stage. For solar metallicity, whereas the helium cores which burn while naked essentially never go into high-mass black holes, those that burn while clothed do so, beginning at ZAMS mass  $\sim 20 M_{\odot}$ , the precise mass depending on the  $^{12}\text{C}(\alpha, \gamma)^{16}\text{O}$  rate as we outline. In this way the SXTs can be evolved, provided that the H envelope of the massive star is removed only following the He core burning.

Whereas this scenario was already outlined in 1998 by Brown et al. [NewA 4 (1999) 313], their work was based on evolutionary calculations of Woosley et al. [ApJ 448 (1995) 315] which employed wind loss rates which were too high. In this article we collect results for lower, more correct wind loss rates, finding that these change the results only little.

We go into the details of carbon burning in order to reconstruct why the low Fe core masses from naked He stars are relatively insensitive to wind loss rate. The main reason is that without the helium produced by burning the hydrogen envelope, which is convected to the carbon in a clothed star, a central  $^{12}\text{C}$  abundance of  $\sim 1/3$  remains unburned in a naked star following He core burning. The later convective burning through  $^{12}\text{C} + ^{12}\text{C}$  reactions occurs at a temperature  $T \sim 80$  keV.

Finally, we show that in order to evolve a black hole of mass  $\gtrsim 10 M_{\odot}$  such as observed in Cyg X-1, even employing extremely massive progenitors of ZAMS mass  $\gtrsim 60 M_{\odot}$  for the black hole, the core must be covered by hydrogen during a substantial fraction of the core burning. In other words, the progenitor must be a WNL star. We evolve Cyg X-1 in an

\*Corresponding author.

E-mail addresses: popenc@nuclear.physics.sunysb.edu (G.E. Brown), alex@ucolick.org (A. Heger), N.Langer@astro.uu.nl (N. Langer), chlee@kias.re.kr (C.-H. Lee), stephan@astro.physik.uni-potsdam.de (S. Wellstein).

analogous way to which the SXTs are evolved, the difference being that the companion in Cyg X-1 is more massive than those in the SXTs, so that Cyg X-1 shines continuously. © 2001 Published by Elsevier Science B.V.

**PACS:** 97.10.Me; 97.10.Tk; 97.60.Jd; 97.60.Lf; 97.80.Jp

**Keywords:** Binaries: close; Stars: neutron; Black hole physics; Stars: evolution; Stars: Wolf–Rayet; Stars: mass-loss

## 1. Introduction

Accepted wisdom in the astrophysical community (for recent relevant articles see Portegies Zwart et al., 1997; Ergma and van den Heuvel, 1998) has been that the results of burning the helium core of a star are independent of whether it is covered by a hydrogen envelope or whether it burns as a ‘naked’ star, the envelope having been removed; e.g. by Roche Lobe overflow in main sequence or red giant stage of the star. This latter situation takes place in binary evolution where the evolution referred to in the above two articles determines the black hole progenitors to have ZAMS mass  $> 40 M_{\odot}$ . Thus, this mass limit for black hole formation is adopted for the evolution of the black hole soft X-ray transients (SXTs) in these articles. On the other hand, as acknowledged in them, the observed number of SXTs requires that many, if not most, stars of ZAMS masses in the interval of  $\sim 20$ – $35 M_{\odot}$  evolve into high-mass black holes, of mass  $\gtrsim 6 M_{\odot}$ .

Brown et al. (1999) offered as explanation of the large number of SXTs that stars in the mass range  $\sim 20$ – $35 M_{\odot}$  could evolve into high-mass black holes as long as their H envelope was lifted off by the companion star only after the He core burning was complete (Case C mass transfer). If it were lifted off earlier, too small an Fe core evolved to make a high-mass black hole. In fact, Timmes et al. (1996) had found that Fe cores of Type Ib supernovae were systematically of lower mass than those of Type II, basically the same concept. The Brown, Lee, and Bethe scenario had been outlined by Brown et al. (1996). Their results were not generally believed, however, because the calculation of Woosley et al. (1995) on which they relied employed mass loss rates that were too large by a factor of 2 to 3 for the naked He stars. Thus, the present work using lower, more correct, rates is necessary.

Weilstein and Langer (1999) recalculated the evolution of naked He cores to their carbon–oxygen

cores with wind loss rates lowered by a factor of 2, and in some cases by larger factors, finding that because of feedbacks, the total wind loss scaled more slowly than linearly with the multiplier on the wind loss rate. In work that we describe here, Alexander Heger, using the Woosley and Weaver computer program Kepler, evolved the CO cores further to their Fe cores. These are not sufficiently massive to end up as high-mass black holes; in fact, the situation is nearly the same as found by Brown et al. (1996) where too large He star mass loss rates had been used. Our main objective in this article is to explain why the masses of the final Fe cores depend only insensitively on the He core mass loss rates. We interpret the Fe core mass as a good indicator of the final compact core mass because binding energy corrections and mass fallback very nearly compensate each other (Brown et al., 1996).

The plan of this paper is as follows:

In Section 2 we discuss the improved measurements of wind losses as deduced from Wolf–Rayet stars. In Section 3 we show that the large differences between ‘naked’ and ‘clothed’ He star burning can be understood in terms of the different ways in which carbon burns in the two cases. In Section 4 we deal with a possible evolution of the  $\sim 10 M_{\odot}$  black hole in Cyg X-1, suggesting that there may be an analogy with the evolution of transient sources; namely that the He core must be kept covered by H during most of its burning. This requires a very massive progenitor, a WNL star. In Section 5 we give simple concluding remarks.

## 2. Wind loss rates in WNE’s (naked He stars)

It is the mass loss rate proposed by Langer (1989a) for hydrogen-free Wolf–Rayet stars, which replaced the previously used assumption of a constant Wolf–Rayet mass loss rate in many evolution-

ary calculations (Schaller et al., 1992; cf. WLW, 1995), which leads to the small final masses of stars above  $\sim 35 M_{\odot}$ . This semi-empirical rate has been criticised as too high, one argument being the existence of massive black hole binaries. Indeed, recently studies of Wolf-Rayet stars show that originally measured wind losses have to be corrected downwards by a factor of 2...3 to account for their 'clumpiness' (Hannam and Koesterke, 1998). This is supported by polarisation measurements of the Thomson scattering, which depend linearly on the wind density (St-Louis et al., 1993; Moffat and Robert, 1994), and is in approximate agreement with the rates that would be deduced from the observed rate of increase in orbital periods for spherical mass loss

$$\frac{\dot{M}}{M} = \frac{\dot{P}}{2P} \quad (1)$$

in Wolf-Rayet binaries. In V444 Cygni  $\dot{P} = 0.202 \pm 0.018$  s yr $^{-1}$  was obtained by Khaliullin et al. (1984) and  $M_{WR} = 9.3 \pm 0.5 M_{\odot}$  by Marchenko et al. (1994), resulting in

$$\dot{M}_{dyn} = 1 \times 10^{-5} M_{\odot} \text{ yr}^{-1}. \quad (2)$$

This is to be compared with the

$$\dot{M} = 0.75 \times 10^{-5} M_{\odot} \text{ yr}^{-1} \quad (3)$$

obtained by St-Louis et al. (1993) from the polarisation measurements. In later work Moffat and Robert (1994) arrive at a mean of  $(0.7 \pm 0.1) \times 10^{-5} M_{\odot} \text{ yr}^{-1}$ . For WNE (defined here as a Wolf-Rayet phase with vanishing surface hydrogen abundance) stars, WLW (1993) used a He wind rate of

$$\dot{M} = 5 \times 10^{-8} (M_{WR}/M_{\odot})^{2.6} M_{\odot} \text{ yr}^{-1} \quad (4)$$

which would give a mass loss rate of  $1.6 \times 10^{-5} M_{\odot} \text{ yr}^{-1}$  for a  $9.3 M_{\odot}$  WR, a factor of 1.6 to 2.1 greater than Eqs. (2) and (3). Given the errors in measurement, we feel safe in saying that the true wind loss rate is not less than 1/3 to 1/2 the WLW value. Furthermore, we should mention that absolute value slope of the mass loss relation Eq. (4) agrees well with the latest empirical mass loss rate determinations of Nugis and Lamers (2000).

However, decreasing the rate to 1/3 does not change our conclusions, as shown by recent calcula-

tions with such a reduced mass loss rates by Wells-tein and Langer (1999), by Fryer et al. (2001), and by the calculations of Heger described later in this paper. These show that stars with initial masses below  $60 M_{\odot}$  are unlikely to produce black holes more massive than  $10 M_{\odot}$ , from naked helium stars, although high mass black holes ( $\sim 5-10 M_{\odot}$ ) may result from the ZAMS mass range  $\sim 20-35 M_{\odot}$  employed by Brown et al. (1999) provided Case C mass transfer occurs in the binary. The relative insensitivity of the final Fe core masses to the wind-loss rate depends somewhat intricately on the way carbon burns in a naked He core, as we describe in the next section.

Before we go on to the next section we discuss, however, the possible evolution of Cyg X-1, a continuously shining black hole binary where the black hole mass is observed to be  $\gtrsim 10 M_{\odot}$ . Even though massive stars are so prone to lose their hydrogen envelopes and then suffer extreme wind mass loss as hydrogen-free Wolf-Rayet stars, Langer (1987) pointed out that the most massive stars may in fact avoid this kind of evolution. The reason for doing so is that these stars, upon core hydrogen exhaustion, produce an extended region of intermediate hydrogen abundance in between the helium core and the hydrogen-rich envelope. While the latter is supposed to be lost quickly in a so-called Luminous Blue Variable (LBV) stage (e.g., Stothers, 2000), this intermediate region may be sufficiently massive that the star does not manage to blow it away during the ensuing so-called WNL stage (defined here as a Wolf-Rayet phase with non-vanishing surface hydrogen abundance).

An example for the resulting evolutionary sequence O star  $\rightarrow$  LBV  $\rightarrow$  WNL  $\rightarrow$  SN is the  $100 M_{\odot}$  sequence of Langer and El Eid (1986). Also WLW (1993) showed that the final masses of stars above  $\sim 80 M_{\odot}$  may increase again for larger initial masses, while stars in the range  $\sim 35 \dots 60 M_{\odot}$  obey the opposite trend. Even using the high mass loss rate, their  $85 M_{\odot}$  WRB star evolved a final He core of  $9.71 M_{\odot}$ . We will estimate in Section 4 that this can be brought up to  $\sim 16 M_{\odot}$  with the lower, more correct, mass loss rates. The reason that the He core of the WLW  $85 M_{\odot}$  star was so much more massive than that of their  $60 M_{\odot}$  star, was not only the higher mass of the progenitor but the fact that the WNL

stage of the  $85 M_{\odot}$  comprised  $\sim 40\%$  of the Wolf-Rayet stage, whereas that of the  $60 M_{\odot}$  star was  $\sim 25\%$ . During the WNL stage the star is covered by hydrogen, so that the mass loss rate is smaller than it would have been for a naked He star, and moreover the helium core mass increases with time.

Although the evolution of star as massive as  $\sim 85 M_{\odot}$  is very uncertain, we believe that the only possibility for them to end up in  $\geq 10 M_{\odot}$  black holes is an extended WNL stage, lasting a major fraction of the core helium burning. As we discuss below, these stars would have an active hydrogen burning shell source, whereas the existence of such a stage in a single massive star would probably not affect its evolution substantially, an expansion of the star due to hydrogen shell burning does give the possibility of common envelope evolution with a companion. In other words, the most massive stars are covered by H during much of their core He burning time, and this may enable them to end up as high-mass black holes.

### 3. ‘Naked’ vs. ‘Clothed’ helium core burning

#### 3.1. Central carbon abundances vs. convective carbon burning

As noted earlier, Timmes et al. (1996) and Brown et al. (1996) showed that compact cores which followed from ‘naked’ helium stars were substantially less massive than followed from those ‘clothed’ during helium burning with hydrogen envelopes. More recently, this has been explored in detail by Wellstein and Langer (1999) and Fryer et al. (2001). An important property of the stellar core that is significantly influenced by this difference is the carbon abundance in the core at central helium depletion. If its abundance is high enough, central carbon burning become exothermic enough to significantly overcome the neutrino losses and burn in a convective core. Otherwise the first convective burning of carbon can only appear in a shell source. The carbon abundance also has significant effect on the location, duration, and extent of the carbon burning shells. For example, as the carbon abundance goes to zero, no convective shell burning occurs either. The lower the carbon abundance, the further out the first

shell forms. The location of the last carbon-burning shell sets the size of the carbon-free core which determines the final evolution of the star and the sizes of the iron core (possible direct black hole formation) and the silicon core (possible black hole formation by fall-back).

There are two major factors that contribute to the resulting carbon abundance after core helium depletion. The first is the dependence on the mass of the helium core and results chiefly from the different behavior of the  $^{12}\text{C}(\alpha, \gamma)^{16}\text{O}$  reaction, as described clearly by Weaver and Woosley (1993). Since this is important to our development, we repeat the argument, supplemented by more recent developments.

Carbon is formed in the triple  $\alpha$ -process; since this is a three-body process, it depends on the density as  $\rho^2$ . Carbon is removed, when possible, by the  $^{12}\text{C}(\alpha, \gamma)^{16}\text{O}$  process, which goes linearly with  $\rho$ . Now, for massive helium cores the central density scales roughly as  $M^{-1/6}$ . With increasing helium core mass this implies that carbon formation, which goes as  $\rho^2$ , will be cut down, compared with  $^{12}\text{C}(\alpha, \gamma)^{16}\text{O}$ , which goes as  $\rho$ . Here, we also note that both the  $^{12}\text{C}(\alpha, \gamma)^{16}\text{O}$  reaction and the triple  $\alpha$ -process have very similar temperature dependence in the temperature regime of central helium burning,  $2 \times 10^8 \text{ K}$ , i.e., the carbon production does not depend on temperature in this regime.

This means that there is a mass above which the post-helium burning central carbon abundance is low enough to skip central carbon burning. Woosley and Weaver (1995) found this transition at a mass of  $19 M_{\odot}$ . However, this limit strongly depends on the still uncertain  $^{12}\text{C}(\alpha, \gamma)^{16}\text{O}$  reaction rate and the stellar evolution model, in particular the prescription used for convection or mixing processes in general, and also mass loss. On this last issue we elaborate later in more detail.

A bare helium star behaves differently from a clothed one in that its helium core does not grow in mass due to hydrogen shell burning, as is the case in clothed stars, but rather shrinks due to mass loss from the surface — the whole star is only a helium core, a Wolf-Rayet star. Generally, the convective core tends to comprise a larger fraction of the helium core, by mass, as the central helium abundance decreases, due to the density-dependence outlined above. This statement depends, again, somewhat on

Table 1

Results of Wellstein and Langer (1999) and Fryer et al. (2001) with reduced WR mass loss rate, 1/2 of WLW (1993) value, at the end of central He burning. The  $M_{Fe}$  is the final Fe core mass

Model	$M_{MS}$	$M_{comp}$	$C_c$	$M_{He}$	$M_{CO}$	$M_{Fe}$	Mass transfer
1s	60	34	0.30		3.37	1.35	B
2s	60	34	0.32	4.07	3.07	1.50	A + AB
5s	40	30	0.33	3.84	2.87	1.49	A + AB
7s	30	24	0.34	3.63	2.71	1.46	A + AB
10s	25	24	0.36	3.39	2.42	1.49	A + AB
17s	20	18	0.36	3.39	2.18	1.56	B

$C_c$ : Central Carbon abundance. Masses: in unit of  $M_\odot$

the description of semiconvection, but probably applies to most models with either Schwarzschild convection (e.g., Schaller et al., 1992), overshooting or fast semiconvection (e.g., Woosley and Weaver, 1995), or rotationally induced mixing (Heger et al., 2000). Additionally, the mass fraction of the convective core also increases with the mass of helium core. However, bare helium cores (early-type Wolf-Rayet stars) experience mass loss rates that are sufficiently strong that the convective core actually tends to shrink in the run of its evolution, in particular also towards the end of helium burning, rather than grow, as it does in a clothed star. This is despite the fact that it still comprises an increasingly larger fraction of the remaining total helium core. The important point is that a growth in mass of the convective core injects new helium into this convective core, while when the mass is constant or decreases, this injection does not occur.

As the triple alpha reaction depends on the third power of the helium mass fraction it loses against the  $^{12}\text{C}(\alpha, \gamma)^{16}\text{O}$  reaction toward the end of central

helium burning, i.e., carbon is mostly burned rather than produced toward the end of central helium burning. That switch typically appears at a central helium mass fraction of around 10–20 %. Most importantly, as can be seen from the central carbon abundances at the end of He burning in Table 1, the He fraction is too low to burn the final carbon, which remains at a central abundance  $\geq 0.30$  for all mass stars. This is roughly double the carbon abundance necessary for convective carbon burning. It is seen in Table 2 that the central carbon abundance in naked He stars goes down only slowly with decreasing mass loss rate, although by the time the mass loss is reduced by a factor of 6 the convective burning stops. In the clothed stars the growth of the core and its accompanied injection of helium after this time thus leads to a further decrease of carbon as compared to the bare helium cores that do not have this additional supply of helium.

In Fig. 1 we show calculations carried out by Tom Weaver (priv. com. 1995) using the KEPLER code (Weaver et al., 1978; Woosley and Weaver, 1995) on

Table 2

Results of Wellstein and Langer (1999) and Fryer et al. (2001) for different mass loss rates of a ZAMS  $60 M_\odot$  star at the end of central He burning. Models '1s#' correspond to the binary system (ZAMS  $60 M_\odot$  and  $34 M_\odot$ ) with Case B mass transfer

Model	Mass loss rate*	$C_c$	$\tau_c$	$M_{CO} [M_\odot]$	$M_{Fe} [M_\odot]$
1s1	1	0.35	3800 yrs	2.35	1.32
1s2	1/2	0.30	1600 yrs	3.37	1.35
1s3	1/3	0.27	700 yrs	4.76	1.61
1s4	1/4	0.25	500 yrs	5.93	1.75
1s6	1/6	0.22	No	8.53	1.50

$C_c$ : Central Carbon abundance.  $\tau_c$ : Convective carbon core burning time.

\* Ratio of WR mass loss to that of WLW, 1993.

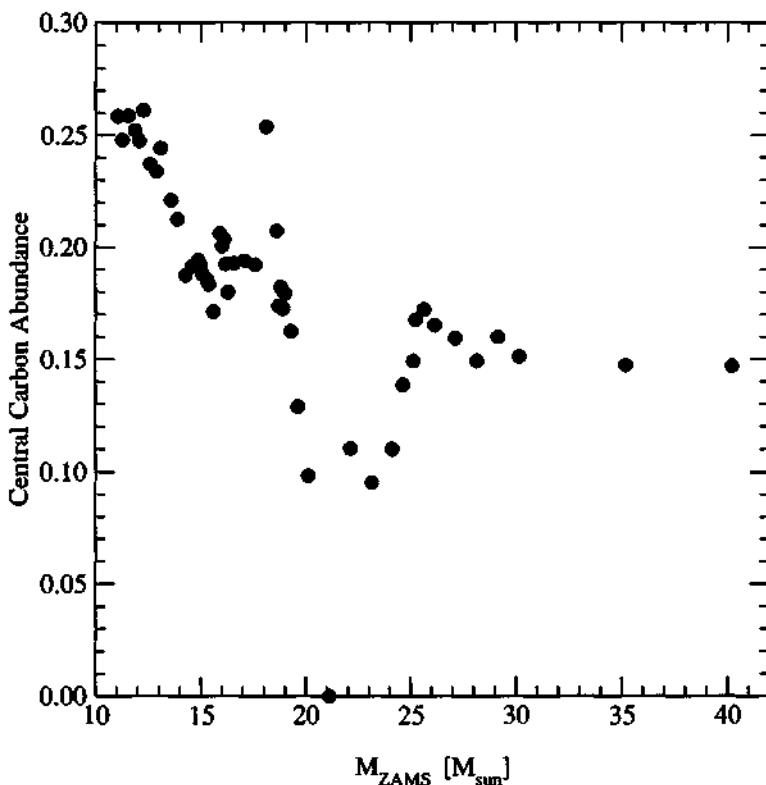


Fig. 1. Central carbon abundance at the end of He core burning for 'clothed' (single) stars as function of ZAMS mass. The calculation was carried out by Tom Weaver (1995, private communication) using the KEPLER code (Weaver et al., 1978; Woosley and Weaver, 1995). The rapid drop in  $C_c$  at ZAMS mass  $M_{ZAMS} \sim 20 M_\odot$  signals the disappearance of convective carbon burning.

the central carbon abundance after core helium depletion for 'clothed' (single) stars without (any) mass loss as function of ZAMS mass. The rapid drop in  $C_c$  at ZAMS mass  $M_{ZAMS} \sim 20 M_\odot$  causes the disappearance of convective carbon burning.

### 3.2. Iron core masses

It is generally known that the central entropy rises with the mass of the star; so that the more entropy is carried off, the less massive will be the final Fe core. This statement can be made more quantitatively by using the Bethe et al. (1979) conclusion that the entropy per nucleon in the Fe core is  $\leq 1$ , in units of the Boltzmann constant. For a ZAMS  $20 M_\odot$ , the central entropy in our Fe core is 0.76 per nucleon when the He core is burned clothed, 0.775 when

burned naked; i.e., essentially the same. Thus, the total entropy of the Fe core is  $\sim N_N$  considering some gradient towards higher entropy at the outer layers of the core, where  $N_N$  is the number of nucleons in it. In this way one can see directly that the entropy that is carried off during the evolution of the Fe core will diminish its mass.

In Fig. 2 we show iron core masses at the time of iron core collapse for a finely spaced grid of stellar masses (Heger et al., 2001). Filled circles and crosses correspond to the core masses of 'clothed' single stars. The circles were calculated by Tom Weaver (1995, priv. com.) using the same physics as in Woosley and Weaver (1995), whereas the crosses employ the improved weak rates by Langanke and Martinez-Pinedo (2000) for electron capture and beta decay. Although the latter are much smaller than the

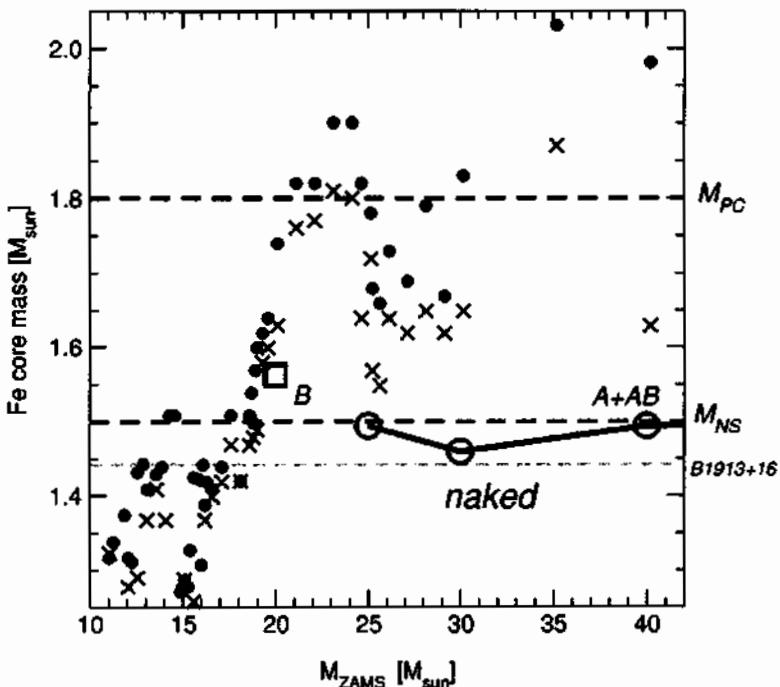


Fig. 2. Comparison of the iron core masses resulting from the evolution of 'clothed' and 'naked' He cores. Filled circles and crosses correspond to the core masses of 'clothed' stars at the time of iron core implosion for a finely spaced grid of stellar masses (Heger et al., 2001). The circular were calculated with the Woosley and Weaver, 1995 code, whereas the crosses employ the vastly improved Langanke and Martinez-Pinedo (2000) rates for electron capture and beta decay. Open circles (square) correspond to the naked He stars in case A + AB (B) mass transfer of Fryer et al. (2001), with reduced WR mass loss rate, see Table 1. If the assembled core mass is greater than  $M_{PC} = 1.8 M_{\odot}$ , where  $M_{PC}$  is the proto-compact star mass as defined by Brown and Bethe (1994), there is no stability and no bounce; the core collapses into a high mass black hole.  $M_{NS} = 1.5 M_{\odot}$  denotes the maximum mass of neutron star (Brown and Bethe, 1994). The mass of the heaviest known well-measured pulsar, PSR B1913 + 16, is also indicated with dashed horizontal line (Thorsett and Chakrabarty, 1999).

rates by Fuller et al. (1985) used by Woosley and Weaver (1995), the final Fe core masses are not much changed, for reasons described by Heger et al. (2001).

For 'clothed' single stars, one sees an anticorrelation between the peak in Fe core masses at ZAMS mass  $\sim 23 M_{\odot}$  and the minimum in  $C_c$  at ZAMS mass  $\sim 21 M_{\odot}$ . In other words, the peak around  $23 M_{\odot}$  in ZAMS masses occurs just where the central carbon abundance at the end of helium core burning is at its minimum of  $\sim 10\%$ , too low for convective core burning or the formation of a carbon-burning convective shell close to the center of the star.

When sufficient carbon remains, it will be burned

convectively in reactions such as  $^{12}\text{C} + ^{12}\text{C} \rightarrow ^{24}\text{Mg}$ ,  $^{20}\text{Ne} + \alpha$  etc., at a temperature of 70–80 keV, several times higher than the  $\sim 20$  keV needed for  $^{12}\text{C}(\alpha, \gamma)^{16}\text{O}$ . If the carbon abundance is lower, the carbon burning phases are reduced and typically bigger carbon-free cores result for otherwise similar stars. Therefore less energy is carried away by neutrinos and the entropy in the core stays higher. More massive silicon and iron cores can form in this case (Boyes et al., 2001). This is clearly seen in Figs. 1 and 2.

Although the central carbon abundance goes up again for ZAMS mass  $\sim 25 M_{\odot}$ , and the Fe core mass goes down, by such high masses the gravitational energy of the stellar envelope is large and it is

difficult for the shock energy after collapse to blow it off, so we believe that these will go into high-mass black holes, as well as the stars of ZAMS mass 20–23 M<sub>⊙</sub>.

### 3.3. Formation of high-mass black holes

Although there is a lot of uncertainty in the literature about the ZAMS masses which end up in high mass black holes, Bethe and Brown (1999) argued that this should occur at the proto compact core mass of

$$M_{\text{PC}} \sim 1.8 \text{ M}_{\odot}. \quad (5)$$

Further uncertainty comes in relating the Fe core mass to the compact core mass. It can, however, be seen from Table 3 of Brown et al. (1996) that the estimated fallback mass following separation in the supernova explosion roughly compensates for the binding energy increase in going from the Fe core to the compact object. Thus, we will assume the compact object masses to be the same as those of the Fe cores. In addition to the uncertainty in this relation, there is the further uncertainty that the fallback mass, which is dependent on the density structure in the layers above the Fe core (Fryer, 1999; Janka, 2001), is unknown, since consistent supernova explosions have not been achieved.

None the less, we believe it useful to move ahead with our estimated  $M_{\text{PC}}$  which indicates that single stars with ZAMS masses  $\gtrsim 20 \text{ M}_{\odot}$  go into high-mass black holes. Brown et al. (2001) recently showed that given presently accepted wind losses (Schaller et al., 1992) the high mass black holes in the transient sources with main sequence companion can only be evolved in the region of ZAMS mass  $\sim 20 \text{ M}_{\odot}$ , so our choice of  $M_{\text{PC}}$  is supported to this extent by evolutionary arguments.

We note briefly that SN 1987A, if it formed a black hole, went into a low mass black hole of mass  $1.5 \text{ M}_{\odot} < M_{\text{BH}} < 1.8 \text{ M}_{\odot}$  according to the Brown and Bethe (1994) estimates. The black hole would have a low mass in this interval because the He envelope was blown off in the delayed explosion. Note from Fig. 2 of main sequence masses that there is only a narrow interval from  $\sim 18 \text{ M}_{\odot}$  to  $20 \text{ M}_{\odot}$  in which this could happen given the above interval.

Schaller et al. (1992) use a  $^{12}\text{C}(\alpha, \gamma)^{16}\text{O}$  S-factor

of  $S(300 \text{ keV}) \sim 100 \text{ keV barns}$  as compared with the Woosley and Weaver 170 keV barns. They bring their central abundance down to 0.16 at the end of core He burning only for a ZAMS  $25 \text{ M}_{\odot}$  star, so presumably this would be the mass at which the Fe cores would begin to rise rapidly in mass with further evolution. We believe the Weaver and Woosley value for the  $^{12}\text{C}(\alpha, \gamma)^{16}\text{O}$  rate to be more correct (Boyes et al., 2001), placing the narrow interval of ZAMS masses from which SN 1987A can be evolved correctly. In fact, most recently experiments including both E1 and E2 components have been carried out (Kunz et al., 2001), obtaining  $S_{\text{tot}}^{300} = (165 \pm 50) \text{ keV barns}$ .

## 4. Possible evolution of Cyg X-1

### 4.1. Space velocity vs. the mass loss in the formation of Cyg X-1

According to the calculations of Wellstein and Langer (1999) and Fryer et al. (2001), it is unlikely that a black hole as massive as the  $10 \text{ M}_{\odot}$  core (Herrero et al., 1995) in Cyg X-1 can be evolved from a naked He star (see Table 1 and Fig. 2). Although the calculations of naked He stars were extended up to only ZAMS  $60 \text{ M}_{\odot}$  stars, the He winds scale with the 2.6 power of the mass, so that higher mass He stars would be expected to behave in the same way. In fact, Wellstein and Langer (1999) and Fryer et al. (2001) were unable to evolve a  $10 \text{ M}_{\odot}$  black hole, whereas the high space velocity of Cyg X-1,  $49 \pm 14 \text{ km s}^{-1}$  (Kaper et al., 1999) indicates substantial mass loss in a Blaauw–Boersma kick (Blaauw, 1961; Boersma, 1961) in the black hole formation, increasing the necessary He core mass. We now estimate this mass loss, following the development of Nelemans et al. (1999). It is reasonable to evolve Cyg X-1 analogously to the soft X-ray black hole transient sources, the difference being in the high mass companion star which presumably makes Cyg X-1 shine continuously.

The high space velocity of Cyg X-1 can be explained by mass loss in the black hole formation, because this loss from the black hole is somewhat off from the center of gravity of the system. From

Nelemans et al. (1999), the runaway velocities from symmetric SNe (See Appendix A for the derivation)

$$v_{\text{sys}} = 213 \times \Delta M \times m \times P_{\text{re-circ}}^{-1/3} \times (M_{\text{BH}} + m)^{-5/3} \text{ km s}^{-1} \quad (6)$$

where masses are in  $M_{\odot}$ ,  $P$  in days, and  $\Delta M = M_{\text{He}} - M_{\text{BH}}$  with  $M_{\text{He}}$  the He star mass of the Black hole progenitor. Here  $P_{\text{re-circ}}$  is the re-circularized orbital period after Blaauw–Boersma kick, and we assume no orbital evolution ( $P_{\text{re-circ}} = P_{\text{obs}}$ ) since the beginning of the mass transfer phase, and neglect small eccentricity as in Nelemans et al. (1999). By putting in average values,

$$v_{\text{sys}} = 8.32 \times \Delta M \text{ km s}^{-1}. \quad (7)$$

In order to obtain the observed velocity, we need  $\Delta M \sim 5.9 M_{\odot}$ , indicating the black hole progenitor mass to be  $M_{\text{He}} \sim 16 M_{\odot}$ . Note that this  $\Delta M$  and  $M_{\text{He}}$  are in the same ballpark as those estimated by Nelemans et al. for Nova Scorpii 1994 (GRO J1655–40). This suggests that the evolution of Cyg X-1 is similar to that of Nova Scorpii 1994, except that the progenitor mass of the black hole must be much higher in Cyg X-1 and the companion is an O-star rather than an F-star.

The current orbital separation of Cyg X-1 is

$$a_{\text{now}} \approx 4.2 R_{\odot} \left( \frac{P_{\text{orb}}}{\text{days}} \right)^{2/3} \left( \frac{M_{\text{BH}} + M_{\odot}}{M_{\odot}} \right)^{1/3} \approx 40 R_{\odot}. \quad (8)$$

Before the explosion, the orbital separation of the black hole progenitor and companion star was

$$a_{\text{preSN}} \approx \frac{a_{\text{now}}}{1 + e_2} \approx 33 R_{\odot} \quad (9)$$

where  $e_2 = \Delta M / (M_{\text{BH}} + M_{\odot}) \approx 0.21$  is the eccentricity right after the supernova explosion. For the derivation, see Appendix A. This is only slightly larger than the  $30 R_{\odot}$  estimated by Bethe and Brown (1999) for the minimum initial separation of the O-star progenitors. If the progenitors were initially closer together they would merge already at this stage of evolution.

As noted by Bethe and Brown (1999), the two O-star progenitors must have initially had a separation of nearly  $33 R_{\odot}$ . Substantial mass loss, espe-

cially by the WR star would have been expected to widen the orbit substantially

$$\frac{a_f}{a_i} = \frac{M_i}{M_f} \sim 2-3 \quad (10)$$

where  $M_i$  and  $M_f$  are the initial and final system masses. In order to tighten the orbit to the  $33 R_{\odot}$  we find before the massive star goes into a black hole, there must have been some period of nonconservative mass transfer.

#### 4.2. Extended WNL stage of black hole progenitor

Langer (1989b) finds that towards the end of core hydrogen burning the radius of a massive star may be much larger than its radius on the zero age main sequence. For example, a  $100 M_{\odot}$  star increases its radius from  $13 R_{\odot}$  at H ignition by a factor  $\sim 4$  to  $53 R_{\odot}$  at core H exhaustion. In the evolutionary phase between central H- and He-burning massive stars develop an intermediate fully convective zone just above the H-burning shell which reaches its maximal spatial extent and  $\sim 18 M_{\odot}$  in mass for the  $100 M_{\odot}$  star (Langer, 1987). There may be further expansion in a hydrogen shell burning stage unless large mass loss sets in. But the latter may begin only later in the LBV stage. In any case, there will be a period before or in the LBV stage where the hydrogen is being propelled outwards, but does not yet have enough velocity to escape. A companion to the very massive star will, if in this region, couple hydrodynamically to the hydrogen in the manner explicitly worked out by Bethe and Brown (1999), and furnish energy to it from a drop in its gravitational binding; i.e., through some mechanism resembling common envelope evolution. We cannot be more explicit, because of the uncertainty in the post main sequence evolution of the very massive stars, but some nonconservative mass transfer seems to be necessary in the evolution of a binary such as Cyg X-1. We admit to being in somewhat of a predicament, however, since we do not want a common envelope evolution which removes the hydrogen before the helium burning is well underway.

It is clear that the black hole in Cyg X-1 cannot have evolved from a naked He star. An important clue to its possible evolution is offered by the evolution of WLW (1993) of a ZAMS  $85 M_{\odot}$  star,

where the WR stage began with a WNL stage which took  $\sim 40\%$  of the total WR time. (Their  $60 M_{\odot}$  star also had a WNL stage, but it took only  $\sim 25\%$  of the WR time.) During this time the WR is covered by hydrogen and the attendant mass loss rate is much lower than that for a naked He star.

Using the times for each WR stage as WLW (1995), we calculate in Table 3 the masses for ZAMS  $60 M_{\odot}$  and  $85 M_{\odot}$  stars. Of course there is feedback on these masses from the change in times of each stage with altered mass loss rate. In order to take the feedback into account, we use the ratios of  $M_{\text{CO}}$  cores calculated by Fryer et al. (2001) for a ZAMS  $60 M_{\odot}$  star for various mass loss rates to the original WLW (1995) one as plotted in Fig. 3.

We note from Table 3 that a ZAMS  $60 M_{\odot}$  star with WR stage mass loss rate reduced by  $1/3$  loses  $1.7 M_{\odot}$  during its WNL stage, at an average rate of  $1.5 \times 10^{-5} M_{\odot} \text{ yr}^{-1}$ . This is an order of magnitude less than it would lose as a WNE (bare) He star, and explain why the final He star mass can be as large as  $13.5 M_{\odot}$ . The WNL mass loss rate is  $\sim 2.25$  times the main sequence mass rate for a ZAMS  $60 M_{\odot}$  star of Castor et al. (1975).

In Fig. 3 we plot our ratios from Table 3 for He star masses alongside the Fryer et al. (2001) ratios of CO core masses. The divergence in our favored range of  $1/2$ – $1/3$  is not large, so we believe our procedure in calculating masses is justified. We see

from Table 3 that we need a mass loss rate of  $\sim 0.4$  times the WLW (1995) one in order to lose  $\sim 6 M_{\odot}$  in the explosion and to be left with an  $\sim 10 M_{\odot}$  black hole.

#### 4.3. Discussion on Cyg X-1 type objects

We want to point out that there is a strong metallicity dependence in our model. As the winds of WNL stars are likely radiation driven (Hamann et al., 2000), we can expect them to weaken with smaller metallicity. Thus, the lower limit of the initial mass range from which the final stellar mass can grow again will be decreasing. Therefore, we expect more Cyg X-1 type systems at lower metallicity. We also would like to point out that there is dependence in mixing processes, in particular rotation, which can increase the lower limit of the initial mass range.

Using standard population synthesis, Bethe and Brown (1999) estimated that there should be  $\sim 7$  Cyg X-1 like objects in the Galaxy, where we see only one (leaving out LMC X-1 and LMC X-3). Although the  $\gamma$ -rays from such an object easily penetrate the Galactic disc, they would not be distinguished from those from gamma ray bursts (Maarten Schmidt, private communication).

As possible WR progenitor of the Cyg X-1 black hole, we suggest ‘WR22’; the most massive Wolf-

Table 3

Mass-loss-rate-dependences of masses (in  $M_{\odot}$ ) at the different stages of WR stars (models 60WRB and 85WRB of WLW93). The WR stage mass loss rates are in units of standard rate of WLW93. The same WR stage times of WLW93 are used for different mass loss rates

	WR stage mass loss rate						$\tau (10^5 \text{ yrs})$
	1	1/2	0.4	1/3	1/4	1/6	
<b><math>60 M_{\odot}</math> WRB</b>							
pre WR stage	26.5	26.5	26.5	26.5	26.5	26.5	
post WNL	22.5	24.2	24.5	24.8	25.1	25.6	1.13 (25%)
post WNE	11.5	15.4	16.5	17.5	18.9	20.8	1.64 (36%)
post WC/WO	6.65	9.60	10.7	11.6	13.2	15.5	1.77 (39%)
after feedback*	6.65	9.53	11.9	13.5	16.8	24.1	
<b><math>85 M_{\odot}</math> WRB</b>							
pre WR stage	45.3	45.3	45.3	45.3	45.3	45.3	
post WNL	34.3	40.0	40.8	41.6	42.5	43.4	1.40 (41%)
post WNE	14.7	20.9	23.0	24.7	27.5	31.2	1.25 (36%)
post WC/WO	9.7	14.3	16.1	17.6	20.2	24.0	0.80 (23%)
after feedback*	9.7	13.9	17.3	19.6	24.5	35.2	

\* Scaled by the numerical results of Fryer et al. (Fig. 3).

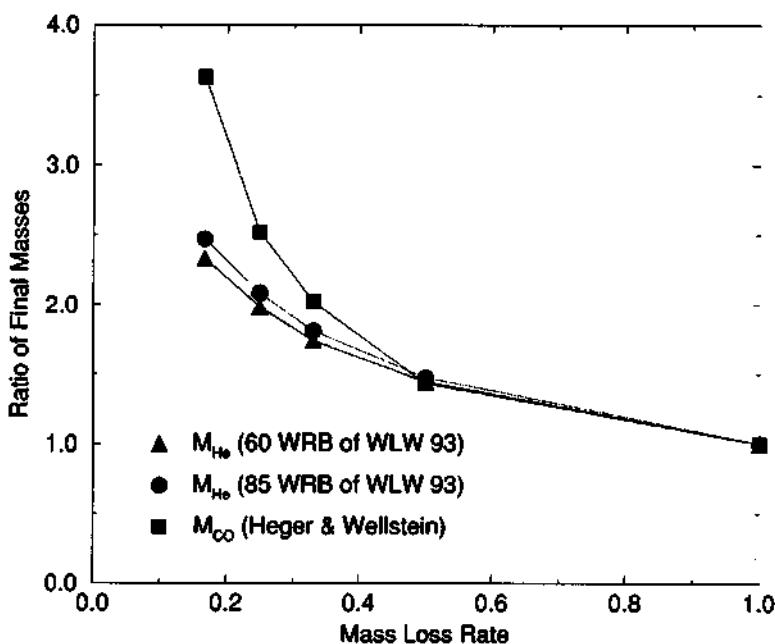


Fig. 3. Mass-loss-rate-dependences of final CO core masses  $M_{\text{CO}}$  (filled square) from the  $60 \text{ M}_\odot$  star discussed in Fryer et al. (2001) and final He core masses (filled triangles and circles for  $60 \text{ M}_\odot$  and  $85 \text{ M}_\odot$ , respectively) after WR stage from Table 1. The masses are scaled by those with standard rate of WLW93.

Rayet ever weighed (Rauw et al., 1996). The minimum Wolf-Rayet mass is  $72 \text{ M}_\odot$  and the mass ratio to O-star is 2.78, giving an O-star mass of  $> 26 \text{ M}_\odot$ . The spectrum of the WN7 exhibits a considerable amount of hydrogen ( $X_H \sim 40\%$ , Hamman et al., 1991, Crowther et al., 1995), strong WR emission lines and absorption lines that belong to the WR component (Niemela, 1973; Moffat and Seggewiss, 1978). Rauw et al. (1996) identify the companion as an O-star. With  $\sim 80$  day period, the binary is, however, much too wide in order to narrow sufficiently in common envelope evolution to produce a Cyg X-1 type binary, according to the Bethe and Brown (1999) estimates.

## 5. Conclusion

Our conclusion to all of the above is simple: In order to evolve high-mass black holes at solar metallicity, the He core of the massive star must be

covered by hydrogen at least most of the time, while it burns.

For SXTs, the stars in the mass range  $\sim 20$ – $35 \text{ M}_\odot$  could evolve into high-mass black holes with the condition that their H envelope was lifted off by the companion star only after the He core burning was complete. For more massive stars like Cyg X-1 type objects, an extended WNL stage could drive the formation of high-mass black holes.

## Acknowledgements

We thank Stan Woosley for many helpful discussions and are indebted to Tom Weaver for supplying us with his grid of stellar evolution models. We wish to thank Maarten Schmidt for communication about the indistinguishability of further Cyg X-1 like objects from gamma ray bursts. This work was partially supported by the U.S. Department of Energy under Grant No. DE-FG02-88ER40388, by the Alexander von Humboldt-Stiftung through Grant

FLF-1065004, by 2000–2001 KIAS Research Fund, and by the Deutsche Forschungsgemeinschaft through Grants La 587/15 and La 587/16.

### Appendix A. Blaauw–Boersma Kick

In this appendix, the details of the Blaauw–Boersma kick (Blaauw, 1961; Boersma, 1961) are summarized. Before the explosion,<sup>1</sup> from Fig. 4(a),  $R_{\text{He}}$ ,  $v_{\text{He}}$ , and  $\omega_1$  are given as

$$\begin{aligned} R_{\text{He}} &= \left( \frac{m}{M_{\text{He}} + m} \right) a_1 \\ v_{\text{He}} &= \frac{1}{a_1} (GmR_{\text{He}})^{1/2} \\ \omega_1 &= \left( \frac{m}{M_{\text{He}} + m} \right) \left( \frac{Gm}{R_{\text{He}}^3} \right)^{1/2}. \end{aligned} \quad (\text{A.1})$$

After the explosion, we assume that  $\Delta M = M_{\text{He}} - M_{\text{BH}}$  was lost from the BH progenitor without interacting with the binary system. The momentum lost from the binary system should be compensated by the space velocity of the new binary system.

$$\Delta p = \Delta M \cdot v_{\text{He}} = (M_{\text{BH}} + m) \cdot v_{\text{sys}} \quad (\text{A.2})$$

which gives the space velocity of the new binary system (c.m. motion of B)

$$v_{\text{sys}} = \left( \frac{\Delta M}{M_{\text{BH}} + m} \right) v_{\text{He}}. \quad (\text{A.3})$$

Right after the explosion, in the new c.m. frame of BH and O-star binary, by defining the semi-major axis  $d_2$  and the eccentricity  $e_2$ , we have

$$\begin{aligned} \omega_2 &= \left( \frac{m}{M_{\text{BH}} + m} \right) \left( \frac{Gm}{d_2^3} \right)^{1/2} \\ R_{\text{BH}} &= \left( \frac{m}{M_{\text{BH}} + m} \right) a_1 \\ v_{\text{BH}} &= v_{\text{He}} + v_{\text{sys}} = \left( \frac{M_{\text{He}} + m}{M_{\text{BH}} + m} \right) v_{\text{He}}. \end{aligned} \quad (\text{A.4})$$

Since the  $v_{\text{BH}}$  is greater than that of the circular motion with radius  $R_{\text{BH}}$

$$\begin{aligned} v_{\text{BH}} &= \left( \frac{M_{\text{He}} + m}{M_{\text{BH}} + m} \right)^{1/2} v_{\text{circ}} > v_{\text{circ}} \\ v_{\text{circ}} &= \left( \frac{Gm^2}{a_1(M_{\text{BH}} + m)} \right)^{1/2}, \end{aligned} \quad (\text{A.5})$$

one can conclude that  $R_{\text{BH}}$  is at the minimum of the orbital separation, i.e.

$$R_{\text{BH}} = d_2(1 - e_2). \quad (\text{A.6})$$

Note that  $v_{\text{BH}}$  and  $R_{\text{BH}}$  are perpendicular right after the explosion as in Fig. 4(c).

The orbital angular momentum of the BH can be expressed using  $a_2$  and  $e_2$

$$\begin{aligned} l_a &= [Gm\mu^2 d_2(1 - e_2^2)]^{1/2} \\ &= [Gm\mu^2 R_{\text{BH}}(1 + e_2)]^{1/2} \end{aligned} \quad (\text{A.7})$$

or

$$\begin{aligned} l_b &= M_{\text{BH}} R_{\text{BH}} v_{\text{BH}} = \mu a_1 v_{\text{BH}} \\ &= \mu (GmR_{\text{BH}})^{1/2} \left( \frac{M_{\text{He}} + m}{M_{\text{BH}} + m} \right)^{1/2} \end{aligned} \quad (\text{A.8})$$

where  $\mu$  is the reduced mass

$$\mu = \frac{M_{\text{BH}} m}{M_{\text{BH}} + m}. \quad (\text{A.9})$$

By equating  $l_a = l_b$  we have

$$e_2 = \frac{\Delta M}{M_{\text{BH}} + m}. \quad (\text{A.10})$$

We assume that there is no angular momentum lost during the circularization process by tidal locking after SN explosion. From angular momentum conservation

$$\begin{aligned} l_{\text{now}} &= (Gm\mu^2 d_{\text{now}})^{1/2} \\ &= [Gm\mu^2 d_2(1 - e_2^2)]^{1/2}. \end{aligned} \quad (\text{A.11})$$

Therefore, from

<sup>1</sup>Subscript 1 (2) indicates the properties before (right after) the explosion, and subscript ‘now’ indicates the current properties observed.

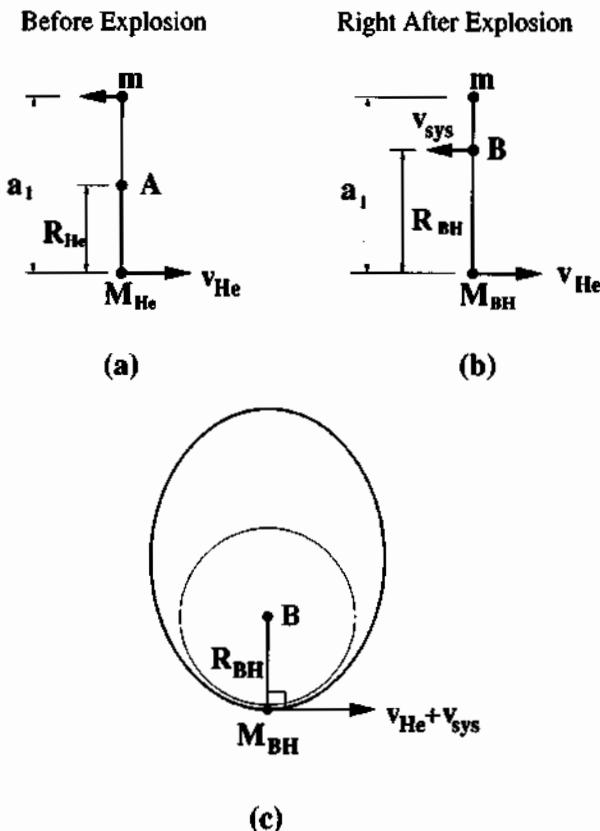


Fig. 4. Coordinates (a) before the explosion in the c.m. frame, (b) right after the explosion in the original c.m. frame, and (c) in the new c.m. frame of black hole and companion right after the explosion.

$$\begin{aligned} d_{\text{now}} &= \left( \frac{m}{M_{\text{BH}} + m} \right) a_{\text{now}} \\ &= d_2 (1 - e_2^2) = R_{\text{BH}} (1 + e_2) \\ &= \left( \frac{m}{M_{\text{BH}} + m} \right) a_1 (1 + e_2), \end{aligned} \quad (\text{A.12})$$

we have the relation of the orbital separations

$$a_{\text{now}} = a_1 (1 + e_2). \quad (\text{A.13})$$

Using this relation, one have

$$\begin{aligned} \omega_1 &= (1 + e_2)^2 \omega_{\text{now}} \\ \omega_2 &= (1 - e_2^{3/2}) \omega_{\text{now}} \end{aligned} \quad (\text{A.14})$$

or the period relations

$$P_{\text{now}} = (1 - e_2^2)^{3/2} P_2 = (1 + e_2)^2 P_1. \quad (\text{A.15})$$

Now Eq. (A.13) can be expressed in terms of observed quantities using

$$\begin{aligned} a_{\text{now}} &= \left( \frac{P_{\text{now}}}{2\pi} \right)^{2/3} [G(M_{\text{BH}} + m)]^{1/3} \\ v_{\text{He}} &= \left( \frac{1 + e_2}{a_{\text{now}}} \right)^{1/2} \left( \frac{Gm^2}{M_{\text{He}} + m} \right)^{1/2} \\ &= \left[ \frac{1}{a_{\text{now}}} \left( \frac{Gm^2}{M_{\text{BH}} + m} \right) \right]^{1/2} \\ &= \frac{(2\pi G)^{1/3} m}{(M_{\text{BH}} + m)^{2/3} P_{\text{now}}^{1/3}}. \end{aligned} \quad (\text{A.16})$$

Finally the Blaauw–Boersma kick velocity becomes

$$\begin{aligned}
 v_{\text{sys}} &= \frac{\Delta M}{M_{\text{BH}} + m} v_{\text{He}} = \frac{(2\pi G)^{1/3} \Delta M m}{(M_{\text{BH}} + m)^{5/3} P_{\text{now}}^{1/3}} \\
 &= 213 \left( \frac{\Delta M}{M_{\odot}} \right) \left( \frac{m}{M_{\odot}} \right) \\
 &\times \left( \frac{P_{\text{now}}}{\text{days}} \right)^{-1/3} \left( \frac{M_{\text{BH}} + m}{M_{\odot}} \right)^{-5/3} \text{ km s}^{-1}.
 \end{aligned} \tag{A.17}$$

## References

- Bethe, H.A., Brown, G.E., 1999. *ApJ* 517, 318.
- Bethe, H.A., Brown, G.E., Applegate, J.H., Lattimer, J., 1979. *Nucl. Phys. A* 324, 487.
- Blaauw, A., 1961. *BAN* 15, 265.
- Boersma, J., 1961. *BAN* 15, 291.
- Boyes, M.M., Heger, A., Woosley, S.E., 2001, in preparation.
- Brown, G.E., Bethe, H.A., 1994. *ApJ* 423, 659.
- Brown, G.E., Lee, C.-H., Bethe, H.A., 1999. *NewA* 4, 313.
- Brown, G.E., Lee, C.-H., Tauris, T., 2001. *NewA* 6, 331.
- Brown, G.E., Weingartner, J.C., Wijers, R.A.M.J., 1996. *ApJ* 463, 297.
- Castor, J.I., Abbott, D.C., Klein, R.I., 1975. *ApJ* 195, 157.
- Crowther, P.A., Hillier, D.J., Smith, L.J., 1995. *A&A* 293, 403.
- Ergma, E., van den Heuvel, E.P.J., 1998. *A&A* 331, L29.
- Fryer, C.L., 1999. *ApJ* 522, 413.
- Fryer, C.L., Heger, A., Wellstein, S., Langer, N., 2001, in preparation.
- Fuller, G.M., Fowler, W.A., Newman, M.J., 1985. *ApJ* 293, 1.
- Hamann, W.-R., Koesterke, L., 1998. *A&A* 335, 1003.
- Hamann, W.-R., Dünnbeil, G., Koesterke, L., Schmutz, W., Wessolowski, U., 1991. *A&A* 249, 443.
- Hamann, W.-R., Koesterke, L., Gräfener, G., 2000. In: Lamers, H., Sapar, A. (Eds.), *ASP Conf. Ser.* 204, *Thermal and Ionization Aspects of Flows from Hot Stars*, p. 197.
- Heger, A., Langer, N., Woosley, S.E., 2000. *ApJ* 528, 368.
- Heger, A., Woosley, S.E., Martinez-Pinedo, G., Langanke, K.-H., 2001, *astro-ph/0011507*.
- Herrero, A., Kudritzki, R.P., Gabler, R., Vilchez, J.M., Gabler, A., 1995. *A&A* 297, 556.
- Janka, H.-Th., 2001. *A&A* 368, 527.
- Kaper, L., Camerón, A., Barziv, O., 1999. In: van der Hucht, K.A., Koenigsberger, G., Eenens R.J. (Eds.), *Wolf-Rayet phenomena in massive stars and starburst galaxies*, IAU Symp. 193, p. 316.
- Khaliullin, K.F., Khaliullina, A.I., Cherepashchuk, A.M., 1984. *SvAL* 10, 250.
- Kunz, R., Jaeger, M., Mayer, A., Hammer, J.W., Starudt, G., Harissopolis, S., Paradellis, T., 2001. *PhRvL* 86, 3294.
- Langanke, K., Martinez-Pinedo, G., 2000. *Nucl. Phys. A* 673, 481.
- Langer, N., 1987. *A&A* 171, L1.
- Langer, N., 1989a. *A&A* 220, 135.
- Langer, N., 1989b. *Rev. Mod. Astron.* 2, 306.
- Langer, N., El Eid, M.F., 1986. *A&A* 167, 265.
- Marchenko, S.V., Moffat, A.F.J., Koenigsberger, G., 1994. *ApJ* 422, 810.
- Moffat, A.F.J., Robert, C., 1994. *ApJ* 421, 310.
- Moffat, A.F.J., Seggewiss, W., 1978. *A&A* 70, 69.
- Nelemans, G., Tauris, T.M., van den Heuvel, E.P.J., 1999. *A&A* 352, L87.
- Niemela, V.S., 1973. *PASP* 85, 220.
- Nugis, T., Lamers, H.J.G.L.M., 2000. *A&A* 360, 227.
- Portegies Zwart, S., Verbunt, F., Ergma, E., 1997. *A&A* 321, 207.
- Rauw, G., Vreux, J.-M., Gosset, E., Hutsemékers, D., Magain, P., Rochowicz, K., 1996. *A&A* 306, 771.
- Schaller, G., Schaerer, D., Meynet, G., Maeder, A., 1992. *A&AS* 96, 269.
- St.-Louis, N., Moffat, A.F.J., Lapointe, L., Efimov, Y.S., Shakhovskoy, N.M., Fox, G.K., Piironen, V., 1993. *ApJ* 410, 342.
- Stothers, R.B., 2000. *ApJ* 530, L103.
- Thorsett, S.E., Chakrabarty, D., 1999. *ApJ* 512, 288.
- Timmes, F.X., Woosley, S.E., Weaver, T.A., 1996. *ApJ* 457, 834.
- Weaver, T.A., Zimmerman, G.B., Woosley, S.E., 1978. *ApJ* 225, 1021.
- Weaver, T.A., Woosley, S.E., 1993. *PhR* 227, 65.
- Wellstein, S., Langer, N., 1999. *A&A* 350, 148.
- Woosley, S.E., Langer, N., Weaver, T.A., 1993. *ApJ* 411, 823.
- Woosley, S.E., Langer, N., Weaver, T.A., 1995. *ApJ* 448, 315.
- Woosley, S.E., Weaver, T.A., 1995. *ApJS* 101, 181.

---

## Chapter 22

# Broad and Shifted Iron-Group Emission Lines in Gamma-Ray Bursts as Tests of the Hypernova Scenario

G.C. McLaughlin, R.A.M.J. Wijers, G.E. Brown and H.A. Bethe

Reprinted with permission from *Astrophysical Journal*, 567 (2002) 454–462

Copyright (2001) by American Astronomical Society

### Commentary by G.E. Brown

In mid-November 2000 I was talking with Hans Bethe on the telephone, as I often did. He had been writing some family history for his children. Hans asserted desire for new scientific endeavors. I referred him to two papers, Amati *et al.* (2000) and Piro *et al.* (2000) in the journal *Science*, which he took (so I did not have to xerox them and fax them to him), which claimed to see Fe lines associated with gamma-ray bursts. A popular article (Schilling 2000) in the same issue of *Science* outlined the supernova model of Vietri and Stella, which purported to explain the Fe lines. In this model a very massive, rapidly spinning star explodes as a supernova, with a remnant rapidly spinning neutron star. A shell of matter with substantial Fe content is expelled by the neutron star. This shell is met some months later by the gamma rays from the gamma-ray burst, which follows the collapse of the neutron star into a black hole. The problem with this scenario is that the collapse of neutron star into black hole should take place in seconds or minutes, but certainly not in months.

Fe is produced only in supernova explosions and in the hypernovae, the very energetic supernova explosions which seem to accompany gamma-ray bursts. Thus, the detection of “Fe” supported the GRB-hypernova association.

Hans immediately grasped the main points: (i) that these lines must be downscattered  $^{56}\text{Ni}$  lines, because  $^{56}\text{Ni}$ , not Fe, is emitted in supernova explosions, decaying by electron capture to  $^{56}\text{Co}$  with a lifetime of six days. The  $^{56}\text{Co}$  decays chiefly by electron capture, but also partly by positron emission, to  $^{56}\text{Fe}$  in 79 days. Since the “Fe” lines should come with the explosion connected with the black hole formation, they must have been emitted as Fe lines.

Fe lines are common in black hole accretion disks. Ralph Wijers immediately saw that the downscattering could take place as the “Fe” lines made their way out, by scattering off

the walls of the funnel responsible for the gamma-ray burst. Gail McLaughlin quickly carried out Monte Carlo calculations which form the meat of Paper 22. My role was to encourage everyone, because the GRB-hypernova association looked correct, for me, as motivated by the discussion in Paper 17. In that paper the GRB was run by the power delivered by the open field lines, whereas the hypernova was powered by the rapidly rotating black hole delivering angular momentum and energy through the closed field lines connecting the black hole to the disk. These closed field lines are frozen in the matter in the disk and rapidly torque it up.

Ralph Wijers has pointed out in work we are pursuing that one should be able to observe  $^{56}\text{Ni}$  lines in hypernovae for many years following the explosion. This is because the ionization wave removing the bound electrons from the  $^{56}\text{Ni}$  goes as  $r^{-2}$ , whereas the electron density drops faster with  $r^{-3}$  in the expansion following the explosion. Thus, in a day or two following the explosion, there are no electrons to capture by the  $^{56}\text{Ni}$  in order to decay into  $^{56}\text{Co}$ . In principle  $^{56}\text{Ni}$  could decay into  $^{56}\text{Co}$  by positron emission ( $^{56}\text{Co}$  decays into  $^{56}\text{Fe}$   $\sim 20\%$  of the time by this path), but the relevant  $f^t$  values are tiny (Gail McLaughlin, private communication).

## References

- (1) Amati, L., et al. (2000), *Science* **290**, 953.
- (2) Piro, L., et al. (2000), *Science* **290**, 955.
- (3) Schilling, G. (2000), *Science* **290**, 926.

## BROAD AND SHIFTED IRON-GROUP EMISSION LINES IN GAMMA-RAY BURSTS AS TESTS OF THE HYPERNOVA SCENARIO

G. C. McLAUGHLIN, R. A. M. J. WIJERS, AND G. E. BROWN

Department of Physics and Astronomy, State University of New York, Stony Brook, NY 11794-3800; [gail@tonic.physics.sunysb.edu](mailto:gail@tonic.physics.sunysb.edu), [rwijers@mail.astro.sunysb.edu](mailto:rwijers@mail.astro.sunysb.edu), [epopescu@nuclear.physics.sunysb.edu](mailto:epopescu@nuclear.physics.sunysb.edu)

AND

H. A. BETHE

Floyd R. Newman Laboratory of Nuclear Studies, Cornell University, Ithaca, NY 14853-5001

Received 2001 June 30; accepted 2001 October 29

### ABSTRACT

In the hypernova/collapsar model of  $\gamma$ -ray bursts, it is natural that radiation is emitted by the inner engine for some time after the burst. This has been discussed as a possible source of the X-ray line emission observed in some afterglows. We show here that the natural geometry of a hypernova—a source of radiation at the bottom of a deep funnel—has very significant consequences for the shape and central energy of the observed emission lines; the lines acquire a very broad scattering wing on the low-energy side and a characteristic second peak 1 Compton wavelength away from the initial energy due to the once- and twice-scattered photons. We suggest that this explains the large width of the observed lines. Furthermore, the downscattering lowers the central line energy (by up to 1 keV in the source rest frame, before the lines become unrecognizable) so that the observed line energies can become consistent with originating from cobalt and nickel, as expected in few-day-old supernova ejecta.

*Subject headings:* gamma rays: bursts — line: profiles — radiative transfer — supernovae: general

### 1. INTRODUCTION

The discovery of  $\gamma$ -ray burst (GRB) afterglows in X-ray (Costa et al. 1997), optical (van Paradijs et al. 1997), and radio (Frail et al. 1997) wavelengths greatly increased the possibilities for studying the objects emitting this radiation. This quickly led to the first measurement of a redshift (Metzger et al. 1997) and the realization that the relativistic fireball blast-wave model for their nature (Rees & Mészáros 1992; Mészáros & Rees 1997) explained the afterglows very well (Wijers et al. 1997). (For recent reviews, see van Paradijs, Kouveliotou, & Wijers 2000 and Piran 1999.) The phenomenal inferred emitted energy is at least that of a supernova; together with rapidly variable  $\gamma$ -ray emission, this suggests a stellar-remnant origin. Leading candidates of this type are mergers of neutron stars and/or black holes (Eichler et al. 1989; Mochkovitch et al. 1993; Janka et al. 1999; Salmonson, Wilson, & Mathews 2001) or core collapses of massive stars (Woosley 1993; Paczynski 1998). The latter are gaining favor for long  $\gamma$ -ray bursts (duration over 2 s; Kouveliotou et al. 1993); these are the only ones for which afterglows have been found), for a number of reasons:

1. Some GRBs have shown evidence of association with a supernova, either by a clear supernova found in their error box (GRB 980425/SN 1998bw; Galama et al. 1998) or by late-time afterglow light curves showing a feature consistent with a supernova. Clearest among the latter are GRB 980326 (Bloom et al. 1999a) and GRB 970228 (Reichart 1999; Galama et al. 2000), although other, less clear cases have been claimed. Given the relative rarity of these cases and the difficulty of detecting a supernova in the presence of a bright afterglow, it is unclear whether this association holds more generally.

2. The small spatial offsets between  $\gamma$ -ray burst counterparts and their host galaxies (e.g., Bulik et al. 1999; Bloom,

Sigurdsson, & Pols 1999b; Bloom, Kulkarni, & Djorgovski 2002) are hard to reconcile with merging neutron stars, given their high expected space velocities and long merger times, which should have them merge far away from where they were born.

3. The X-ray column density and extinction, as well as the prevalence of strong Mg I absorption lines in many optical spectra of afterglows, indicate that  $\gamma$ -ray bursts often occur in dense regions and behind much larger column densities than expected from random locations in galaxies (Galama & Wijers 2001). This suggests that they occur amidst the large amounts of dust and gas that one expects to find in star-forming regions.

4. There is increasing evidence that  $\gamma$ -ray bursts are associated with star formation on a more global scale, both because the  $\gamma$ -ray burst rate as a function of redshift is consistent with following the star formation rate (Totani 1997; Wijers et al. 1998; Krumholz, Thorsett, & Harrison 1998; Kromers et al. 2000; Schmidt 2001) and because the average star formation rate in  $\gamma$ -ray burst host galaxies is significantly above that of samples of galaxies at the same redshift (Fruchter 2002).

An emission line has recently been found in some X-ray afterglows of  $\gamma$ -ray bursts (Piro et al. 1999; Antonelli et al. 2000; Piro et al. 2000) whose energy is roughly consistent with Fe K $\alpha$  at the redshift of the host. This line has been adduced as evidence that the environment of the burst is heavily enriched in iron and thus the result of a recent supernova explosion. If this interpretation is correct, it would provide strong support for the connection between  $\gamma$ -ray bursts and stellar explosions. This makes the lines of considerable importance to unveiling the nature of  $\gamma$ -ray bursts. In this paper, we investigate some problems with the interpretation of the lines and suggest a mechanism that may help solve them. We begin by discussing models for the

formation of iron-group lines in  $\gamma$ -ray burst afterglows and their difficulties in accounting for all the observations (§ 2). Then we discuss our model for electron scattering of line photons in a hypernova funnel (§ 3) and the types of emission-line profiles that it predicts (§ 4). Finally, we discuss the relevance of our findings to current and future observations and some possible complications (§ 5) and summarize our conclusions (§ 6).

## 2. MOTIVATION

The general thrust of models for the X-ray lines is that X-rays from the burst or afterglow irradiate cooler material, ionizing it almost completely. Recombination of the material causes emission lines, and around 6–8 keV the iron-group elements are the dominant contributors. The very large iron line luminosity, about  $10^{52} (d\Omega/4\pi)$  photons  $s^{-1}$ , sustained for a day or so, requires  $10^{57} (d\Omega/4\pi)$  emitted photons ( $d\Omega$  is the solid angle illuminated by the iron line source). If every atom emitted only once (e.g., if the ionization happened in a flash, and every atom recombined only once, or if the recombination time were of the order of the observed emission time), one would need of the order of  $50 (d\Omega/4\pi) M_\odot$  of pure iron-group material to take part. Therefore, one needs a dense enough medium to have many recombinations per atom, but even then detailed models often require the presence of large amounts of this material. This gives rise to the claim that the irradiated medium is heavily enriched in iron and thus must have been enriched by a supernova quite recently. The reported BeppoSAX WFC discovery of a transient absorption feature in the prompt burst spectrum of GRB 990705 (Amati et al. 2000) would, if typical of hypernovae, also set stringent constraints on the amount of enriched ejecta and their distance from the  $\gamma$ -ray burst.

The models for the origin of the line come in two basic kinds. In one type of model, the supranova, the medium is a shell ejected in an explosion some months prior to the  $\gamma$ -ray burst (Vietri et al. 2001). When the burst flux hits the shell, it emits its lines; as one needs many photons per atom, the recombination time should be much less than the burst duration, so the density needs to be high. However, the observed duration of the line emission, about a day, is much longer. One can get this by choosing the shell size to be about a light-day so that the duration of the emission is set by the light crossing time of the shell. The emission is naturally isotropic, so  $10^{57}$  line photons must be produced. Weak points of this model are a few: first, the observed supernova “bumps” in  $\gamma$ -ray burst afterglow light curves are consistent with the supernova going off at the same time as the  $\gamma$ -ray burst, rather than the required few months before the burst. Second, the required time delay is created by having a supernova first make a massive neutron star that is partly supported by rotation. When it slows down, its centrifugal support gives way at some point and it collapses to a black hole, giving a  $\gamma$ -ray burst. The problem with this may be that there is little reason for the delay between supernova and  $\gamma$ -ray burst to be even approximately constant, which in turn makes it hard to understand why the iron shell should be about a light-day across. Third, the lines have thus far been identified as iron lines and in cases with an independently known redshift that does appear the best in terms of the observed line energy of 7 keV (Piro et al. 1999, 2000). However, a supernova primarily produces nickel, not iron. In Figure 1 we show the total

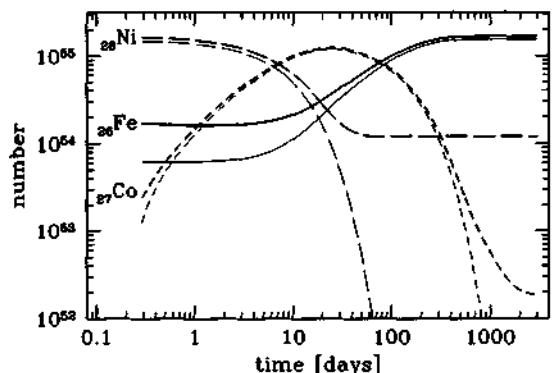


FIG. 1.—Abundances of nickel, cobalt, and iron as a function of time since the supernova for a  $40 M_\odot$  star (model S40C,  $16 M_\odot$  He core), as modeled by Woosley & Weaver (1995). The thick lines give the sum over all isotopes; the thin ones represent only the  $A = 56$  isotope of each.

abundance of iron, cobalt, and nickel as a function of time since the supernova. Iron does not become the most abundant element until 3 months after the supernova, and it takes over half a year for it to really dominate.

The other type of model is different than that described above in two respects: the irradiated material is presumed to be closer to the source (e.g., Böttcher 2000), most likely leftover material from the disrupted star. Also, the irradiation source is long-lived. It could be either the afterglow itself or residual accretion onto the compact remnant of the hypernova. The fact that the line flux and afterglow flux evolve differently with time favors the latter (e.g., in GRB 000214; Antonelli et al. 2000). A slight variation on this is the model by Rees & Mészáros (2000), who assume that the compact object emits a relativistic wind, which gives rise to heating and radiation when it hits the disrupted star. This makes little difference since in both cases the net effect is that the surface of the ejecta produces the line radiation due to bombardment with X-rays. In this model the density of the line-emitting material can be much higher; thus, the number of recombination photons per atom is much larger. Also, since this model is naturally anisotropic and thus requires fewer photons produced to begin with, much less iron-group material is needed, but enrichment by supernova ejecta still seems required (Böttcher 2000). This model does not require a delay between supernova and  $\gamma$ -ray burst and thus is more consistent with the observed supernova signatures in  $\gamma$ -ray bursts. However, the lack of delay aggravates the line identification problem: in the first few days the ejecta really are dominated by nickel, yet the line energy as observed is usually more consistent with iron.

Thus, while both types of model have problems, we feel that the disagreement in timing between the observed signs of supernovae in a few  $\gamma$ -ray bursts and the predictions of the supranova model are very hard to overcome. In this paper, we investigate the possibility of remedying the line energy problem in the context of hypernova models, in which we use the idea from Rees & Mészáros (2000) that we see a line emitted from the walls of a funnel in the progenitor star that has just been created by the burst. The wall is energized by residual emission of energy from the compact object formed during the  $\gamma$ -ray burst.

### 3. MODEL AND CALCULATIONS

We model the funnel and the scattering and emission within it as follows: the funnel is taken to be a cone with some fixed opening angle,  $\theta_0$  (to be precise,  $\theta_0$  is the angle between the cone axis and the wall). The wall is taken to have a high enough density that the mean free path of photons within the wall material is negligible compared to the cone size so that all scattering within the wall takes place in a skin layer of negligible depth. Absorption processes are parametrized by assuming a fixed, energy-independent ratio for the scattering to absorption mean free path. Photons are emitted isotropically from a ring on the cone surface some fixed distance from the apex of the funnel. The ratio,  $h_{\text{cone}}$ , of the total height of the cone to the emission height is another parameter of the problem.

We use Monte Carlo calculations to compute the photon propagation and energy change. Since the scattering of the photons is by free electrons, the angular information of scattering events must be retained in order to calculate the direction of the outgoing photon and to determine the photon energy loss. The calculations are undertaken as follows: A photon is emitted on the surface of the cone, and its direction is chosen randomly in spherical coordinates ( $\phi$ ,  $\cos \theta$ ). In this coordinate system,  $\theta = 0$  points directly out of the cone and  $\theta = \pi$  points directly down;  $\phi = 0$  points to the axis of the cone. Depending on the direction, the photon may either escape entirely, cross the cone and enter the wall at a new location, or enter the wall at its current location. One condition to describe the escape of the photon is

$$\theta < \theta_0. \quad (1)$$

The condition that describes whether the photon will cross the cone is given by

$$\cos \phi > \tan \theta_0 / \tan \theta. \quad (2)$$

If the photon crosses the cone, the change in height of the photon is given by

$$\frac{\Delta h}{h} = 2 \frac{(\cos \phi + \tan \theta_0 / \tan \theta)}{\tan \theta / \tan \theta_0 - \tan \theta_0 / \tan \theta}. \quad (3)$$

If the new height,  $h_{\text{new}} = h + \Delta h$ , is greater than the total height of the cone, then the photon escapes. If  $h_{\text{new}} < h_{\text{cone}}$ , the photon will enter the cone wall at the new point, and we then need the direction of the photon at the cone coordinates of the new point. The angle  $\theta$  is unchanged at this new position. However, the coordinate system has rotated in  $\phi$ , since  $\phi = 0$  always points back to the center of the cone. The new  $\phi$  is given by

$$\phi_{\text{new}} = \pi - \arcsin \left( \frac{\sin \phi}{1 + \Delta h/h} \right). \quad (4)$$

If the photon enters the wall then, for bookkeeping purposes, rotation matrices are used to describe its direction in coordinates in which  $\theta = 0$  parallel to the edge of the wall. For convenience, we will refer to these as the wall coordinates. The wall coordinates are given in terms of the cone coordinates as

$$\cos \theta_w = -\cos \phi \sin \theta \sin \theta_0 + \cos \theta \cos \theta_0, \quad (5)$$

$$\sin \phi_w = \sin \phi \sin \theta / \sin \theta_w. \quad (6)$$

Inside the wall the photon scatters off an electron at a distance randomly determined by the photon mean free

path  $\lambda_{\text{scat}}$ , such that the probability of the photon traveling a distance  $d$  before scattering is  $P_{\text{scat}} \propto \exp(-d/\lambda_{\text{scat}})$ . Given this distance  $d$ , the photon has a probability of being absorbed of  $P_{\text{abs}} = 1 - \exp(-d/\lambda_{\text{abs}})$ . In our calculations, we take  $\lambda_{\text{abs}} = n \lambda_{\text{scat}}$  and examine several different values of  $n$ , as discussed in § 4. Rather than throw away photons that fail the absorption test, we track a photon intensity that begins as  $I = 1$  and is multiplied by the absorption probability after each scattering. Every photon that has not escaped the cone by the time its intensity has diminished to less than a percent is discarded in our calculations. For  $\lambda_{\text{abs}} \approx 100$ , these photons have typically diminished in energy, so they are between 0.8 and 1.2 keV. If the photon is not absorbed, we calculate the distance from the edge of the wall at which the scattering occurs using  $d$  and the two angles in the wall coordinate system,  $x = d \sin \theta_w \cos \phi_w$ .

The scattered photon has a direction relative to the photon before scattering that is chosen randomly according to the Thomson cross section  $\sigma_{\text{Th}} \propto 1 + \cos^2 \theta_{\text{out}}$ , where  $\theta_{\text{out}}$  is the angle of the photon coming out of the scattering event relative to the angle of the photon going into the scattering event  $\theta_w$ . The energy of the photon is degraded according to this angle:  $E + \delta E = E/[1 + (E/m_e c^2)(1 - \cos \theta_{\text{out}})]$ , where  $m_e$  is the mass of the electron.

In order to determine whether the photon can escape the wall,  $\theta_{\text{out}}$  and  $\phi_{\text{out}}$  must be converted to the wall coordinate system:

$$\cos \theta_{\text{wall}} = -\sin \theta_w \sin \theta_{\text{out}} \cos \phi_{\text{out}} + \cos \theta_w \cos \theta_{\text{out}}, \quad (7)$$

$$\begin{aligned} \sin \phi_{\text{wall}} = & (\sin \theta_w \cos \phi_{\text{out}} \cos \theta_w \sin \theta_{\text{out}} \\ & + \cos \phi_w \sin \phi_{\text{out}} \sin \theta_{\text{out}} \\ & + \sin \phi_w \sin \theta_w \cos \theta_{\text{out}}) / \sin \theta_{\text{wall}}. \end{aligned} \quad (8)$$

Once the new photon direction is described in the wall coordinate system, we determine whether it escapes the wall. Again, we randomly choose a distance traveled,  $d$ , according to the mean free path. We update the distance from the edge of the wall to be  $x + \Delta x$ , in which  $\Delta x = d \sin \theta_{\text{wall}} \cos \phi_{\text{wall}}$ . We test to see if the photon path has taken it outside the wall. If not, then we continue to allow the photon to scatter in the wall until it leaves the wall or its intensity has decreased to less than 1%. Once the photon leaves the wall, we use rotation matrices to describe it again in the cone coordinate system. It then either escapes the cone entirely or crosses to a new point on the cone and enters the wall at a new height, calculated as starting with equation (3), with the exception that the initial path length into the wall,  $d$ , is not chosen randomly. It is the path length from the photon exit, diminished by the amount needed to exit the wall.

In summary, the photon bounces around in the cone and in the walls of the cone until either it escapes or its intensity becomes negligible. The intensities and energies of all escaping photons are recorded and used to generate the line profiles discussed in § 4.

### 4. RESULTS

We present line profiles calculated from various cone heights and opening angles and absorption mean free paths. We make contact with observation by convolving our profiles with Gaussian functions and adding power-law spectra. We also examine the effect of the finite lifetime of various

isotopes produced in a supernova-like explosion. In particular, we look at the time rate of change of nickel, cobalt, and iron abundances and the effect on the line profiles.

Before presenting the results of the Monte Carlo simulations, it is worth pointing out that there is an analytic expression for the average number of scatterings a photon will have on the way out of the cone, in a certain limiting case. If the photon does not wander into the wall and only stays on the surface, there is no absorption, and the funnel is infinite, then the average number of bounces that a photon will have on its way out of the funnel is

$$\bar{N} = (1 + \cos \theta_0) / (1 - \cos \theta_0), \quad (9)$$

where  $\theta_0$  is again the opening angle of the cone. One can see that for small cone opening angles the average number of scatterings is quite large. For example, for  $\theta_0 = 10^\circ$  the average number of bounces before escape is 130, approximately half of which are bounces "backward," i.e., bounces that do not result in cone crossings. Reducing the cone height to  $h_{\text{cone}} = 2$ , but keeping the rest of the simplifications, this number is reduced to 27. This is due to the long tail in the distribution of number of scatterings; e.g., in the infinite cone limit with  $\theta_0 = 10^\circ$ , the median number of scatterings is 89 (for a height of 2, the median is 16), whereas the mode is 0 for any opening angle.

These long tails are also present in all of the line profiles presented here, which are calculated by the method described in § 3. For example, in Figure 2 extended line profiles occur for all the cone heights studied. This figure shows results for cone heights of  $h_{\text{cone}} = 2$  and  $h_{\text{cone}} = 10$ . The line photons are always emitted from a height of 1, so the cone heights should be interpreted as ratios. Both figures plot results for cone opening angles of  $\theta_0 = \pi/16$ ,  $\theta_0 = \pi/8$ , and  $\theta_0 = \pi/4$ . One can see that increasing the opening angle has an effect that is similar to decreasing the height of the cone, and this is demonstrated in Figure 3. The tail length is strongly influenced by the absorption mean free path, since the tail contains photons that have scattered many times. Figure 4 shows cases where the mean free path for absorption has been reduced to  $\lambda_{\text{abs}} = 10\lambda_{\text{scatt}}$  and  $\lambda_{\text{abs}} = 2\lambda_{\text{scatt}}$ . Since the ratio is mostly a function of temperature, with colder material being more absorbing, this illustrates the effect of changing the wall from hot to cold; for X-ray photons, the wall retains considerable absorption opacity until the temperature reaches the keV range. Mostly we use a hot wall with  $\lambda_{\text{abs}} = 100\lambda_{\text{scatt}}$  because a radiation bath that suffices to ionize iron and nickel all the way to a hydrogen-like ion will also reduce the X-ray absorption opacity to very much below that of a cold gas.

The curves that have the initial photons emitted from or close to the surface of the wall have several features. There is an initial spike at the energy of the original line. These are all of the photons that escape immediately without any scatterings. There is also a pair of peaks separated by about  $\Delta E/E = 2E/m_e c^2$ , with the higher energy one almost coincident with the escape spike. The lower energy peak comes primarily from the photons that scatter once in the opposite direction of the opening angle and then scatter back, but also partially from twice-scattered photons. Figure 5 shows the decomposition of one of the line profiles into one, two, three, and four times scattered photons. The completely unscattered photons are not plotted as a separate curve since they simply comprise the initial spike. The effect of the angular dependence of the Thomson cross section is very

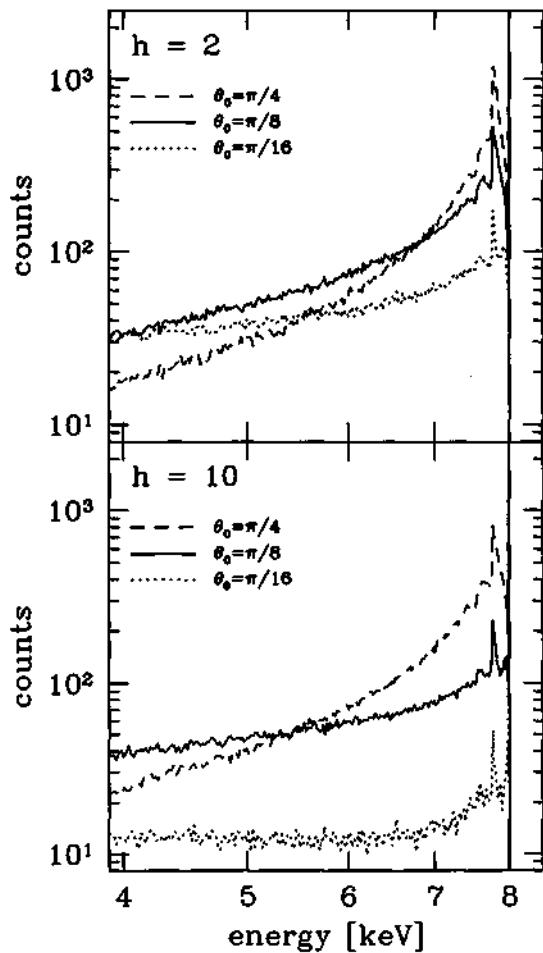


FIG. 2.—Line profiles for cone heights of 2 (top) and 10 (bottom). Photons of energy 8 keV are emitted at height 1. In each panel we show the effect of decreasing opening angle:  $\theta_0 = \pi/4$  (dashed line);  $\theta_0 = \pi/8$  (solid line); and  $\theta_0 = \pi/16$  (dotted line). As the opening angle decreases, the peak of few-times scattered photons near 8 keV decreases, and the low-energy tail becomes flatter. The depression is stronger for greater height at all opening angles.

clearly seen in the once-scattered curve. This is very similar to the energies shown in Illarionov et al. (1979) of photons scattered off free electrons but without any geometrical constraints, such as in our study.

We also show the lines convolved with Gaussians that mimic the spectral resolutions of the CCD and the High Resolution Transmission Grating (HETG) of the *Chandra X-Ray Observatory*. For the former, we use a FWHM of 0.1 keV, and for the latter we use a FWHM of 0.0033 keV; these values are appropriate near  $E = 4$  keV, where the lines are seen in our (observer) frame. A power-law spectrum with photon index of  $-2.2$  is added to the profiles with 10 times the number of counts (between 2 and 10 keV) as in the peak of the line, to mimic the background against which the line in the GRB 991216 spectrum is seen. We define the "peak" as the area within twice the full width at half

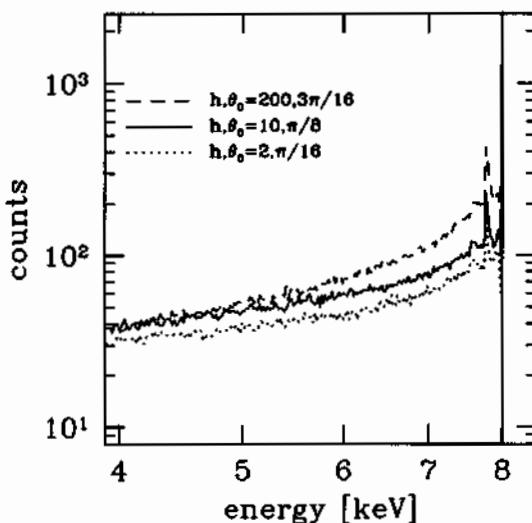


FIG. 3.—Increasing the opening angle of the cone. This has an effect that is very similar to decreasing the height, as illustrated by the similarity of three cases. Top curve:  $(h, \theta_0) = (200, 3\pi/16)$ ; middle curve:  $(10, \pi/8)$ ; bottom curve:  $(2, \pi/16)$ .

the initial peak height maximum after convolution. We have also given our line profiles a redshift of  $z = 1$ , since recently detected lines have occurred at roughly this redshift (Piro et al. 2000). In Figure 6 we show the effect of the CCD resolution and of the high-resolution grating. The lower resolution smears out most of the features of the line near the peak, but the large scattering wing remains clearly visible.

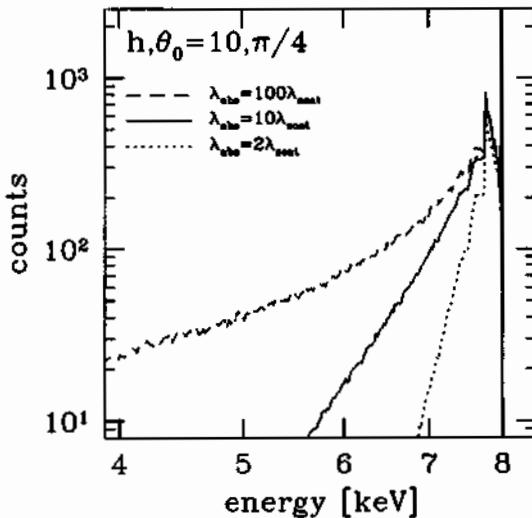


FIG. 4.—Effect of increasing the relative importance of photon absorption. Top curve:  $\lambda_{\text{abs}} = 100\lambda_{\text{scatt}}$ ; middle curve:  $\lambda_{\text{abs}} = 10\lambda_{\text{scatt}}$ ; bottom curve:  $\lambda_{\text{abs}} = 2\lambda_{\text{scatt}}$ . The opening angle of the cone is  $\pi/4$ , and the height is 10.

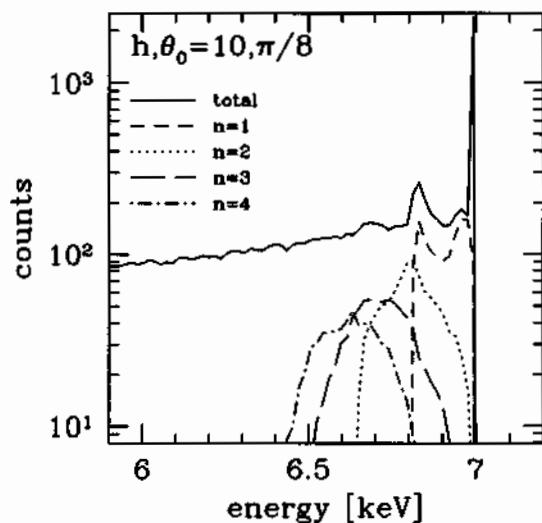


FIG. 5.—Decomposition of a line profile with a cone height of 10 and opening angle of  $\pi/8$  into photons emitted with one, two, three, and four bounces. These are the bottom curves with number of bounces increasing from right to left. The overall line profile is a solid line.

The coarser CCD resolution, however, is adequate for studying the time variation of the signal due to decay of the nickel and cobalt isotopes. In Figure 7 we show examples of these time profiles. Here we have used the results of supernova nucleosynthesis abundance calculations of Woosley & Weaver (1995) as initial conditions (model S40C) and allowed the unstable isotopes to decay, keeping track of all abundances. Most of the change in the line profiles we present is due to the decay of the  $A = 56$  isotopes of nickel, cobalt, and iron. We have used line energies of 8 keV for nickel, 7.4 keV for cobalt, and 6.9 keV for iron. Nickel-56 has a relatively short half-life of 6.1 days as compared with cobalt-56, which has a half-life of 78.8 days. In Figure 7 we show line profiles at 0.3, 1, 3, 10, and 100 days after the explosion. Even shortly after the explosion, the line profile clearly changes on a timescale of days as nickel-56 decays to cobalt-56. At later times, one can clearly see the buildup of the iron abundance on timescales of months.

One further curious point should be noted here: the decay of nickel-56 is almost exclusively by electron capture. This means that fully ionized nickel-56 has a decay time many orders of magnitude longer than neutral nickel-56! The wall material we see is dominated by fully ionized material, and therefore in principle nickel could dominate the observed emission even longer. Hydrogen-like nickel-56 will decay only half as fast as neutral nickel-56 because only one of the capturable K-shell electrons is present. Cobalt-56 decays by both electron capture and positron emission, so it will decay in a fully ionized state, but 5 times more slowly than normal. Of course, the ionized part of the wall is very thin and thus contains almost no mass. Therefore, any amount of mixing in the material that makes up the funnel will cause a given atom to spend little time in the fully ionized state, and thus the influence of ionization on the decay rate may be small in practice.

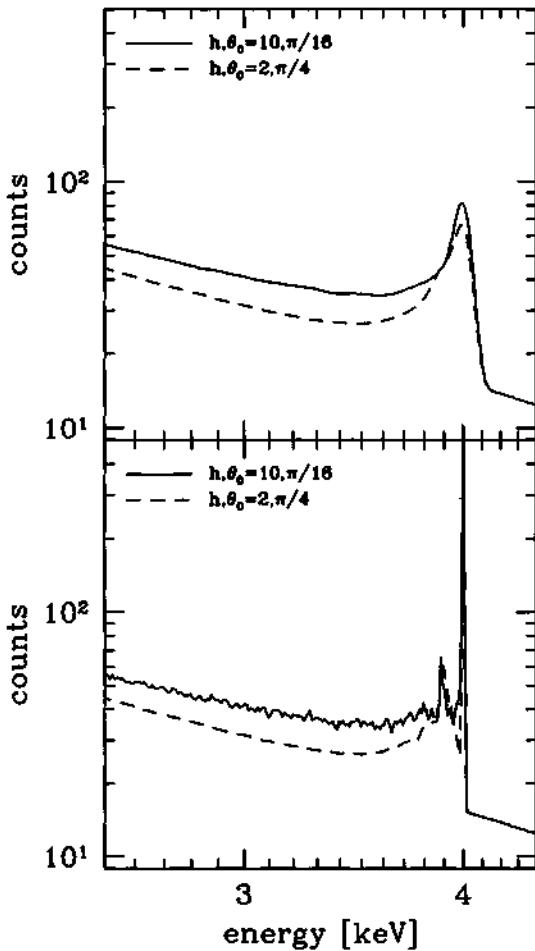


FIG. 6.—Line profiles redshifted to  $z = 1$  and convolved with a Gaussian to simulate detector resolution. A power-law spectrum has also been added with 10 times as many counts as in the narrow part of the line, so the overall spectrum resembles that of the observed case of GRB 991216 (Piro et al. 2000). Bottom line: profile for a height of 2 and an angle of  $\pi/4$ . Top line: profile for a height of 10 and an angle of  $\pi/16$ . Top panel: resolution of 0.1 keV around the line energy, mimicking the *Chandra* CCD. Bottom panel: mimics the *Chandra* HETG with a resolution of 0.003.

#### 4.1. Connection with Real Sources

Thus far we have phrased our results in terms of the minimal number of (mostly dimensionless) parameters and kept them general. However, we do primarily have the physical setting of a hypernova GRB in mind, so we briefly sketch a set of numbers to which our model applies that represent the likely situation in the aftermath of a GRB powered by a collapsing massive star. It is very like the situation sketched by Rees & Mészáros (2000). In order to have the required  $10^{52}$  photons  $s^{-1}$  in the observed line, we require perhaps  $10^{53}$  photons  $s^{-1}$  emitted above the ionization threshold. This in turn implies an X-ray luminosity of  $L_X \sim 10^{45} \beta$  ergs  $s^{-1}$ , where  $\beta > 1$  converts the ionizing luminosity into the total source luminosity. This is well in

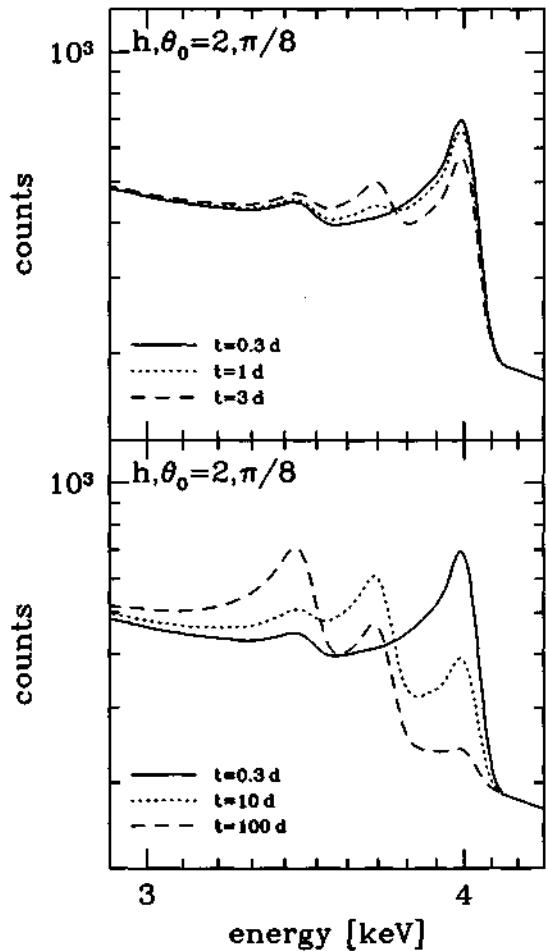


FIG. 7.—Evolution of the redshifted line profile at CCD resolution (0.1 keV) as it decays from primarily nickel isotopes to iron, by way of cobalt. Top panel: line after 0.3 days (solid line), 1 day (dotted line), and 3 days (dashed line). Bottom panel: delays of 0.3 (solid line), 10 (dotted line), and 100 days (dashed line). In this figure the funnel height is 2, and the cone angle is  $\pi/8$ .

excess of the Eddington luminosity of a stellar mass black hole, and thus cannot be emitted by the vicinity of the black hole in photons. This poses no problems to the model because, as suggested by Rees & Mészáros (2000), it can be emitted as a particle wind and converted to X-rays in shocks near the funnel wall.

If we take  $h \sim 2$  and take the total funnel size to be of order the initial size of the star ( $2 \times 10^{11}$  cm), then the line photons will be primarily generated at  $r \sim 10^{11}$  cm. This implies a very high local X-ray flux at the reprocessing site:  $F_X = 10^{22} \beta$  ergs  $cm^{-2} s^{-1}$ . Under such circumstances, the temperature in the funnel wall will be set by ionization/recombination physics to a value of a few times less than the ionization threshold energy, which in this case is about  $10^8$  K. The density in the funnel wall will be set by the fact that the wall gas pressure must balance the impinging

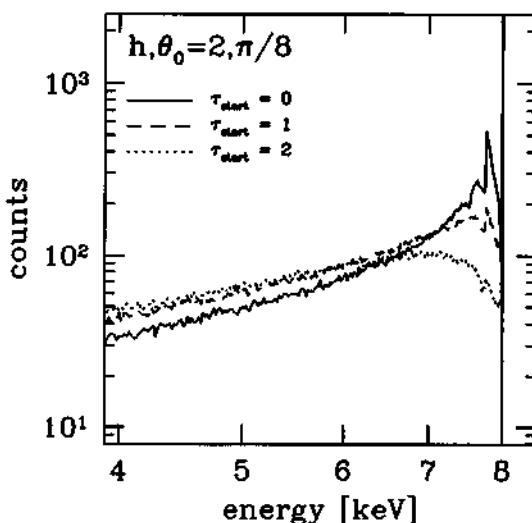


FIG. 8.—Effect of emission of the line photons at finite scattering optical depth. Solid curve: photons initially emitted at the surface of the wall. Dashed line: photons emitted at an optical depth of 1. Dotted line: photons emitted at an optical depth of 2. The initial spike heights for the three curves are 22:4:1. The plot was made for an opening angle of  $\pi/8$  and a cone height of 2.

radiation pressure, which implies  $n_e = 10^{19} \text{ cm}^{-3}$ . With this, we can assess the great importance of radiative ionization in the wall: the ionization parameter,  $\xi = L/m^2 = 10^4$ , indicating that even iron-group elements are largely ionized. It could even be so high that Thomson scattering becomes important: the scattering mean free path is only  $10^5 \text{ cm}$ , illustrating that the wall is indeed very thin. Within this layer, the electron temperature could become dominated by scattering and set to the Compton temperature,  $kT_C \sim E/4$ , which is likely of the order of 1 keV for a typical X-ray spectrum. (Note that whether it is set by Compton scattering or ionization balance, the temperature is well below the line energy, justifying our approximation of a cold wall in which only Compton downscattering of the photons is important. Also, the thermal width of any iron-group lines will be of the order of 1 keV, negligible relative to the scattering width.)

The high degree of ionization of nickel means that the mean free path due to the tiny fraction of hydrogen-like Ni atoms can be orders of magnitude smaller than the scattering mean free path. This implies (Fig. 8) that the line may easily be scattered to a width that effectively makes it invisible. Thus, the transient nature of the observed lines, as well as their total absence in some GRBs, might well be due to an ionizing flux that is too high instead of one that is too low.

##### 5. DISCUSSION

Our study of the influence of the funnel in hypernovae on emerging X-ray lines shows that the shape of the lines is very strongly affected by scattering off the funnel walls. This is true even if the funnel has an opening angle as large as 45 degrees and the line photons are produced halfway between the bottom and top of the funnel. Quite generically, the lines can be broadened to 0.5–1 keV FWHM and their centers

shifted down in energy by up to 1 keV before they become too wide to recognize as lines. An important characteristic is that the line remains fairly sharp on the blue side, with a broad red wing being the main reason for the increased width and line-center shift. If the lines are shifted down even further, then they tend to become so smeared out that they may escape detection against a background power-law spectrum from the source.

In Figure 9 we show an example of how a line may appear to become redshifted due to Compton scattering. In this figure, 8 keV nickel lines were emitted at an optical depth of 2 and  $\lambda_{\text{cone}} = 10\lambda_{\text{wall}}$  from a cone with  $h = 10$  and  $\theta_0 = \pi/8$ . The peak of this line occurs at around 7 keV, which happens to be the same energy as iron K $\alpha$  lines. We see that absorption effects, both as the photons make their way out of the wall and also on the tail of the distribution, conspire to create a peak about 1 keV below the original line energy.

The shape of the lines in the case of perfect scattering (i.e., no photons are lost by absorption) is mostly set by the direct-escape probability from the emission point. Since this probability reflects a combination of the funnel opening angle and the funnel height relative to the emission point, it follows that those two parameters are very difficult to disentangle from realistic data. The observed line wing is quite sensitive to absorption, although as the importance of absorption is increased, most photons never escape the funnel. Since the escaping ones are always those that have scattered few times, their energies have not changed very much. As is illustrated by Figures 2–5, the lines always have very extended wings in the perfect scattering case, even if the cone is wide and the escape probability is high; there is always a significant number of photons that scatter many times. The only exception to this is absorption, which

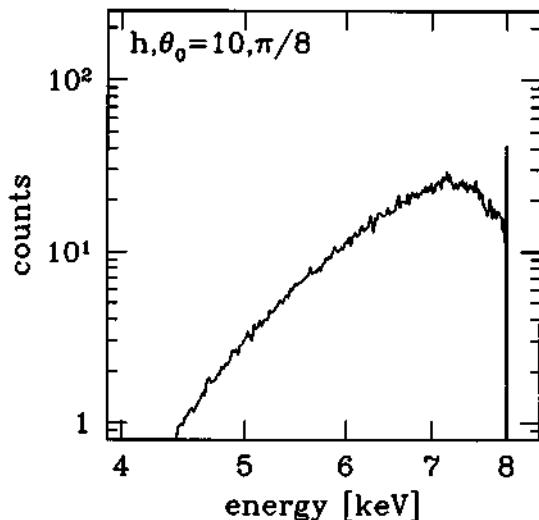


FIG. 9.—Shift of an 8 keV line to 7 keV in a cone with  $h = 10$  and  $\theta_0 = \pi/8$ . The photons were emitted at an optical depth of 2, and the ratio of scattering to absorption is 10. The shift of the line peak is due to the combined effect of starting at finite optical depth, which depresses the blue side, and absorption of the tail, which depresses the red side.

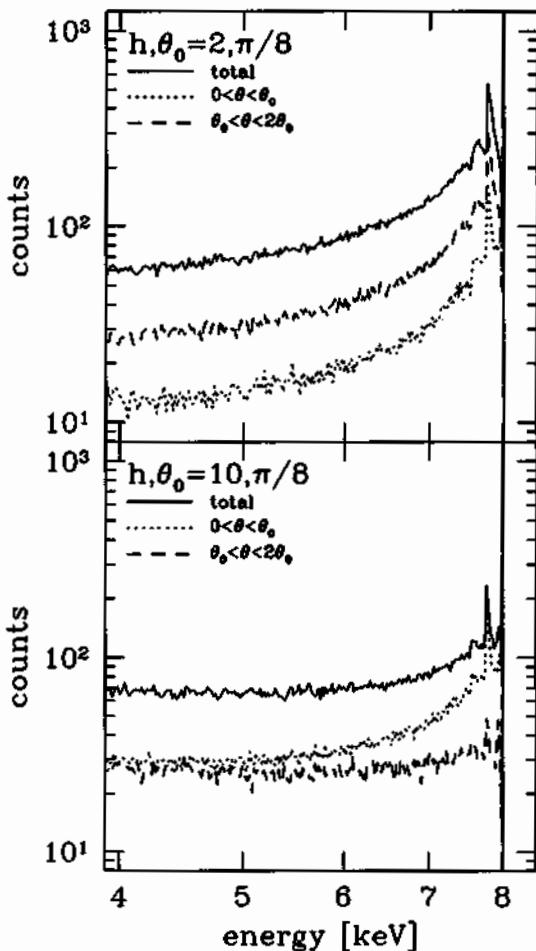


FIG. 10.—Variation in intensity and shape of the line profile with emission angle (or the observer viewing angle relative to the cone axis) for an opening angle of  $\pi/8$  and cone heights of 2 (top) and 10 (bottom). Top curve (solid in both panels): sum over all emission angles. Dashed curve: photons emitted between angles between  $\theta_0$  and  $2\theta_0$ . Dotted curve: photons emitted between 0 and  $\theta_0$ . Note how the intensities of the on-axis and off-axis components are reversed between the shallow cone (top) and the deep cone (bottom).

ensures that no photons scatter many times. Therefore, narrowness of the line core due to wide funnels can be distinguished from narrowness due to absorption by inspecting the red wing of the line.

A most interesting feature of the profiles is the structure near the original line energy: the escape peak, a sign of unscattered flux, and two or three bumps near it that represent the 1–3 times scattered photons. The bumps are very clear signs of a Compton scattering origin of the line and can only be seen with grating resolution. Unfortunately, their presence is not guaranteed, as can be seen in Figure 8. Rees & Mészáros (2000) showed that close to the compact object, the required X-ray flux is so large that the ionized surface layer of the funnel becomes optically thick to Thomson scattering (since the atomic edge cross section for

nickel or iron is more than 100 times the Thomson cross section, this requires extreme ionizing fluxes). This means that a typical line photon is not created at the wall surface, but at a few scattering optical depths into the wall. As a result, the number of photons that escape with few scatterings becomes minimal very quickly, removing the typical signatures of Compton scattering. (Note that especially for narrow funnels, the escape direction from the cone is nearly perpendicular to the wall normal, so even a small optical depth into the wall reduces the direct-escape probability to nearly zero.)

Another way in which the line shape can be affected is by the angle of exit. In particular, the flux that escapes the funnel from near the top, at an angle greater than the cone opening angle,  $\theta_0$ , has a markedly different spectrum than the total flux; the direct-escape spike and the backscatter spike (i.e., the red spike of once-scattered photons plus the peak of the twice-scattered ones) are markedly smaller (Fig. 10).

## 6. CONCLUSION

We have shown that the funnel geometry associated with a hypernova naturally produces some features of the emission lines observed in afterglows. In particular, the large width of the line seen in GRB 991216 is a natural result of the large scattering wings induced by repeated Compton recoils as a photon bounces around many times in the funnel prior to escape. Therefore, it is not necessary to invoke very high velocities of dense gas clumps or shells to explain the wide lines. At the same time, this causes the center energy of the line to be lowered by large amounts, depending on the funnel depth and opening angle. Widths of 0.5–1 keV, as observed, can be obtained reasonably well. However, for significantly greater widths the line becomes unrecognizable.

The far red wing of the line provides a diagnostic of the relative importance of scattering and absorption opacities in the funnel wall and thus of the temperature of the wall upon which this ratio chiefly depends. Furthermore, at grating resolutions one can recognize sharp spikes in the line due to the photons that have scattered 0–2 times prior to escape. The magnitude of these lines is a measure of the escape probability from the bottom of the funnel (a combination of the opening angle and the depth).

The shape of the lines is further affected by the composition of the ejecta. Contrary to most earlier papers on the subject, the ejecta cannot predominantly emit iron lines unless they are older than half a year. Any younger ejecta are dominated by nickel and/or cobalt. The only other time at which the ejecta are strongly dominated by only one iron-group element is the first 1–2 days, but then the dominant element is nickel, of which the K $\alpha$  line is at 8 keV. The discrepancy between this and the observed line energy of 7 keV can be explained by Compton downscattering.

In short, we feel that there is evidence in the lines of scattering origin in a hypernova funnel. Future high-resolution spectra (with long exposure times) will easily confirm or deny this proposition. If confirmed, the shape of the lines provides direct diagnostics of the properties of the hypernova remnants in  $\gamma$ -ray bursts.

G. C. M. and G. E. B. are supported by the US Department of Energy under grant DE-FG02-88ER40388. R. A. M. J. W. is supported in part by NASA (award 21098).

## REFERENCES

- Amati, L., et al. 2000, *Science*, 290, 953  
 Antonelli, L. A., et al. 2000, *ApJ*, 545, L39  
 Bloom, J. S., et al. 1999a, *Nature*, 401, 453  
 Bloom, J. S., Kulkarni, S. R., & Djorgovski, S. G. 2002, *AJ*, in press (astro-ph/0010176)  
 Bloom, J. S., Sigurdsson, S., & Polk, O. R. 1999b, *MNRAS*, 305, 763  
 Böttcher, M. 2000, *ApJ*, 539, 102  
 Bulik, T., Belczynski, K., & Zbijewski, W. 1999, *A&AS*, 138, 483  
 Costa, E., et al. 1997, *Nature*, 387, 783  
 Eichler, D., Livio, M., Piran, T., & Schramm, D. N. 1989, *Nature*, 340, 126  
 Fraaij, D. A., Kulkarni, S. R., Nicastro, L., Feroci, M., & Taylor, G. B. 1997,  
*Nature*, 389, 261  
 Fruchter, A. S. 2002, in Proc. Jan van Paradijs Memorial Symp. (2001  
 June), From X-Ray Binaries to Gamma-Ray Bursts, ed. E. P. J. van den  
 Heuvel, L., Kaper, & E. Kol (San Francisco: ASP), in press  
 Galama, T. J., et al. 1998, *Nature*, 395, 670  
 ———. 2000, *ApJ*, 536, 185  
 Galama, T. J., & Wijers, R. A. M. J. 2001, *ApJ*, 549, L209  
 Illarionov, A., Kallman, T., McCray, R., & Koos, R. 1979, *ApJ*, 228, 279  
 Janka, H.-Th., Eberl, T., Ruffert, M., & Fryer, C. 1999, *ApJ*, 527, L39  
 Kommers, J. M., Lewin, W. H. G., Kouveliotou, C., van Paradijs, J.,  
 Pendleton, G. N., Meegan, C. A., & Fishman, G. J. 2000, *ApJ*, 533,  
 696  
 Kouveliotou, C., Meegan, C. A., Fishman, G. J., Bhat, N. P., Briggs, M. S.,  
 Koshut, T. M., Paciesas, W. S., & Pendleton, G. N. 1993, *ApJ*, 413,  
 L101  
 Krumholz, M. R., Thorsett, S. E., & Harrison, F. A. 1998, *ApJ*, 506, L81  
 Mészáros, P., & Rees, M. J. 1997, *ApJ*, 476, 232  
 Metzger, M. R., Djorgovski, S. G., Kulkarni, S. R., Steidel, C. C.,  
 Adelberger, K. L., Fraaij, D. A., Costa, E., & Frontera, F. 1997, *Nature*,  
 387, 879  
 Mochkovitch, R., Hernanz, M., Isern, J., & Martin, X. 1993, *Nature*, 361,  
 236  
 Paczyński, B. 1998, *ApJ*, 494, L45  
 Piran, T. 1999, *Phys. Rep.*, 314, 575  
 Piro, L., et al. 1999, *ApJ*, 514, L73  
 ———. 2000, *Science*, 290, 955  
 Rees, M. J., & Mészáros, P. 1992, *MNRAS*, 258, 41P  
 ———. 2000, *ApJ*, 545, L73  
 Reichart, D. E. 1999, *ApJ*, 521, L111  
 Salmonson, J. D., Wilson, J. R., & Mathews, G. J. 2001, *ApJ*, 553, 471  
 Schmidt, M. 2001, *ApJ*, 552, 36  
 Totani, T. 1997, *ApJ*, 486, L71  
 van Paradijs, J., et al. 1997, *Nature*, 386, 686  
 van Paradijs, J., Kouveliotou, C., & Wijers, R. A. M. J. 2000, *ARA&A*, 38,  
 379  
 Vietri, M., Ghisellini, G., Lazzati, D., Fiore, F., & Stella, L. 2001, *ApJ*, 550,  
 L43  
 Wijers, R. A. M. J., Bloom, J. S., Bagla, J. S., & Natarajan, P. 1998,  
*MNRAS*, 294, L13  
 Wijers, R. A. M. J., Rees, M. J., & Mészáros, P. 1997, *MNRAS*, 288, L51  
 Woosley, S. E. 1993, *ApJ*, 405, 273  
 Woosley, S. E., & Weaver, T. A. 1995, *ApJS*, 101, 181

This page is intentionally left blank

---

## Chapter 23

# Discovery of a Black Hole Mass-Period Correlation in Soft X-Ray Transients and Its Implication for Gamma-Ray Burst and Hypernova Mechanisms

C.-H. Lee, G.E. Brown and R.A.M.J. Wijers

Reprinted with permission from *Astrophysical Journal*, 575 (2002), 996-1006

Copyright (2002) by American Astronomical Society

### **Commentary**

Stan Woosley's (1993) model of a Collapsar in which a Wolf-Rayet rotates sufficiently rapidly so that the outer part of the star is centrifugally supported while the middle falls into a black hole is our starting point for modeling GRBs (gamma-ray bursts). MacFadyen & Woosley (1999) added magnetohydrodynamic effects. In Paper 17 we showed how these combined so that the GRB is powered by open magnetic field lines from the rapidly rotating black hole and the hypernova explosion in the disk powered through the closed field lines which thread the black hole and couple it to the disk. Our idea in the present paper is to show that some of the outer matter in the He star, the center of which is falling into a black hole, is supported during this infall by centrifugal force. We suggest that this support lasts for the viscous time  $\sim 100$  sec, during which time angular momentum and energy supplied to the inner disk through the closed magnetic field lines from the black hole reach the supported matter. The matter is then ejected.

We evolved the black holes in a binary system so that the massive star was first brought into sufficient rotation, as we describe in Paper 23. Spruit & Phinney (1998) showed that the viscosity from magnetic braking brought a single Wolf-Rayet into nearly uniform rotation, especially if it had evolved through the red giant. If the He core is in uniform rotation it will collapse inwards without forming an accretion disk (Mineshige et al. 1997). A poster by Alexander Heger and Stan Woosley at the 2001 AAS meeting says that "if we include current approximation of angular momentum transport by magnetic fields, the resulting spin rates (of bare He cores) become too low to form centrifugally supported disks in the inner part of the core".

In the evolution of the black hole transient binaries, in the removal of the H envelope in common envelope evolution the outer part of the He core, which is left, is brought into

corotation with the companion — at least this is what we surmise from an extrapolation of the results of Rasio & Livio (1996). Furthermore, tidal interactions in the He-star, companion system will be strong, because of the close proximity of the stars, and further tend to isochronize them. This has been verified in the Heger & Woosley (2001) calculation. In any case, nearly all of the angular momentum initially resides in the outer part of the He star, which should be well isochronized with the companion. With the assumption in the paper that angular momentum is conserved, we need only know the initial angular momentum of the He core. (This assumption should be good because relatively little angular momentum is advected off the outer part of the He core.)

In addition to the difficulty in having enough angular momentum in single Wolf-Rayets, there is the problem of mass loss. In Paper 21 we showed that a “naked” He star would blow away sufficiently that a high-mass black hole could not be formed. Common envelope evolution can, therefore, only be initiated after He core burning has been completed (Case C mass transfer).

For the above reasons we evaluated GRBs from the massive star in a binary in Paper 17. In Paper 23 we show that there is an empirical relation between black-hole mass and orbital period in a binary. This empirical relationship is already evident in terms of the present orbital periods, but much of the increase in black hole mass with period results from mass transfer widening the orbit following the explosion and the much more meaningful relationship should be between the binary periods preceding the explosion in which the black hole is born and the black hole mass produced then. Our problem then, which is partially resolved in Paper 23, is to work back from the present orbits to the preexplosion ones. This involves undoing the effects of stable mass transfer, which have widened the orbit.

### *Preexplosion Properties of Soft X-Ray Transients*

Far and away the most interesting binary to date is Nova Scorpii. The large amount of mass delivered onto the companion in the explosion and the high system velocity indicate a large mass loss in the explosion from the He star, nearly half of the system mass, as discussed in Paper 17. In that work we were still an order of magnitude low on the mass deposited on the companion, suggesting that the asymmetry in the hypernova explosion would increase this amount. Our view of Nova Scorpii has been substantially changed by the work of Beer & Podsiadlowski (2002), who show that at the time of explosion the F-star companion was much closer, only  $\sim 6R_\odot$  away, from the He star and that it first widened its orbit in the explosion and then evolved to its present  $\sim 15R_\odot$  separation only later through conservative stable mass transfer from the donor to the black hole. Obviously, the initial closeness of the companion to the He star aids both in isochronism and in deposition of a large amount of material from the He star onto the companion in the explosion.

The short preexplosion period of 0.4 days implies sufficient angular momentum that  $\sim 2M_\odot$  of the outer material of the He star is supported centrifugally. Because of the

extreme mass loss  $\sim 5\text{--}6M_{\odot}$ , nearly half of the system mass, we would need here help from the magnetohydrodynamic effects to arrive before the dynamical time is over; i.e., while the GRB central engine is still operating. This help must arrive in  $\sim 5$  sec, in order to produce the greater mass loss. From the preliminary results of Orosz (2002), the measured  $^{24}\text{Mg}$  abundance indicates that there was substantial mass loss, but probably somewhat less than in Nova Scorpii, in IL Lupi (4U 1543-47). Some GRB central engines operate for only a few seconds, so it is quite possible for them to supply magnetohydrodynamic energy to the disk in that time. Knowing only the present period of the binary, we must have additional information from the metal anomalies in the companion from which we can estimate the mass loss if we are to calculate the preexplosion period of the binary.

The 14 SXTs (soft X-ray transient sources) in Table 1 of Paper 23 separate nearly equally into those with unevolved main sequence and those with evolved companions. This immediately tells us that there are many unobserved black hole binaries, since the lifetime for the evolved companions is nearly two orders of magnitude shorter than that for the main sequence ones. The former are the “silent partners” referred to in Paper 14.

The binaries with unevolved main sequence companions have much of their history erased by magnetic braking and gravitational waves which take away angular momentum and narrow the orbit. We argue in Paper 23 that these binaries initially have companions of mass  $M_d \lesssim 2.5M_{\odot}$ . With final masses of  $M_d \simeq 0.5M_{\odot}$ , this means that up to  $\sim 2M_{\odot}$  could have accreted onto the black holes. We believe that this explains why some of the black holes seem to be somewhat more massive than the  $\sim 7M_{\odot}$  that would come from the most copious ZAMS  $\sim 20M_{\odot}$  stars. The point that puzzles us most is that all of the He star masses which are progenitors for black holes in the binaries with evolved companions come from ZAMS masses in the upper part of our assumed  $20\text{--}30M_{\odot}$  interval.

We can reconstruct in more detail the preexplosion orbits of the binaries with evolved companions, also Nova Scorpii and IL Lupi, which are near the end of their main sequence lifetime. In our reconstruction of the Nova Scorpii initial period we find that the period has been increased  $\sim 75\%$  by  $0.46M_{\odot}$  mass transfer. It is best to begin with V4641 Sgr as “fiducial” binary, where the same amount of mass transfer would change the period only  $\sim 12\%$ , because of the much smaller (percentage) difference in black hole and companion masses. Thus, the present separation cannot be far from that at preexplosion.

As outlined in Paper 14, higher mass donors can supply the drop in gravitational binding energy to expel the H-envelope in the giant black-hole progenitor already at a larger final separation  $a_f$  than if the donor mass is less. Indeed, we find that  $a_f$  is roughly linear in  $M_d$ ,  $a_f \propto M_d$ . Since, as outlined in Papers 20 and 23, the pre-common-envelope separation is constrained\* to a very narrow interval about  $a_i \sim 1500R_{\odot}$ , this means that the linear scaling of  $a_f$  with  $M_d$  should work pretty well, a small spread coming because of the possible

\*This constraint results from the very narrow interval in radius available for Case C mass transfer. In fact, if the mass loss is decreased to half that we used in Paper 21, then high-mass black holes can be formed by fallback (Fryer *et al.* 2001). However, in our evolution of GRBs through binaries, the MHD help from the black hole could prevent this accretion.

$\sim 20\text{--}30M_{\odot}$  band in the initial giant ZAMS masses, the spread going as ZAMS masses to the 0.55 power. Thus a  $5.33R_{\odot}$  preexplosion separation in Nova Scorpii with ZAMS  $1.91M_{\odot}$  companion would imply an  $\sim 18R_{\odot}$  preexplosion separation in V4641 Sgr, not much less than its present  $21R_{\odot}$ . It could have widened the remaining  $3R_{\odot}$  by a small amount of mass loss in the explosion and by accretion from the companion B-star onto the black hole. The latter cannot be greater than  $\sim 1M_{\odot}$ . (We identify the preexplosion radius with the  $a_f$  of common envelope evolution, because the He star lifetime is too short for substantial changes in separation.) Since the mass loss in the explosion is given by

$$a_{\text{postexpl.}} = a_{\text{preexpl.}} \left( 1 + \frac{\Delta M}{M_{\text{syst,f}}} \right),$$

we immediately see that the mass loss  $\Delta M$  in the explosion in V4641 Sgr has been less than 16% of the final system mass  $M_{\text{syst,f}}$ , or  $< 14\%$  of the initial system mass, since the initial companion mass cannot have been greater than the initial black hole mass.

In Paper 23 we show that the binary 1915+105, in which the black hole mass is  $14.4M_{\odot}$ , would have looked like V4641 Sgr at an earlier time,  $\sim 4.6M_{\odot}$  having accreted from the donor onto the black hole in the meantime.

At least roughly, we see that part of the most massive black holes were formed by accretion, consistent with most of the original black holes coming from the more copious ZAMS stars in the lower part of the  $20\text{--}30M_{\odot}$  ZAMS region. This seems to be especially true for the AMLs.

As is clear from the Acknowledgments in Paper 23, we added the evolution of GRS 1915+105, as following after a V4641, only after the paper had first been submitted. These binaries, the first with the highest black hole mass and the second with the highest companion mass, are extremely interesting, although in ways not directly relevant for Paper 23, so we discuss them here.

We received a very interesting e-mail from Krzysztof Belczynski pointing out that to transfer  $4M_{\odot}$  from the  $6.5M_{\odot}$  giant to the black hole would require a rate of  $\sim 4 \times 10^{-5}M_{\odot}\text{ yr}^{-1}$ . He took the transfer time to be from the end of the Hertzsprung gap to the beginning of He core burning. If we assume the time scale for the transfer,  $\sim 4 \times 10^5$  yr for a  $6.5M_{\odot}$  giant (Schaller *et al.* 1992), we would need a transfer rate of  $1.2 \times 10^{-5}\text{ yr}^{-1}$ , roughly 250 times Eddington for a  $10M_{\odot}$  black hole. Krzysztof could not accept such a high rate. We agree that the rate is extremely high.

We would make the point, however, that hypercritical inflow, in which the photons are carried in with the adiabatic inflow, is much easier in black holes where the matter can easily go over the event horizon, than onto neutron stars, where accretion is onto the surface. The trapping radius for hypercritical accretion was derived in Paper 5, Eq. (2.10):

$$r_{\text{tr}} \simeq 0.6R_s \dot{m},$$

where  $R_s$  is the Schwarzschild radius and  $\dot{m} = \dot{M}/\dot{M}_{\text{Edd}}$ . Although this looks quite small, for accretion onto the neutron star  $r_{\text{tr}}$  must be taken to be much larger. That is, as described

in Paper 18 in some detail, the accreting matter is initially of too low temperature to cool by neutrinos, and it simply accumulates on the neutron star, forming an accretion shock, as described in the Commentary to Paper 18. The accretion shock has to pile up a lot of  $\gamma = 4/3$  matter so that the central region gets hot enough to emit neutrino pairs. Now the incoming matter can only be trapped outside of the accretion shock so that  $r_{\text{tr}} \gtrsim r_{\text{sh}}$ , where (Eq. (2.23) of Paper 5)

$$r_{\text{sh}} \simeq 2.6 \times 10^8 \text{ cm} (\dot{M}/M_{\odot} \text{ yr}^{-1})^{-0.37}.$$

It is this shock radius, which is relatively large compared with the dimensions of the neutron star, that makes the minimum rate for hypercritical accretion as large as

$$\dot{m} \simeq 10^4 \quad \text{for a neutron star.}$$

In the case of the black hole accretor the matter can flow adiabatically across the event horizon, once its angular momentum has been brought down as described above by the accretion disk.

Hypocritical accretion in an accretion disk is somewhat different from the spherical situation and we may have to be more careful than in Paper 18, where the accretion rate of  $\sim 1 M_{\odot} \text{ yr}^{-1}$  meant  $\dot{M} \sim 10^8 \dot{M}_{\text{Edd}}$  and the photons were clearly trapped and carried in with the adiabatic inflow. Here we consider  $100-1000 \dot{M}_{\text{Edd}}$ . With an accretion disk present, the photons will be trapped when the time to carry them into the center of the accretion disk is less than the time for the photons to diffuse out nearly spherically, since the disk will be both geometrically and optically thick, with nearly equal  $H$  and  $R$ .

In Paper 18 we found that the inward radial velocity in the disk was

$$|v_r| = (3/7)\alpha v_K,$$

where  $v_K$  was the Keplerian velocity  $\sqrt{GM/r}$ , a factor of  $\sqrt{2}$  less than the free fall velocity assumed for the photons. The viscous disk time is, thus,

$$\tau_{\text{visc}} = \frac{r}{|v_r|}.$$

For the electron diffusion the optical depth is

$$\tau_{es} = \int_r^{\infty} \rho \kappa_{es} dr = \frac{2R\dot{m}}{\sqrt{r}r_s},$$

where  $r_s = 2GM/c^2$  and  $\dot{M}_{\text{Edd}} = 4\pi c R / \kappa_{es}$ . This gives, in random walk, a dynamical time for diffusion of

$$\tau_{\text{diff}} \simeq \frac{h}{c} \frac{h}{\lambda_{es}} = \frac{h^2}{c} \kappa_{es} \rho.$$

We obtain the trapping radius by setting

$$\tau_{\text{visc}} = \tau_{\text{diff}},$$

where we have used the random walk diffusion time, and we find the trapping radius

$$r_{\text{tr}} \simeq 0.5\dot{m}R.$$

This is the result for a thick accretion disk found by Begelman & Meier (1982), who say, "It seems to be generally agreed that radiation pressure exerted through Thomson scattering cannot halt spherical accretion into a black hole," and give references. Once  $\dot{m}$  is sufficiently large we would expect hypercritical accretion to set in and the disk to become thick. We return to this point below. This should be compared with  $r_{\text{tr}} = 2R\dot{m}$  for the spherical situation in Eq. (3.8) of Paper 3, which was lowered to  $r_{\text{tr}} = 0.6R_s\dot{m}$ , as noted earlier in Paper 5. The decrease came from replacing random walk by a solution of the diffusion equation.

We note, however, that with the very much higher  $\dot{m} \sim 10^{14}$  of MacFadyen & Woosley (1999) the angular momentum of the matter cannot be removed quickly enough for accretion of the matter directly into the black hole, as in Paper 18 for the neutron star, and matter piles up until the inner accretion disk, which becomes a thin disk, is able to cool by neutrino pair emission.

In Papers 5 and 12 we pointed out that in common envelope evolution a neutron star can accrete  $\sim 1M_\odot \text{ yr}^{-1}$ , because in hypercritical accretion the photons are swept in by the adiabatic inflow. However, for hypercritical accretion a lower limit of  $\dot{M} = 10^4 M_{\text{Edd}}$  is needed. This results from the pressure of the radiation, as does the usual limit of  $M_{\text{Edd}}$ . However, if there is no surface onto which the material accretes and radiates away its binding energy but just the event horizon as on a black hole, this lower limit need not be applicable. The matter can simply be swept over the event horizon in the adiabatic inflow.

The matter must first be "processed", i.e., have most of its angular momentum removed, before it can go into the black hole. This "processing" should be carried out by the accretion disk. One might think that the luminosity from the disk should not exceed Eddington; otherwise the pressure might remove the outer parts of the disk. (As we note below, because the disk is porous, a thin disk can accrete up to at least  $\sim 10M_{\text{Edd}}$  (Begelman 2002).) In Grimm *et al.* (2001) the average luminosity for 1915+105 is that of a hydrogen accretion rate of  $\sim 3 \times 10^{-8} M_\odot \text{ yr}^{-1}$  onto a neutron star, whereas that for V4641 Sgr is  $\sim 100$  times less. In both cases there are jets, indicating super-Eddington accretion at times. We are out of the cone of the jet in 1915+105, so the total luminosity may be substantially super-Eddington.

We interpret the jets to result from the accretion onto the accretion disk being so great at times that the accretion disk cannot accept it all. That 1915+105 is much more active than V4641 Sgr is quantified by the  $\sim 100$  times greater average luminosity. However, at high accretion rates we would expect the disk to be optically thick and that the photons would be carried by the matter over the event horizon, so there may be some limit to the disk luminosity, with increasing accretion.

As Greiner *et al.* wrote (2001), the black hole in 1915 may be spinning rapidly. In Paper 17 we show that generally  $\sim 10^{53}$  ergs is available in the spin energy of the black

hole, but that only  $\sim 10^{52}$  ergs is used in the GRB and hypernova explosion, mostly in the latter.

We believe that in the short time  $\sim 5\text{--}10$  sec central engines enough energy will be transferred from the rapidly rotating black hole to the accretion disk to power the hypernova explosion. The latter dismantles the accretion disk so that no further rotational energy can be transferred, leaving the black hole in rapid rotation.

An interesting byproduct of our black hole evolution in binaries is their possible and promising use to explain ultraluminous X-ray sources (ULX) in external galaxies. These have been extensively observed by the Chandra X-ray observatory (see Makishima *et al.* 2000). We follow here the discussion of King *et al.* (2001). “*A key to understand their nature may be that they appear to occur preferentially, although not exclusively, in regions of star formation.*” The initial explanation of these ULXs was that they came from  $\sim 100M_{\odot}$  black holes; the disadvantage of this explanation was that such objects are not seen in our galaxy and no one knows how to evolve them.

King *et al.* (2001) suggest that the ULXs originate from the same sort of binaries we have evolved in this volume, the black hole soft X-ray transients. The high luminosity arises when they cross the Herzsprung gap, as GRS 1915+105 does at this time and as V4641 Sgr is beginning to do. The time scale for this is the thermal time scale

$$\tau_{\text{th}} = \frac{3 \times 10^7}{(M/M_{\odot})^2} \text{ yrs},$$

i.e.  $\sim 7 \times 10^5$  yrs for the  $6.5M_{\odot}$  companion in V4641 Sgr, which becomes something like GRS 1915+105 after crossing much of the Herzsprung gap according to our scenario in Paper 23.

The initial problem with the SXTs as the explanation was that luminosities of  $\sim 10$  times Eddington were needed. (Here we use Eddington accretion to mean that limit for a neutron star with  $\sim 20\%$  efficiency, midrange for the efficiencies of 6 to 42% for nonrotating and rapidly rotating black hole, respectively.) The initially proposed  $\sim 100M_{\odot}$  black holes were brought down to  $\sim 10M_{\odot}$  ones by the realization that accretion could proceed at  $\dot{M} \sim 10\dot{M}_{\text{Edd}}$  because of instabilities in the accretion disk which make the disk “porous” (Shaviv 1998, 2000; Begelman 2002). Since the Eddington limit goes linearly with the Schwarzschild radius  $R_s$  and  $R_s$  linearly with the mass,  $M_{BH}$ , a  $10M_{\odot}$  star accreting at  $10\dot{E}_{\text{Edd}}$  would look like a  $100M_{\odot}$  black hole accreting at Eddington.

Taking the Schaller *et al.* (1992) models, we have estimated the mass loss in going from V4641 Sgr to GRS 1915+105 in assumed conservative mass transfer as in Paper 23. The assumed  $5M_{\odot}$  companion goes out of equilibrium in the mass transfer, which should be calculated in an evolutionary code. We hope to get an estimate by simply assuming the mass going over the Roche Lobe to be transferred to the black hole, and readjusting the Roche Lobe according to the changed mass ratio. As noted earlier, we find the mass transfer to be at a rate of  $\dot{M} \sim 10^{-5}M_{\odot}$  per year, about 200 times the  $\sim 5 \times 10^{-8}M_{\odot} \text{ yr}^{-1}$  assumed for Eddington.

Belloni *et al.* (1997) find the emission from 1915+105 to be rather complicated in detail. They interpret the  $\sim 1$  sec appearance and disappearance of emission from an optically thick inner accretion disk as coming from fluctuations in the inner radius of the disk from  $\sim 20\text{--}80$  km, the lower radius during bursts and the upper during quiescence.

We cannot make a detailed model, but support that the accretion is hypercritical most of the time, as would be necessary to build up the black hole mass of 1915+105 from that V4641 Sgr over  $4\text{--}5 \times 10^5$  yrs.

King *et al.* (2001) emphasize the importance of beaming in building up the high apparent luminosities which are observed. Our evolution of the SXTs as fossil remnants of GRBs involves initially large effects of beaming in order to produce the gamma-ray bursts, and, as noted earlier, the binaries should be left in considerable rotation. In fact, jets are seen in 1915+105, as well as evidence of rotation.

Our evolution in Paper 23 where the separation  $a_f$  between donor (companion) and black hole following common envelope evolution is proportional to  $m_d$  (Eq. (5) of Paper 23), is helpful in understanding which binaries are good progenitors of ULXs. The “silent partners” of Paper 14 become SXTs only after they evolve, because only then can their outer matter cross the Roche Lobe. Their companions are the more “massive” ones and consequently pour more matter into the black hole.

We can obtain a limit on the companion mass necessary for a ULX from the evolution of Nova Scorpii by Beer & Podsiadlowski (2002). Their F-star companion was initially of  $2.5M_\odot$ , and began evolving in main sequence, with resulting loss of mass to the black hole. The calculation of Beer & Podsiadlowski is a fully evolutionary one. None the less we find that we can reproduce their results by conservative mass transfer, in which the period of the binary goes as the cube of the reduced mass. This is not surprising, in that the quiescent luminosity in Nova Scorpii is  $\sim 2.5 \times 10^{32}$  ergs, a factor of  $\sim 10^6$  less than Eddington (Garcia *et al.* 1997). During bursts the luminosity is higher,  $\sim 10^{38}$  ergs. What we learn from this is that rapid mass transfer — although less than Eddington — is going on silently between companion star and black hole while the observable effects are small, except during outbursts. We extend this scenario now to the Herzsprung gap, where both mass transfer and outbursts are much greater. The calculated rate of Beer & Podsiadlowski (2002) for the (future) crossing of the Herzsprung gap is  $\sim 4 \times 10^{-8} M_\odot \text{ yr}^{-1}$ , slightly below  $\dot{M}_{\text{Edd}}$  for a  $10M_\odot$  black hole. (Their black hole mass at the present time is  $5.3M_\odot$ .) According to Paper 23, the reason that Nova Scorpii (and IL Lupi) can begin transferring mass in main sequence is that they are kicked into relatively large orbits in the explosion which forms the black hole due to large mass losses. (Companions in both binaries show substantial metal abundances that would have come from the explosion.)

In general we found in Paper 23 that binaries with initial companion mass  $< 2.5M_\odot$  stay in main sequence. We interpret Nova Scorpii and IL Lupi as having companions in between the main sequence ones and those of the silent partners, as the companion mass increases. Thus, we estimate  $m_d \sim 3M_\odot$  as minimum mass for the “silent partner” companion.

A  $3M_{\odot}$  companion could be stripped down to its  $0.4M_{\odot}$  He core in transferring matter after expansion to the Herzsprung gap. Using the thermal time scale of  $3.3 \times 10^6$  yrs, this means a maximum average rate of mass transfer of  $0.9 \times 10^{-6} M_{\odot} \text{ yr}^{-1}$  or  $\sim 20\dot{M}_{\text{Edd}}$  for a  $10M_{\odot}$  black hole. Thus an actual rate of  $\dot{M} \sim 10\dot{M}_{\text{Edd}}$  is reasonable for such a star, and progenitor binaries with  $3M_{\odot}$  companion should give ULXs.

The highly super-Eddington (hypercritical) accretion with more massive companions, as in V4641 Sgr and 1915+105, is unlikely to produce higher luminosity, because the high density infalling matter can very efficiently trap the photons and carry them adiabatically inwards with the adiabatic inflow. We propose to divide the accretion disk into the outer thin part with  $\dot{m} \lesssim 10$  and the inner thick part with  $\dot{m} \gtrsim 10$ . (We basically extend  $\dot{M}_{\text{Edd}}$  to  $10\dot{M}_{\text{Edd}}$  to take into account the effects of inhomogeneities, assuming a tenfold increase.) Photons in the outer thin disk are emitted; those in the inner thick disk are trapped and carried inwards. The two regions are effectively separated by (approximately) stationary matter.

Given the complexities of accretion disks, this may seem to be an oversimplified approximation. Indeed, the complication induced by inhomogeneities in the thin disk has forced an increase in  $\dot{M}_{\text{Edd}}$  to an effective  $10\dot{M}_{\text{Edd}}$ , where the factor 10 is an estimate from Begelman (2002). The concept of photon trapping is, however, an exact translation of neutrino trapping in supernova explosions, which is discussed in Paper 1. There are, however, instabilities which lead to oscillations in the accretion, which we mention below. A major assumption is that the superposition of these on our above simple model does not change the average accretion rate.

Whereas SXTs such as 1915+105 may well produce many of the observed ULXs, this cannot be the whole story, since they are also observed in elliptical galaxies. King *et al.* (2001) also suggest unstable accretion during a thermal time scale of a radiative star onto a neutron star such as occurs in Her X-1, the companion being more massive than the neutron star, as a source. Whereas the companions may be sufficiently massive in spiral galaxies to produce the ULXs, it is not clear that they can do so in elliptical galaxies. We have not dealt with the relevant binaries in this volume.

### *Evolution of Cyg X-1*

We earlier estimated the  $a_f = a_{\text{preexpl.}}$  of V4641 Sgr to be  $\sim 18R_{\odot}$ . From Table 1 of Paper 23 we take  $M_d = 17.8M_{\odot}$  and  $M_{\text{BH}} = 10.1M_{\odot}$  for Cyg X-1, which gives a present  $P \simeq 40$  days. We needed the maximum giant ZAMS mass  $M = 30M_{\odot}$ , which we allowed ourselves in order to evolve V4641 Sgr, and we will also need the maximum mass to obtain the high-mass black hole in Cyg X-1. Using Eq. (5) of Paper 23 to scale from V4641 Sgr to Cyg X-1, we obtain

$$a_f = a_{\text{preexpl.}} = \left( \frac{17.8M_{\odot}}{6.53M_{\odot}} \right) 18R_{\odot} = 49R_{\odot},$$

somewhat larger than the  $40R_\odot$  obtained from the parameters in Table 1 of Paper 23.

In Paper 21 we suggested that the only way to evolve the binary Cyg X-1 with black hole mass of  $\gtrsim 10M_\odot$  was to begin with a black hole progenitor which was a WNL, covered with hydrogen during almost all of its core helium burning time. Trouble was encountered, however, in getting the binary separation in Cyg X-1 down to the observed  $\sim 40R_\odot$ . Some kind of common envelope evolution, which we could not accurately describe for the high-mass stars dealt with there, had to be invoked.

We believe now that Cyg X-1 can be evolved similarly to the other black hole transient sources, it being different because the massive ( $\sim 18M_\odot$ ) companion is more highly evolved, as well as more massive, than other companions, so that it shines continuously. (Roughly speaking, a certain fraction  $f$  of the Eddington limit is necessary for continuous shining.) The Herrero *et al.* (1995) values of masses in Table 1 of Paper 23 are somewhat smaller than earlier measurements, but on the other hand, one would expect some mass losses, especially from the donor, which would widen the orbit. Given uncertainties,  $49R_\odot$  is a ballpark estimate.

Furthermore, Cyg X-1 has been shown to have a substantial system velocity indicating mass loss in the explosion which produces the black hole (Kaper *et al.* 1999). The value of  $50 \text{ km s}^{-1}$  used by Nelemans *et al.* (1999) may be too high (Lex Kaper 2001) but in this the evolution of Cyg X-1 resembles that of the transient sources.

### *Construction of Schematic Model*

Under our assumption that angular momentum is conserved following formation of the rotating He core, we could argue from the angular momentum of the initial core in which the outer part was tidally synchronized, and did not need to carry out the subsequent differential rotation. This initial core was taken from Woosley's  $25M_\odot$  star immediately following He core burning, since the H envelope had to be removed in Case C mass transfer. From Eq. (9) of Paper 23 we see that matter in the original rotating core beyond

$$R_c \equiv \left( \frac{\hat{l}(\hat{a})GM_c}{c\Omega} \right)^{1/2}$$

is centrifugally supported. The  $\hat{l}(\hat{a})$  is the dimensionless specific angular momentum for Kerr parameter  $\hat{a}$ , and  $M_c$  is the total mass inside the cylinder of radius  $R_c$ .

The curve of  $\hat{l}(\hat{a})$  is given in the top part of Fig. 5 of Paper 17. For a maximally rotating black hole  $\hat{l}(\hat{a} = 1) = 2$ , whereas for the schematic model of this reference with  $\hat{a} = 0.8$ ,  $\hat{l}(\hat{a}) \approx 3$ .

We term the material between two adjacent cylinders as a "band" of material. We show in what follows that the material inside of  $R_c$  falls into the black hole in a dynamical time, which is estimated to be 5–10 sec. This is the dynamical time of some of the shortest,

most powerful GRBs, so it is not surprising that Nova Scorpii and IL Lupi are helped by magnetohydrodynamical effects. We have verified by numerical calculation that the dynamical time for the unsupported matter is changed only negligibly by taking into account its angular momentum. Thus, a reasonable approximation is a “sudden” one for the infall of the unsupported matter.

The band of centrifugally supported matter, with least support (because its initial  $r$  is given by  $R_c$ ), then finds itself with specific angular momentum  $\gtrsim \hat{l}$ , not far outside the black hole in space. It has been brought there by the differential rotation which went on between the uncovering of the He core and formation of the black hole. Closed field lines threading the black hole couple to the matter in the disk; in fact, they are frozen in it.

The first general discussion of how these field lines torque up the inner accretion disk was given by van Putten (1999), and a general review in van Putten (2001). Since the black hole rotates several times faster than the accretion disk, it supplies energy rapidly to torque the latter up. Within the Blandford-Znajek (1977) mechanism, a detailed discussion of this was given by Li-Xin Li (2000), who finds the efficiency for the black hole to torque up the accretion disk to be nearly double the efficiency with which the black hole delivers power into the GRB. He does not, however, use the force-free region of Blandford-Znajek, which we favor.

The outer part of the initial He core, which contains nearly all of the angular momentum, is radiative. It is satisfying that the radiative polytrope curves fit the black hole mass, period relation better than the convective ones (Fig. 11 of Paper 23).

The magnetic fields on the inner disk must be very strong, comparable with those in magnetars ( $\sim 10^{15}$  G) in order to power the GRBs with shorter central engine times ( $\sim 5$  sec) (Paper 17). With such strong fields perpendicular to the disk the  $\alpha$  in the  $\alpha$ -viscosity will be not much less than unity, but the inner disk is quite thin,  $h \sim 0.2 \times r$  (MacFadyen & Woosley 1999), so that the viscous time scale

$$\tau_{\text{vis}} \sim \frac{1}{\alpha} \left( \frac{r}{h} \right)^2 \tau_{\text{dyn}}$$

is much longer than the dynamical time scale. Thus the conversion of the rotational energy into heat is relatively slow. Further out in the disk,  $(r/h)^2$  is no longer large, but  $\alpha$  drops. Thus, the power (Li 2000)

$$P_{\text{H,D}} = T_{\text{H,D}} \Omega_D,$$

where  $T_{\text{H,D}}$  is the torque supplied by the more rapidly rotating black hole, goes chiefly into spinning the matter in the disk. The conversion of the increased kinetic energy into heat takes a longer time.

Only  $\sim 10\%$  of the  $\sim 10^{53}$  ergs available to be deposited in the disk through its being torqued up is needed for the hypernova explosion. Thus, the explosion, which needs  $\sim 10^{52}$  ergs, will take place once this amount of the energy is changed into heat. The explosion is that of a rotating thick plate, definitely aspherical.

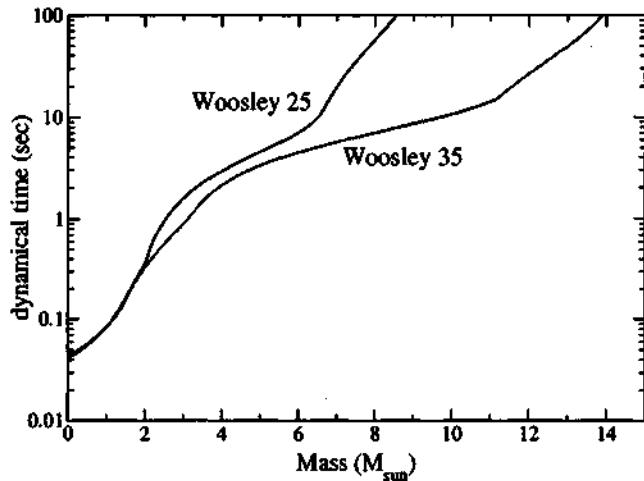


Figure. The dynamical times for Woosley's  $25M_{\odot}$  and  $35M_{\odot}$  stars during the collapse.

### Dynamical Time

The dynamical time for infall is often calculated for the collapse of a spherically symmetric mass distribution, of radius  $R_0$ . Then it is easily found that

$$\frac{dr}{dt} = - \sqrt{\frac{2GM}{r} - \frac{2GM}{R_0}}$$

or

$$\begin{aligned} t &= - \int_{R_0}^0 \frac{dr}{\sqrt{\frac{2GM}{r} - \frac{2GM}{R_0}}} \\ &= \frac{\pi}{2\sqrt{2}} \left( \frac{R_0^3}{2GM} \right)^{1/2} \\ &\approx 1766 \left( \frac{R}{R_{\odot}} \right)^{3/2} \left( \frac{M}{M_{\odot}} \right)^{-1/2} \text{ sec}, \end{aligned}$$

$\sim 10$  minutes for an  $8M_{\odot}$  He core.

Woosley's He cores, at the time of collapse, are centrally condensed, the center having burned to heavy metals. Therefore, the much denser central core drops into the black hole in seconds. We show results of the numerical calculation of this dynamical time in the figure.

Once the part of the core which is not centrifugally supported starts falling in, it is very hard to stop it. This is the reason why the black holes in the transient sources are chiefly composed of this part.

The gamma-ray burster will thus be run by the rapidly rotating black hole, which delivers some of its rotational energy up the open field lines. The energy carried by the closed magnetic field lines which thread the black hole and are frozen in the matter of the disk powers the hypernova explosion.

### Addendum

After this commentary was written there appeared a beautiful example of the GRB-hypernova association (Bloom *et al.* 2002). In this paper the light curve showed a bump, reddened by the many metal lines, rising on that from the GRB afterglow after several days. This bump was fit well by redshifting the known hypernova light curve from SN1998bw, previously interpreted as a hypernova explosion, to the relevant redshift (0.36). The GRB and accompanying hypernova have a common explanation in the model of Papers 17 and 23. Both are powered by the rotational energy of the black hole, the hypernova through the closed field lines coupling the black hole to the accretion disk.

In a review of the Woods Hole meeting “Gamma-Ray Burst and Afterglow Astronomy 2001” by Tsvi Piran (2002), it is pointed out that the actual GRB energy is narrowly distributed around a “mere”  $\sim 10^{51}$  ergs. Hypernova calculations which reproduce the spectra of 1998bw seem to need  $\sim 10^{52}$  ergs, ten times more (see the results in Fig. 4 of Paper 17 of Nomoto *et al.* (2000)).

Van Putten (2000, 2001) has listed many channels through which the rotational energy of the black hole can be dissipated.

We note that in the scenario we developed here, the black hole rotational energy will depend only on the binary period just before black hole formation, which, in turn, is determined chiefly by companion mass through Eq. (5) of Paper 23. However, the power, the rate at which the energy is delivered, is proportional to  $B^2$ , which will vary from star to star. If  $B^2$  is low, so that the central engine must work for a long time, then neutrino losses will be somewhat increased.

Finally, we note that an extensive numerical calculation, “Extraction of Black Hole Rotational Energy by a Magnetic Field and the Formation of Relativistic Jets”, by Koide *et al.* (2002), supports in a general way that the rotational energy can be extracted from black holes as in Blandford-Znajek, although the energy comes out in somewhat different forms.

### References

- (1) Beer, M.E., and Podsiadlowski, Ph. (2002), *Monthly Notices of the Royal Astronomical Society* **331**, 351.

- 
- (2) Begelman, M.C., and Meier, D. L. (1982), *Astrophysical Journal* **253**, 873.
  - (3) Begelman, M.C. (2002), *Astrophysical Journal* **568**, L97.
  - (4) Belloni, T., Mendez, M., King, A.R., van der Klis, M., and van Paradijs, J. (1997), *Astrophysical Journal Letters* **479**, L145.
  - (5) Blandford, R.D., and Znajek, R.L. (1977), *Monthly Notices of the Royal Astronomical Society* **179**, 433.
  - (6) Bloom, J.S., et al. (2002), *Astrophysical Journal* **572**, L45.
  - (7) Fryer, C.L., Heger, A., Langer, N., and Wellstein, S. (2001), astro-ph/0112539.
  - (8) Garcia, M.R., McClintock, J.E., Narayan, R., and Callanan, R.J. (1997), Proceedings of the 13th North American Workshop on CVs, eds. S. Howell, E. Kuulkers and C. Woodward (San Francisco: ASP), 506.
  - (9) Greiner, J., et al. (2001), *Nature* **414**, 522.
  - (10) Grimm, H.-J., Gilfanov, M., and Sunyaev, R. (2001), astro-ph/0109239.
  - (11) Heger, A., and Woosley, S.E. (2001), poster at the 2001 AAS meeting.
  - (12) Herrero, A., Kudritzki, R.P., Gabler, R., Vilchez, J.M., and Gabler, A. (1995), *Astronomy & Astrophysics* **297**, 556.
  - (13) Kaper, L. (2001), private communication.
  - (14) Kaper, L., Camerón, A., and Barziv, O. (1999), in *Wolf-Rayet Phenomena in Massive Stars and Starburst Galaxies*, IAU Symp. 193, eds. K.A. van der Hucht, G. Koenigsberger and R.J. Eenens, p. 316.
  - (15) King, A.R., Davies, M.B., Ward, M.J., Fabbiano, G., and Elvis, M. (2001), *Astrophysical Journal* **552**, L109.
  - (16) Koide, S., Shibata, K., Kudoh, T., and Meier, D.L. (2002), *Science* **295**, 1688.
  - (17) Li, L.X. (2000), *Astrophysical Journal* **533**, L115.
  - (18) MacFadyen, A.I., and Woosley, S.E. (1999), *Astrophysical Journal* **524**, 262.
  - (19) Makishima, K., et al. (2000), *Astrophysical Journal* **535**, 632.
  - (20) Mineshige, S., Nomura, H., Hirose, M., Nomoto, K., and Suzuki, T. (1997), *Astrophysical Journal* **489**, 227.
  - (21) Nelemans, G., Tauris, T.M., and van den Heuvel, E.P.J. (1999), *Astronomy & Astrophysics* **352**, L87.
  - (22) Orosz, J.A. (2002), private communication.
  - (23) Piran, T. (2002), *Science* **295**, 917.
  - (24) Rasio, F.A., and Livio M. (1996), *Astrophysical Journal* **471**, 366.
  - (25) Schaller, et al. (1992), *Astronomy & Astrophysics Suppl.* **96**, 269.
  - (26) Shaviv, N.J. (1998), *Astrophysical Journal* **494**, L193.
  - (27) Shaviv, N.J. (2000), *Astrophysical Journal* **532**, L137.
  - (28) Spruit, H., and Phinney, E.S. (1998), *Nature* **393**, 139.
  - (29) van Putten, M.H.P.M. (1999), *Science* **284**, 115.
  - (30) van Putten, M.H.P.M. (2000), *Astrophysical Journal* **555**, L41.
  - (31) van Putten, M.H.P.M. (2001), *Physics Reports* **345**, 1.
  - (32) Woosley, S.E. (1993), *Astrophysical Journal* **405**, 273.

## DISCOVERY OF A BLACK HOLE MASS-PERIOD CORRELATION IN SOFT X-RAY TRANSIENTS AND ITS IMPLICATION FOR GAMMA-RAY BURST AND HYPERNOVA MECHANISMS

C.-H. LEE,<sup>1,2,3</sup> G. E. BROWN,<sup>3</sup> AND R. A. M. J. WIJERS<sup>3,4</sup>

*Received 2001 October 17; accepted 2002 April 23*

### ABSTRACT

We investigate the soft X-ray transients with black hole primaries, which may have been the sources of gamma-ray bursts (GRBs) and hypernovae earlier in their evolution. For systems with evolved donors, we are able to reconstruct the pre-explosion periods and find that the black hole mass increases with the orbital period of the binary. This correlation can be understood in terms of angular momentum support in the helium star progenitor of the black hole, if the systems with shorter periods had more rapidly rotating primaries prior to their explosion; centrifugal support will then prevent more of its mass from collapsing into the black hole on a dynamical time. This trend of more rapidly rotating stars in closer binaries is usual in close binaries and in the present case can be understood in terms of spin-up during spiral-in and subsequent tidal coupling. We investigate the relation quantitatively and obtain reasonable agreement with the observed mass-period correlation. An important ingredient is the fact that the rapidly rotating new black hole powers both a GRB and the hypernova explosion of the remaining envelope, so that the material initially prevented from falling into the black hole will be expelled rather than accreted. For systems in which the donor is now and will remain in main sequence, we cannot reconstruct the pre-explosion period in detail, because some of their history has been erased by angular momentum loss through magnetic braking and gravitational waves. We can, however, show that their periods at the time of black hole formation were most likely 0.4–0.7 days, somewhat greater than their present periods. Furthermore, their black holes would have been expected to accrete  $\sim 1 M_{\odot}$  of material from the donor during their previous evolution. Comparison with predictions suggests that little mass will be lost in the explosion for the relatively high pre-explosion periods of these binaries. A natural consequence of the He star rotation is that black holes formed in the shorter period (before explosion) soft X-ray transients acquire significant Kerr parameters. This makes them good sources of power for GRBs and hypernovae, via the Blandford-Znajek mechanism, and thus supports our model for the origin of GRBs in soft X-ray transients.

*Subject headings:* accretion, accretion disks — binaries: close — black hole physics — gamma-rays: bursts

### 1. INTRODUCTION

Recent observations strongly suggest a connection between gamma-ray bursts (GRBs) and supernovae, with indications that the supernovae in question are especially energetic and of Type Ib/c, i.e., core collapses of massive stars that have lost their hydrogen envelope (see van Paradijs, Kouveliotou, & Wijers 2000 and references therein). This supports suggestions by Woosley (1993) and Paczynski (1998) for the origin of GRBs in stellar core collapses. The hydrodynamics of a jet escaping from a star and causing its explosion was explored in detail by MacFadyen & Woosley (1999), who showed that contrary to accepted wisdom, a fairly baryon-free, ultrarelativistic jet could plow through the collapsing star and emerge with large Lorentz factors. The powering of the outflow by coupling of high magnetic fields to the rotation of the black hole (Blandford & Znajek 1977), first suggested by Paczynski (1998) in the context of GRBs, was worked out in detail by van Putten

(1999, 2001). Li has also discussed the deposition of energy from a black hole into the accretion disk in a recent series of papers (e.g., Li 2000a, 2000b, 2000c, 2000d, 2002).

Building on these thoughts, we have modeled both the powering of a GRB by black hole rotation and the stellar evolution pathways that set up favorable conditions for that mechanism (Brown et al. 2000). An essential ingredient in this model is a rapidly rotating black hole, and it is this aspect that we focus on in the present paper. A single star initially in uniform rotation will tend to develop a differential rotation, because the core contracts strongly during evolution, and angular momentum conservation will therefore increase its angular velocity. However, given enough time, viscous stresses will even out these differences, and thus the net result is a loss of angular momentum of the innermost regions of the star. Spruit & Phinney (1998) argued that magnetic-field-mediated coupling is strong enough in single stars during the giant phase to make the cores very slow; so slow, in fact, that they required an asymmetric kick in the birth of pulsars to get their spin frequencies up to observed values. Livio & Pringle (1998) subsequently used observations of novae to argue for a weaker coupling, but still their coupling strength would lead to spin energies of black holes that are negligible as power sources for GRBs.

However, as suggested by MacFadyen & Woosley (1999), a massive star in a close binary will spin faster for a number of reasons: first, when the hydrogen envelope is lifted off by spiral-in, it will cease to serve as a sink of angular momentum for the core. Second, the tidal friction concomitant to

<sup>1</sup> School of Physics, Seoul National University, Seoul 151-747, Korea; chlee@kias.re.kr.

<sup>2</sup> School of Physics, Korea Institute for Advanced Study, Seoul 130-012, Korea.

<sup>3</sup> Department of Physics and Astronomy, State University of New York, Stony Brook, NY 11794; popeno@nuclear.physics.sunysb.edu; twijers@mail.astro.sunysb.edu.

<sup>4</sup> Current address: Astronomical Institute "Anton Pannekoek," University of Amsterdam, Kruislaan 403, 1098 SJ Amsterdam, The Netherlands; rwijers@science.uva.nl.

the spiral-in process will spin up the inner region, giving it a larger angular momentum than the same region in a single star (Rasio & Livio 1996). Third, tidal coupling in the close binary will tend to bring the primary into corotation with the orbital period. This latter process is not very efficient in the short post-spiral-in life of the binaries we consider, but its effect does probably matter to the outer layers of the helium star, which can be important for our work. With its more rapid rotation, the helium star then forms a black hole with a large Kerr parameter, which immediately after its formation (in a few seconds) begins to input power into its surroundings at a very high rate. This, then, powers both a GRB (e.g., Brown et al. 2000) and the expulsion of the material that was centrifugally prevented from falling into the black hole. In fact, van Putten (1999, 2001) estimates that the power input into that material exceeds that into the GRB, and Li (2000b) also finds that more energy can be extracted by the disk than by the GRB. It should be noted that an initially less rapidly rotating black hole could be spun up by disk accretion quite rapidly and start a similar process after some accretion has taken place (MacFadyen & Woosley 1999; Brown et al. 2000). Some implications of such more complicated sequences of events are discussed by Lee, Lee, & van Putten (2001).

In § 2 we present the data on known soft X-ray transients (SXTs), showing the relation between present orbital period and black hole mass. Since theory predicts a relation between pre-explosion orbital period and black hole mass, we then consider carefully the pre- and postexplosion evolution of the systems (§ 3) and use this to reconstruct the pre-explosion orbit for as many systems as possible (§ 4). Then we develop our model for the mass and spin of black holes in SXTs and use it to explain the mass-period correlation (§ 5). We summarize our conclusions in § 6.

## 2. AN EMPIRICAL CORRELATION BETWEEN ORBITAL PERIOD AND BLACK HOLE MASS

We have collected data from the literature on black hole binaries in our Galaxy. In Table 1, we collect data of those for which the mass function is known and some manner of mass estimate for both the black hole and the companion can be given. In Table 2, we list the properties of two key systems in more detail. In Figure 1, we show the masses of the black holes as a function of orbital period. While the ranges of black hole masses for main-sequence and evolved systems overlap, the latter tend to have higher masses; the exception is Nova Sco 1994, which we see later is a natural but rare case of the general evolution scenario that we describe in this paper. In Figure 2, we show the donor masses as a function of orbital period. They show a more obvious trend of more massive donors in evolved systems. As we see, this is a natural consequence of the fact that only evolved systems can come into Roche contact in wide binaries, and more massive donors are more likely to come into contact via nuclear evolution. (The various curves are explained in § 3.)

In the following sections, we argue that the correlation between black hole mass and period also has physical meaning: the shorter the orbital period, the more rapidly rotating the helium star progenitor to the black hole. Rapid rotation centrifugally prevents some fraction of the helium star from collapsing into a black hole, resulting in a smaller black hole mass. The correlation in Figure 1 is weak because evolution

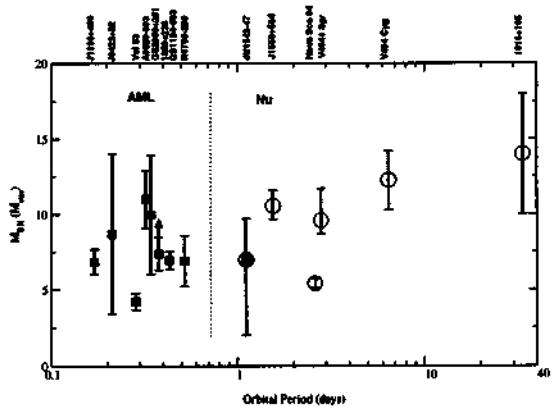


FIG. 1.—Black hole mass as a function of present orbital period of 14 SXTs. Note that the orbital period is on a logarithmic scale. SXTs with sub-giant or giant companions are indicated by big open circles (denoted as “Nu” for nuclear evolution). Filled squares indicate SXTs with main-sequence companions (denoted as “AML” for angular momentum loss). The vertical dotted line indicates the possible existence of different classes according to evolutionary path of the binary, as discussed in § 3. 4U 1543–47 is marked with both symbols, since we believe it to be right on the borderline between main-sequence and evolved; for the purpose of modeling, it can be treated as evolved.

of the binary since formation of the black hole has washed out the relation. Properly, we should consider the correlation between pre-explosion orbital period and postexplosion black hole mass. Much of our work presented here is concerned with understanding the evolution of these binaries and using this knowledge to find the systems for which we can reconstruct those parameters. Using that subset, we show much better agreement between our model predictions and the observed relation between reconstructed period and mass; this supports our evolutionary model and has ramifications for the origin of GRBs.

## 3. THE EVOLUTION OF SOFT X-RAY TRANSIENTS

### 3.1. Prior to the Formation of the Black Hole

Following on the work of Brown, Weingartner, & Wijers (1996), who showed the importance of mass loss of helium stars in binaries in determining the final outcome of binary evolution, Brown, Lee, & Bethe (1999), Wellstein & Langer (1999), and Brown et al. (2001a) showed that massive helium stars could evolve into high-mass black holes only if they were covered with hydrogen during most of their helium core burning era (case C mass transfer in binaries). In cases A or B mass transfer in binaries (Roche lobe overflow in main-sequence or red giant stage), the Fe core that was left was too low in mass to go into a high-mass black hole. Brown et al. (2001a) showed that high-mass black holes could be formed only if the mass was taken off the black hole progenitor after helium core burning was finished; i.e., case C mass transfer. Brown, Lee, & Tauris (2001c) showed that with the Schaller et al. evolution, this could happen only in the neighborhood of zero-age main sequence (ZAMS) mass  $20 M_{\odot}$ , definitely not at  $25 M_{\odot}$  and higher; because of the wind losses, the mass transfer would begin as Roche lobe overflow only in case B mass transfer,

TABLE I  
PARAMETERS OF BLACK HOLE BINARIES IN OUR GALAXY WITH MEASURED MASS FUNCTIONS

X-Ray Name	Companion Type	$P_{\text{orb}}$ (day)	$K_{\text{opt}}$ ( $\text{km s}^{-1}$ )	$f(M_{\text{bh}})$ ( $M_{\odot}$ )	$i$ (deg)	$M_{\text{BH}}$ ( $M_{\odot}$ )	$M_{\text{opt}}$ ( $M_{\odot}$ )	$d$ (kpc)	Other Name(s)	References
XTE J1118+480	K7 V-M0 V	0.169930(4)	701(10)	6.1(3)	81(2)	6.0–7.7	0.09–0.5	1.9(4)	KV Ursae Majoris	1, 2
XN Per 92	M0 V	0.2127(7)	380.6(65)	1.15–1.27	28–45	3.4–14.0	0.10–0.97	...	GRO J0422+32, V518 Persei	3
XN Vel 93	K6–M0	0.2852	475.4(59)	3.05–3.29	~78	3.64–4.74	0.50–0.65	...	MM Velorum	4
XN Mon 75	K4 V	0.3230	433(3)	2.83–2.99	40.75(300)	11.0(19)	0.68(18)	1.164(114)	A0620–003, V616 Monocerotis, N Mon 1917	3, 5
XN Vul 88	K5 V	0.3441	520(16)	5.01(12)	47–75	6.04–13.9	0.26–0.59	2	GS 2000+251, QZ Vulpeculae	3, 6
XTE 1859+226	...	0.380(3)	570(27)	7.4(11)	...	...	...	...	V406 Vulpeculae	7
XN Mus 91	K4 V	0.4326	406(7)	2.86–3.16	54.0(15)	6.95(6)	0.56–0.90	3.0	GS 1124–683, GU Muscae	3, 8
XN Oph 77	K3 V	0.5213	420(30)	4.44–4.86	60–80	5.2–8.6	0.3–0.6	5.5	H1705–250, V2107 Ophiuchi	3
4U 1543–47	A2 V	1.1164	129.6(18)	0.252(11)	~22	2.0–9.7	1.3–2.6	9.1(11)	MX 1543–475, JL Lupi	9, 10
XTE J1550–564	G8 IV–K4 III	1.552(10)	349(12)	6.86(71)	70.8–75.4	$10.56^{+1.02}_{-0.38}$	$1.31^{+0.33}_{-0.37}$	4.7–5.9 (?)	V381 Normae	11
XN Sco 94	F6 III	2.6127(8)	227(2)	2.64–2.82	67–71	5.1–5.7	1.1–1.8	3.2	GRO J1655–40, V1033 Scorpii	3, 12
V4641 Sagittarii	B9 III	2.817	211.0(31)	2.74(12)	...	$9.61^{+2.08}_{-0.88}$	$6.53^{+1.6}_{-1.05}$	$9.59^{+2.72}_{-2.19}$	XTE J1819–254, SAX J1819.3–2525	13
Cyg X-1	O9.7 Iab	5.5996	74.7(10)	0.25(1)	...	~10.1	~17.8	2.5	1956+350, V1357 Cyg, HDE 226868	14
XN Cyg 89	K0 IV	6.4714	208.5(7)	6.02–6.12	52–60	10.3–14.2	0.57–0.92	2.2–3.7	GS 2023+338, V404 Cygni, N Cyg 1938, 1959	3, 15, 16
GRS 1915+105	K–M III	33.5(15)	140(15)	9.5(30)	70(2)	14(4)	1.2(2)	12.1(8)	V1487 Aquilae	17

Note.—Binaries are listed in order of increasing orbital period. All systems except Cyg X-1 (steady X-ray source) are SXTs. XN indicates X-ray nova. Earlier observations (Greene, Bailyn, & Orosz 2001) gave the black hole mass in Nova Sco as  $6.3 \pm 0.5 M_{\odot}$ . New analyses of the light curve by Beer & Podsiadlowski 2002 give a somewhat smaller mass  $5.4 \pm 0.3 M_{\odot}$  and  $1.45 \pm 0.35 M_{\odot}$  for the companion.

REFERENCES.—(1) McClintock et al. 2001; (2) Wagner et al. 2001; (3) Bailyn et al. 1998; (4) Filippenko et al. 1999; (5) Gelino et al. 2001b; (6) Harlaftis, Horne, & Filippenko 1996; (7) Filippenko & Chornock 2001; (8) Gelino et al. 2001a; (9) Orosz et al. 1998; (10) Orosz 2002; (11) Orosz et al. 2002; (12) Beer & Podsiadlowski 2002; (13) Orosz et al. 2001; (14) Herrero et al. 1995; (15) Shahbaz et al. 1994; (16) Shahbaz et al. 1996; (17) Greiner et al. 2001.

TABLE 2  
PARAMETERS FOR NOVA SCORPII (BEER & PODSIADLOWSKI 2002) AND V4641 SGR (OROSZ ET AL. 2001)

Parameter	Nova Scorpii	V4641 Sgr
Orbital period (days)	2.623	2.817
Black hole mass ( $M_{\odot}$ )	$5.4 \pm 0.3$	$9.61^{+0.08}_{-0.08}$
Companion mass ( $M_{\odot}$ )	$1.45 \pm 0.35$	$6.53^{+0.03}_{-0.03}$
Total mass ( $M_{\odot}$ )	6.85	$16.19^{+0.08}_{-0.04}$
Mass ratio	0.27	$1.50 \pm 0.13$
Orbital separation ( $R_{\odot}$ )	15.2	$21.33^{+0.25}_{-0.02}$
Companion radius ( $R_{\odot}$ )	4.15	$7.47^{+0.03}_{-0.02}$
Distance (kpc)	3.2	$9.59^{+0.02}_{-0.19}$

with these higher main-sequence masses, because the giant radii of these stars exceed their radii in the supergiant phase. Below we see that the high black hole masses in a few binaries require us to extend the range upward to about  $30 M_{\odot}$ .

A major uncertainty in the evolution of all compact X-ray binaries is the phase of spiral-in that occurred in their evolution; these binaries are initially very wide, and when the primary fills its Roche lobe and transfers mass to the secondary, the mass transfer leads to instability, resulting in

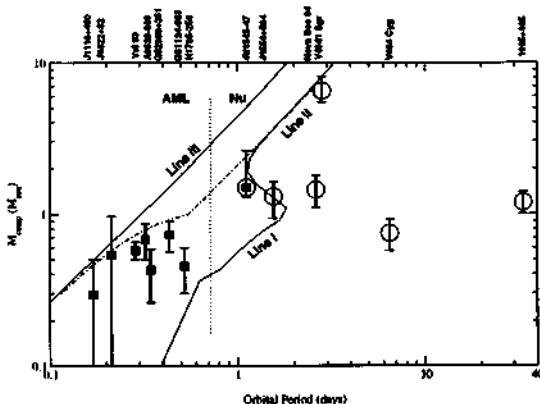


FIG. 2.—Companion mass as a function of present orbital period of 13 SXTs. XTE 1859+226 is not included because the companion mass is not well determined (Filippenko & Chornock 2001). Symbols of SXTs are the same as in Fig. 1. Line III indicates the orbital period for which a companion of that mass fills its Roche lobe on the ZAMS. No system can exist above and to the left of this line for a significant duration. Lines I and II are the upper-period limit for systems that can come into contact while the donor is on the main sequence. For high masses (line II), this limit is set by the period at which the evolution time of the companion is too short to allow the orbit to shrink significantly before it leaves the main sequence. For low masses (line I), where the donor never evolves off the main sequence within a Hubble time, the limit is set by the period for which the shrinking timescale of the orbit equals the Hubble time. The dot-dashed line indicates the point at which a system that starts its life on lines I/II comes into Roche contact. For very low masses, this equals line III, because the donor never moves significantly away from its ZAMS radius, whereas for very high masses it equals line II, because the orbit cannot shrink before the companion evolves off the main sequence. At intermediate masses, the companion expands somewhat, while the orbit shrinks and fills its Roche lobe at a larger period than line III. Systems that become SXTs with main-sequence donors within a Hubble time must start between line III and line I/II. At the start of mass transfer, they must lie in the narrow strip between line III and the dot-dashed line.

the secondary plunging into the primary's envelope. Next, dissipation of orbital energy of the secondary causes the primary's envelope to be ejected and the orbit to shrink. Following the original work by Webbink (1984), Brown et al. (2001c) write the standard formula for common-envelope evolution as

$$\frac{GM_p M_e}{\lambda R} = \frac{GM_p M_e}{\lambda r_L a_f} = \alpha_{ce} \left( \frac{GM_{He} M_d}{2a_i} - \frac{GM_p M_d}{2a_f} \right), \quad (1)$$

where  $M_p$  is the total mass of the black hole progenitor star just before the common envelope forms,  $M_e$  is the mass of its hydrogen envelope,  $M_{He}$  is the mass of its core,  $a_i$  and  $a_f$  are the initial and final separation, before and after the common envelope, respectively, and  $r_L \equiv R_L/a$  is the dimensionless Roche lobe radius. This equation essentially relates the loss of orbital energy of the secondary to the binding energy of the ejected envelope. The parameter  $\lambda$  is a shape parameter for the density profile of the envelope. It can vary greatly between stars (Tauris & Dewi 2001), but for the extended, deeply convective giants we deal with in case C mass transfer, it is always close to  $7/6$ . (See also Appendix C of Brown et al. 2001b.) The parameter  $\alpha_{ce}$  accounts for the efficiency with which orbital energy is used to expel the envelope, and may also account for some other effects such as extra energy sources and the possibility that each mass element of the envelope receives more than the minimum energy needed to escape (see, e.g., Bhattacharya & van den Heuvel 1991 and references therein).

Given the parameters of the system at first Roche contact, when spiral-in starts, the final separation is determined by the product of  $\lambda$  and  $\alpha_{ce}$ , the efficiency of the energy conversion. In general, these parameters are only the simplest recipe prescription for the complex hydrodynamical interaction during spiral-in. While we therefore cannot predict the value of  $\lambda\alpha_{ce}$  from first principles, we can try to find its value from constraints in some systems and then assume it is the same for all similar systems. Brown, Lee, & Tauris found a great regularity in the evolution of SXTs with main-sequence companions, all but one of which are K or M stars, which constrained the efficiency  $\lambda\alpha_{ce}$  to be 0.2–0.5. However, these authors did not include mass loss in the explosion, which we do here in our evolution of SXTs with evolved companions. Since mass loss substantially widens the orbits, including it in the common-envelope evolution must bring the (pre-explosion)  $a_f$  to a smaller value: if  $M_{post}$  is the black hole plus companion mass and  $\Delta M$  the mass lost in the formation of the black hole, we have

$$a_{f,post} = a_{f,pre} (1 + \Delta M / M_{post}) \quad (2)$$

(after the orbit has been recircularized). Therefore, there has been an extra widening since the explosion by a factor of up to about 1.5. We found  $\lambda\alpha_{ce}$  to be in the lower part of the interval found by Brown et al. (2001c),  $\lambda\alpha_{ce} \sim 0.2$ .

We can achieve a more precise “calibration” of the value of  $\lambda\alpha_{ce}$  if we manage to find some systems in which we can estimate both the initial and the final separation. To estimate the initial separation (at the onset of spiral-in), we need to know the mass and radius of the black hole progenitor and combine this with the Roche lobe-filling condition. The helium star progenitors in at least three of the evolved binaries seem to be too massive for the 20–23  $M_{\odot}$  ZAMS progenitors used by Brown et al. (2001c); the black hole in V404 Cyg is probably at least  $10 M_{\odot}$  (Shahbaz et al. 1994,

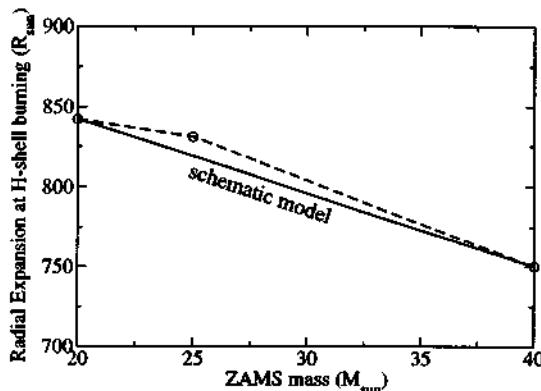


FIG. 3.—Radial expansion at the end of the giant branch. The data are taken from Schaller et al. (1992). Radial expansions are 842, 831, and  $750 R_{\odot}$  for 20, 25, and  $40 M_{\odot}$ , respectively. The radial expansion shows an almost linear dependence on the ZAMS mass.

1996; Bailyn et al. 1998), and the black hole in Nova Scorpii is of mass  $\sim 5.4 \pm 0.3 M_{\odot}$  (Beer & Podsiadlowski 2002), and the mass loss in black hole formation is  $\gtrsim 5 M_{\odot}$  (Nelemans, Tauris, & van den Heuvel 1999), so that the progenitor of the helium star must have been  $\sim 11 M_{\odot}$ . From Table 2, the black hole in V4641 Sgr is of mass  $9.61^{+2.08}_{-4.88} M_{\odot}$  (Orosz et al. 2001). The tentative conclusion from the above is that at least these binaries with evolved companions seem to have come from helium cores of  $\sim 11 M_{\odot}$  or ZAMS mass  $\sim 30 M_{\odot}$ . With high wind mass-loss rates as proposed by Schaller et al. (1992), such massive stars have larger radii as giants than as supergiants, thus making case C mass transfer impossible. However, since radii and mass-loss rates of evolved stars are very uncertain, we take the view that the need for  $\sim 11 M_{\odot}$  helium cores implies that their progenitors,  $30 M_{\odot}$  main-sequence stars, do expand enough to allow case C mass transfer.

In Figure 3, we summarize the radius at the end of the branch as a function of ZAMS mass (Schaller et al. 1992). The ZAMS mass dependence of this final giant radius is adequately represented by a linear function in the region of  $20$ – $40 M_{\odot}$ . We assume that the radial expansion during the helium burning can be scaled to the case of a  $20 M_{\odot}$  star using this linear relation as follows:

$$R(M; t) = \left[ -\frac{842 R_{\odot} - 750 R_{\odot} (M - 20 M_{\odot})}{842 R_{\odot}} + 1 \right] \times R(20 M_{\odot}; t). \quad (3)$$

Further, we took the mass-loss rate of  $20 M_{\odot}$  as standard and scaled the mass-loss rate in proportion to the ZAMS mass. The allowed range of case C mass transfer with ZAMS mass  $20 M_{\odot}$  is  $971 R_{\odot} < R < 1185 R_{\odot}$ , that of Schaller et al. (1992). In Figure 4 are given the possible initial orbital separations for case C mass transfer for the  $1.91 M_{\odot}$  companion appropriate for Nova Sco (see § 4) and for the  $6.53 M_{\odot}$  companion appropriate for V4641 Sgr (Orosz et al. 2001).

Now if we look at equation (1), we see that  $a_f$  scales almost linearly with the donor (companion) mass  $M_d$ . The envelope mass  $M_e$  is roughly  $0.7 M_{\text{giant}}$  (Bethe & Brown

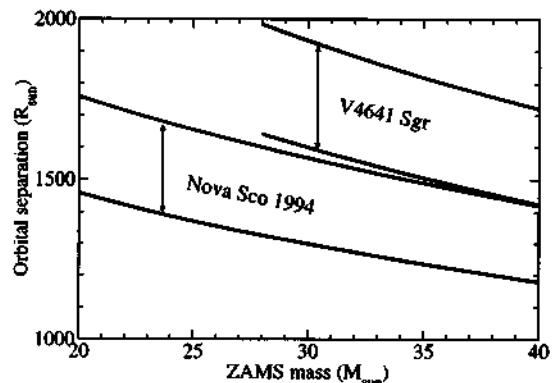


FIG. 4.—Possible initial orbital separations for case C mass transfer with  $1.91 M_{\odot}$  (Nova Sco) and  $6.53 M_{\odot}$  (V4641 Sgr) main-sequence companions. For Nova Sco, the estimated companion mass  $1.91 M_{\odot}$  during the common-envelope evolution is used as discussed in § 3. The x-axis is the ZAMS mass of the black hole progenitor.

1998), and we use

$$M_{\text{He}} = 0.08 (M_{\text{giant}}/M_{\odot})^{1.45} M_{\odot}, \quad (4)$$

so that

$$a_f \propto \frac{M_d}{M_{\odot}} \left( \frac{M_{\text{giant}}}{M_{\odot}} \right)^{-0.55} a_i, \quad (5)$$

assuming  $\lambda \alpha_{ce}$  to be constant and with neglect of the small term in  $a_i^{-1}$  in the right-hand side of equation (1). From our curves (Fig. 4), we see that the 20% possible variation in  $a_i$  results in the same percentage variation in  $a_f$ . Because the actual ZAMS mass can be anywhere in the range  $20$ – $30 M_{\odot}$ , there can be an additional  $\sim 25\%$  variation in  $a_f$  with giant mass, as compared with the linear dependence on  $M_d$ . In view of the modest size of these variations at a given donor mass, we make the approximation in the rest of the paper that the pre-explosion orbital separation depends only on  $M_d$  and scales linearly with  $M_d$ . This simple scaling and the modest amount of scatter around it are partly the result of the weak dependences on initial parameters in equation (5) but chiefly the result of the fact that our model uses case C mass transfer. This constrains the Roche contact to first occur when the radius of the star is in a very narrow range, between the maximum radius in the giant phase and the maximum radius in the supergiant phase.

To complete our calibration of the spiral-in efficiency, we must find systems in which we can also estimate the orbital separation just after spiral-in well. This is complicated by the fact that mass transfer has taken place since the spiral-in. Most SXTs have small mass ratios, and for such small mass ratios the orbital separation is fairly sensitive to the amount of mass transferred, making it hard to derive the post-spiral-in separation from the present one. The exception is V4641 Sgr, in which the present mass ratio is close to 1. Since the initial mass ratio could not have been significantly greater than 1 (since that would result in unstable mass transfer), and furthermore the orbital period changes very little with mass transfer for nearly equal masses, we can fairly approximate the post-spiral-in separation by the

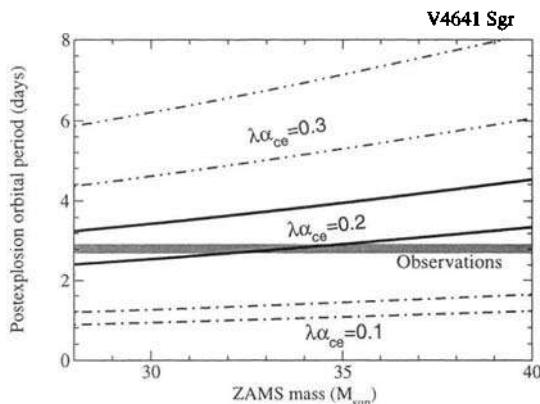


FIG. 5.—Postexplosion orbital periods of V4641 Sgr after black hole formation for various common-envelope efficiencies. The x-axis is the ZAMS masses of black hole progenitor. The width of each band is determined by the initial band of possible case C mass transfer given in Fig. 4. Here  $\lambda\alpha_{ce} \sim 0.2$  is consistent with current observations.

present one. In Figure 5, we show the predicted ranges of post-spiral-in orbital periods for different values of  $\lambda\alpha_{ce}$ . Clearly, a value quite close to 0.2 is indicated. For 4U 1543–47 (IL Lup), we find that it is near the boundary between evolved and main-sequence evolution. To place it there, as discussed in § 4, we find from the reconstructed orbital period in Figure 6 that  $\lambda\alpha_{ce} \sim 0.2$  is also consistent with the properties of this system.

In short, the general properties of SXTs and the specific cases of V4641 Sgr and IL Lup favor  $\lambda\alpha_{ce} \sim 0.2$ , which we therefore adopt as a general efficiency for the evolution of other transient sources. This then makes it possible to make quite specific predictions for the prior evolution of many of the other SXTs.

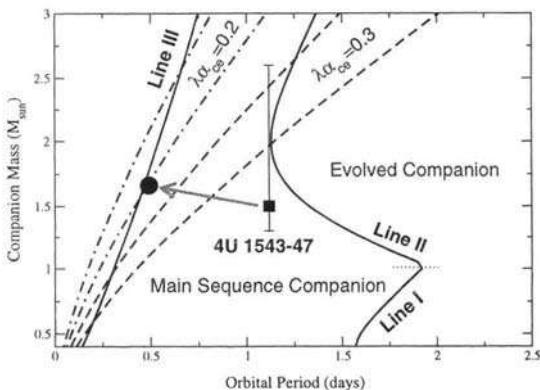


FIG. 6.—Enlargement of the central region of Fig. 2, indicating the present location of IL Lup. If we adopt the same value  $\lambda\alpha_{ce} = 0.2$  for the common-envelope parameter, we see that the reconstructed post-spiral-in period places the system right on the main-sequence line, implying a fairly tight constraint on the initial parameters of this system. (We took the ZAMS mass of the black hole progenitor to be  $30 M_\odot$ , which corresponds to  $M_p \sim 25 M_\odot$  in the beginning of case C mass transfer, and  $M_{He} = 11 M_\odot$ .)

### 3.2. Expected Regularities

From the above theory, certain regularities follow for the system behavior as a function of companion mass. First, the binding energy relation for spiral-in (eq. [1]) shows that very nearly  $a_f \propto M_d$ , with not much variation due to other aspects of the systems (see § 3.1). Furthermore, the relation between Roche lobe radius and donor mass when  $M_d \ll M_{BH}$  implies that  $R_L/a_f \propto M_d^{1/3}$  (e.g., Eggleton 1983). As a result, the Roche lobe radius of the donor just after spiral-in will scale with donor mass as  $R_L \propto M_d^{4/3}$ . On the other hand, the donor radius itself depends on its mass only as  $R_d \propto M_d^{0.8}$ . Therefore, a low-mass donor overfills its Roche lobe immediately after spiral-in. In the donor mass range we consider ( $M_d \gtrsim 0.7 M_\odot$ ) it does not overfill its Roche lobe by much, so we assume that the system adjusts itself quickly by transfer of a small amount of mass to the He star, which widens the orbit until the donor fills its Roche lobe exactly. Above this minimum mass, there will be a range of donor masses that are close enough to filling their Roche lobes after spiral-in that they will be tidally locked and will come into contact via angular momentum loss (AML). Above this, there will be a range of mixed evolution, where both AML and nuclear evolution (Nu) play a role. Finally, for the most massive donors,  $M_d > 2 M_\odot$ , the post-spiral-in orbits will be too wide for AML to shrink them much, so mass transfer will be initiated only via nuclear expansion of the donor. Of course, the ranges of case C radii of stars and variations of primary masses will ensure that the boundaries between these regions are not sharp; near the boundaries the fate of the system depends on its precise initial parameters.

## 4. RECONSTRUCTING THE PRE-EXPLOSION ORBITS

1. Nova Sco 94 (GRO J1655–40): The most extensive evolutionary studies have been made for Nova Sco. Starting from the work of Regos, Tout, & Wickramasinghe (1998), who make the case that the companion is in late main-sequence evolution, Beer & Podsiadlowski (2002) carry out extensive numerical calculations of the evolution, starting with a pre-explosion mass of  $2.5 M_\odot$  and separation of  $\sim 6 R_\odot$ . More schematically, we arrived at a pre-explosion mass of  $1.91 M_\odot$  and separation of  $5.33 R_\odot$ . We consequently have a 0.4 day pre-explosion period. With  $\sim 6 M_\odot$  mass loss in the explosion (Nelemans et al. 1999), nearly half the system mass, the binary period increases to 1.5 days, well beyond the period gap. This is also the period required if the common-envelope efficiency in this binary were again 0.2 (Fig. 7). This explains why Nova Sco is the only system with a giant donor and a black hole mass in the lower end of the range; its evolution really places it among the narrow-orbit systems. Generally, the mass loss during explosion is mild and does not change which category a system belongs to. But in those exceptional cases in which the mass loss comes close to half the total mass, the orbit widens very much and converts an AML system to a nuclear evolution system. We discuss in § 6 that help in expelling the mass may come from early onset of the GRB mechanism. After explosion, the binary evolves to its present period by nearly conservative mass transfer. Our estimate is that  $0.41 M_\odot$  is transferred from the donor to the black hole. Brown et al. (1999b) first made the case that Nova Sco was the relic of a GRB.

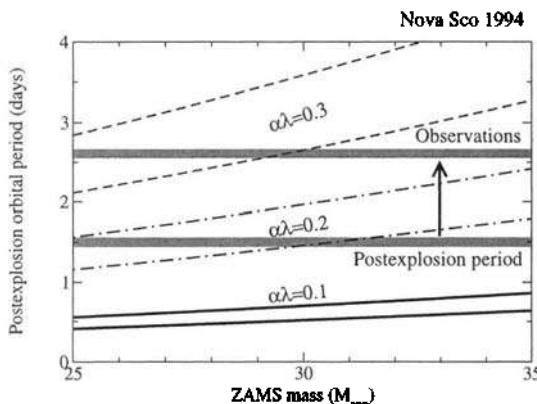


FIG. 7.—Postexplosion orbital periods of Nova Sco after black hole formation for various common-envelope efficiencies. For our calculation, we use a postexplosion black hole mass  $M_{\text{BH}} \sim 4.94 M_{\odot}$  and a companion mass  $M_d \sim 1.91 M_{\odot}$ . The  $x$ -axis is the ZAMS mass of black hole progenitor. The width of each band is determined by the initial band of possible case C mass transfer separations given in Fig. 4. A postexplosion period of about 1.5 days is consistent with  $\alpha\lambda \sim 0.2$ . By conservative mass transfer of  $0.46 M_{\odot}$  from the companion star to the black hole, the orbital period can be evolved to the currently observed orbital period.

2. V4641 Sgr: As we discussed in § 3, this system is our calibrator for the spiral-in efficiency, and we assume that its present state is very close to the one immediately following spiral-in.

3. GRS 1915+105: Recently, Greiner et al. (2001) determined the period and black hole mass of GRS 1915+105 to be 33.5 days and  $14 \pm 4 M_{\odot}$ . Interestingly, we can evolve a system with properties very close to this by simply starting from V4641 Sgr and following its future evolution with conservative mass transfer ( $P_{\text{orb}} \propto \mu^3$ , where  $\mu$  is the reduced mass); allowing for  $4.6 M_{\odot}$  to be transferred from the donor to the black hole, we have

$$P_{1915} = \left( \frac{9.61 \times 6.53}{14.21 \times 1.93} \right)^3 P_{4641} = 33.7 \text{ days} . \quad (6)$$

This would give a companion mass of  $1.93 M_{\odot}$ , as compared with the Greiner et al. (2001) mass of  $M_d = 1.2 \pm 0.2 M_{\odot}$ . However, the mass transfer cannot be completely conservative because of loss by jets, etc., as evidenced by the microquasar character of this object. Furthermore, the above  $M_d$  is viewed as a lower limit by Greiner et al. because the donor is being cooled by rapid mass loss, but its mass is estimated by comparison with noninteracting stars. We thus believe our evolution to be reasonable. We position the pre-explosion period and black hole mass of GRS 1915+105 at the same point as V4641 Sgr. Since mass transfer and widening of the orbit always occur together, the effect of this post-explosion evolution is to introduce a weak secondary correlation between orbital period and companion mass in the long-period regime, where such a correlation is not expected to arise from the pre-explosion evolution.

4. IL Lup: Recently, overabundances of Mg in the companion star of IL Lup have been observed (J. A. Orosz 2002, private communication). In analogy with the case of the overabundances in Nova Sco (Israelian et al. 1999; Brown et al. 2000), this indicates that there was an explosion at the

time of black hole formation in this system, in which some of the material ejected from the core of the helium star progenitor to the black hole ended up on the companion. Based on these observations and our given efficiency  $\lambda\alpha_{\infty} = 0.2$ , one can start with the

$11 M_{\odot}$  He star and  $1.7 M_{\odot}$  companion as a possible progenitor of IL Lup. From the lower boundary of the curve with  $\lambda\alpha_{\infty} = 0.2$  in Figure 6, the period would be 0.5 days. By losing  $4.2 M_{\odot}$  during the explosion, the binary orbit would be widened to 1.12 days. The period had to be shortened to 0.8 days by magnetic braking and gravitation wave radiation before the mass transfer started. Conservative transfer of  $0.23 M_{\odot}$  from the companion to the black hole would bring the period from 0.8 days to the present 1.1164 days.

5. V404 Cyg: The black hole in V404 Cyg appears to be somewhat more massive than in IL Lup, so we begin with a similar mass companion, but a  $10 M_{\odot}$  black hole, which would have a period of 0.63 days. Again, we neglect mass loss in the explosion, although a small correction for this might be made later. Conservative transfer of  $1 M_{\odot}$  from the donor to the black hole then brings the period to

$$0.63 \text{ days} \left( \frac{1.7 M_{\odot} \times 10 M_{\odot}}{0.7 M_{\odot} \times 11 M_{\odot}} \right)^3 = 6.7 \text{ days} , \quad (7)$$

close to the present 6.47 day period. Here we take 11 and  $0.7 M_{\odot}$  as current masses in V404 Cyg (Orosz 2002). The black hole in V404 Cyg seems to be somewhat more massive than the others in the transient sources, with the exception of that in GRS 1915+105. In both cases, we achieve the relatively high black hole masses and periods by substantial accretion onto the black hole.

6. GRO J1550–564: The high-mass black hole in J1550–564,  $10.56 M_{\odot}$  (Orosz et al. 2002), is slightly less massive than the assumed black hole mass of V404 Cyg, and the companion is more massive than V404 Cyg with short period, 1.552 days. So, we start from the same initial conditions just derived for V404 Cyg (Fig. 11) and end up with the present system via simple conservative mass transfer.

7. Cygnus X-1: Cyg X-1 is usually not considered to have come from the same evolutionary path as the SXTs, since it is a persistent X-ray source with a much more massive donor. But with the discovery of objects with relatively massive donors in the SXT category, such as V4641 Sgr, it is worth considering the implications of our model for it. Cyg X-1 has been shown to have an appreciable system velocity (Kaper et al. 1999), although it may be only one-third the  $50 \text{ km s}^{-1}$  given there, depending on the O star association (L. Kaper 2001, private communication). The evolution of Cyg X-1 may have been similar to that of the transient sources, the difference being in the copious mass loss from the companion O9 I star, causing the black hole to accrete and emit X rays continuously. If we scale to Nova Sco to obtain the initial binary separation, we find

$$a_f = \frac{17.8 M_{\odot}}{1.91 M_{\odot}} \times 5.33 R_{\odot} = 50 R_{\odot} , \quad (8)$$

somewhat larger than the present binary separation of  $40 R_{\odot}$ . (We would obtain  $38 R_{\odot}$  if we scaled from the Beer & Podsiadlowski [2002] companion mass of  $2.5 M_{\odot}$  for Nova Sco.) Given uncertainties in the mass measurements, we believe it possible for Cyg X-1 to be accommodated in this scheme. Some sort of common-envelope evolution seems to

be necessary to narrow the orbit in the evolution involving the necessarily very massive progenitor stars (Brown et al. 2001a).

#### 4.1. Problems with the Close (AML) Systems

Reconstruction of the AML binaries is more complicated, because they have lost angular momentum through magnetic braking and gravitational waves, so that their present positions as plotted in Figure 1 are not those at pre-explosion time. As with the evolved companions, matter will have been accreted onto the black hole, so the black hole masses will be somewhat greater than just following the explosion. As noted earlier, the binaries with less massive companions with separation  $a_f$  at the end of common-envelope evolution overfill their Roche lobes. The outer part of the companion, down to the Roche lobe  $R_L$ , is transferred onto the He star. This mass transfer widens the orbit to  $R_L$ , possibly overshooting. Unless much mass is lost in the explosion when the black hole is formed, the Roche lobe radius is unchanged by the formation of the black hole and corresponds to line III in Figure 2.

Brown et al. (2001c) explored the evolution of ZAMS  $1.25 M_\odot$  stars under magnetic braking, gravitational waves, and mass transfer to the black hole. We adapt the same methods to make a more detailed study of the AML. First of all, we construct (Fig. 2) the lower limit on the companion mass for evolution in a Hubble time, giving the dashed line there. All binaries with companions in the main sequence at the beginning of mass transfer must lie between

the dashed line and line III in that figure. The fact that the AMLs tend to lie below the dashed line implies both mass loss from the companion and accretion onto the black hole. Therefore, all these systems have shrunk their orbits and increased their black hole mass since the formation of the black hole by amounts that cannot be determined well. In Figure 10, however, we show where the four shortest period AMLs would have come from, had they lost  $0.7 M_\odot$  from an initial  $1.5 M_\odot$ . From our earlier discussion about the  $a_f$  following common-envelope evolution, we saw that binaries with companions that stayed in the main sequence were favored to come from companion masses less than  $2 M_\odot$ , and from Figure 2 we see that they would chiefly have companion ZAMS mass greater than  $\sim 1 M_\odot$ , so that most of them would initially have periods of 0.4–0.7 days (which follows from the separations obtained from our eq. [5]). In trying to understand the detailed evolution of the AML, we begin from a binary with a  $2 M_\odot$  companion that just fills its Roche lobe following common-envelope evolution. We then follow its evolution under the two different assumptions made in Brown et al. (2001c): (1) that its time of evolution is always given by its initial  $2 M_\odot$  mass, i.e., ignoring effects of mass loss on the internal evolution time (Fig. 8, dashed lines, and Fig. 9, right dashed line), and (2) that the evolution of the star proceeds according to its adjusted mass (Fig. 8, solid lines, and Fig. 9, left dashed line). Since mass loss drives the companion out of thermal equilibrium, these two extremes bracket the outcome of a full stellar model calculation.

In summary, the AML systems have had the information on their postexplosion parameters partly erased by subse-

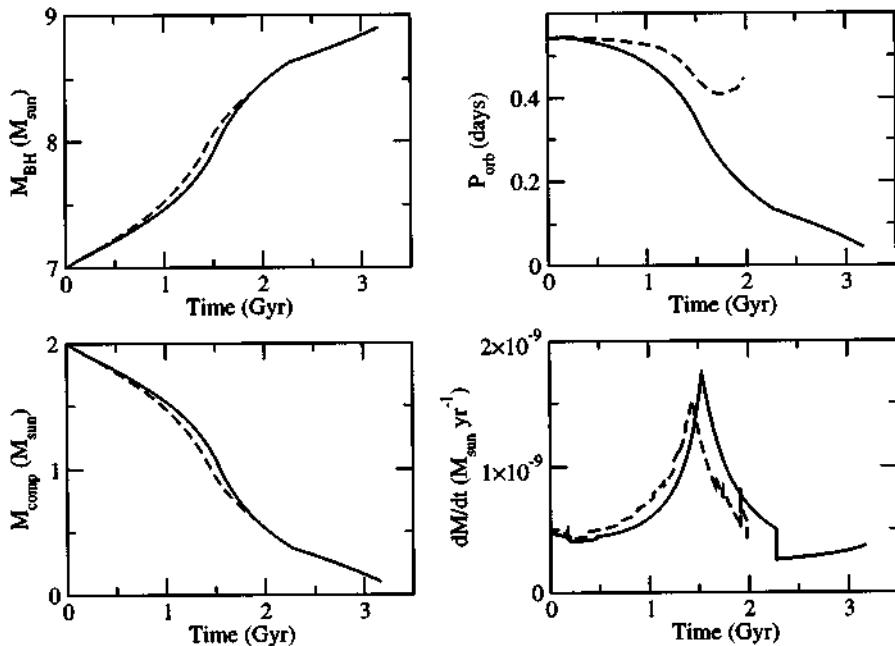


FIG. 8.—Evolution of a binary with  $7 M_\odot$  black hole and  $2 M_\odot$  companion for the initial period of 0.54 days. The solid line marks the evolution in case the companion star adjusts itself as it loses mass; the dashed line traces the evolution in case the mass loss does not affect the internal timescale of the companion star, so that it follows the same time evolution as an undisturbed  $2 M_\odot$  star.

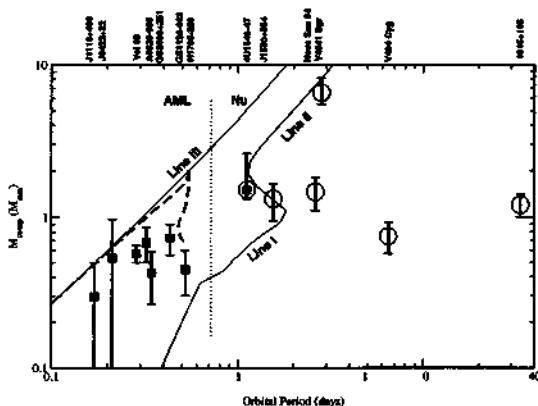


FIG. 9.—Evolutionary tracks of a binary with  $9 M_{\odot}$  black hole and  $2 M_{\odot}$  companion for the initial period of 0.54 days. The two evolution possibilities are as in Fig. 8. Left (right) dashed line corresponds to the solid (dashed) lines in Fig. 8.

quent evolution, in a manner that we cannot undo. Therefore, they can only provide a crude consistency check on the mass-period relation for black holes in SXTs, rather than provide precise constraints.

### 5. ANGULAR MOMENTUM AND ITS CONSEQUENCES FOR THE MASS AND SPIN OF THE BLACK HOLE

It is, in general, a difficult and unsolved problem to calculate the angular momentum of a stellar core at any given time. Even if we make the usual assumption that the rotation is initially solid body and not very far away from the maximal stable rotation frequency, the viscous coupling between the various layers of the star as it evolves is poorly known, and thus it is hard to be very quantitative. The general trend, however, is that the core will shrink and the envelope expand. In absence of viscous coupling, every mass element retains its angular momentum, and hence the core spins up as the envelope spins down, setting up a strong gradient in rotation frequency between the core and the envelope. Viscosity will then act to reduce this gradient, transporting angular momentum from the core to the envelope, but the efficiency of this process is very uncertain (Spruit & Phinney 1998; Livio & Pringle 1998).

As we noted in § 1, in our scenario, a number of effects will increase the angular momentum of the core relative to a similar core of a single star: (1) during spiral-in, the matter somewhat inside the orbit of the secondary is spun up by tidal torques (Rasio & Livio 1996), (2) the removal of the envelope halts the viscous slowdown of the core by friction with the envelope, and (3) during the post-spiral-in evolution, tidal coupling will tend to spin the helium star up even closer to the orbital period than was achieved by the first effect. This will not be a very strong effect because the duration of this phase is short, but it will affect the outer parts of the helium star somewhat, and this is the most important part (see below).

The net result of all these effects will be that the helium star will spin fairly rapidly, especially its envelope. The core

is not so crucial to our argument about the fraction of the star that can fall into the black hole, since the few solar masses in it will not be centrifugally supported even in quite short orbits. For the purpose of a definite calculation, we therefore make the following assumptions: (1) the helium star corotates with the orbit before explosion and is in solid-body rotation, and (2) the mass distribution of the helium star with radius is given by a fully radiative zero-age helium main-sequence star. This latter approximation is, of course, not extremely good. However, what counts is the angular momentum as a function of mass, so the fact that the mass distribution has changed from helium ZAMS to explosion would be entirely inconsequential if no redistribution of angular momentum had taken place in the interim. As we saw above, any redistribution of angular momentum would take the form of angular momentum transport toward the outer layers. This means that relative to our ideal calculations below, a better calculation would find more angular momentum in the outer layers and therefore somewhat smaller black hole masses than the ones we calculate.

We now investigate how much mass will be prevented from falling into the black hole by the angular momentum of the He star, under the above assumptions of solid-body rotation with a period equal to that of the binary. If we assume that angular momentum is conserved during the collapse, we can get the cylindrical radius  $R_c$ , within which matter is not centrifugally prevented from falling into the black hole:

$$R_c^2 \Omega = \bar{l}(\hat{a}) \frac{GM_c}{c}, \quad (9)$$

where  $\bar{l}(\hat{a})$  is the dimensionless specific angular momentum of the marginally bound orbit for a given Kerr parameter  $\hat{a}$ , and  $M_c$  is the total mass inside the cylinder of radius  $R_c$ . The Kerr parameter becomes

$$\hat{a} = \frac{I_c \Omega}{GM_c^2/c} = k^2 \bar{l}(\hat{a}), \quad (10)$$

where  $I_c$  is the total moment of inertia inside the cylinder of radius  $R_c$ ,  $I_c = k^2 M_c R_c^2$ . Here  $M_c$  gives an estimate of the final black hole mass. Combining these relations with a profile of angular momentum and mass versus radius using the assumptions listed above, we can calculate the expected black hole mass and Kerr parameter as a function of SXT period before explosion.

In Figure 10, we show the predicted relation between orbital period and black hole mass for different helium star masses in our model. We compare these with the present properties of all SXTs for which the required parameters are known. The properties are consistent with the theoretical relations but do not confirm it very strongly because of the evolutionary changes discussed in § 4. Specifically, the AML systems lie above and to the left of the curves, because their orbits shrunk and their black holes accreted mass since the formation of the black hole. However, as we saw in § 4.1, plausible amounts of conservative mass transfer since the explosion would place the systems among the theoretical postexplosion curves (Fig. 10, open squares and arrows).

To test the theory more strongly, we show in Figure 11 only those systems for which the pre-explosion properties could be reconstructed (§ 4). We compare the observed points with ideal polytropic helium stars of  $7$  and  $11 M_{\odot}$  and with a full-model calculation obtained from Woosley

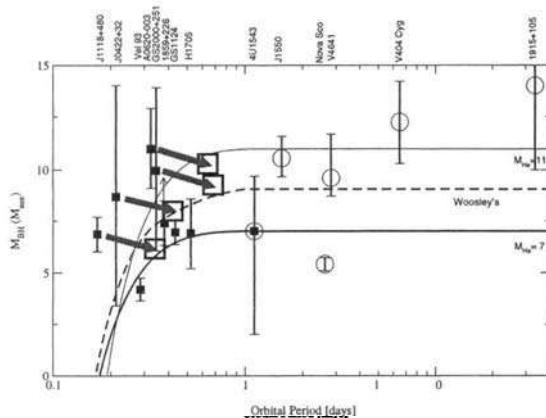


FIG. 10.—Present orbital period vs. black hole masses of SXTs. The deviations from the theoretical curves are substantial because of the post-explosion evolution of the binaries. The arrows on the AML systems point to an approximate postexplosion location if the donor mass was initially  $1.5 M_{\odot}$ , and  $0.7 M_{\odot}$  has now been transferred to the black hole. The solid lines indicate the possible ranges of black hole masses with polytropic index  $n = 3$  (radiative) for given pre-explosion spin periods that are assumed to be the same as the pre-explosion orbital period. Here we used  $R_{\text{He}} = 0.22(M_{\text{He}}/M_{\odot})^{0.6} R_{\odot}$ . For comparison, the results with a “scaled” He core ( $9.15 M_{\odot}$ ) of Woosley’s  $25 M_{\odot}$  star at the beginning of  $^{12}\text{C}$  burning with  $T_c = 5 \times 10^8$  K, appropriate for case C mass transfer, are plotted as a dashed line (S. E. Woosley 2001, private communication). In this plot, we scaled the radius of Woosley’s core,  $R_{\text{Woosley}} \sim 3 \times 10^{10}$  cm, by a factor of 2. As can be seen, the AML systems can plausibly originate from systems within the curves and thus are consistent with our theory. However, since they could have originated anywhere between the open square and their current location, they do not strongly test the theory.

(2001). By coincidence, the curves converge near the region of the shortest period observed systems, so that the uncertainty in helium star mass is not of great importance to the outcome. A helium star mass in the lower end of the range ( $7-9 M_{\odot}$ ) may be somewhat preferred for these systems. For periods above 1 day, angular momentum support is not

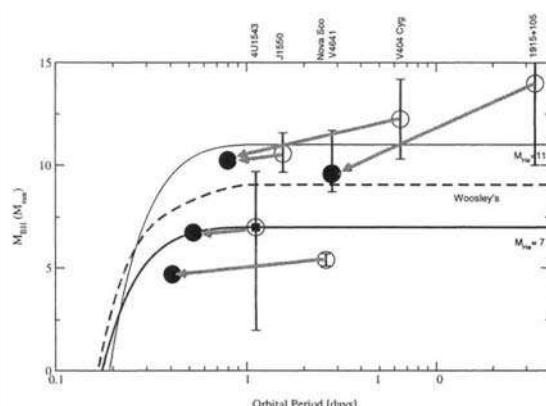


FIG. 11.—Reconstructed pre-explosion orbital period vs. black hole masses of SXTs with evolved companions. The reconstructed pre-explosion orbital periods and black hole masses are marked by filled circles, and the current locations of binaries with evolved companions are marked by open circles.

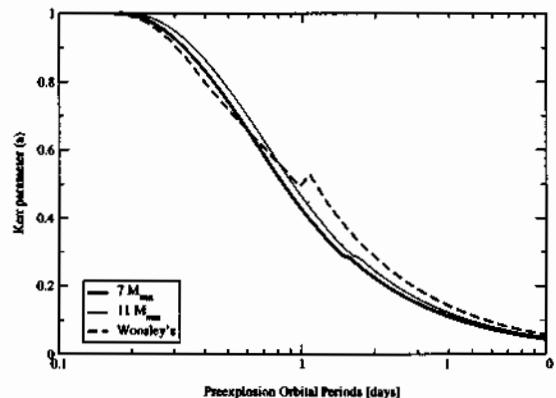


FIG. 12.—Kerr parameter of the black hole resulting from the collapse of a helium star synchronous with the orbit, as a function of orbital period. The conditions are the same as in Fig. 10, as is the meaning of the three curves. Woosley’s helium core is of mass  $9.15 M_{\odot}$  from a ZAMS  $25 M_{\odot}$  star (S. E. Woosley 2001, private communication). Note that the result depends very little on the mass of the helium star or on whether we use a simple polytrope or a more sophisticated model. The plot illustrates that rapidly rotating black holes needed for powering GRBs originate only from originally short-period SXTs.

important, and the mass of the final black hole will be very close to that of the helium star and thus varies somewhat from system to system. As we can see, the reconstructed pre-explosion properties lie much closer to the theoretical predictions.

As a corollary, we find that systems with very large velocities, like Nova Sco, will be rare; at the shortest pre-explosion orbits, where much mass is ejected, the companion mass tends to be small. Then the center of mass of the binary is close to that of the helium star, which strongly limits the systemic velocity induced by the mass loss. On the other hand, for the widest systems, where the companion tends to be massive enough to allow a significant systemic velocity induced by mass loss, the mass loss itself becomes too small to induce much of a systemic velocity.

An important result for our proposed relation between SXTs and hypernovae and GRBs is shown in Figure 12. This figure shows the expected Kerr parameter of the black hole formed in our model. We see that for the short-period systems, this Kerr parameter is very large, 0.7–0.9. This means that we are justified in adding only the mass that immediately falls into the black hole, because as soon as the rapidly rotating black hole is formed, it will drive a very large energy flux in the manner described by Brown et al. (2000). This both causes a GRB and expels the leftover stellar envelope. The systems with longer orbital periods do not give rise to black holes with large Kerr parameters and thus are presumably not the sites of GRBs.

## 6. CONCLUSIONS

We have shown that there is an observed correlation between orbital period and black hole mass in SXTs. We have modeled this correlation as resulting from the spin of the helium star progenitor of the black hole; if the pre-explosion orbit has a short period, then the helium star spins rapidly. This means that some part of its outer envelope is

centrifugally prevented from falling into the black hole that forms at the core. This material is then expelled swiftly, leading to a black hole mass much less than the helium star mass. As the orbital period is lengthened, the centrifugal support wanes, leading to a more massive black hole. The reason for swift expulsion of material held up by a centrifugal barrier is the fact that black holes formed in our scenario naturally have high Kerr parameters (Fig. 12). This implies that they input very high energy fluxes into their surrounding medium via the Blandford-Znajek mechanism, and thus power both a GRB and the expulsion of the material that does not immediately fall in.

However, because the correlation is induced between the orbital period before explosion and the black hole mass, its manifestation in the observed correlation between black hole mass and present orbital period is weakened because of postexplosion evolution of the binaries. We therefore considered the evolution in some detail, and for a subset of the systems, we were able to reconstruct the pre-explosion orbital periods. The correlation between pre-explosion period and black hole mass (Fig. 11) is in much better agreement with our model than the original one between present period and black hole mass (Fig. 1). We developed a quantitative model for the relation between period and mass and showed that it fits the subset of reconstructible SXT orbits.

Nova Sco stands out as the most extreme case of mass loss, nearly half of the total system mass, and therefore a great widening in the orbit, which gets its period well beyond the gap between shrinking and expanding orbits. From Figure 11, we see that its black hole mass is far below the polytropic line for its  $M_{He} = 11 M_{\odot}$  progenitor. We believe that in the case of this binary, a short central engine time of several seconds was able to furnish angular momen-

tum and energy to the disk quickly enough to stop the infall of some of the interior matter not initially supported by centrifugal force; i.e., the angular momentum was provided in less than a dynamical time. In other words, the Blandford-Znajek mechanism that drives the GRB not only expelled the matter initially supported for a viscous time by angular momentum, but actually stopped the infall within a dynamical time.

Since we can also compute the Kerr parameters of the black holes formed via our model, we find that the short-period systems should have formed black holes with Kerr parameters in the range 0.7–0.9. This makes them prime candidates for hypernovae and GRBs and thus provides further support for our earlier study in which we posited that SXTs with black hole primaries are the descendants of GRBs. We can now also refine this statement: SXTs with short orbital periods before the formation of the black hole have given rise to a GRB in the past.

We would like to thank J. Orosz and S. Woosley for useful advice and information. Hans Bethe critically read the manuscript and offered helpful suggestions. We are very grateful to the referee for a number of helpful suggestions, which enabled us to improve the manuscript. In particular, the referee brought to our attention the Greiner et al. (2001) observations of 1915+105, which we could evolve as V4641 Sgr viewed at a later time. C. H. L. was in part supported by the BK21 project of the Korea Ministry of Education. We were in part supported by the US Department of Energy under grant DE-FG02-88ER40388. R. A. M. J. W. also acknowledges partial support from NASA under grant NAG5-10772.

#### REFERENCES

- Bailyn, C. D., Jain, R. K., Coppi, P., & Orosz, J. A. 1998, ApJ, 499, 367
- Beer, M., & Podsiadlowski, Ph. 2002, MNRAS, 331, 351
- Bethe, H. A., & Brown, G. E. 1998, ApJ, 506, 780
- Bhattacharya, D., & Van den Heuvel, E. P. J. 1991, Phys. Rep., 203, 1
- Blandford, R. D., & Znajek, R. L. 1977, MNRAS, 179, 433
- Brown, G. E., Heger, A., Langer, N., Lee, C.-H., Wellstein, S., & Bethe, H. A. 2001a, NewA, 6, 457
- Brown, G. E., Lee, C.-H., & Bethe, H. A. 1999a, NewA, 4, 313
- Brown, G. E., Lee, C.-H., Lee, H. K., & Bethe, H. A. 1999b, AIP Conf. Proc. 526, Gamma-Ray Bursts, ed. R. M. Kippen, R. S. Mallozzi, & G. J. Fishman (New York: AIP), 628
- Brown, G. E., Lee, C.-H., Portegies Zwart, S. F., & Bethe, H. A. 2001b, ApJ, 547, 345
- Brown, G. E., Lee, C.-H., & Tauris, T. 2001c, NewA, 6, 331
- Brown, G. E., Lee, C.-H., Wijers, R. A. M. J., Lee, H. K., Israeliyan, G., & Bethe, H. A. 2000, NewA, 5, 191
- Brown, G. E., Weintraub, J., & Wijers, R. A. M. J. 1996, ApJ, 463, 297
- Eggleton, P. P. 1983, ApJ, 268, 362
- Filippenko, A. V., & Chornock, R. 2001, IAU Circ., 7644, 2
- Filippenko, A. V., Leonard, D. C., Matheson, T., Li, W., Moran, E. C., & Riess, A. G. 1999, PASP, 111, 969
- Gelino, D. M., Harrison, T. E., & McNamara, B. J. 2001a, AJ, 122, 971
- Gelino, D. M., Harrison, T. E., & Orosz, J. A. 2001b, AJ, 122, 2668
- Greene, J., Bailyn, C. D., & Orosz, J. A. 2001, ApJ, 554, 1290
- Greiner, J., et al. 2001, Nature, 414, 522
- Hartfus, E. T., Horne, K., & Filippenko, A. V. 1996, PASP, 108, 762
- Herrero, A., Kudritzki, R. P., Gabler, R., Vilchez, J. M., & Gabler, A. 1995, A&A, 297, 556
- Israeliyan, G., Rebolo, R., Basri, G., Casares, J., & Martin, E. L. 1999, Nature, 401, 142
- Kaper, L., Cameron, A., & Barziv, O. 1999, in IAU Symp. 193, Wolf-Rayet Phenomena in Massive Stars and Starburst Galaxies, ed. K. A. van der Hucht, G. Koenigsberger, & R. J. Eenens (San Francisco: ASP), 316
- Lee, H.-K., Lee, C. H., & van Putten, M. H. P. M. 2001, MNRAS, 324, 781
- Li, L. X. 2000a, ApJ, 531, L111
- \_\_\_\_\_. 2000b, ApJ, 533, L115
- \_\_\_\_\_. 2000c, ApJ, 540, L17
- \_\_\_\_\_. 2000d, ApJ, 544, 375
- \_\_\_\_\_. 2002, ApJ, 567, 463
- Livio, M., & Pringle, J. E. 1998, ApJ, 505, 339
- MacFadyen, A. I., & Woosley, S. E. 1999, ApJ, 524, 262
- McClintock, J. E., Garcia, M. R., Caldwell, N., Falco, E. E., Garnavich, P. M., & Zhao, P. 2001, ApJ, 551, L147
- Nelemans, G., Tauris, T. M., & van den Heuvel, E. P. J. 1999, A&A, 352, L87
- Orosz, J. A., Jain, R. K., Bailyn, C. D., McClintock, J. E., & Remillard, R. A. 1998, ApJ, 499, 375
- Orosz, J. A., et al. 2001, ApJ, 555, 489
- \_\_\_\_\_. 2002, ApJ, 568, 845
- Paczynski, B. 1998, ApJ, 494, L45
- Rasio, F. A., & Livio, M. 1996, ApJ, 471, 366
- Regis, E., Tout, C. A., & Wickramasinghe, D. 1998, ApJ, 509, 362
- Schaller, G., Schaefer, D., Meynet, G., & Maeder, A. 1992, A&AS, 96, 269
- Shahbaz, T., Bandyopadhyay, R., Charles, P. A., & Naylor, T. 1996, MNRAS, 282, 977
- Shahbaz, T., Ringwald, F. A., Bunn, J. C., Naylor, T., Charles, P. A., & Casares, J. 1994, MNRAS, 271, L10
- Spruit, H., & Phinney, E. S. 1998, Nature, 393, 139
- Tauris, T. M., & Dewi, J. D. M. 2001, A&A, 369, 170
- van Paradijs, J., Kouveliotou, C., & Wijers, R. A. M. J. 2000, ARA&A, 38, 379
- van Putten, M. H. P. M. 1999, Science, 284, 115
- \_\_\_\_\_. 2001, Phys. Rep., 345, 1
- Wagner, R. M., Folitz, C. B., Shahbaz, J., Casares, J., Charles, P. A., Starrfield, S. G., & Hewett, P. 2001, ApJ, 556, 42
- Webbink, R. F. 1984, ApJ, 277, 355
- Wellstein, S., & Langer, N. 1999, A&A, 350, 148
- Woosley, S. E. 1993, ApJ, 405, 273

This page is intentionally left blank

## Commentary on Appendices A-D

Although kaon condensation was an integral part of Paper 5, *a scenario for a large number of black holes in the Galaxy* by Brown and Bethe, we cannot claim it to be definitely realized in nature, although it is an interesting possibility. We have, therefore, relegated the papers to the Appendix, where the relevant formulae will be available to interested research workers.

In Fig. 1 of Paper 19 we showed that the neutron star mass measurement in all 19 binaries with degenerate companions was consistent with  $M_{\text{NS}}(\text{max}) = 1.5M_{\odot}$ , the limit of Paper 5. (In the case of nondegenerate companions, major corrections must be made for the modification in structure of the companion star, so interpretation of the measurements is much more difficult.)

Whereas in Paper 19 we showed Fig. 1 as a support for the  $1.5M_{\odot}$  upper limit for neutron star mass, we somewhat undid this argument in Paper 21. There we showed that if Roche Lobe overflow occurs before completion of He core burning, the Fe core which results from burning the “naked” He star is unlikely to give a compact object of mass  $\gtrsim 1.5M_{\odot}$  (gravitational). In the evolution of the compact binaries shown in Fig. 1 of Paper 19, Roche Lobe overflow almost always takes place in the red giant phase, preceding He core burning, so the binary evolution itself would limit the neutron star mass to  $\sim 1.5M_{\odot}$ . Case C mass transfer, i.e., mass transfer following He core burning, can, however, occur for stars with ZAMS masses  $\sim 20M_{\odot}$ . In Paper 19 we pinpointed the (assumed by us) low-mass black hole that resulted from SN1987A in the ZAMS interval  $18-20M_{\odot}$  and below  $18M_{\odot}$  we should be able to evolve neutron stars in binaries with masses  $> 1.5M_{\odot}$  were they to exist. (We are justified in requiring similar fates for the collapse of an assumed single star progenitor of SN1987A and the companion in a binary which undergoes Case C mass transfer. This is because in the latter case the companion evolves as a single star, there not being time after Case C mass transfer for winds, etc., to respond to the new-formed nakedness of the He star.)

It is clear to us that neutron stars must have a large strangeness content. This follows from the delayed collapse scenario of Prakash (1995), which gives an explanation for the existence of a neutron star for  $\sim 10$  sec in SN1987A, during which time neutrinos were

emitted, and then the sudden cutoff of any signal. This is most simply explained by the neutron star dropping into a (low-mass) black hole.

The above scenario is fulfilled not because of the strangeness, but because particles with negative strangeness ( $K^-$ ,  $\Lambda$ ,  $\Sigma$ ,  $\Xi$ ) are negatively charged, and therefore each such particle requires a proton to neutralize its charge. Thus, the presence of a substantial component of negative strangeness requires an equal additional component of protons.

Now, in the standard theory of neutron stars, without inclusion of strangeness, once a neutron star that exceeds the limit of  $M_{\text{NS}}(\text{max})$ , which we take to be  $1.5M_\odot$ , is formed, then as neutrinos leave they decrease the neutrino degeneracy pressure. As this happens protons change into neutrons, which increases the pressure through the increased nuclear symmetry energy and accompanying increased pressure. This latter increase is at least as large as the former decrease in neutrino degeneracy pressure. Thus, the neutron star is more unstable at time  $t = 0$  than at later times, and if it is going to drop into a black hole it will do so immediately. This is the standard scenario.

However, when substantial strangeness is present the protons necessary to balance the negative charge of the strangeness keep the symmetry pressure from rising so much with time, the neutronization being less. In kaon condensation electrons are replaced by  $K^-$ -mesons, which are bosons, and can therefore all go into a zero-momentum state. At higher densities there are more antikaons than there would be electrons, where Fermi energies would have to be paid. Therefore, more protons are needed to neutralize the charge.

In the competing scenario, championed by Glendenning (revised in Glendenning 2001, and also worked out in many papers by Prakash & Lattimer), some of the neutrons in the star are replaced by hyperons,  $\Lambda$  and  $\Sigma^-$ . At intermediate densities it is favorable to do this because the neutron chemical potential  $\mu_N$  is very high, due to strong repulsion from the vector mean field. In the case of the  $\Sigma^-$  both an electron and a neutron can be replaced by the hyperon, with low chemical potential.

The two scenarios, kaon condensation or hyperon introduction, give nearly the same results; i.e., they both give a maximum neutron star mass  $M_{\text{NS}}(\text{max}) \sim 1.5M_\odot$ . (See Glendenning 2001 for this limit with hyperonic matter; the limit is given in Paper 5 for kaon condensation.)

In Paper D it is suggested that as long as the vector mean field felt by the neutrons is strongly repulsive, scaling roughly linearly with density, hyperonic matter may be favored over kaon condensation. However, the vector mean field should drop at the chiral restoration transition, its role being taken over by perturbative gluon exchange. At this stage, kaon condensation should take over. In any case, the various transitions take place at relatively high temperatures,  $T \sim 50$  MeV, so one should really consider how to minimize the free energy. Clearly some pieces of both phases discussed above will be present at each density. Indeed, an energetic neutron at the top of its Fermi sea plus  $K^-$  meson can make a transition to a  $\Sigma^-$ . Thus, the excitation  $(\Sigma^- N^{-1})^{S=0}$  has the same quantum numbers as the  $K^-$ . Combining the kaon condensation and hyperonic matter scenarios, one might say that electrons are replaced by “kaesobars” at the phase transition (Brown *et al.* 1998) where the

kaesobar is

$$\propto (\Sigma^- N^{-1})^{S=0} + \sqrt{1 - \alpha^2} |K^-|. \quad (1)$$

We believe the important point is that there must be substantial strangeness present in neutron stars.

Paper A, the chiral perturbation theory calculation of kaon condensation by C.-H. Lee, is the first systematic calculation, order by order up to and including one-loop corrections, of kaon condensation. The loops do not, in fact, contribute much. The important term, which other workers in the field have not yet included (even though there have been a multitude of papers, mostly with results at variance with our work), is the range term, the role of which is most simply seen as an  $\omega$ -dependent renormalization of the  $KN$  sigma term; i.e.,

$$\Sigma_{KN} \rightarrow \left(1 - 0.37 \frac{\omega_{K^-}}{m_K^2}\right) \Sigma_{KN}. \quad (2)$$

The important point is that other workers set the value of the  $\Sigma_{KN}$  term from fits to the  $KN$  scattering amplitudes. At zero density  $\omega_{K^-} = m_K$ , so that they actually fit the quantity  $(1 - 0.37)\Sigma_{KN} = 0.63\Sigma_{KN}$ , rather than  $\Sigma_{KN}$  itself, according to us.

Equation (24) of Paper D shows that the kaon effective mass is given by

$$m_K^* \approx \left(1 - \rho \frac{(\Sigma_{KN})_{\text{eff}}}{f^2 m_K^2}\right), \quad (3)$$

where the approximate equality is used because the  $f$  and  $m_K$  in the denominator are set at their  $\rho = 0$  values (whereas it might be expected that these should be replaced by their *in-medium* values). The  $\omega_K$  is

$$\omega_K = \pm \frac{\omega_\pm}{m_K} V_K + \sqrt{k^2 + m_K^{*2}}, \quad (4)$$

where the first term on the right is the vector mean field, attractive for the  $K^-$  meson. In any case, we see that the  $m_K^*$  is substantially raised by using  $(\Sigma_{KN})_{\text{eff}}$  rather than  $\Sigma_{KN}$ .

What do we expect to happen at high densities? With temperature, lattice gauge calculations show that the vector repulsion drops an order of magnitude as the temperature goes through the chiral restoration phase transition (Brown & Rho 1996). Given that this occurs also for finite densities, the hyperonic matter phase would be disfavored.

Of course, with the drops in vector mean field the negative first term on the right hand side of Eq. (4) would decrease, disfavoring the kaon condensation. But one would expect the  $f$  in the denominator of Eq. (3) to go to zero, since  $f$  is the order parameter of the chiral restoration transition. Thus, we expect kaon condensation to occur at a density somewhat below that of chiral restoration. Our estimates in mean field approximation (Brown *et al.* 1998; Kim *et al.* 2000) are that chiral restoration occurs at  $\rho \sim 3\rho_0$ .

Paper C, by Li & Brown, is somewhat more empirical, arriving at a density dependence of the  $\omega_{K^-}(\omega = 0)$  as a function of density  $\rho$ . In that paper predictions were made for the

$K^-/K^+$  ratio in  $Ru + Ru$  at 1.69 GeV per nucleon. The experiment was carried out only later (Wiśniewski *et al.* 2001). It is clear from the experimental results that kaon medium effects are needed.

## References

- (1) Brown, G.E., Lee, C.-H., and Rapp, R. (1998), *Nuclear Physics A* **639**, 455c.
- (2) Brown, G.E., Li, G.Q., Rapp, R., Rho, M., and Wambach, J. (1998), *Acta Physica Polonica B* **9**, 2309.
- (3) Brown, G.E., and Rho, M. (1996), *Physics Reports* **269**, 1.
- (4) Glendenning, N.K. (2001), *Physics Reports* **342**, 393, 333.
- (5) Kim, Y., Rapp, R., Brown, G.E., and Rho, M. (2000), *Physical Review C* **62**, 1.
- (6) Prakash, M., Cooke, J.R., and Lattimer, J.M. (1995), *Physical Review D* **52**, 661.
- (7) Wiśniewski *et al.* (2001), *Eur. Phys. J. A* **9**, 515.

---

## Appendix A

# Kaon Condensation in Dense Stellar Matter

C.-H. Lee

Reprinted with permission from *Physics Reports*, 275 (1996) 255–341  
Copyright (1996) by Elsevier Science B.V.

### **Commentary by C.-H. Lee**

In the autumn of 1992, after returning from my six-month military service, I got a short message from Prof. G.E. Brown, whom I had never met before: “I heard that you’re working on kaon condensation in chiral perturbation theory in one-loop order. I would like to see the results.” In fact, at that time, I hadn’t even started my research on kaon condensation yet. I was just planning to do. This was how I started to work with Prof. G.E. Brown.

After writing a few papers together, finally I met G.E. Brown at INT, Seattle in the spring of 1995, just before finishing my Ph.D. thesis. Immediately, we realized that we had something in common, and he has been calling me “Thin and Hungry Boy” since then. I liked the nickname. In the autumn of 1995, I sent a copy of my Ph.D. thesis to G.E. Brown, and was invited to write a Physics Report based on my Ph.D. thesis.

This page is intentionally left blank

---

# KAON CONDENSATION IN DENSE STELLAR MATTER

**Chang-Hwan LEE**

*Center for Theoretical Physics, Seoul National University, Seoul 151-742, South Korea*



AMSTERDAM – LAUSANNE – NEW YORK – OXFORD – SHANNON – TOKYO



# Kaon condensation in dense stellar matter\*

Chang-Hwan Lee

*Center for Theoretical Physics, Seoul National University, Seoul 151-742, South Korea*

Received December 1995; editor: G.E. Brown

## Contents:

1. Introduction	258	6.2. Essence of kaon condensation	306
2. Kaon condensation in dense matter	260	6.3. "Scaled" chiral Lagrangians	308
2.1. Observed neutron star masses	260	6.4. $\rho_s$ versus $\rho_B$	313
2.2. Lowering maximum neutron star mass	260	7. Kaon condensation in other approaches	314
2.3. Baym's theorem	263	7.1. Mean-field approach of Brown and Rho	314
3. Effective chiral Lagrangian	266	7.2. Relativistic approach of Muto et al.	316
3.1. Chiral perturbation theory	266	7.3. Kaons on the hypersphere	318
3.2. Chiral transformation	268	8. Discussion	321
3.3. Renormalization group flow analysis	270	Appendix A. Neutron star in the core of supernova	322
3.4. Heavy baryon ChPT and chiral counting	271	A.1. Neutron star story	322
3.5. Effective chiral Lagrangian	273	A.2. Maximum neutron star mass	325
4. KN scattering amplitudes	275	Appendix B. Vertices and Feynman rules	329
4.1. On-shell amplitudes	275	Appendix C. Mass and wave function	
4.2. Off-shell amplitudes	281	renormalizations	332
4.3. Interpolating fields	285	Appendix D. Renormalization of $O(Q^3)$ counter	
4.4. The role of $A^*$ in kaon proton interactions	290	terms	334
5. Kaon self-energy and kaonic atom	293	Appendix E. $1/m$ corrections of baryon octet	335
5.1. Kaon self-energy	293	Appendix F. Decay width	337
5.2. Kaonic atom	296	Appendix G. $K^- p$ transition amplitudes	337
6. Kaon condensation	301	References	339
6.1. Kaon effective mass and critical density	301		

\* In part based on the Ph.D. thesis (August 95) of Seoul National University.

---

**Abstract**

Kaon-nucleon scattering, kaonic atom and kaon condensation are treated on the same footing by means of the chiral perturbation theory to the next-to-next-to-leading order ("N<sup>2</sup>LO"). Constraining the low-energy constants in the chiral Lagrangian by on-shell  $KN$  scattering lengths and kaonic atom data, the off-shell s-wave scattering amplitude up to one-loop order corresponding to N<sup>2</sup>LO and the critical density of kaon condensation up to *in-medium* two-loop order are computed. The effects on kaon-proton scattering of the quasi-bound  $\Lambda(1405)$  and on kaonic atoms and kaon condensation of  $\Lambda(1405)$ -proton-hole excitations through four-Fermi interactions are studied to all orders in density within the *in-medium* two-loop approximation. It is found that the four-Fermi interaction terms in the chiral Lagrangian play an essential role in providing attraction for kaonic atoms thereby inducing condensation, but the critical density is insensitive to the strength of the four-Fermi interaction. The prediction for the critical density is robust against changes of the parameters in the chiral Lagrangian, and gives – for "natural" values of the four-Fermi interactions – a rather low critical density,  $\rho_c \lesssim 4\rho_0$ , required to explain the maximum neutron star mass and the recent data of SN1987A. Once the BR (Brown-Rho) scaling sets in, the critical density,  $\rho_c < 3\rho_0$ , is completely insensitive to the parameters in which possible uncertainties of the theory lie.

---

## 1. Introduction

In a series of recent works [1,2], we have discussed kaon–nucleon and kaon–nuclear interactions in terms of a chiral perturbation expansion with the objective to predict within the framework of chiral effective Lagrangians the onset of kaon condensation in dense hadronic matter relevant to compact stars that are formed from the gravitational collapse of massive stars. This research was given a stronger impetus by a recent suggestion of Brown and Bethe [3] that if kaon condensates develop at a matter density  $\rho \lesssim 4 \rho_0$  (where  $\rho_0 \approx 0.16/\text{fm}^3 \approx 2.5 \times 10^{14} \text{ g/cm}^3$  is normal nuclear matter density) in the collapse of large stars, then low-mass black holes are likely to form in place of neutron stars of the mass greater than 1.5 times the solar mass  $M_\odot$ .

Ever since the first paper of Kaplan and Nelson [4], there have been numerous investigations on kaon condensation in dense neutron-star matter as well as in nuclear matter based both on effective chiral Lagrangians [1,5–8] and on phenomenological off-shell meson–nucleon interactions [9,10]. The two ways of addressing the problem gave conflicting results with the chiral Lagrangian approaches generally predicting a relatively low critical density,  $\rho_c \sim (2\text{--}4)\rho_0$ , while the phenomenological approaches based more or less on experimental inputs giving results that tend to exclude condensation at a low enough density to make it relevant in the collapse process.

Our objective in this review is to discuss the results of a consistent and systematic calculation in the context of chiral perturbation theory. This effort is largely motivated from the strong-interaction point of view by the recent success in confronting, in terms of chiral dynamics, the classic nuclear physics problems such as nuclear forces [11,12] and exchange currents [13–15]. This review addresses the problem of applying chiral perturbation theory to multi-hadron systems that contain the strange-quark flavor. While the standard problems of nuclear physics such as nuclear forces and exchange currents involve the chiral quarks  $u$  and  $d$  for which the mass scale involved is small compared with the typical QCD chiral symmetry breaking scale  $\Lambda_\chi \sim 1 \text{ GeV}$  so that a simultaneous expansion in derivatives and quark mass matrix is justified, here the strange-quark mass which is not small renders the expansion in the quark mass a lot more delicate and hence the low-order expansion highly problematic. This caveat has to be kept in mind in assessing the validity of the procedure we will adopt.

Another problem of potential importance is that both kaonic atom and kaon condensation introduce an additional scale, namely the matter density  $\rho$  or more precisely the Fermi momentum  $k_F$ . So far in low-order chiral perturbation calculations, the result depended on the order of density dependence included in the calculation. In fact, one of the important differences between the chiral Lagrangian approach and the phenomenological model approach arose at the order  $\rho^2$ . It is thus clear that one has to be consistent in the chiral counting, not only with respect to the usual expansion parameters practiced in free space, but also with respect to the density expansion. This then raises the question of how to treat the dynamics involved in the non-strange sector as well as in the strange sector. So far the dynamics in the non-strange sector is assumed to be given by what we know from nuclear phenomenology that is mostly given in terms of meson-theoretic approaches combined with many-body techniques, and perturbations in the strange direction are treated in terms of chiral Lagrangian at tree order or at most one-loop order. The problem with this is that there is no consistency between the two sectors as regards chiral symmetry and other constraints of QCD. Indeed so far nobody has been able to describe correctly nuclear ground state (including nuclear matter) starting from chiral Lagrangians, so one is justified to wonder how a low-order chiral Lagrangian calculation of kaon condensation without regard to the normal matter can be trusted. We cannot offer a solution to this

problem here but we will make an effort to point out the salient points that are closely related to this issue.

Following the recent development of nuclear chiral dynamics [11–16], we incorporate spontaneously broken chiral symmetry using the Jenkins–Manohar heavy-baryon chiral Lagrangian [17] as extended in [1] to  $\mathcal{O}(Q^3)$  to describe s-wave kaon nucleon scattering to one loop order in chiral perturbation theory (ChPT). In addition to the usual octet and decuplet baryons and the octet pseudo-Goldstone fields, the  $A(1405)$  was found to figure importantly in the kaon–nucleon process. This is because as is well known, the  $A(1405)$  (which we denote  $A^*$  for short) influences strongly the amplitude of the  $K^-p$  scattering near threshold and hence kaon–nuclear interactions in kaonic atom [18] and kaon condensation involving protons as in “nuclear stars.” We introduce this state as an elementary field as discussed in [1]. The reason for this is that first of all, the  $A^*$  is a bound state and hence cannot be described by a finite chiral perturbation expansion and secondly in the Callan–Klebanov skyrmion description [19], it is a configuration of a  $K^-$  wrapped by an  $SU(2)$  soliton and hence is as “elementary” as the even-parity  $A(1115)$  of the octet baryon.

In addition to these terms operating in the single-baryon sector, we need terms that involve multi-baryon fields in the Lagrangian for describing many-body systems. There have been discussions of four-Fermi interaction terms in non-strange sectors [11,12,14,20]. We find that in the s-wave kaon–nuclear sector, two such four-Fermi interaction terms involving  $A^*$  can intervene. In p-wave kaon–nuclear interactions, there can be more four-Fermi interactions as they can involve the entire battery of the octet and decuplet but we will not be concerned with them in this review.

By a straightforward extension of an amplitude whose parameters are fixed by on-shell kaon–nucleon scattering, we are able to predict an off-shell kaon–nucleon amplitude relevant for kaonic atom as well as kaon-condensation phenomena. The predicted off-shell amplitude was found to be in fair agreement with the phenomenological fit [21]. This off-shell amplitude provides the kaon self-energy in linear density approximation, equivalent to the usual optical potential approximation. The critical density obtained in this approximation is a bit higher than that obtained before at tree order but still in the regime quite relevant to the stellar collapse. If one goes beyond the linear density approximation which would be required in a simultaneous expansion in all the scales involved, four-Fermion interactions come into play. For the s-wave kaon condensation process, there are two independent four-Fermi interactions with arbitrary constants. In order to fix these parameters, we appeal to the recent data on kaonic atoms [18]. Since the experimental data are quite uncertain, we cannot pin them down unambiguously. However the physical quantities that we are interested in turn out to be rather insensitive to the free parameter.

The four-Fermi interactions – which are higher order in density – play an important role for giving rise to an attraction for kaonic atoms. This attraction certainly comes in for pushing the system toward condensation. However, they remain “irrelevant” and become suppressed at the kinematic regime in which condensation occurs [22,23]. As a consequence, their influence on the critical density is quite weak: The strength of the four-Fermi interactions, which cannot be pinned down precisely at present, does not figure importantly in the condensation phenomena. Using the renormalization group flow analysis of Lee, Rho and Sin [24], one can make a general argument on the stability in the strangeness direction of nuclear matter at high density.

Once the BR scaling [25,26] sets in in the Weinberg–Tomozawa term, the critical density is completely insensitive to the parameters in which possible uncertainties of the theory lie [1]. We found that the scaling subsumes higher order effects, thus suppressing higher chiral order effects of

the scaled Lagrangian. With the BR scaling, the Weinberg-Tomozawa term is more important than the scalar interaction, the KN sigma term. Even without the KN sigma term, we have the kaon condensation with the critical density  $\rho_c < 3\rho_0$ . Recently, Tatsumi et al. [27] obtained more or less the same results in a model-independent approach based on current algebra and PCAC.

This review is organized as follows. The observed neutron star masses and the basic properties of s-wave kaon condensation are discussed in Section 2. The effective chiral Lagrangian to  $\mathcal{O}(Q^3)$  in the chiral counting, consisting of the octet pseudo-Goldstone bosons and the octet and decuplet baryons that figure in our calculation, is given in Section 3. The renormalization group flow analysis [24] is also given in Section 3. In Section 4, we calculate to one-loop order, corresponding to N<sup>2</sup>LO, both on-shell and off-shell KN scattering amplitudes. Some issues regarding Adler's soft-meson conditions in chiral perturbation theory are also discussed. The analysis of kaonic atom data is treated in Section 5. Kaon condensation in our approach is summarized in Section 6. In Section 7, other approaches to kaon condensation are discussed, including the mean-field approach of Brown and Rho [28], the relativistic approach of Muto et al. [27,29] and the recent results of bound-state approach to the Skyrme model [30]. In Section 8, we mention some of the unsolved open issues in the problem as a conclusion. The formation of neutron star, which is believed to be the astrophysical implications of kaon condensation, and detailed formulas are collected in the appendices.

## 2. Kaon condensation in dense matter

It is believed that the low mass of well measured neutron star is the astrophysical implication of kaon condensation. In this Section, we review the observed neutron star masses and the possibilities of pion and kaon condensation. Baym's theorem, which is the basis of s-wave kaon condensation, is summarized at the end of this Section.

### 2.1. Observed neutron star masses

In Fig. 2.1, the measured neutron star masses are plotted. Most of all the neutron star masses are below  $1.5M_\odot$  except for Vela X-1 and 4U 1700-37. But by the recent analysis [31], the lower limit on the mass of Vela X-1 has been brought below  $1.5M_\odot$ . Furthermore there is an argument that 4U 1700-37 is a low mass black hole [31]. If this is the case, it is striking that all well measured neutron star masses lie below  $1.5M_\odot$ . This calls for theoretical arguments to lower the maximum neutron star mass. In Table 2.1, the maximum allowed mass of a neutron star is given for various equations of state [32]. Pion and Kaon condensations are good candidates to explain the low maximum neutron star mass.

### 2.2. Lowering maximum neutron star mass

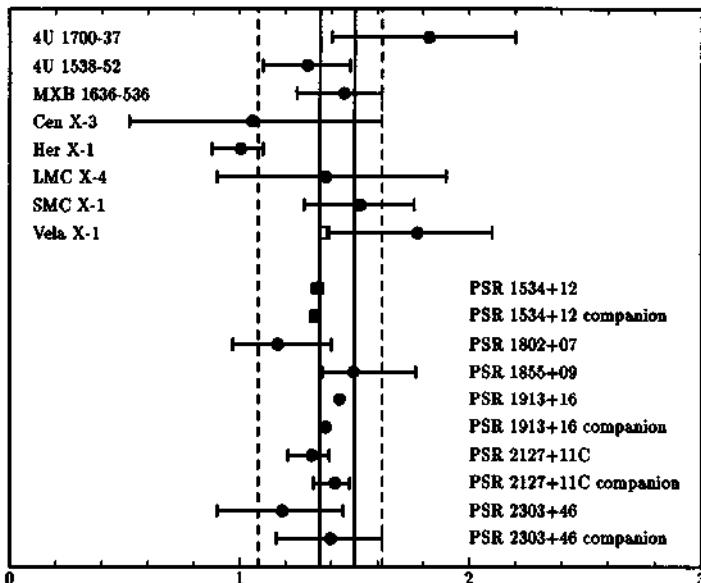
#### 2.2.1. Pion condensation

Since the pioneering work by Bahcall and Wolf [33], it has been known that a pion condensate, if it exists, will enhance tremendously the emissivity of the neutrinos from within the core of a neutron star. They considered the reaction

Table 2.1

The maximum mass of neutron star for various equations of state cited in [32].

Equation of state	Maximum mass ( $M_{\odot}$ )
$\pi$ or K condensate	< 1.5
R	1.6
BJ	1.9
TNI	2.0
TI	2.0
MF	2.7

Fig. 2.1. Neutron star mass in unit of solar mass ( $M_{\odot}$ ). Empty box of Vela X-1 is given by the recent analysis of van Kerwijk [31].

and its inverse reaction, in which pions were treated as real particles existing with a certain probability inside the neutron star. Subsequently Maxwell et al [34] carried out a detailed calculation on the neutrino emissivity in the presence of pion condensates. The reaction mechanism may be written symbolically as



where “ $\pi$ ” represent the pion condensate built in the quasiparticle states. The main conclusion of [34] was that even a small amount of pion condensates will cause a dramatic enhancement of neutrino emissivity, making the equation of state softer than standard calculations would predict. This soft equation of state could then lower the maximum neutron star mass as in Table 2.1.

However, the much-discussed p-wave pion condensation is now considered to be rather unlikely to take place at low density. Based on renormalization group flow equation, Lee et al. [24] showed that the Yukawa coupling term, responsible for p-wave condensation,  $\bar{\psi}\tau \cdot \pi\gamma_5\psi$ , becomes irrelevant after radiative corrections and cannot induce instability needed for a phase transition. Furthermore, the axial vector coupling constant  $g_A$  in nuclear medium is effectively quenched, roughly down to one. [35–37] The quenched axial-vector coupling then push the critical density to a higher density ( $> 5\rho_0$ ),

$$\rho_c \propto \frac{1}{g_A^2 - 1}, \quad (2.3)$$

where most of the approximations associated with the effective hadronic Lagrangians must have been broken down.

An s-wave pion condensation is also unlikely to occur since chiral symmetry protects pion mass (PCAC);  $\pi^-N$  interaction is repulsive. Even if s-wave condensation occurs, the effect would be negligible.

### 2.2.2. Kaon condensation

The next candidate process is the kaon ( $K^-$ ) condensation. (The kaon mass in free space is  $\sim 495$  MeV). It has been recently suggested that if kaon condensation sets in in the neutron star at about 3 or 4 times normal nuclear matter density, the new equation of state of dense matter is softer than conventional ones [3]. With this soft equation of state, the maximum mass of compact object formed in the collapse of large star is only about  $1.5 M_\odot$ , just the detected neutron-star mass bound. Their arguments extend the estimated range of main sequence masses, for which stars go into black holes, down to  $\sim 18M_\odot$ . Interestingly, this possibly includes SN1987A, with progenitor mass of  $18 \pm 2M_\odot$  (refer to Appendix A).

According to Kaplan and Nelson [4], the attraction needed for kaon condensation mainly comes from the KN sigma term

$$\delta M_K^2 \approx -\frac{\Sigma_{KN}}{f^2} \langle N^\dagger N \rangle + \dots, \quad (2.4)$$

where  $\Sigma_{KN}$  comes from the explicit chiral symmetry breaking,

$$\Sigma_{KN} \approx \frac{1}{2} (\hat{m} + m_s) \langle N | \bar{u}u + \bar{s}s | N \rangle. \quad (2.5)$$

In the collapse of large stars, a large negative chemical potential  $\mu_-$  is established in the core of the star. SN1987A is an example of such a star, whose progenitor mass was  $18 \pm 2M_\odot$  and the corresponding compact core mass  $1.5M_\odot$ . (See Appendix A.) During the collapse of the star, the neutrinos formed in the electron capture reaction



are trapped. So the chemical equilibrium leads to the following condition:

$$\mu_e + \mu_p = \mu_n + \mu_\nu. \quad (2.7)$$

However, as neutrinos leave the star, the neutrino chemical potential can be neglected. The final equilibrium condition is

$$\mu_e + \mu_p = \mu_n. \quad (2.8)$$

Then the electron chemical potential  $\mu_e = \mu_-$  is determined by this chemical equilibrium. For densities of  $2 \sim 4\rho_0$  it lies between

$$200 \text{ MeV} < \mu_e < 300 \text{ MeV}. \quad (2.9)$$

Note that  $\mu_e$  is large and increasing with density; this will be discussed in more detail in Section 6.

Since  $\epsilon_{K^-} \approx \sqrt{M_K^2 + \delta M_K^2}$ , Eq. (2.4), decreases with density, there will be some critical density  $\rho_c$  at which

$$\epsilon_{K^-}(\rho_c) = \mu_e(\rho_c). \quad (2.10)$$

Above this density it will be energetically favorable for neutrons to decay into  $(p, K^-)$  instead of  $(n, e^-)$ , or electrons to change into  $K^-$  through the reaction



As before, the neutrinos leave the star because the kaon condensation will occur after the era of neutrino trapping.

Considering the ambiguities in  $\Sigma_{KN}$ , they obtained the critical density for kaon condensation,

$$2.3\rho_0 \leq \rho_c \leq 3\rho_0. \quad (2.12)$$

After the pioneering work of Kaplan and Nelson, Thorsson et al. [7,8] obtained the desired critical densities of kaon condensation, using tree order chiral perturbation in heavy baryon limit at zero temperature. But because the tree order Lagrangian they used does not satisfy the KN scattering data, their result seems to be incomplete. In this review, we overcome the defect using chiral perturbation theory up to one loop order [1], which satisfies the constraints from the KN scattering data and kaonic atom data.

### 2.3. Baym's theorem

Early in 1973, G. Baym [38] obtained the equilibrium thermodynamic conditions of pion condensed system. The main point was that the condensed wave function of pion,  $\langle\phi(r, t)\rangle$ , must depend on time as

$$\langle\phi(r, t)\rangle \propto e^{-i\mu_\pi t}, \quad (2.13)$$

where  $\mu_\pi$  is the chemical potential of  $\pi^-$ . Since the condensed  $K^-$  wave function that we are interested in has a similar time dependence, we summarize Baym's theorem applied to the condensed  $K^-$  wave function. In this case,  $\langle\phi(r, t)\rangle$  corresponds to the kaon wave function.

In a condensed phase, the ground state energy  $E = \langle H \rangle$ , or the free energy  $F = E - TS$ , is a functional of  $\langle\phi(r, t)\rangle$  and  $\langle\pi(r, t)\rangle$ , and their complex conjugate  $\langle\phi^\dagger(r, t)\rangle = \langle\phi(r, t)\rangle^*$  and  $\langle\pi^\dagger(r, t)\rangle = \langle\pi(r, t)\rangle^*$ , where  $\pi$  is the momentum conjugate to  $\phi$ ,

$$\pi = \dot{\phi} + \frac{\delta \mathcal{L}_{int}}{\delta \dot{\phi}}. \quad (2.14)$$

In terms of the charge density,

$$\rho = i(\phi^\dagger \pi - \pi^\dagger \phi), \quad (2.15)$$

the ground state energy  $\langle H \rangle$  must be a minimum under an arbitrary variation

$$\delta\langle H \rangle = \mu_K \delta(\rho) = i\mu_K (\langle \phi^\dagger \rangle \delta(\pi) + \delta(\phi^\dagger) \langle \pi \rangle + \text{c.c.}). \quad (2.16)$$

The ground state energy  $\langle H \rangle$  is also a minimum under variation of  $\langle \phi \rangle$  and  $\langle \pi \rangle$ , hence

$$\delta\langle H \rangle = \langle \delta H / \delta \pi \rangle \delta(\pi) + \langle \delta H / \delta \phi \rangle \delta(\phi) + \text{c.c.} \quad (2.17)$$

Comparing the coefficients of  $\delta(\phi^\dagger)$  and  $\delta(\pi^\dagger)$  between Eqs. (2.16) and (2.17), we have

$$\langle \delta H / \delta \pi^\dagger \rangle \equiv \langle \dot{\phi} \rangle = -i\mu_K \langle \phi \rangle, \quad \langle \delta H / \delta \phi^\dagger \rangle \equiv -\langle \dot{\pi} \rangle = i\mu_K \langle \pi \rangle. \quad (2.18)$$

These conditions determine the time dependence of the condensed wave functions as

$$\langle \phi(r, t) \rangle = \exp(-iu_K t) \langle \phi(r) \rangle, \quad \langle \pi(r, t) \rangle = \exp(-iu_K t) \langle \pi(r) \rangle. \quad (2.19)$$

For neutral meson ( $\pi^0$ ), the chemical potential vanishes, so the condensed neutral meson field is constant in time.

The net charge for the condensed  $K^-$  is

$$\langle \rho_K \rangle_{cond} = -2\text{Im}(\langle \phi \rangle^* \langle \pi \rangle). \quad (2.20)$$

Neglecting terms containing  $\dot{\phi}$  in the interaction Lagrangian, the momentum conjugate  $\pi$  is simply  $\pi = \dot{\phi}$ , and thus

$$\langle \rho_K \rangle_{cond} = 2\mu_K |\langle \phi(r) \rangle|^2. \quad (2.21)$$

The equilibrium under the reaction



implies that  $\mu_K = \mu_n - \mu_p$ , the difference of the n and p chemical potentials. In neutron rich matter  $\mu_K > 0$ , hence the electric charge density associated with the kaon field is negative. This implies  $K^-$  condensation.

The ground state Hamiltonian density with relevant particles is

$$\mathcal{H}' = \mathcal{H} - \mu_n \rho_n + \mu_p \rho_p - \mu_e \rho_e - \mu_K \rho_K - \mu_\mu \rho_\mu. \quad (2.23)$$

Prior to kaon condensation, proton, neutron and electron are in equilibrium,



with the electron chemical potential  $\mu_e = \mu_n - \mu_p$ . Hence, in an equilibrium state, the kaon chemical potential is equal to that of the electron,  $\mu_K = \mu_e$ . In the same way,  $\mu_\mu = \mu_e$  when muon is present. From these conditions, the ground state energy depends on  $\mu_n$ ,  $\mu_K$  and  $\langle \phi \rangle$ ,

$$E' = E - \mu_n n_b + \mu_K n_q, \quad (2.25)$$

where the baryon density and the net charge density are given by

$$n_b = \sum_i B_i \langle \rho_i \rangle, \quad n_q = \sum_i Q_i \langle \rho_i \rangle, \quad (2.26)$$

where the sum is over all species  $i$ , and  $B_i$  and  $Q_i$  are the baryon and electric charge of species  $i$ . At fixed condensed kaon fields, we have

$$\left( \frac{\partial E'}{\partial \mu_n} \right)_{\mu_K} = -n_b, \quad \left( \frac{\partial E'}{\partial \mu_K} \right)_{\mu_n} = n_q. \quad (2.27)$$

Hence the charge neutrality condition is

$$\left( \frac{\partial E'}{\partial \mu_K} \right)_{\mu_n} = 0. \quad (2.28)$$

The differential of effective energy density is

$$\delta E'(\mu_n, \mu_K, |\langle \phi \rangle|^2) = -\delta(n_b \mu_n) + n_q \delta \mu_K + \frac{\partial E'}{\partial |\langle \phi \rangle|^2} \delta |\langle \phi \rangle|^2. \quad (2.29)$$

Note that the first term of RHS is related to beta equilibrium, and the second to charge neutrality. Without kaon condensation, these two conditions determine the chemical potential of electron and the proton fraction in dense matter.

From the kaon Lagrangian,

$$\mathcal{L} = \partial_\mu \phi^\dagger \partial^\mu \phi - M_K^2 \phi^\dagger \phi, \quad (2.30)$$

one can get the field equation of the condensed kaon by

$$(\mu_K^2 - M_K^2 + \nabla^2) \langle \phi(r) \rangle - J(r) = 0, \quad (2.31)$$

where the source of the condensed kaon field is

$$J(r) = -\langle \delta L_{int} / \delta \phi^\dagger(r) \rangle - i \mu_K \langle \delta L_{int} / \delta \dot{\phi}(r) \rangle. \quad (2.32)$$

This can be rewritten as

$$J(r) = \int d^3 r' \Pi(r, r'; \omega = \mu_K) \langle \phi(r') \rangle + \mathcal{O}(|\langle \phi \rangle|^2), \quad (2.33)$$

where  $\Pi$  is the  $K^-$  self-energy in medium.

The third term in Eq. (2.29) is simply related to the field equation,

$$\frac{\partial E'}{\partial |\langle \phi \rangle|^2} = -(\mu_K^2 - M_K^2 - \Pi(\mu_K)). \quad (2.34)$$

Hence, the condition of minimizing  $E'$  with respect to  $|\langle \phi \rangle|^2$  is equivalent to guaranteeing that  $\langle \phi \rangle$  obeys the field equation Eq. (2.31).

The  $K^-$  condensation threshold is given by the solution

$$\int D^{-1}(r, r', \mu_K) \langle \phi(r') \rangle d^3 r' = 0, \quad (2.35)$$

where

$$D^{-1}(\mathbf{r}, \mathbf{r}', \mu_K) = (\mu_K^2 - M_K^2 + \nabla^2) \delta(\mathbf{r}, \mathbf{r}') - \Pi(\mathbf{r}, \mathbf{r}', \mu_K). \quad (2.36)$$

For  $\langle \phi \rangle = 0$  prior to condensation,  $D$  has pole at the frequency  $\mu_K$ . Above threshold, the higher order terms in Eq. (2.33) cannot be neglected. Hence, one should solve the full equation Eq. (2.31),  $\mathcal{O}(\langle \phi \rangle^2)$ , in order to get the amplitude  $\langle \phi \rangle$ . In this review, since we are calculating the critical densities for kaon condensation, we need only the meson self-energy in medium, Eq. (2.36). In order to calculate the equation of state of dense matter beyond the critical density, we need the  $\mathcal{O}(\langle \phi \rangle^2)$  terms in Eq. (2.33).

### 3. Effective chiral Lagrangian

In this section, the building blocks of effective chiral Lagrangians are discussed. The basic properties of chiral symmetry and the QCD effective Lagrangian [39–42] are discussed in the first subsection, and the chiral transformation is summarized in the next subsection. The heavy baryon chiral theory and the associated chiral counting are discussed in the third subsection. In the last subsection, Subsection 3.5, the effective Lagrangian relevant to KN scattering is discussed.

#### 3.1. Chiral perturbation theory

##### 3.1.1. Basis to chiral symmetry

Consider the 3-flavor  $SU(3)$  quark fields,

$$q(x) = \begin{pmatrix} u(x) \\ d(x) \\ s(x) \end{pmatrix}. \quad (3.1)$$

The relevant QCD Lagrangian is

$$\mathcal{L}_{QCD} = i\bar{q}\gamma_\mu \mathcal{D}^\mu q - \bar{q}\mathcal{M}q - \frac{1}{2g^2} \text{Tr } G_{\mu\nu}G^{\mu\nu}. \quad (3.2)$$

where  $G_{\mu\nu}$  is the field strength of the gluon field,

$$G_{\mu\nu} = \partial_\mu G_\nu - \partial_\nu G_\mu - i[G_\mu, G_\nu], \quad (3.3)$$

and  $\mathcal{D}_\mu$  denotes the covariant derivatives,

$$\mathcal{D}_\mu q(x) = \partial_\mu q(x) - iG_\mu(x)q(x), \quad (3.4)$$

and  $\mathcal{M}$  is the quark mass matrix,

$$\mathcal{M} = \begin{pmatrix} m_u & 0 & 0 \\ 0 & m_d & 0 \\ 0 & 0 & m_s \end{pmatrix}. \quad (3.5)$$

In order to discuss the chiral symmetry, consider the massless case. The state of a free relativistic quark is completely characterized by its energy  $E$ , momentum  $\mathbf{p}$  and helicity  $\hat{h} = \boldsymbol{\sigma} \cdot \mathbf{p}/|\mathbf{p}|$  which

is identical to chirality with  $\gamma_5$ , the chirality operator. If we decompose the spinor into a right and left-handed components,

$$q = \frac{1}{2}(1 - \gamma_5)q + \frac{1}{2}(1 + \gamma_5)q = P_L q + P_R q = q_L + q_R, \quad (3.6)$$

where  $P_{L,R}$  are projection operators,  $P_{L,R}^2 = P_{L,R}$ ,  $P_L \cdot P_R = 0$ ,  $P_L + P_R = 1$  with the property

$$\frac{1}{2}\hat{h}q_{L,R} = \pm\frac{1}{2}q_{L,R}. \quad (3.7)$$

This shows that the states  $q_{L,R}$  are helicity eigenstates. In terms of those fields, the quark part of the Lagrangian equation (3.2) takes the form

$$\mathcal{L}(q_L, q_R) = i\bar{q}_L \gamma_\mu \partial^\mu q_L + i\bar{q}_R \gamma_\mu \partial^\mu q_R. \quad (3.8)$$

Note that the left and right-handed fermion modes do not communicate. One can apply separate  $SU(3)_{L,R}$  transformations which leave the Lagrangian invariant,

$$q_L \longrightarrow e^{ie_L + ie_L T^a} q_L, \quad q_R \longrightarrow e^{ie_R + ie_R T^a} q_R, \quad (3.9)$$

with  $T^a$  the general  $SU(3)$  generator. The Noether currents associated with these symmetries are

$$J_\mu^{0L,R} = \bar{q}_{L,R} \gamma_\mu q_{L,R}, \quad J_\mu^{aL,R} = \bar{q}_{L,R} \gamma_\mu T^a q_{L,R}. \quad (3.10)$$

Equivalently, one can construct the vector and axial-vector currents,

$$V_0^\mu = \bar{q} \gamma^\mu q, \quad A_0^\mu = \bar{q} \gamma^\mu \gamma_5 q; \quad V_a^\mu = \bar{q} \gamma^\mu T_a q, \quad A_a^\mu = \bar{q} \gamma^\mu \gamma_5 T_a q, \quad (3.11)$$

with  $J_i^{\mu L,R} = (V_i^\mu \pm A_a^\mu)/2$ . Taking the generally accepted picture, one of these currents is anomalous, i.e. the singlet axial current  $A_0^\mu$  fails to be conserved

$$\partial_\mu A_0^\mu = \frac{N_f}{8\pi^2} \text{Tr } G_{\mu\nu} \tilde{G}^{\mu\nu}. \quad (3.12)$$

Thus, for the massless 3-flavor QCD, the vacuum is invariant under the group

$$G = \underbrace{SU(3)_L \otimes SU(3)_R}_{\text{chiral group } G} \otimes U(1)_V, \quad (3.13)$$

where the quark number symmetry  $U(1)_V$  is realized as the baryon number, which we will not consider in what follows.

Chiral symmetry means that for massless fermions chirality is a constant of motion. To make chiral symmetry a viable concept for massive fermions, the corresponding eigenvalues of the mass matrix have to be small compared to a typical energy scale of the system under consideration. The typical scale of chiral symmetry breaking is  $\Lambda_\chi = 4\pi f_\pi = \mathcal{O}(1 \text{ GeV})$ . However, from hadron spectrum, it is generally accepted that the chiral symmetry is spontaneously broken down to its vectorial subgroup,

$$SU(3)_L \otimes SU(3)_R \longrightarrow SU(3)_V. \quad (3.14)$$

This spontaneous symmetry breaking generates 8 massless Goldstone bosons.

A fermion mass term explicitly break this symmetry since it mixes the left- and right-handed components,

$$\bar{q}\mathcal{M}q = \bar{q}_L\mathcal{M}q_R + \bar{q}_R\mathcal{M}q_L. \quad (3.15)$$

The divergence of the currents are

$$\partial_\mu V_a^\mu = \frac{1}{2}i\bar{q}(\mathcal{M}T_a - T_a\mathcal{M})q, \quad \partial_\mu A_a^\mu = \frac{1}{2}i\bar{q}(\mathcal{M}T_a + T_a\mathcal{M})q. \quad (3.16)$$

Note that for  $m_u = m_d = m_s$ , the vector currents are conserved.

### 3.1.2. Low energy effective theory

Starting from  $\mathcal{L}_{QCD}$ , one can write the generating functional,

$$\exp\{iZ[j, \eta, \bar{\eta}]\} = \int [dq][d\bar{q}][dG] \exp\left\{i \int d^4x \mathcal{L}_{QCD}[j, \eta, \bar{\eta}]\right\}, \quad (3.17)$$

where  $j$ ,  $\eta$  and  $\bar{\eta}$  are the external sources corresponding to mesons, baryons and anti-baryons, respectively. In principle, low energy physics can be described by this QCD generating functional. Practically, however, it is impossible to calculate the properties of the low energy physics using QCD Lagrangian.

Following S. Weinberg [43], one needs to construct effective Lagrangian of  $\mathcal{L}_{QCD}$ , which can describe the low energy physics. The starting point is the Goldstone nature of the pseudoscalar bosons. If the effective Lagrangian has the same symmetries as QCD Lagrangian and is constructed most generally, one can get the same low-energy generating functional,

$$\exp\{iZ[j, \eta, \bar{\eta}]\} = \int [d\xi][dB][d\bar{B}] \exp\left\{i \int d^4x \mathcal{L}_{eff}[\xi, j, B, \bar{B}, \eta, \bar{\eta}]\right\}, \quad (3.18)$$

where  $\xi$  is SU(3) Goldstone bosons and  $B$  is SU(3) baryon octet family. The small masses of Goldstone bosons make the power expansion of  $\mathcal{L}_{eff}$  possible,

$$\mathcal{L}_{eff} = \mathcal{L}_1 + \mathcal{L}_2 + \mathcal{L}_3 + \dots, \quad (3.19)$$

where  $\mathcal{L}_n$  depend on  $(Q/\Lambda)^n$ . Here  $Q$  is the characteristic momentum of Goldstone bosons ( $\sim M_\pi$ ), and  $\Lambda = 4\pi f_\pi \simeq 1\text{GeV}$ . Since the characteristic momentum  $Q$  is much smaller than the symmetry breaking scale  $\Lambda$ , one can expect that the higher power terms are much less important than lower terms. So the physics can be described by some low order terms.

In the following subsections, we discuss the explicit construction of a low energy effective theory, Chiral Perturbation Theory (ChPT).

### 3.2. Chiral transformation

The relatively small masses of pseudo-scalar mesons( $\bar{q}q$ ), compared with typical baryon( $qqq$ ) masses, imply their Goldstone nature, which can be most easily analyzed on the basis of an effective Lagrangian. Our basic assumption is the pattern of spontaneous chiral symmetry breaking

$$G \equiv SU(3)_L \otimes SU(3)_R \xrightarrow{SCSB} H \equiv SU(3)_V. \quad (3.20)$$

Let us denote  $\phi^a$  ( $a = 1, \dots, 8$ ) the coordinates describing the Goldstone fields in the coset space  $G/H$ . The change of the Goldstone coordinates under a chiral transformation  $g \equiv (g_L, g_R) \in G$  is given by

$$U(\phi) \xrightarrow{G} g_L U(\phi) g_R^\dagger. \quad (3.21)$$

Note that we construct the Lagrangian which is invariant under the global chiral transformation  $g$ .

The usual non-linear representation of  $U$  is given by

$$U = \exp(\sqrt{2}iM/f), \quad (3.22)$$

where the meson octet field is

$$M = \begin{pmatrix} \frac{1}{\sqrt{2}}\pi^0 + \frac{1}{\sqrt{6}}\eta & \pi^+ & K^+ \\ \pi^- & -\frac{1}{\sqrt{2}}\pi^0 + \frac{1}{\sqrt{6}}\eta & K^0 \\ K^- & \bar{K}^0 & -\sqrt{\frac{2}{3}}\eta \end{pmatrix}. \quad (3.23)$$

Note that  $U(\phi)$  transforms linearly under the chiral group, but the induced transformations on the Goldstone bosons are highly non-linear.

The lowest order chiral Lagrangian with  $U(\phi)$  which satisfy the chiral symmetry is

$$\mathcal{L} = \frac{1}{4}f^2 \text{Tr } \partial U \partial U^\dagger. \quad (3.24)$$

Because of the finite quark and Goldstone boson masses, the explicit chiral symmetry breaking mass term is introduced by imposing the transformation

$$\mathcal{M} \xrightarrow{G} g_L M g_R^\dagger. \quad (3.25)$$

The chiral Lagrangian with mass term is

$$\mathcal{L} = \frac{1}{4}f^2 \text{Tr } \partial U \partial U^\dagger + \frac{1}{2}f^2 r \text{Tr } [\mathcal{M}(U + U^\dagger - 2) + \text{h.c.}] \quad (3.26)$$

Before introducing the baryon fields, we can decompose the meson field  $U$

$$U(\phi) \equiv \xi_L(\phi) \xi_R^\dagger(\phi). \quad (3.27)$$

Without loss of generality, we can take a choice of coset representative such that  $\xi_L(\phi) = \xi_R^\dagger(\phi) \equiv \xi(\phi)$ , where  $\xi(\phi) \equiv (\xi_L(\phi), \xi_R(\phi)) \in G$ . The change of the Goldstone coordinates under a chiral transformation  $g \equiv (g_L, g_R) \in G$  is given by

$$\xi_L(\phi) \xrightarrow{G} g_L \xi_L(\phi) h^\dagger(\phi, g), \quad \xi_R(\phi) \xrightarrow{G} g_R \xi_R(\phi) h^\dagger(\phi, g), \quad (3.28)$$

where  $h(\phi, g) \in H$  is a compensating local transformation. Note that the same transformation  $h(\phi, g)$  occurs in the left and right sectors, making  $U(\phi)$  independent of  $h$  field.

There are many choices of incorporating baryon fields in chiral effective Lagrangian [44]. In this review we take  $SU(3)$  representation of baryon fields,

$$B = \begin{pmatrix} \frac{1}{\sqrt{2}}\Sigma^0 + \frac{1}{\sqrt{6}}\Lambda & \Sigma^+ & p^+ \\ \Sigma^- & -\frac{1}{\sqrt{2}}\Sigma^0 + \frac{1}{\sqrt{6}}\Lambda & n^0 \\ \Xi^- & \Xi^0 & -\sqrt{\frac{2}{3}}\Lambda \end{pmatrix}. \quad (3.29)$$

The transformation property of the baryon field is

$$B \xrightarrow{G} hBh^\dagger. \quad (3.30)$$

In order to construct a chiral-invariant Lagrangian with baryon fields, we need meson fields which transform the same way as the baryons. The possible fields and their transformation properties are

$$\begin{aligned} \mathcal{M}_\pm &= \xi_L^\dagger \mathcal{M} \xi_R \pm \xi_R^\dagger \mathcal{M}^\dagger \xi_L \xrightarrow{G} h \mathcal{M}_\pm h^\dagger, & V_\mu &= \frac{1}{2} \left\{ \xi_L^\dagger \partial_\mu \xi_L + \xi_R^\dagger \partial_\mu \xi_R \right\} \xrightarrow{G} h V_\mu h^\dagger + h \partial_\mu h^\dagger, \\ D_\mu B &= \partial_\mu B + [V_\mu, B] \xrightarrow{G} h(D_\mu B)h^\dagger, & A_\mu &= \frac{1}{2} i \left\{ \xi_L^\dagger \partial_\mu \xi_L - \xi_R^\dagger \partial_\mu \xi_R \right\} \xrightarrow{G} h A_\mu h^\dagger. \end{aligned} \quad (3.31)$$

Then baryon fields can couple with meson field through  $A_\mu$  or  $V_\mu$ , making the Lagrangian chirally invariant. Both  $A_\mu$  and  $V_\mu$  have derivatives on meson fields; only the derivative couplings with baryon fields are possible in a chiral invariant Lagrangian. We can now construct the most general chiral-invariant Lagrangian involving baryon fields, using  $B, \mathcal{M}_\pm, D_\mu, A_\mu$  at a given order. The lowest chiral order general Lagrangian involving baryon fields is

$$\begin{aligned} \mathcal{L} = & +\text{Tr } \bar{B}(i\not{\partial} - m_B)B + i\text{Tr } \bar{B}\gamma^\mu[V_\mu, B] \\ & + \text{DTr } \bar{B}\gamma^\mu\gamma_5\{A_\mu, B\} + \text{FTr } \bar{B}\gamma^\mu\gamma_5[A_\mu, B]. \end{aligned} \quad (3.32)$$

### 3.3. Renormalization group flow analysis

Given such an effective field theory, we can make a general argument on the stability in various flavor directions of nuclear matter at high density [24]. This can be done along the line of arguments developed for condensed matter physics by Shankar [22] and Polchinski [23] using renormalization group flow. We sketch the essential argument following Lee, Rho and Sin [24].

What we are interested in is whether the system in question develops instability along the direction of strangeness and if so, by which physical mechanism. This analysis will not give us the critical density. The critical density will be calculated by using chiral perturbation theory. For this purpose we will focus on the kaon frequency near the electron chemical potential. By Baym's theorem [38], one can identify the kaon chemical potential associated with charge conservation,  $\mu_K$ , with the electron chemical potential,  $\mu_e$ , which we shall simply write  $\mu$  in what follows. This means that we will be looking at the vicinity of  $\omega \sim \mu$  in the kaon dispersion formula. We shall assume that

$$|\omega - \mu| \ll \mu. \quad (3.33)$$

We assume that nucleons in nuclear matter are in Fermi-liquid state with the Fermi energy  $\mu_F$  and the Fermi momentum  $k_F$ . Define  $\psi$  as the nucleon field fluctuating around the Fermi surface such that the momentum integral has a cut-off  $\Lambda_N$ ,

$$k_F - \Lambda_N < |k| < k_F + \Lambda_N. \quad (3.34)$$

Kaons can interact with the nucleons through three-point functions of the  $KN\bar{N}$  type (Yukawa interaction) and through four-point interactions of the  $KK\bar{N}\bar{N}$  type. We shall consider S-wave kaon-nucleon interactions, for which the Yukawa interaction can be ignored. A generic action involving the nucleon field  $\psi$  and the kaon field  $\Phi$  can then be written, schematically, as

$$\begin{aligned} S = & \int d\omega d^3q \Phi^*(\omega, q) (\omega - q^2/2\mu_K) \Phi(\omega, q) - \int d\omega d^3q \tilde{M}_K \Phi^* \Phi \\ & + \int (d\omega d^3q)^2 (d\epsilon d^3k)^2 h \Phi^* \Phi \psi^\dagger \psi \delta^4(\omega, \epsilon, q, k) \\ & + \int d\epsilon d^3k \psi^\dagger (\epsilon - \epsilon(k)) \psi + g \int (d\epsilon d^3k)^4 \psi^\dagger \psi^\dagger \psi \psi \delta^4(\epsilon, k), \end{aligned} \quad (3.35)$$

where  $\tilde{M} = (M_K^2 - \mu^2)/2\mu$  and  $h$  and  $g$  are constants. The four-Fermi interaction with the coefficient  $g$  stands for Fermi-liquid interactions in nuclear matter. (In nuclear matter, one can have four such terms because of the nucleon spin and isospin degrees of freedom. We need not specify them for our purpose.) This is a toy action but it is generic in that the results of ChPT we will obtain below can be put into this form.

The renormalization group flow of this action can be analyzed in the following way. Since we are assuming that nuclear matter is a Fermi-liquid fixed point, fluctuations in the non-strange direction in the nucleon sector are stable: The four-Fermi interaction  $g$  is irrelevant or at best marginal. Fluctuations in the strange direction involve the kaon field  $\Phi$ . Suppose we have integrated out all the high-frequency modes above the cut-off  $\Lambda$  measured with respect to  $\mu$ . We are interested in the stability of the system under the renormalization group transformation  $\Lambda \rightarrow s\Lambda$  ( $s < 1$ ) as  $s \rightarrow 0$ . A scaling analysis shows that the interaction term  $h$  is irrelevant while the “mass term”  $\tilde{M}$  is relevant. The renormalization group-flow of the “mass term” and the interaction term  $h$  can be readily written down and solved [24] (with  $t = -\ln s$ ),

$$\tilde{M}(t) = \left( \tilde{M}_0 - \frac{Dh_0}{1+a} \right) e^t + \frac{Dh_0}{1+a} e^{-at}, \quad (3.36)$$

with

$$h(t) = h_0 e^{-at}, \quad h_0 \geq 0, \quad (3.37)$$

where  $D = (3(1+\alpha^2)\alpha/2\mu)\rho_N > 0$ ,  $\alpha = \Lambda/k_F > 0$  and  $a = 1/2$ . We see from Eq. (3.36) that as  $s \rightarrow 0$  for which  $h \rightarrow 0$ ,  $\tilde{M}$  changes sign for some  $(\tilde{M}_0, h_0 \geq 0)$ . Thus although irrelevant, an attractive interaction  $h_0$  determines the direction of the mass flow whereas it is the “mass term” that drives the system to instability.

### 3.4. Heavy baryon ChPT and chiral counting

When baryons are present, ChPT in relativistic form is not as firmly formulated as when they are absent [45]. The reason is that the baryon mass  $m_B$  is  $\sim \Lambda_K \sim 1$  GeV, the chiral symmetry breaking scale. It is more expedient, therefore, to redefine the baryon field so as to remove the mass from the baryon propagator

$$B_\nu = e^{im_B \gamma^\nu v^\mu x_\mu} P_+ B, \quad (3.38)$$

where  $P_+ = (1 + \gamma \cdot v)/2$  and write the baryon four-momentum

$$p_\mu = m_B v_\mu + k_\mu, \quad (3.39)$$

where  $k_\mu$  is the small residual momentum indicating the baryon being slightly off-shell. When acted on by a derivative, the baryon field  $B_v$  yields a term of  $O(k)$ . The propagator of the baryon octet fields is simplified to

$$\frac{i}{v \cdot k}. \quad (3.40)$$

The loop calculation is simplified since there are no gamma matrices. The spin operator  $S_v^\mu$  is defined by

$$v \cdot S_v = 0, \quad S_v^2 B_v = -\frac{3}{4} B_v, \quad \{S_v^\mu, S_v^\nu\} = \frac{1}{2}(v^\mu v^\nu - g^{\mu\nu}), \quad [S_v^\mu, S_v^\nu] = i\epsilon^{\mu\nu\alpha\beta} v_\alpha (S_v)_\beta. \quad (3.41)$$

In the baryon rest frame, the spin operator  $S_v$  reduces to the usual spin operator  $\sigma/2$ .

Chiral perturbation theory in terms of  $B_v$  and Goldstone bosons ( $\xi$ ) is known as “heavy-baryon (HB) ChPT” [46]. HBChPT consists of making chiral expansion in derivatives on Goldstone boson fields,  $\partial_M/\Lambda_\chi$ , and on baryon fields,  $\partial_{B_v}/m_B$ , and in the quark mass matrix,  $\kappa M/\Lambda_\chi^2$ . In the meson sector, this is just what Gasser and Leutwyler [47] did for  $\pi\pi$  scattering. In the baryon sector, consistency with this expansion requires that the chiral counting be made with  $B^\dagger(\dots)B$ , not with  $\bar{B}(\dots)B$ . This means that in medium, it is always the baryon density  $\rho(r)$  that comes in and *not* the scalar density  $\rho_s(r)$ . This will be further discussed later.

Following Weinberg [11], we organize the chiral expansion in power  $Q^\nu$  where  $Q$  is the characteristic energy/momentum scale we are looking at ( $Q \ll \Lambda_\chi$ ) and

$$\nu = 4 - N_n - 2C + 2L + \sum_i \Delta_i, \quad (3.42)$$

with the sum over  $i$  running over the vertices that appear in the graph and

$$\Delta_i = d_i + \frac{1}{2}n_i - 2. \quad (3.43)$$

Here  $\nu$  gives the power of small momentum (or energy) for a process involving  $N_n$  nucleon lines,  $L$  number of loops,  $d_i$  number of derivatives (or powers of meson mass) in the  $i$ th vertex,  $n_i$  number of nucleon lines entering into  $i$ th vertex and  $C$  is the number of separate connected pieces of the Feynman graph. Chiral invariance requires that  $\Delta_i \geq 0$ , so that the leading power is given by  $L = 0$ ,  $\nu = 4 - N_n - 2C$ .

As an example, consider  $KN$  scattering. The leading term here is the tree graph with  $\nu = 1$  and with  $N_n = C = 1$ . The next order terms are  $\nu = 2$  tree graphs with  $\Delta = 1$  that involves two derivatives or one factor of the mass matrix  $M$ . From  $\nu = 3$  on, we have loop graphs contributing together with appropriate counter terms.

In considering kaon-nuclear interactions as in the case of kaon condensation, we need to consider the case with  $N_n \geq 2$  and  $C \geq 2$ . In dealing with many-body systems, one can simply fix  $4 - N_n$  and consider  $C$  explicitly. For instance if one has two nucleons (for reasons mentioned below, this is sufficient, with multinucleon interactions being suppressed), then we have  $4 - N_n = 2$  but  $C$  can be 2 or 1, the former describing a kaon scattering on a single nucleon with a spectator nucleon propagating without interactions and the latter a kaon scattering irreducibly on a two-nucleon complex. Thus intrinsic  $n$ -nucleon processes are suppressed compared with  $(n - 1)$ -nucleon processes by at least  $O(Q^2)$ . This observation will be used later for arguing that four-Fermi interactions are negligible in

kaon condensation. This is somewhat like the suppression of three-body nuclear forces [11] and of three-body exchange currents [14] in chiral Lagrangians.

### 3.5. Effective chiral Lagrangian

We start by writing down the effective chiral Lagrangian that we shall use in the calculation. Let the characteristic momentum/energy scale that we are interested in be denoted  $Q$ . The standard chiral counting orders the physical amplitude as a power series in  $Q$ , say,  $Q^\nu$ , with  $\nu$  an integer. To leading order, the kaon–nucleon amplitude  $T^{KN}$  goes as  $\mathcal{O}(Q^1)$ , to next order as  $\mathcal{O}(Q^2)$  involving no loops and to next to next order (*i.e.*, N<sup>2</sup>LO) at which one-loop graphs enter as  $\mathcal{O}(Q^3)$ . Following Jenkins and Manohar [17], we denote the velocity-dependent octet baryon fields  $B_v$ , the octet meson fields  $\exp(i\pi_a T_a/f) \equiv \xi$ , the velocity-dependent decuplet baryon fields  $T_v^\mu$ , the velocity four-vector  $v_\mu$  and the spin operator  $S_v^\mu$  ( $v \cdot S_v = 0$ ,  $S_v^2 = -3/4$ ), the vector current  $V_\mu = [\xi^\dagger, \partial_\mu \xi]/2$  and the axial-vector current  $A_\mu = i\{\xi^\dagger, \partial_\mu \xi\}/2$ , and write the Lagrangian density to order  $Q^3$ , relevant for the low-energy s-wave scattering, as

$$\begin{aligned} \mathcal{L}^{(1)} &= \text{Tr } \bar{B}_v (iv \cdot \mathcal{D}) B_v - \bar{T}_v^\mu (iv \cdot \mathcal{D} - \delta_T) T_{v,\mu} + 2D \text{Tr } \bar{B}_v S_v^\mu \{A_\mu, B_v\} \\ &\quad + 2F \text{Tr } \bar{B}_v S_v^\mu [A_\mu, B_v] + C(\bar{T}_v^\mu A_\mu B_v + \bar{B}_v A_\mu T_v^\mu) + 2H \bar{T}_v^\mu (S_v \cdot A) T_{v,\mu}, \\ \mathcal{L}^{(2)} &= a_1 \text{Tr } \bar{B}_v \chi_+ B_v + a_2 \text{Tr } \bar{B}_v B_v \chi_+ + a_3 \text{Tr } \bar{B}_v B_v \text{Tr } \chi_+ \\ &\quad + d_1 \text{Tr } \bar{B}_v A^2 B_v + d_2 \text{Tr } \bar{B}_v (v \cdot A)^2 B_v + d_3 \text{Tr } \bar{B}_v B_v A^2 + d_4 \text{Tr } \bar{B}_v B_v (v \cdot A)^2 \\ &\quad + d_5 \text{Tr } \bar{B}_v B_v \text{Tr } A^2 + d_6 \text{Tr } \bar{B}_v B_v \text{Tr } (v \cdot A)^2 + d_7 \text{Tr } \bar{B}_v A_\mu \text{Tr } B_v A^\mu \\ &\quad + d_8 \text{Tr } \bar{B}_v (v \cdot A) \text{Tr } B_v (v \cdot A) + d_9 \text{Tr } \bar{B}_v A_\mu B_v A^\mu + d_{10} \text{Tr } \bar{B}_v (v \cdot A) B_v (v \cdot A), \\ \mathcal{L}^{(3)} &= c_1 \text{Tr } \bar{B}_v (iv \cdot \mathcal{D})^3 B_v + c_2 \text{Tr } \bar{B}_v (iv \cdot \mathcal{D}) (i^2 \mathcal{D}^\mu \mathcal{D}_\mu) B_v \\ &\quad + g_1 \text{Tr } \bar{B}_v A_\mu (iv \cdot \vec{\mathcal{D}}) A^\mu B_v + g_2 \text{Tr } B_v A_\mu (iv \cdot \vec{\mathcal{D}}) A^\mu \bar{B}_v + g_3 \text{Tr } \bar{B}_v v \cdot A (iv \cdot \vec{\mathcal{D}}) v \cdot A B_v \\ &\quad + g_4 \text{Tr } B_v v \cdot A (iv \cdot \vec{\mathcal{D}}) v \cdot A \bar{B}_v + g_5 (\text{Tr } \bar{B}_v A_\mu \text{Tr } (iv \cdot \vec{\mathcal{D}}) A^\mu B_v - \text{Tr } \bar{B}_v A_\mu (iv \cdot \vec{\mathcal{D}}) \text{Tr } A^\mu B_v) \\ &\quad + g_6 (\text{Tr } \bar{B}_v v \cdot A \text{Tr } B_v (iv \cdot \vec{\mathcal{D}}) v \cdot A - \text{Tr } \bar{B}_v v \cdot A (iv \cdot \vec{\mathcal{D}}) \text{Tr } B_v v \cdot A) \\ &\quad + g_7 \text{Tr } \bar{B}_v [v \cdot A, [iD^\mu, A_\mu]] B_v + g_8 \text{Tr } B_v [v \cdot A, [iD^\mu, A_\mu]] \bar{B}_v \\ &\quad + h_1 \text{Tr } \bar{B}_v \chi_+ (iv \cdot \mathcal{D}) B_v + h_2 \text{Tr } \bar{B}_v (iv \cdot \mathcal{D}) B_v \chi_+ + h_3 \text{Tr } \bar{B}_v (iv \cdot \mathcal{D}) B_v \text{Tr } \chi_+ \\ &\quad + l_1 \text{Tr } \bar{B}_v [\chi_-, v \cdot A] B_v + l_2 \text{Tr } \bar{B}_v B_v [\chi_-, v \cdot A] + l_3 [\text{Tr } \bar{B}_v \chi_-, \text{Tr } B_v v \cdot A], \end{aligned} \quad (3.44)$$

where the covariant derivative  $\mathcal{D}_\mu$  for baryon fields is defined by

$$\begin{aligned} \mathcal{D}_\mu B_v &= \partial_\mu B_v + [V_\mu, B_v], \\ \mathcal{D}_\mu T_{v,abc}^\nu &= \partial_\mu T_{v,abc}^\nu + (V_\mu)_a^d T_{v,dbc}^\nu + (V_\mu)_b^d T_{v,adc}^\nu + (V_\mu)_c^d T_{v,abd}^\nu, \end{aligned} \quad (3.45)$$

$\delta_T$  is the  $SU(3)$  invariant decuplet–octet mass difference, and

$$\chi_\pm \equiv \xi \mathcal{M} \xi \pm \xi^\dagger \mathcal{M} \xi^\dagger, \quad (3.46)$$

with  $\mathcal{M} = \text{diag}(m_u, m_d, m_s)$  the quark mass matrix that breaks chiral symmetry explicitly. The explicit decuplet fields are given in Appendix B. There are many other terms involving the decuplet

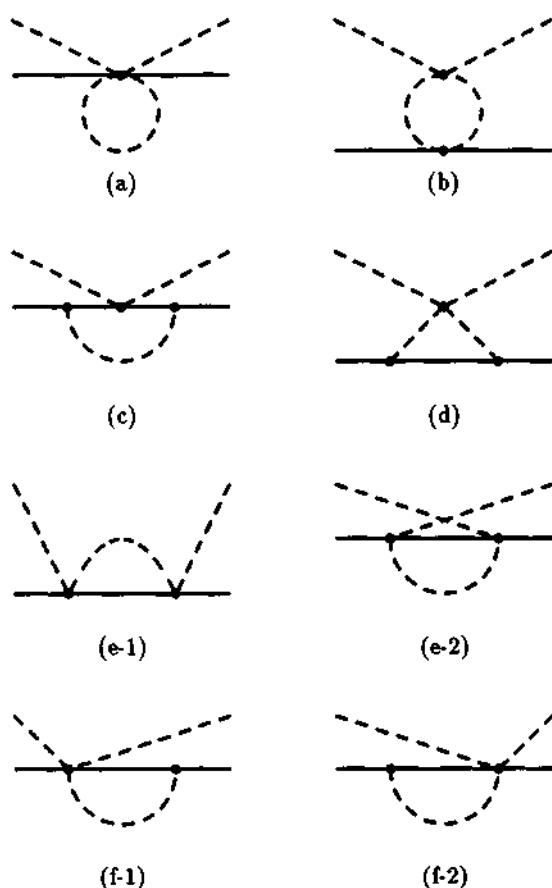


Fig. 3.1. One-loop Feynman diagrams contributing to  $K^\pm N$  scattering: The solid line represents baryons (nucleon for the external and octet and decuplet baryons for the internal line) and the broken line pseudo-Goldstone bosons ( $K^\pm$  for the external and  $K$ ,  $\pi$  and  $\eta$  for the internal line). There are in total thirteen diagrams at one loop, but for s-wave  $KN$  scattering, for reasons described in the text, we are left with only six topologically distinct one-loop diagrams.

that one can write down but we have written only those that enter in the calculation. Among the many parameters that figure in the Lagrangian, a few can be fixed right away. For instance, we will simply fix the constants  $F$  and  $D$  at tree order since to  $\mathcal{O}(Q^3)$  that we will be interested in, they are not modified. We shall use  $D = 0.81$  and  $F = 0.44$ . The constant  $C$  can also be fixed at this stage from the decay process  $A(1230) \rightarrow N\pi$ . We shall use  $|C|^2 (\approx 2.58)$ . Of course, the flavor  $SU(3)$  can be substantially broken as we will discuss later, so one cannot take this value too seriously. The determination of all other constants  $a_i, \dots, l_i$  (or more precisely the combinations thereof) will be described below.

The number of parameters may appear daunting to some readers but the situation turns out to be much simpler than what it looks. As we will see later, once the constants are grouped into an appropriate form, there remain only four parameters for on-shell  $K^\pm N$  amplitudes. These parameters can be fixed on-shell by the four s-wave scattering lengths. Off-shell, however, one parameter remains

free but the off-shell amplitude turns out to be rather insensitive to the one free parameter. This drastic simplification can be understood easily as follows. First of all, the heavy-fermion formalism (in short HFF) makes those subleading terms (*i.e.*, terms with  $\nu \geq 2$ ) involving the spin operator  $S_\mu$  vanish, since they are proportional to  $S \cdot q$ ,  $S \cdot q'$ , or  $S \cdot q S \cdot q'$ , all of which are identically zero. As a consequence, there are no contributions to the *s*-wave meson–nucleon scattering amplitude from one-loop diagrams in which the external meson lines couple to baryon lines through the axial vector currents. This leaves only six topologically distinct one-loop diagrams, Fig. 3.1 (out of thirteen in all), to calculate for the *s*-wave meson–nucleon scattering, apart from the usual radiative corrections in external lines. Since we are working to  $\mathcal{O}(Q^3)$ , only  $\mathcal{L}^{(1)}$  enters into the loop calculation. Loops involving other terms can contribute at  $\mathcal{O}(Q^4)$  or higher. The next term  $\mathcal{L}^{(2)}$  contributes terms at order  $\nu = 2$ , that is, at tree order. These will be determined by the  $KN$  sigma term and terms that could be calculated by resonance saturation. There are some uncertainties here as we shall point out later, but they turn out to be quite insignificant in the results. The next terms in  $\mathcal{L}^{(3)}$  remove the divergences in the one-loop contributions and involve two finite counter terms – made up of two linear combinations of the many parameters appearing in the Lagrangian – that are to be determined empirically. As we will mention later, these constants are determined solely by isospin-odd amplitudes, the loop contribution to isospin-even amplitudes being free of divergences.

The vertices of the Feynman graph and renormalization properties, relevant to  $KN$  scattering and kaon condensation, are summarized in the appendices.

## 4. KN scattering amplitudes

### 4.1. On-shell amplitudes

The complete on-shell *s*-wave  $KN$  scattering amplitudes calculated to N<sup>2</sup>LO ( $\mathcal{O}(Q^3)$ ) [1] read

$$\begin{aligned} a_0^{K^\pm p} &= \frac{\eta}{f^2} \left[ \mp M_K + (\bar{d}_s + \bar{d}_v) M_K^2 + \{(L_s + L_v) \pm (\bar{g}_s + \bar{g}_v)\} M_K^3 \right] + \delta a_{A^*}^{K^\pm p}, \\ a_0^{K^\pm n} &= \frac{\eta}{f^2} \left[ \mp \frac{1}{2} M_K + (\bar{d}_s - \bar{d}_v) M_K^2 + \{(L_s - L_v) \pm (\bar{g}_s - \bar{g}_v)\} M_K^3 \right], \end{aligned} \quad (4.1)$$

where  $\eta$  is  $m_B/(4\pi(m_B + M_K))$ ,  $\bar{d}_s$  is the t-channel isoscalar contribution of  $\mathcal{O}(Q^2)$ , and  $\bar{d}_v$  is the t-channel isovector one of  $\mathcal{O}(Q^2)$ :

$$\begin{aligned} \bar{d}_s &= -\frac{1}{2B_0}(a_1 + 2a_2 + 4a_3) + \frac{1}{4}(d_1 + d_2 + d_7 + d_8) + \frac{1}{2}(d_3 + d_4) + d_5 + d_6, \\ \bar{d}_v &= -\frac{1}{2B_0}a_1 + \frac{1}{4}(d_1 + d_2 + d_7 + d_8), \end{aligned} \quad (4.2)$$

with  $B_0 = M_K^2/(\hat{m} + m_s)$  where  $M_K$  is the kaon mass and  $\hat{m} = (m_u + m_d)/2$ . Here  $\delta a_{A^*}^{K^\pm p}$  is the contribution from the  $A^*$  to be specified below and  $L_s(L_v)$  is the finite crossing-even t-channel isoscalar (isovector) one-loop contribution,

$$L_s M_K = \frac{1}{128\pi f^2 M_K^2} \left( \frac{1}{3}(D - 3F)^2 (M_\pi^2 + 3M_\eta^2) M_\eta - 9M_K^2 \sqrt{M_\eta^2 - M_K^2} \right) \approx -0.109 \text{ fm},$$

$$\begin{aligned}
L_v M_K = \frac{1}{128\pi f^2 M_K^2} & \left( -\frac{1}{3}(D+F)(D-3F)(M_\pi^2 + 3M_\eta^2)(M_\pi + M_\eta) - 3M_K^2 \sqrt{M_\eta^2 - M_K^2} \right. \\
& \left. - \frac{1}{6}(D+F)(D-3F)(M_\pi^2 + 3M_\eta^2)(M_\pi^2 + M_\eta^2) \int_0^1 \frac{1}{\sqrt{(1-x)M_\pi^2 + xM_\eta^2}} \right) \\
& \approx +0.021 \text{ fm}, \tag{4.3}
\end{aligned}$$

where  $f = 93 \text{ MeV}$  and physical masses are used to obtain the numbers. The quantity  $\bar{g}_s(\bar{g}_v)$  is the crossing-odd t-channel isoscalar (isovector) contribution from one-loop plus counter terms which after the dimensional regularization specified in Appendix D, takes the form

$$\bar{g}_{s,v} = \alpha'_{s,v} + \beta'_{s,v} + \frac{1}{32\pi^2 f^2 M_K^2} \left( \gamma_{s,v} + \sum_{i=\pi,K,\eta} \delta_{s,v}^i \ln \frac{M_i^2}{\mu^2} \right), \tag{4.4}$$

where  $\mu$  is the arbitrary scale parameter that enters in the dimensional regularization and  $\alpha$  and  $\beta$  are contributions from the counter terms in  $\mathcal{O}(Q^3)$ , and  $\gamma$  and  $\delta$  are finite loop contributions. On-shell, one can combine  $\alpha$  and  $\beta$  into one set of parameter to be determined from experiments. Off-shell, however, they are multiplied by a different power of the frequency  $\omega$  as indicated in subsection 4.2.2 and hence represent two independent parameters. This introduces one unfixed parameter in the off-shell case. However it turns out that the off-shell amplitudes are rather insensitive to the precise values of these constants, so we set (somewhat arbitrarily)  $\alpha'_{s,v} \approx \beta'_{s,v}$  in our calculation, Fig. 4.1.

The explicit forms of  $\alpha$ ,  $\beta$ ,  $\gamma$  and  $\delta$  are given in Appendix D. It should be noted that while  $\alpha$  and  $\beta$  are  $\mu$ -dependent,  $\bar{g}$  is scale-independent. ( $\gamma$  and  $\delta$  are scale-independent numbers.) Thus if one fixes  $\bar{g}$  from experiments, then for a given  $\mu$ , one can fix  $\alpha + \beta$  at a fixed  $\mu$ . Equivalently, we can separate out the specific  $\mu$ -dependent terms so as to cancel the  $\ln \mu$  term in Eq. (4.4), thereby defining  $\mu$ -independent constants  $\alpha'$  and  $\beta'$ ,

$$\alpha'_{s,v} + \beta'_{s,v} = \alpha'_{s,v} + \beta'_{s,v} - \frac{1}{32\pi^2 f^2 M_K^2} \sum_{i=\pi,K,\eta} \delta_{s,v}^i \ln \frac{M_i^2}{\mu^2}, \tag{4.5}$$

and determine  $\alpha'_{s,v}$  and  $\beta'_{s,v}$  from experiments. From now on when we go off-shell, we will drop the primes understanding that we are dealing with the  $\mu$ -independent parameters. (On-shell, this subtlety is not relevant since we can work directly with  $\bar{g}$  of Eq. (4.4).)

#### 4.1.1. $A^*$ contribution

To understand the role of the  $A^*$ , we observe that the measured scattering lengths are repulsive in all channels except  $K^-n$  [48,49].

$$\begin{aligned}
a_0^{K^+p} &= -0.31 \text{ fm}, & a_0^{K^-p} &= -0.67 + i0.63 \text{ fm}, \\
a_0^{K^+n} &= -0.20 \text{ fm}, & a_0^{K^-n} &= +0.37 + i0.57 \text{ fm}. \tag{4.6}
\end{aligned}$$

Although the experimental  $K^-N$  scattering lengths are given with error bars, the available  $K^+N$  data are not very well determined. Since both are used in fitting the parameters of the Lagrangian, we do not quote the error bars here and shall not use them for fine-tuning. For our purpose, we do not need great precision in the data as the results are extremely robust against changes in the parameters.

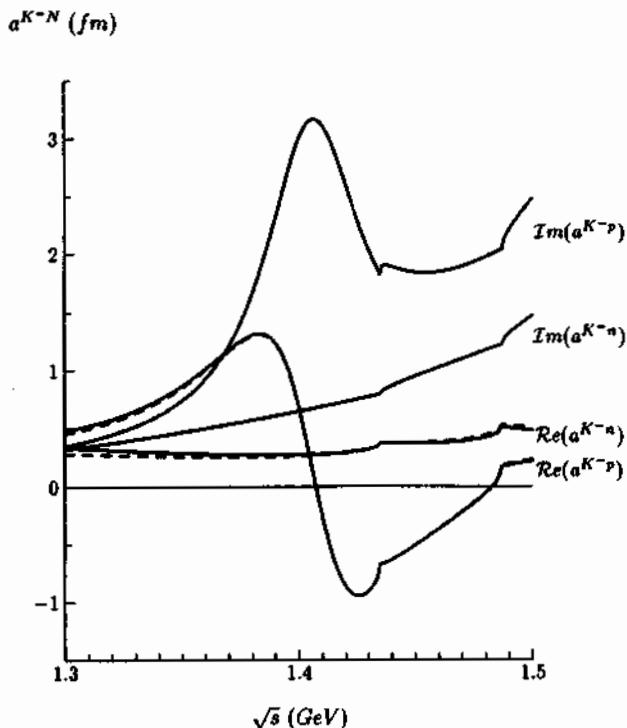


Fig. 4.1.  $K^-N$  amplitudes as function of  $\sqrt{s}$ : These figures correspond to Eqs. (23) and (24) with  $\tilde{g}_{A^*}^2 = 0.25$ ,  $\Gamma_{A^*} = 50$  MeV and  $\alpha_{p,n} = \beta_{p,n}$ , fixed in the way described in the text. The first kink corresponds to the  $KN$  threshold and the second around 1.5 GeV to  $\sqrt{s} = m_B + M_\eta$ , for  $M_\eta \approx 547$  MeV. The solid line of real part corresponds to  $Z = -0.5$  and the dashed line corresponds to  $Z = 0.15$ , and the imaginary parts are independent of  $Z$  value.

The repulsion in  $K^-p$  scattering cannot be explained from Eq. (4.1) without the  $A^*$  contribution. In fact it is well known that the contribution of the  $\Lambda(1405)$  bound state gives the repulsion required to fit empirical data for s-wave  $K^-p$  scattering [1,50]. As mentioned, we may introduce the  $A^*$  as an elementary field. To leading order in the chiral counting, it takes the form

$$\mathcal{L}_{A^*} = \bar{A}_v^*(iv \cdot \partial - m_{A^*} + m_B) A_v^* + (\sqrt{2} g_{A^*} \text{Tr}(\bar{A}_v^* v \cdot AB_v) + \text{h.c.}) . \quad (4.7)$$

The coupling constant  $g_{A^*}$  can be fixed by the decay width  $A^* \rightarrow \Sigma\pi$  [1] if one ignores  $SU(3)$  breaking

$$g_{A^*}^2(pK^-) \approx g_{A^*}^2(\Sigma\pi) \approx 0.15. \quad (4.8)$$

This is what one would expect at tree order. The corresponding width comes out to be

$$\Gamma_{A^*} \approx 50 \text{ MeV}, \quad (4.9)$$

which is in agreement with the empirical width of  $\Lambda(1405)$ . This together with perturbative unitarity suggests that the amplitude (4.13) with  $m_{A^*}$  replaced by a complex mass could be used to take into account the  $\mathcal{O}(Q^3)$  effect.

If one wants to go to one-loop order [51] corresponding to  $\mathcal{O}(Q^3)$  at which  $SU(3)$  breaking enters, then we encounter two counter terms  $h_{1,2}^*$ ,

$$\mathcal{L}^{*3} = h_1^* \sqrt{2} \bar{A}_v^* \text{Tr} (\chi_v v \cdot A B_v) + h_2^* \sqrt{2} \bar{A}_v^* \text{Tr} (\chi_v B_v v \cdot A) + \text{h.c.} \quad (4.10)$$

and the renormalized coupling to  $\mathcal{O}(Q^3)$  will take the form

$$\begin{aligned} g_{A^*}(\Sigma\pi) &= g_{A^*} + \sum_{i=\pi, K, \eta} \alpha_i^{\Sigma\pi} \ln \frac{M_i^2}{\mu^2} + \beta^{\Sigma\pi} + 2(h_1^{*\prime} + h_2^{*\prime}) \hat{m}, \\ g_{A^*}(pK^-) &= g_{A^*} + \sum_{i=\pi, K, \eta} \alpha_i^{pK^-} \ln \frac{M_i^2}{\mu^2} + \beta^{pK^-} + 2(h_1^{*\prime} m_s + h_2^{*\prime} \hat{m}), \end{aligned} \quad (4.11)$$

where  $\alpha_i$  and  $\beta$  are calculable loop contributions. In [51], Savage notes that if one ignores the counter terms, then the finite log terms would imply that  $g_{A^*}(pK^-)$  would come out to be considerably smaller than  $g_{A^*}(\Sigma\pi)$ , presumably due to an  $SU(3)$  breaking. In our approach we choose to pick the constants from experiments since we see no reason to suppose that the counter terms are zero and furthermore the presently available data [49] give

$$(g_{A^*}^2(pK^-))_{exp} \approx 0.25, \quad (4.12)$$

which is *bigger* than the  $SU(3)$  value (4.8). If we take the complex mass for  $m_{A^*}$ , the real part of Eq. (4.13) becomes smaller. Hence, in order to compensate for the reduced amplitude, we found that it was necessary to take a larger effective coupling,  $\tilde{g}_{A^*}^2 \approx 0.25$  at  $\mathcal{O}(Q^3)$ . This difference may perhaps be justified by the fact that  $SU(3)$  breaking which enters at one-loop order could induce the coupling constants to differ by as much as 30% [52]. We will take this value in our calculation.

It should also be mentioned that the Callan–Klebanov Skyrme predicts a value close to (4.8) [53]. The  $A^*$  contribution to the kaon–proton scattering amplitude is now completely determined,

$$\delta a_{A^*}^{K^\pm p} = -\frac{m_B}{4\pi f^2(m_B + M_K)} \left[ \frac{\tilde{g}_{A^*}^2 M_K^2}{m_B \mp M_K - m_{A^*}} \right]. \quad (4.13)$$

To one-loop order, the  $A^*$  mass picks up an imaginary part through the graph  $A^* \rightarrow \Sigma\pi \rightarrow A^*$ . In our numerical work we will take  $m_{A^*}$  to be complex. The presence of the imaginary part explains that the empirical coupling constant (4.12) is bigger than the  $SU(3)$  value (4.8).

#### 4.1.2. On-shell constraints

The data for  $K^\pm n$  are not well determined, so we cannot pin down the parameters in a quantitative way. Nonetheless, we can use these data to constrain the low-energy constants in our effective chiral Lagrangian. To tree order, the amplitudes are real, the imaginary parts of the amplitudes appearing at  $\mathcal{O}(Q^3)$  involving loop graphs. In Eq. (4.1),  $f$  is the meson decay constant in chiral limit and the difference between  $f_\pi$  and  $f_K$  is of order  $Q^2$ . Therefore we are allowed to simply take  $f \approx f_\pi \approx 93$  MeV, the physical value. Now requiring consistency with the data at tree level leads to

$$(\bar{d}_s - \bar{d}_v)_{emp} \approx (0.05-0.06) \text{ fm}, \quad (\bar{d}_s + \bar{d}_v)_{emp} \approx 0.13 \text{ fm}, \quad \tilde{g}_{A^*}^2 = 0.15, \quad (4.14)$$

with  $m_{A^*} = 1.405$  GeV. Note that  $\tilde{g}_{A^*}^2 = 0.15$  is consistent with empirical value Eq. (4.8), which corresponds to  $g_{A^*}^2/4\pi \approx 0.3$  in the conventional notation [49].

**Table 4.1**

Scattering lengths from three leading order contributions for the empirical value of the constant  $g_A^2 = 0.25$ . Also shown is the contribution from  $A^*$ .

$g_{A^*}^2 = 0.25$	$\mathcal{O}(Q)$	$\mathcal{O}(Q^2)$	$\mathcal{O}(Q^3)$	$A^*$
$a^{K^+ p}$ (fm)	-0.588	0.316	-0.114	0.076
$a^{K^- p}$ (fm)	0.588	0.316	-0.143	-1.431
$a^{K^+ n}$ (fm)	-0.294	0.277	-0.183	0.000
$a^{K^- n}$ (fm)	0.294	0.277	-0.201	0.000

Let us call these “tree-order empirical.” Although it is difficult to make a precise statement due to the uncertainty in the data and the  $\Lambda(1405)$  parameters, the above values provide a persuasive indication that *the net contribution of order  $Q^2$  or higher for  $K^\pm p$  ( $K^\pm n$ ) is attractive and amounts to  $\approx 33\%$  (26–31%) of the strength given by the leading-order vector coupling.* This feature will be reconfirmed at one-loop order.

In one-loop order, we are left with four parameters in (4.1),  $d_s$ ,  $d_v$ ,  $\bar{g}_s$  and  $\bar{g}_v$ , which we can determine with the four experimental (real parts of) scattering lengths (4.6). The results are

$$d_s \approx 0.201 \text{ fm}, \quad d_v \approx 0.013 \text{ fm}, \quad \bar{g}_s M_K \approx 0.008 \text{ fm}, \quad \bar{g}_v M_K \approx 0.002 \text{ fm}. \quad (4.15)$$

The scattering amplitudes in each chiral order are given in Table 4.1. One sees that while the order  $Q$  and order  $Q^2$  terms are comparable, the contribution of order  $Q^3$  is fairly suppressed relative to them. As a whole, the subleading chiral corrections are verified to be consistent with the “naturalness” condition as required of effective field theories. Using other sets of values of  $f$ ,  $D$  and  $F$  does not change  $L_s$  and  $L_v$  significantly and leave unaffected our main conclusion. As expected, the  $A^*$  plays a predominant role in  $K^- p$  scattering near threshold. This indicates that it will be essential in describing kaon–nuclear interactions, e.g., kaonic atoms.

#### 4.1.3. Imaginary part of $a^{K^- p}$

By introducing the complex  $\Lambda(1405)$  mass, we found that the chiral series for the real part of the scattering amplitudes converges more rapidly, but the resulting imaginary  $K^- p$  scattering amplitude is too big at  $\mathcal{O}(Q^3)$  compared to the experimental value. This raises questions as to whether the imaginary mass of  $A^*$  induces the double counting of the imaginary  $K^- p$  scattering amplitudes. In this section we show that a large imaginary part of the  $K^- p$  scattering amplitude can be understood simply by introducing the  $A^*$  as a fundamental field [2]. Furthermore the presence of a large imaginary part does not affect the convergence of the real part.

The imaginary parts of  $K^- N$  scattering can be obtained from the three diagrams of Fig. 4.2. They can be generated from the integral

$$\begin{aligned} & \frac{1}{i} \int \frac{d^4 k}{(2\pi)^4} \frac{1}{v \cdot k - \delta m} \frac{1}{(k - \omega)^2 - M^2} \\ &= i \frac{1}{4\pi} \sqrt{(\omega - \delta m)^2 - M^2} \otimes \theta(\omega - \delta m - M) + \dots \end{aligned} \quad (4.16)$$

where the ellipsis represents the real part. Noting that three different channels contribute to Fig. 4.2a:  $(\Sigma^+ \pi^-)$ ,  $(\Sigma^0 \pi^0)$  and  $(\Lambda \pi^0)$ , we find the imaginary parts from this graph to be unambiguously given by

$$\begin{aligned} a_{(a)}^{K^- p} &= i\eta \frac{5}{4} f(\tilde{\omega}_\Sigma) + i\eta \frac{3}{4} f(\tilde{\omega}_\Lambda) \approx +i 0.191 \text{ fm}, \\ a_{(a)}^{K^- n} &= i\eta f(\tilde{\omega}_\Sigma) + i\eta \frac{3}{2} f(\tilde{\omega}_\Lambda) \approx +i 0.272 \text{ fm}, \end{aligned} \quad (4.17)$$

where

$$\begin{aligned} \eta &\equiv \frac{1}{4\pi(1+M_K/m_p)}, \quad f(\omega) \equiv \frac{(\omega+M_K)^2}{16f^4} \frac{1}{4\pi} \sqrt{\omega^2 - M_\pi^2}, \\ \tilde{\omega}_{\Sigma,\Lambda} &= (M_K + m_p - m_{\Sigma,\Lambda}). \end{aligned} \quad (4.18)$$

Here the physical masses of  $p$ ,  $\Sigma$  and  $\Lambda$  (1115 MeV) are used. Comparing with the empirical values, Eq. (4.6), we see that they amount to only about 20% for  $K^- p$  and about 50% for  $K^- n$ .<sup>1</sup> We clearly need other contributions to explain the large imaginary amplitudes.

When the  $A^*$  is introduced as an elementary field, we have additional contributions from Fig. 4.2b and c. Two channels contribute in each diagram of Fig. 4.2b:  $\pi^0 \Sigma^0$ ,  $\pi^- \Sigma^+$  and three channels contribute to the diagram of Fig. 4.2c:  $(\Sigma^\pm \pi^\mp)$ ,  $(\Sigma^0 \pi^0)$ , with the same coupling. The resulting imaginary parts are

$$\begin{aligned} a_{(b)}^{K^- p} &= -i\eta \frac{3g^2}{16\pi f^4} \frac{M_K(M_K + \tilde{\omega}_\Sigma)\tilde{\omega}_\Sigma}{M_K + m_p - m_{A^*}} \sqrt{\tilde{\omega}_\Sigma^2 - M_\pi^2} \approx -i 0.710 \text{ fm}, \\ a_{(c)}^{K^- p} &= i\eta \frac{3g_{A^*}^4}{4\pi f^4} \frac{M_K^2 \tilde{\omega}_\Sigma^2}{(M_K + m_p - m_{A^*})^2} \sqrt{\tilde{\omega}_\Sigma^2 - M_\pi^2} \approx +i 2.294 \text{ fm}. \end{aligned} \quad (4.19)$$

Here  $a_{(c)}^{K^- p}$  is related with the decay width of  $A^*$ ,

$$\Gamma(\tilde{\omega}_\Sigma) = \frac{6g_{A^*}^2}{4\pi f^2} \tilde{\omega}_\Sigma^2 \sqrt{\tilde{\omega}_\Sigma^2 - M_\pi^2} \approx 93 \text{ MeV}. \quad (4.20)$$

Substituting  $\tilde{\omega}_\Sigma$  into  $\tilde{\omega}_0 = (m_{A^*} - m_\Sigma)$ , we obtain  $\Gamma(\tilde{\omega}_0) \approx 57 \text{ MeV}$ , as expected for the coupling  $g_{A^*}^2 = 0.15$  needed to produce the experimental decay width<sup>2</sup>.

The total imaginary amplitude then is

$$a_{tot}^{K^- p} = a_{(a)}^{K^- p} + a_{(b)}^{K^- p} + a_{(c)}^{K^- p} \approx +i 1.771 \text{ fm}. \quad (4.21)$$

This is somewhat too big compared with the experimental scattering amplitudes, Eq. (4.6). However we are not doing a fine-tuning here. What is noteworthy is that the large empirical imaginary amplitude cannot be understood without  $A^*$ .

<sup>1</sup> In [1], we neglected the mass differences of baryon octet. If we neglect the mass differences between  $p$ ,  $\Sigma$ ,  $\Lambda$ , the imaginary parts become  $a_{(a)}^{K^- p} = +i0.635 \text{ fm}$  and  $a_{(a)}^{K^- n} = i0.794 \text{ fm}$ , which are nearly the same as empirical values Eq. (4.6).

<sup>2</sup> The difference between the calculated decay width 57 MeV and the empirical one originates from the difference between the relativistic and nonrelativistic approach (heavy baryon limit). Comparison with the calculation that uses a relativistic chiral Lagrangian is discussed in Appendix F.

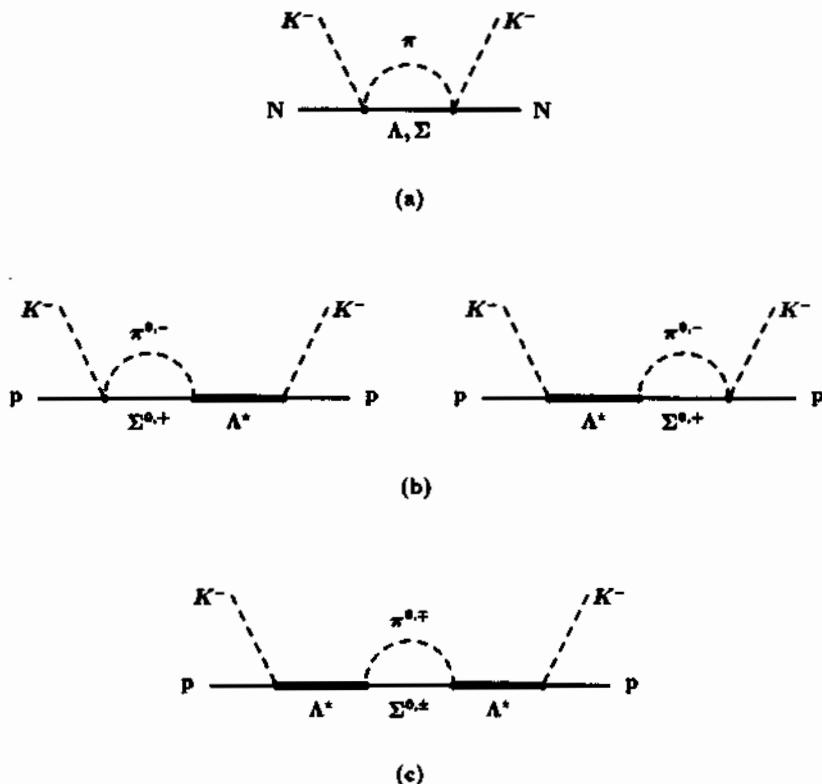


Fig. 4.2. Diagrams contributing to the imaginary parts of  $K^- N$  scattering.

This total imaginary amplitude is nearly the same as that with the imaginary  $\Lambda^*$  mass at  $\mathcal{O}(Q^3)$ . In conclusion, the introduction of the imaginary mass of  $\Lambda^*$  generates no serious problem. In this review, we keep the imaginary mass of  $\Lambda^*$  at  $\mathcal{O}(Q^3)$ .

#### 4.2. Off-shell amplitudes

We now turn to off-shell s-wave  $K^-$  forward scattering from static nucleons. The kinematics involved are  $t = 0$ ,  $q^2 = q'^2 = \omega^2$ ,  $s = (m_B + \omega)^2$  with an arbitrary (off-shell)  $\omega$ .

In going off-shell, we need to separate different kinematic dependences of the constant  $\bar{d}_{s,v}$  of Eq. (4.2) which consists of what would correspond to the KN sigma term  $\Sigma_{KN}$  at tree order,  $\sigma_{KN} (\approx \Sigma_{KN}) \equiv -\frac{1}{2}(\bar{m} + m_1)(a_1 + 2a_2 + 4a_3)$  involving the quark mass matrix  $\mathcal{M}$ , and the  $d_i$  terms containing two time derivatives. The  $\sigma_{KN}$  term, which gives an attraction, is a constant independent of the kaon frequency  $\omega$ , whereas the  $d_i$  terms which are repulsive are proportional to  $\omega^2$  in s-wave. In kaonic atoms and kaon condensation, the  $\omega$  value runs down from its on-shell value  $M_K$ . This means that as  $\omega$  goes down, the attraction stays unchanged and the repulsion gets suppressed. Thus while for on-shell amplitudes, they can be obtained independently of any assumptions, they need to be separated for off-shell amplitudes that we are interested in. If we could determine the KN sigma term from experiments, there would be no ambiguity. The trouble is that the sigma term extracted

from experiments is not precise enough to be useful. The presently available value ranges

$$\Sigma_{KN} \sim (200\text{--}400) \text{ MeV}. \quad (4.22)$$

Here we choose to separate the two components by estimating the contributions to  $d_i$  from the leading  $1/m_B$  corrections with the octet and decuplet intermediate states in the relativistic Born graphs. As suggested by the authors of Ref. [54] for  $\pi N$  scattering, we might assume that the counter terms  $d_i$  could equally be saturated by such intermediate states.

#### 4.2.1. $1/m_B$ correction

In writing down the off-shell amplitudes with the parameters determined on-shell, we have assumed that the  $\mathcal{O}(Q^2)$  terms quadratic in  $\omega$  can be calculated by saturating the Born graphs with the octet and decuplet intermediate states.

The derivation of the  $1/m_B$  corrections by the baryon octet is summarized in Appendix E, and the results are given below, Eq. (4.24). In calculating the decuplet contributions in the lowest order in  $1/m_B$  from the Feynman graphs in relativistic formulation, one encounters the usual off-shell non-uniqueness characterized by a factor  $Z$  in the decuplet–nucleon–meson vertex when the decuplet is off-shell,

$$\mathcal{L} = C \bar{T}_\mu (g^{\mu\nu} - (Z + \frac{1}{2}) \gamma^\mu \gamma^\nu) A_\nu B + \text{h.c.} \quad (4.23)$$

This gives the  $Z$  dependence in the  $1/m_B$  corrections to  $\bar{d}_{s,v}$ ,

$$\begin{aligned} \bar{d}_{s,\pm} &= -\frac{1}{48m_B} ((D + 3F)^2 + 9(D - F)^2) - \frac{|C|^2}{12m_B} (Z + \frac{5}{2})(\frac{1}{2} - Z), \\ \bar{d}_{v,\pm} &= -\frac{1}{48m_B} ((D + 3F)^2 - 3(D - F)^2) + \frac{|C|^2}{36m_B} (Z + \frac{5}{2})(\frac{1}{2} - Z), \end{aligned} \quad (4.24)$$

where the first part of RHS is the contribution from the baryon octet contributions, and second from the decuplet contribution. We shall fix  $Z$  from  $\pi N$  scattering ignoring the  $SU(3)$  breaking which occurs at higher chiral order than we need. The precise value of  $Z$  turns out not to be important for both kaonic atoms and kaon condensation. Now our chiral Lagrangian gives, at tree order, the isoscalar  $\pi N$  scattering length

$$a_{\pi N}^{(+)} = \frac{1}{4\pi f^2(1 + M_\pi/m_B)} (2\bar{D}_{\pi N} M_\pi^2 + \Sigma_{\pi N}), \quad (4.25)$$

where  $\Sigma_{\pi N}$  is the  $\pi N$  sigma term ( $\approx 45 \text{ MeV}$ ) and  $\bar{D}_{\pi N} = \frac{1}{4}(d_1 + d_2) + \frac{1}{2}(d_5 + d_6)$ . If we take the empirical value of  $a_{\pi N}^{(+)} = -0.01M_\pi^{-1}$ , we obtain

$$\bar{D}_{\pi N}^{\text{emp}} \approx -1.29/m_B \approx -0.27 \text{ fm}. \quad (4.26)$$

The  $1/m$  correction to  $\pi N$  scattering is

$$\bar{D}_{\pi N}^{\frac{1}{m}} = -\frac{(D + F)^2}{16} \frac{1}{m_B} - \frac{2}{9}|C|^2 \frac{1}{m_B} (Z + \frac{5}{2})(\frac{1}{2} - Z). \quad (4.27)$$

For the constants  $D$ ,  $F$  and  $C$  used in this review, this formula reproduces the experimental value (4.26) for  $Z \simeq -0.5$ . This will be used in our calculation of the off-shell amplitudes.

Alternatively one could fix  $Z$  at one-loop order as in [54]. This however affects our results insignificantly. To see this, take  $Z \approx 0.15$  obtained in Ref. [54]. The resulting off-shell  $KN$  amplitudes are given by the dotted lines in Fig. 4.1. This represents less than 1 % change in the kaon self-energy, an uncertainty that can be safely ignored in the noise of other uncertainties inherent in heavy-baryon chiral perturbation theory.

This is somewhat like saturating the dimension-four counter terms  $L_i$  in the chiral Lagrangian by resonances, a prescription which turns out to be surprisingly successful. The reason for believing that this might be justified in the present case is that the  $\mathcal{O}(Q^2)$  terms are not affected by chiral loops, so must represent the degrees of freedom that are integrated out from the effective Lagrangian. But there are no known mechanisms that would contribute to  $d_i$  other than the baryon resonances. When computed by resonance saturation, the contributions go like  $1/m_B$ . However this is not to be taken as  $1/m_B$  corrections that arise as relativistic corrections to the static limit of a relativistic theory. The HFF as used in chiral perturbation theory does not correspond merely to a non-relativistic reduction although at low orders, they are equivalent. To be more specific, imagine starting with the following relativistic Lagrangian density

$$\mathcal{L}_{rel} = \dots + e_i \text{Tr} [\bar{B} \gamma_\mu A^\mu \gamma_\nu A^\nu B] + \dots + f_i \text{Tr} [D_\mu \bar{B} A^\mu A^\nu D_\nu B] + \dots \quad (4.28)$$

In going to the heavy-baryon limit, we get the  $\mathcal{O}(Q^2)$  terms of the form

$$\mathcal{L}_v = \dots + d_i \text{Tr} [\bar{B}_v v \cdot A^2 B_v] + \dots, \quad (4.29)$$

with  $d_i = (d_{1/m} + e_i + m_B^2 f_i + \dots)$  where  $d_{1/m}$  is the calculable  $1/m_B$  correction from the relativistic leading-order Lagrangian. Clearly the  $e_i$  and  $f_i$  terms cannot be computed [51]. Thus if one imagines that the constants  $d_i$  are infested with the terms of the latter form, there is no way that one can estimate these constants. While this introduces an element of uncertainty in our calculation, it does not seriously diminish the predictability of the theory: Much of the uncertainty are eliminated in our determination of the parameters by experimental data. From Eq. (4.15) and the  $d_i$  terms estimated, we can extract the parameter

$$\sigma_{KN} \approx 2.83 M_\pi, \quad (4.30)$$

which is not the *sigma term*  $\Sigma_{KN} = \frac{1}{2}(\bar{m} + m_s)\langle P|\bar{u}u + \bar{s}s|P\rangle$ . There are loop corrections to be added to this value. This parameter will be used for off-shell  $KN$  scattering amplitudes and kaon self-energy.

#### 4.2.2. Off-shell amplitudes

In terms of the low-energy parameters fixed by the on-shell constraints, the off-shell  $K^- N$  scattering amplitude is given by

$$\begin{aligned} a^{K^- p} = & \frac{1}{4\pi(1 + \omega/m_B)} \left\{ T_v^{K^- p}(\omega = M_K) - \frac{\omega^2}{f^2} \left( \frac{\tilde{g}_{\Lambda^*}^2}{\omega + m_B - m_{\Lambda^*}} \right) + \frac{1}{f^2} (\omega - M_K) \right. \\ & + \frac{1}{f^2} (\omega^2 - M_K^2) \left( \tilde{d}_s - \frac{\sigma_{KN}}{M_K^2} + d_v + \frac{a_1}{2B_0} \right) + \frac{1}{f^2} (L_p^+(\omega) - L_p^+(M_K)) \\ & \left. - \frac{1}{f^2} (L_p^-(\omega) - L_p^-(M_K)) \right\}, \end{aligned} \quad (4.31)$$

$$\begin{aligned} a^{K^-n} = & \frac{1}{4\pi(1+\omega/m_B)} \left\{ T_v^{K^-n}(\omega = M_K) + \frac{1}{2f^2}(\omega - M_K) \right. \\ & + \frac{1}{f^2}(\omega^2 - M_K^2) \left( \bar{d}_s - \frac{\sigma_{KN}}{M_K^2} - \bar{d}_v - \frac{a_1}{2B_0} \right) + \frac{1}{f^2}(L_n^+(\omega) - L_n^+(M_K)) \\ & \left. - \frac{1}{f^2}(L_n^-(\omega) - L_n^-(M_K)) \right\}. \end{aligned} \quad (4.32)$$

Here  $B_0 = M_K^2/(\bar{m} + m_s)$  and

$$\begin{aligned} L_p^+(\omega) = & \frac{\omega^2}{64\pi f^2} \left\{ [2(D-F)^2 + \frac{1}{3}(D+3F)^2]M_K + \frac{3}{2}(D+F)^2M_\pi + \frac{1}{2}(D-3F)^2M_\eta \right. \\ & - \frac{1}{3}(D+F)(D-3F)(M_\pi + M_\eta) - \frac{1}{6}(D+F)(D-3F)(M_\pi^2 + M_\eta^2)\mathcal{I}_{\pi\eta} \Big\} \\ & + \frac{\omega^2}{8f^2} (4\Sigma_k^{(+)}(-\omega) + 5\Sigma_k^{(+)}(\omega) + 2\Sigma_\pi^{(+)}(\omega) + 3\Sigma_\eta^{(+)}(\omega)), \\ L_p^-(\omega) = & \alpha_p M_K^2 \omega + \beta_p \omega^3 + \frac{1}{4f^2} \omega^2 \left\{ -\frac{1}{2}\Sigma_k^{(-)}(\omega) - \Sigma_\pi^{(-)}(\omega) - \frac{3}{2}\Sigma_\eta^{(-)}(\omega) \right\}, \\ L_p^-(M_K) = & (\bar{g}_s + \bar{g}_v) M_K^3, \\ L_n^+(\omega) = & \frac{1}{64\pi f^2} \omega^2 \left\{ [\frac{5}{2}(D-F)^2 + \frac{1}{6}(D+3F)^2]M_K + \frac{3}{2}(D+F)^2M_\pi + \frac{1}{2}(D-3F)^2M_\eta \right. \\ & + \frac{1}{3}(D+F)(D-3F)(M_\pi + M_\eta) + \frac{1}{6}(D+F)(D-3F)(M_\pi^2 + M_\eta^2)\mathcal{I}_{\pi\eta} \Big\} \\ & + \frac{\omega^2}{8f^2} (2\Sigma_k^{(+)}(-\omega) + \Sigma_k^{(+)}(\omega) + \frac{5}{2}\Sigma_\pi^{(+)}(\omega) + \frac{3}{2}\Sigma_\eta^{(+)}(\omega)), \\ L_n^-(\omega) = & \alpha_n M_K^2 \omega + \beta_n \omega^3 + \frac{1}{4f^2} \omega^2 \left\{ \frac{1}{2}\Sigma_k^{(-)}(\omega) - \frac{5}{4}\Sigma_\pi^{(-)}(\omega) - \frac{3}{4}\Sigma_\eta^{(-)}(\omega) \right\}, \\ L_n^-(M_K) = & (\bar{g}_s - \bar{g}_v) M_K^3, \quad \mathcal{I}_{\pi\eta} = \int_0^1 dx \frac{1}{\sqrt{(1-x)M_\pi^2 + xM_\eta^2}}, \end{aligned} \quad (4.33)$$

where

$$\begin{aligned} \Sigma_i^{(+)}(\omega) = & -\frac{1}{4\pi} \sqrt{M_i^2 - \omega^2} \times \theta(M_i - |\omega|) + \frac{i}{2\pi} \sqrt{\omega^2 - M_i^2} \times \theta(\omega - M_i), \\ \Sigma_i^{(-)}(\omega) = & -\frac{1}{4\pi^2} \sqrt{\omega^2 - M_i^2} \ln \left| \frac{\omega + \sqrt{\omega^2 - M_i^2}}{\omega - \sqrt{\omega^2 - M_i^2}} \right| \times \theta(|\omega| - M_i) \\ & - \frac{1}{2\pi^2} \sqrt{M_i^2 - \omega^2} \sin^{-1} \frac{\omega}{M_i} \times \theta(M_i - |\omega|). \end{aligned} \quad (4.34)$$

Note that  $\Sigma_i^{(\pm)}(\omega)$  come from the diagrams of Fig. 3.1e (see Appendix B), where the intermediate states can go on-shell giving rise to an imaginary part of the amplitude. The functions  $L_{p,n}^-(\omega)$  contain four parameters  $\alpha_{p,n}$  and  $\beta_{p,n}$ . Owing to the constraints at  $\omega = M_K$ ,  $L_{p,n}^-(M_K)$ , they reduce to two. These two cannot be fixed by on-shell data. However since the off-shell amplitudes are rather

insensitive to the precise values of these constants, we will somewhat arbitrarily set  $\alpha_{p,n} \approx \beta_{p,n}$  in calculating Fig. 4.1 and the critical densities. This ambiguity will be taken into account in considering the critical densities. (See Table 6.5)

The predicted off-shell  $K^-p$  and  $K^-n$  scattering amplitudes are shown in solid line in Fig. 4.1 for the range of  $\sqrt{s}$  from 1.3 GeV to 1.5 GeV with  $g_{A^*}^2 = 0.25$  and  $\Gamma_{A^*} = 50$  MeV taken from experiments. (The dotted lines in Fig. 4.1 are explained in Subsection 4.2.1.)

The  $K^-n$  scattering is independent of the  $A^*$  and so the amplitude varies smoothly over the range involved. [1] Our predicted  $K^-p$  amplitude is found to be in fairly good agreement with the empirical fit of Ref. [21]. The striking feature of the real part of the  $K^-p$  amplitude, repulsive above and attractive below  $m_{A^*}$  (1405 MeV) as observed here, and the  $\omega$ -independent attraction of the  $K^-n$  amplitude are relevant to kaonic atoms [18] and to kaon condensation in “nuclear star” matter. The imaginary part of the  $K^-p$  amplitude is somewhat too high compared with the empirical fits. This may have to do with putting the experimental  $A^*$  decay width for the imaginary part of the mass. Self-consistency between loop corrections and the imaginary part of the mass would have to be implemented to get the correct imaginary part of the  $K^-p$  amplitude.

#### 4.2.3. Adler's soft-meson conditions

The off-shell amplitude calculated here does not satisfy Adler's soft-meson conditions that follow from the usual PCAC assumption that the pseudoscalar meson field  $\pi$  interpolate as the divergence of the axial current. The chiral Lagrangian used here does not give the direct relation  $\pi^i \sim \partial_\mu J_5^{i\mu}$  where  $J_5^{i\mu}$  is the axial current with flavor index  $i$ . Therefore in the soft-meson limit which corresponds in the present case to setting  $\omega$  equal to zero, the  $\pi N$  amplitude is not given by  $-\Sigma_{\pi N}/f^2$  as it does in the case of Adler's interpolating field [55]. In fact, it gives  $\Sigma_{\pi N}/f^2$  which has the opposite sign to Adler's limit. This led several authors to raise the possibility that a different physics might be involved in the chiral perturbation description of the off-shell processes that take place near  $\omega = 0$  [9,10,56].

A simple answer to this objection is that physics should not depend upon the interpolating field for the Goldstone bosons  $\pi$  [57]. The physics is equivalent whether one uses the  $\pi$  field as defined by the chiral Lagrangian used here or the  $\pi' \propto \partial_\mu J_5^{i\mu}$  field that gives Adler's conditions. Both are interpolating fields and they are just field-redefinitions of each other. This is natural since the Goldstone boson field is an auxiliary field in QCD.

An explicit illustration of the equivalence in the case considered in this review is given in the next section. Let it suffice here to say that if one wishes, one could rewrite the effective Lagrangian in such a way that Ward-Takahashi identities, to which Adler's conditions belong, are satisfied, without changing the physics involved. See Ref. [58].

#### 4.3. Interpolating fields

In this section, we show explicitly that to the chiral order we are concerned with, physics does not depend on the way the kaon field  $K$  (or in general the Goldstone boson field  $\pi$ ) interpolates. There is nothing new in what we do below: It is a well-known theorem [45]. But to those who are not very familiar with the modern notion of effective field theories, it has been a bit of a mystery that an off-shell amplitude which does not obey Adler's soft-pion theorem such as in the case of our Lagrangian could give the same physics with the amplitude which does in a situation which involves

the equation of state, not just an on-shell S-matrix. To have a simple idea, we start with the toy model of Manohar<sup>3</sup>.

#### 4.3.1. Adler's conditions in a toy model

Consider the simple Lagrangian,

$$\mathcal{L} = \frac{1}{2}(\partial K)^2 - \frac{1}{2}M_K^2(1 + \epsilon\bar{B}B)K^2, \quad (4.35)$$

where  $K$  is the charged kaon field which will develop a condensate in the form  $\langle K^- \rangle = v_K e^{-i\mu t}$ . As written, this Lagrangian can give only a linear density dependence in the kaon self-energy at tree order. The effective energy-density for s-wave kaon condensation linear in  $v_K^2$  is

$$\tilde{\epsilon} = (\mu^2 - M_K^2(1 + \epsilon\rho_N))v_K^2. \quad (4.36)$$

Thus the chemical potential in the ground state is

$$\mu^2 = M_K^2(1 + \epsilon\rho_N). \quad (4.37)$$

We shall now redefine the kaon field. To do this, we have, from the Noether construction and the equation of motion,

$$j_A^\mu = f\partial^\mu K, \quad \partial_\mu j_A^\mu = fM_K^2(1 + \epsilon\bar{B}B)K. \quad (4.38)$$

This invites us to redefine the kaon field as

$$\tilde{K} = K(1 + \epsilon\bar{B}B) = \frac{\partial_\mu j_A^\mu}{fM_K^2}, \quad (4.39)$$

thus giving the divergence of the axial current as the interpolating field of the kaon field. This kaon field will then satisfy the Adler soft-meson theorem. To see this, we rewrite the original Lagrangian in terms of the new field  $\tilde{K}$ ,

$$\begin{aligned} \mathcal{L} &= \frac{1}{(1 + \epsilon\bar{B}B)^2} \left( \frac{1}{2}(\partial\tilde{K})^2 - \frac{1}{2}M_K^2(1 + \epsilon\bar{B}B)\tilde{K}^2 \right) + \mathcal{O}(\partial B) \\ &= \frac{1}{2}(1 - 2\epsilon\bar{B}B) \left( \partial\tilde{K} \right)^2 - \frac{1}{2}M_K^2(1 - \epsilon\bar{B}B)\tilde{K}^2 + \mathcal{O}(\tilde{K}^2(\bar{B}B)^{n \geq 2}) + \mathcal{O}(\partial B). \end{aligned} \quad (4.40)$$

We can immediately read off the  $KN$  scattering amplitude and verify that it satisfies the Adler theorem. Note, however, that to obtain the same kaon self-energy that enters into the energy-density of the matter system as the one given by Eq. (4.35), it would be necessary to keep multi-baryon terms, which means that in medium, higher density dependence needs to be taken into account.

Now to see that the interpolating fields  $K$  and  $\tilde{K}$  give the same physics in medium, consider the critical point at which condensation sets in. From Eq. (4.39), we have

$$|\langle \tilde{K} \rangle| = \tilde{v}_K = v_K(1 + \epsilon\rho_N). \quad (4.41)$$

Substituting this into Eq. (4.35) or directly from Eq. (4.40), we get the effective energy-density

<sup>3</sup> A.V. Manohar, private communication.

$$\tilde{\epsilon} = (\mu^2 - M_K^2(1 + \epsilon\rho_N)) \frac{\tilde{v}_K^2}{(1 + \epsilon\rho_N)^2}. \quad (4.42)$$

Clearly (4.36) and (4.42) give the same critical point given by the vanishing of the energy density. This is of course quite trivial. If one keeps terms up to linear in  $\rho_N$ , Eq. (4.42) reads

$$\tilde{\epsilon} = (\mu^2 - M_K^2 - (2\mu^2 - M_K^2)\epsilon\rho_N)\tilde{v}_K^2. \quad (4.43)$$

This is the kaon self-energy in linear density approximation, which satisfies Adler's consistency condition. As it stands, it looks very different from Eq. (4.36), but to the linear order in density, this gives the same pole position,

$$\begin{aligned} (\mu^2 - M_K^2 - (2\mu^2 - M_K^2)\epsilon\rho_N) &= 0 \longrightarrow (1 - 2\epsilon\rho_N)\mu^2 = M_K^2(1 - \epsilon\rho_N) \\ &\longrightarrow \mu^2 = M_K^2(1 + \epsilon\rho_N). \end{aligned} \quad (4.44)$$

Beyond the linear-density approximation, one must use Eq. (4.42) instead of Eq. (4.43) for the physics of the toy-model Lagrangian: Consistency requires that *all* kaon–multi-nucleon scattering terms (that is, terms higher order in  $\rho_N$ ) be included.

We now turn to the real issue. We shall compare the approach used here (called “Kaplan–Nelson” (KN)) to the Gasser–Sainio–Svarc (GSS) approach which implements Adler's conditions by means of external fields [58].

#### 4.3.2. KN approach

The KN approach corresponds to introducing the source field  $p_i$  in the Lagrangian

$$\mathcal{L}_{\text{source}} = p^+ K^- + p^- K^+ \dots. \quad (4.45)$$

The self-energy is then obtained from

$$\begin{aligned} \langle \tau(x_1, x_2) \rangle &= \langle T(K^+(x_2)K^-(x_1)) \rangle_\rho \\ &= - \left\langle \frac{\delta^2}{\delta p^+(x_1)\delta p^-(x_2)} \int \mathcal{D}[K, B] e^{i \int (\mathcal{L}_{\text{orig}} + \mathcal{L}_{\text{source}})} \Big|_{p^\pm=0,\dots=0} \right\rangle_\rho \\ &= i\Delta_K^{(0)}(x_1 - x_2) + i\Sigma_K \int \Delta_K^{(0)}(x_1 - z)\Delta_K^{(0)}(z - x_2) + \mathcal{O}(\Sigma_K^2), \end{aligned} \quad (4.46)$$

where  $\langle \dots \rangle_\rho$  represents the expectation value in the ground state of dense matter. Here  $\Delta_K^{(0)}$  is the free kaon propagator

$$\Delta_K^{(0)}(x - y) = \frac{1}{(2\pi)^4} \int \frac{e^{-ik(x-y)}}{k^2 - M_K^2 + i\epsilon} d^4 k \quad (4.47)$$

and  $\Sigma_K$  is the kaon self-energy calculated to order  $\nu = 3$  with the Lagrangian

$$\mathcal{L}_0 = \partial K^+ \partial K^- - M_K^2 K^+ K^-, \quad \mathcal{L}_{\text{int}} = \mathcal{L}_{\text{int}}^{\nu=1} + \mathcal{L}_{\text{int}}^{\nu=2} + \dots \quad (4.48)$$

Summing the series (4.46), we have

$$\tau(x_1, x_2) = \frac{i}{(2\pi)^4} \int \frac{e^{ik \cdot (x_1 - x_2)}}{k^2 - M_K^2 - \Sigma_K + i\epsilon} d^4 k. \quad (4.49)$$

The propagator to order  $\nu = 3$  in momentum space is

$$\Delta_K = \frac{i}{k^2 - M_K^2 - \Sigma_K + i\epsilon} \quad (4.50)$$

and the effective energy density to order  $v_K^2$  is

$$\epsilon = -i\Delta_K^{-1}|v_K|^2 = -(\mu^2 - M_K^2 - \Sigma_K)|v_K|^2. \quad (4.51)$$

#### 4.3.3. GSS approach

The GSS approach corresponds to introducing into the Lagrangian the source field  $p_i$  of the form

$$\mathcal{L}_{\text{source}} = p^+ K^- (1 + \alpha_p \bar{p} p + \alpha_n \bar{n} n) + p^- K^+ (1 + \alpha_p \bar{p} p + \alpha_n \bar{n} n) + \dots, \quad (4.52)$$

where  $\alpha_p = (1/f^2 M_K^2)(\sigma_{KN} + C_{KN})$  and  $\alpha_n = (1/f^2 M_K^2)(\sigma_{KN} - C_{KN})$  are related with the parameters of the Lagrangian (3.44) by

$$\sigma_{KN} = -\frac{1}{2}(\hat{m} + m_s)(a_1 + 2a_2 + 4a_3)$$

and

$$C_{KN} = -\frac{1}{2}(\hat{m} + m_s)a_1.$$

As shown in Ref. [58], this way of introducing the source field reproduces Adler's conditions (soft-pion, self-consistency etc.). Now the self-energy is obtained from the two-point function

$$\begin{aligned} \langle \tau(x_1, x_2) \rangle &= - \left\langle \frac{\delta^2}{\delta p^+(x_1) \delta p^-(x_2)} \int \mathcal{D}[K, B] e^{i \int (\mathcal{L}_{\text{orig}} + \mathcal{L}_{\text{source}})} \Big|_{p^\pm = 0, \dots} \right\rangle_\rho \\ &= \langle T(K^+(1 + \alpha_p \bar{p} p + \alpha_n \bar{n} n)(x_2) K^-(1 + \alpha_p \bar{p} p + \alpha_n \bar{n} n)(x_1)) \rangle_\rho \\ &= (1 + \alpha_p \rho_p + \alpha_n \rho_n)^2 \langle T(K^+(x_2) K^-(x_1)) \rangle_\rho \\ &= (1 + \alpha_p \rho_p + \alpha_n \rho_n)^2 \frac{i}{(2\pi)^4} \int \frac{e^{ik \cdot (x_1 - x_2)}}{k^2 - M_K^2 - \Sigma_K + i\epsilon} d^4 k, \end{aligned} \quad (4.53)$$

where  $\Sigma_K$  is identical to that of the KN approach since the  $\alpha_p$  and  $\alpha_n$  are  $\nu = 2$  terms that are unaffected by loop corrections (of order  $\nu = 3$ ). Finally the full propagator is given by

$$\Delta'_K = (1 + \alpha_p \rho_p + \alpha_n \rho_n)^2 \frac{i}{k^2 - M_K^2 - \Sigma_K + i\epsilon}, \quad (4.54)$$

which can be rewritten in the form of Eq. (4.50),

$$\Delta'_K = \frac{i}{k^2 - M_K^2 - \Sigma'_K + i\epsilon} \quad (4.55)$$

with a redefined self-energy

$$\Sigma_K = \frac{1}{(1 + \alpha_p \rho_p + \alpha_n \rho_n)^2} \{ (k^2 - M_K^2) [(1 + \alpha_p \rho_p + \alpha_n \rho_n)^2 - 1] + \Sigma_K \}. \quad (4.56)$$

One can verify that to the linear order in density, this has the structure mentioned in the main text, namely, it changes from an attraction on-shell to a repulsion off-shell. Finally the effective energy density linear in  $\tilde{v}_K^2$  is

$$\tilde{\epsilon} = -i(\Delta'_K)^{-1} |\tilde{v}_K|^2 = -i\Delta_K^{-1} \left| \frac{\tilde{v}_K}{1 + \alpha_p \rho_p + \alpha_n \rho_n} \right|^2, \quad (4.57)$$

where  $\tilde{v}_K$  is the kaon expectation value in terms of the GSS field.

#### 4.3.4. Effective energy-density and physical observables

As in the toy model, we can write  $\tilde{v}_K$  in terms of  $v_K$ ,

$$\tilde{v}_K = v_K (1 + ax + b), \quad (4.58)$$

where  $a = (\alpha_p - \alpha_n)\rho$  and  $b = \alpha_n\rho$ . Thus the energy-densities to all orders in  $v_K$  and  $\tilde{v}_K$  can be written as

$$\epsilon = \sum_{n \geq 2} \alpha_n(\mu, x, \rho) |v_K|^n + \beta(\mu, x, \rho), \quad (4.59)$$

$$\tilde{\epsilon} = \sum_{n \geq 2} \alpha_n(\mu, x, \rho) \left| \frac{\tilde{v}_K}{1 + ax + b} \right|^n + \beta(\mu, x, \rho), \quad (4.60)$$

where  $\alpha_n(\mu, x, \rho)$  and  $\beta(\mu, x, \rho)$  are given in Eq. (6.2). By differentiating the effective energy density with respect to  $\mu, x, v_K$  ( $\tilde{v}_K$ ), one can get the equation of state as a function of density  $\rho$ . In the GSS approach, we have three equations of state,

$$\begin{aligned} 0 &= \frac{\partial \tilde{\epsilon}}{\partial \mu} = \sum_{n \geq 2} \frac{\partial \alpha_n(\mu, x, \rho)}{\partial \mu} \left( \frac{\tilde{v}_K}{1 + ax + b} \right)^n + \frac{\partial \beta(\mu, x, \rho)}{\partial \mu}, \\ 0 &= \frac{\partial \tilde{\epsilon}}{\partial x} = \sum_{n \geq 2} \frac{\partial \alpha_n(\mu, x, \rho)}{\partial x} \left( \frac{\tilde{v}_K}{1 + ax + b} \right)^n + \frac{\partial \beta(\mu, x, \rho)}{\partial x} \\ &\quad + \frac{a}{1 + ax + b} \sum_{n \geq 2} n \alpha_n(\mu, x, \rho) \left( \frac{\tilde{v}_K}{1 + ax + b} \right)^n, \\ 0 &= \frac{\partial \tilde{\epsilon}}{\partial \tilde{v}_K} = \frac{1}{\tilde{v}_K} \sum_{n \geq 2} n \alpha_n(\mu, x, \rho) \left( \frac{\tilde{v}_K}{1 + ax + b} \right)^n. \end{aligned} \quad (4.61)$$

Now using the third equation, we see that the last term in the second equation vanishes. To see that the two ways give the same physics, it suffices to note that the KN approach gives the same set of equations except for the replacement  $\tilde{v}_K = v_K(1 + ax + b)$  which does not affect anything.

An identical conclusion is reached by Thorsson and Wirzba [59] by a slightly different reasoning.

#### 4.4. The role of $\Lambda^*$ in kaon proton interactions

In this section, as a side remark, s-wave  $K^-p$  scattering into various channels near threshold is analyzed at tree order in heavy-baryon chiral perturbation theory with the  $\Lambda(1405)$  introduced as an independent field [2]. We show that chiral perturbation expansion treating the  $\Lambda(1405)$  as elementary is consistent with all threshold data including a double-charge-exchange process suppressed at leading order of chiral expansion in the absence of the  $\Lambda(1405)$ . We also discuss s-wave  $K^+p$  scattering phase shifts at low energy.

##### 4.4.1. $K^-p$ scattering

The explicit symmetry breaking  $a_i$  terms in the chiral Lagrangian Eq. (3.44) can be determined by the baryon mass splitting and the  $\pi N$  sigma term. Here we use the results of Kaplan and Nelson [4],

$$\begin{aligned} a_1 &= -0.28, \quad a_2 = 0.56, \quad a_3 = -1.1, \\ m_u &= 6 \text{ MeV}, \quad m_d = 12 \text{ MeV}, \quad m_s = 240 \text{ MeV}. \end{aligned} \quad (4.62)$$

These parameters give the sigma term

$$\Sigma_{KN} = -\frac{1}{2}(m_u + m_s)(a_1 + 2a_2 + 4a_3) \simeq 438 \text{ MeV}. \quad (4.63)$$

Given the large error bar ( $\pm 0.3$ ) in  $a_3$ , this  $\Sigma_{KN}$  is consistent with the recent lattice calculations of Dong & Liu [60] (see also Fukugita et al. [61]),

$$\Sigma_{KN} = \frac{(m_u + m_s)\langle N|(\bar{u}u + \bar{s}s)|N\rangle}{(m_u + m_d)\langle N|(\bar{u}u + \bar{d}d)|N\rangle} \Sigma_{\pi N} \simeq 450 \pm 30 \text{ MeV}. \quad (4.64)$$

For s-wave scattering, there is no distinction between  $A^2$  and  $(v \cdot A)^2$  in (3.44). So the combination

$$\bar{d}_i = d_i + d_{i+1} \quad (i : \text{odd number}) \quad (4.65)$$

enters in the scattering amplitudes. This means that there are five independent parameters at  $\mathcal{O}(Q^2)$  for s-wave  $KN$  scattering. These parameters need to be fixed for the equation of state of “nuclear” stars with kaon condensation. The transition matrix elements for various channels are summarized in Appendix G.

Now requiring consistency with the experimental  $KN$  scattering lengths, Eq. (4.6), we get two constraints on  $\bar{d}_i$  at  $\mathcal{O}(Q^2)$  [1] (refer to Eq. (4.14)),

$$(\bar{d}_s - \bar{d}_v)_{emp} \approx (0.05-0.06) \text{ fm}, \quad (\bar{d}_s + \bar{d}_v)_{emp} \approx 0.13 \text{ fm}, \quad (4.66)$$

with the isoscalar constants  $d_s$  and the isovector constants  $d_v$  defined by

$$\begin{aligned} d_s &= -\frac{1}{2M_K^2}(m_u + m_s)(a_1 + 2a_2 + 4a_3) + \frac{1}{4}(\bar{d}_1 + 2\bar{d}_3 + 4\bar{d}_5 + \bar{d}_7), \\ \bar{d}_v &= -\frac{1}{2M_K^2}(m_u + m_s)a_1 + \frac{1}{4}(\bar{d}_1 + \bar{d}_7). \end{aligned} \quad (4.67)$$

We are therefore left with three independent parameters to be determined from other experimental data.

**Table 4.2**  
Leading-order contributions to  $K^- p$  scattering.

Channel	$f^2 T_{\pi^\pm}$	$f^2 T_{A^*}$
$\pi^- \Sigma^+$	$(2M_K + m_B - m_{\Sigma^+})/4$	$g(m_{\Sigma^+})$
$\pi^+ \Sigma^-$	-	$g(m_{\Sigma^-})$
$K^- p$	$M_K$	$g(m_p)$
$\bar{K}^0 n$	$M_K/2$	$g(m_n)$
$\pi^0 \Sigma^0$	$(2M_K + m_B - m_{\Sigma^0})/8$	$g(m_{\Sigma^0})$
$\pi^0 \Lambda$	$\sqrt{3}(2M_K + m_B - m_A)/8$	-
$\eta \Sigma^0$	$\sqrt{3}(2M_K + m_B - m_{\Sigma^0})/8$	-
$\eta \Lambda$	$3(2M_K + m_B - m_A)/8$	$g(m_\Lambda)$
$K^+ \Xi^-$	-	$g(m_{\Xi^-})$
$K^0 \Xi^0$	-	$g(m_{\Xi^0})$

$g(m) = -g_{A^*}^2 \frac{M_K(M_K + m_B - m)}{m_B + M_K - m_{A^*}}$

#### 4.4.2. Threshold branching ratios

Once the constants are fixed from some experiments, we could then make predictions for other physical quantities in low-energy  $KN$  scattering. In particular, the following threshold branching ratios are of particular interest [62]:

$$\gamma = \frac{|\mathcal{T}_{\pi^+ \Sigma^-}|^2}{|\mathcal{T}_{\pi^- \Sigma^+}|^2}, \quad R_c = \frac{\sum_{i=\pi^\pm \Sigma^\mp} |\mathcal{T}_i|^2}{\sum_{j=\pi^0 \Lambda, \pi^0 \Sigma^0, \pi^\pm \Sigma^\mp} |\mathcal{T}_j|^2}, \quad R_n = \frac{|\mathcal{T}_{\pi^0 \Lambda}|^2}{\sum_{i=\pi^0 \Lambda, \pi^0 \Sigma^0} |\mathcal{T}_i|^2}. \quad (4.68)$$

The empirical values are [63]

$$\gamma^{exp} = 2.36 \pm 0.04, \quad R_c^{exp} = 0.664 \pm 0.011, \quad R_n^{exp} = 0.19 \pm 0.02. \quad (4.69)$$

Since we still have three constants left unfixed, we cannot compare theory directly with experiments for these quantities. However for the chiral perturbation approach to be viable, the leading  $\mathcal{O}(Q^1)$  chiral order which involves no unknown counterterms should dominate. In other words, higher-order terms should be suppressed according to the counting rule,  $(Q/4\pi f)^n$ . To verify this, we calculate the branching ratios with the scattering amplitudes computed at  $\mathcal{O}(Q^1)$ . The results are summarized in Table 4.2. The numerical results are

$$\gamma = 1.93, \quad R_c = 0.64, \quad R_n = 0.11. \quad (4.70)$$

Here only the Weinberg-Tomozawa term and the  $A^*$  contribution in leading order are taken into account. Clearly the leading tree contributions play a dominant role for  $\gamma$  and  $R_c$  while  $R_n$  apparently requires some higher order corrections. Note that without the  $A^*$ , the transition to  $\pi^+ \Sigma^-$  would be suppressed at the leading order, so we would have  $\gamma = 0$ . Furthermore the enhancement of the  $\pi^+ \Sigma^-$  channel over the  $\pi^- \Sigma^+$  channel is principally due to  $A^*$ : while only the  $A^*$  term contributes to the  $\pi^+ \Sigma^-$  channel, both the Weinberg-Tomozawa term and the  $A^*$  term contribute to the channel  $\pi^- \Sigma^+$  but with an opposite sign, giving rise to the enhancement of the ratio  $\gamma$ .

From the point of view of chiral perturbation theory for kaon condensation, the most meaningful outcome of the present exercise is that we are now able to determine the three parameter combinations

**Table 4.3**  
Best-fit parameters in  $fm$  for the branching ratios.

$\bar{d}_1 + \bar{d}_7$	$\bar{d}_3 + \bar{d}_7$	$\bar{d}_7 + \bar{d}_9$	$\gamma$	$R_c$	$R_n$
0.039	-3.35	-5.90	2.36	0.63	0.19

**Table 4.4**  
Comparison of the predicted phase shifts (in deg) with experimental data. Exp A and Exp B correspond to the data of [64] and [65], respectively.  $P_{lab}$  and  $\omega_{K,cm}$  are given in MeV.

$P_{lab}$	$\omega_{K,cm}$	Exp A	Exp B	Our result
145	504	$-8.2 \pm 0.9$	-	-9.05
175	508	$-10.2 \pm 0.3$	-	-11.2
178	509	-	$-10.1 \pm 0.8$	-11.4
205	513	$-11.4 \pm 0.4$	-	-13.4
235	518	$-13.3 \pm 0.4$	-	-15.8
265	525	$-15.0 \pm 0.2$	$-16.0 \pm 0.4$	-18.3

that appear at  $\mathcal{O}(Q^2)$ . This determination would allow us to calculate to  $\mathcal{O}(Q^2)$  the equation of state needed for describing the properties of the kaon condensed state. The best-fit parameters using  $(\bar{d}_s - \bar{d}_v)_{emp} = 0.055 fm$  are given in Table 4.3.

#### 4.4.3. S-wave phase shift for $K^+ p$ scattering

While the  $\Lambda(1405)$  contributes unimportantly to the s-wave phase shifts for  $K^+ p$  scattering (since it enters in the crossed term with a large energy denominator), it is however important to check whether or not the chiral perturbation approach to kaon–nuclear interactions with the large  $KN$  sigma term,  $\Sigma_{KN} \approx 438$  MeV, is consistent with the well-measured  $K^+ p$  phase shifts [64,65]. The s-wave  $K^+ p$  phase shift is related to the scattering amplitude

$$\delta(P_{lab}) = \tan^{-1}(q_{cm}(P_{lab})a_0^{K^+ p}(P_{lab})), \quad (4.71)$$

where  $q_{cm}$  is the meson momentum in c.m. frame, and  $P_{lab}$  is the meson laboratory momentum.

We have computed the phase shifts to  $\mathcal{O}(Q^2)$  chiral order for which there are no unknown constants once the scattering lengths are fit. The results are given in Table 4.4 and Fig. 4.3. Although there are no experimental data to compare with for  $P_{lab} < 145$  MeV, we see that the  $\mathcal{O}(Q^2)$  chiral perturbation theory is consistent with the low-energy part of the phase shifts up to  $P_{lab} \sim m_\pi$ .

#### 4.4.4. The role of $\Lambda^*$

We have shown that chiral perturbation theory with the  $\Lambda(1405)$  introduced as an elementary matter field can satisfactorily describe low-energy s-wave  $K^\pm p$  scattering. In particular, the OZI-suppressed double-charge-exchange process  $K^- p \rightarrow \pi^+ \Sigma^-$  is found to be enhanced relative to the OZI-allowed process  $K^- p \rightarrow \pi^- \Sigma^+$  when  $\Lambda(1405)$  is introduced. We have also shown that a large  $KN$  sigma term,  $\Sigma_{KN} \approx 430$  MeV, is consistent with the experimental phase shifts for  $K^+ p$  scattering at low energy,  $P_{lab} \lesssim m_\pi$ . Our results provide evidence that the  $\Lambda(1405)$  as an elementary field is a phenomenologically viable concept, in a way that resembles the  $\Delta$  in the nonstrange sector.

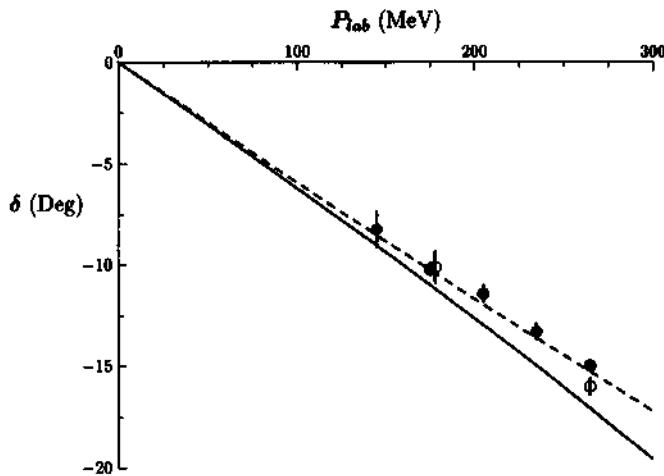


Fig. 4.3. The s-wave phase shifts (in deg) for  $K^+p$  scattering. The filled and empty circles correspond to the data of [64] and [65], respectively. Solid line corresponds to our results and the dashed line corresponds to the extrapolation of experimental scattering amplitude,  $\delta = \tan^{-1} q_{cm} a_{exp}^{K^+p}$ .

## 5. Kaon self-energy and kaonic atom

### 5.1. Kaon self-energy

If one were to limit oneself to linear density approximation which would be reliable in dilute systems, then what we have obtained so far is sufficient for studying kaon–nucleon interactions in many-body systems. The off-shell KN amplitude calculated to  $\mathcal{O}(Q^3)$  in impulse approximation gives the optical potential, which, expressed as a self-energy, is depicted by Fig. 5.1a,

$$\Pi_K^{imp}(\omega) = - \left( \rho_p T_{free}^{K^-p}(\omega) + \rho_n T_{free}^{K^-n}(\omega) \right), \quad (5.1)$$

where  $T^{KN}$  is the off-shell s-wave KN transition matrix. The amplitude  $T^{KN}$  taken on-shell, i.e.,  $\omega = M_K$ , and the scattering length  $a^{KN}$  are related by  $a^{KN} = [1/4\pi(1 + M_K/m_B)]T^{KN}$ . For the purpose of studying kaon condensation, the linear density approximation may not be reliable enough and one would have to study the *effective action* (or *effective potential* in translationally invariant systems).

To go beyond the linear density approximation, there are two major effects to be considered. The first is the Pauli correction and the second is many-body correlations.

#### 5.1.1. Pauli correction

The Pauli effect can be most straightforwardly taken into account in the self-energy by modifying the nucleon propagator in the loop graphs contributing to the KN scattering amplitude to one appropriate in medium

$$G^0(k) \simeq \frac{i}{v \cdot k + i\epsilon} - 2\pi\delta(k_0)\theta(k_F - |k|). \quad (5.2)$$

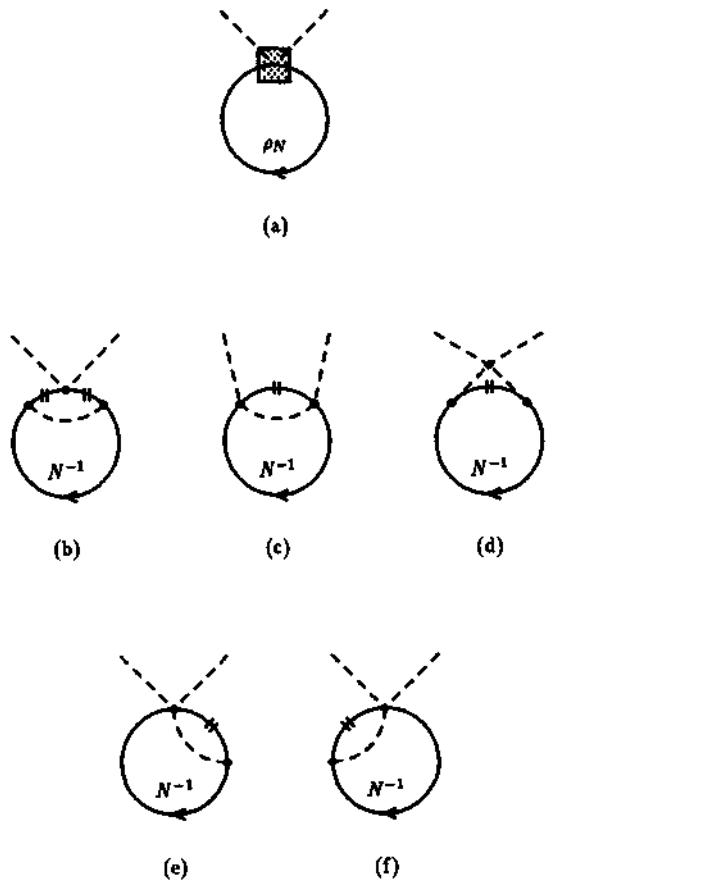


Fig. 5.1. (a) The linear density approximation to the kaon self-energy in medium,  $\Pi_K$ . The square blob represents the off-shell  $K^-N$  amplitude calculated to  $\mathcal{O}(Q^3)$ ; (b)–(f) medium corrections to  $T^{KN}$  of Fig. (a) with the free nucleon propagator indicated by a double slash replaced by an in-medium one, Eq. (5.2). The loop labeled  $\rho_N$  represents the in-medium nucleon loop proportional to density,  $N^{-1}$  the nucleon hole ( $n^{-1}$  and/or  $p^{-1}$ ), the external dotted line stands for the  $K^-$  and the internal dotted line for the pseudoscalar octet  $\pi, \eta, K$ .

where  $k_F$  is the nucleon Fermi momentum related to density  $\rho_N$  by the usual relation  $\rho_N = (\gamma/6\pi^2)k_F^3$ , with the degeneracy factor  $\gamma = 2$  for neutron and proton in nuclear matter.

We write down the explicit forms of the quantities that enter in Eq. (5.9).

$$\begin{aligned}
 \delta T_{\rho_N}^{K^-p} = & -\frac{1}{12f^4} \left\{ (D+F)^2 [(M_\pi^2 + M_K^2 - \omega^2) D_{4,\pi\pi}^p + D_{6,\pi\pi}^p] + \frac{1}{2}(D+F)^2 \right. \\
 & \times [(M_\pi^2 + M_K^2 - \omega^2) D_{4,\pi\pi}^p + D_{6,\pi\pi}^p] + \frac{1}{2}(D-3F)^2 [(M_K^2 - \frac{1}{3}M_\pi^2 - \omega^2) D_{4,\eta\eta}^p + D_{6,\eta\eta}^p] \\
 & - (D+F)(D-3F) \left[ (\frac{1}{3}M_\pi^2 - \frac{1}{3}M_K^2 - \omega^2) D_{4,\pi\eta}^p + D_{6,\pi\eta}^p \right] \Big\} + \frac{\omega^2}{4f^4} (8\Sigma_K^p(\omega) + \Sigma_K^n(\omega)) \\
 & + \frac{1}{12f^4} \{ 2(D+F)FG_\pi^p + (D+F)^2 G_\pi^n - 2(D-3F)FG_\eta^p \}, \tag{5.3}
 \end{aligned}$$

$$\begin{aligned} \delta T_{\rho_N}^{K^-n} = & -\frac{1}{12f^4} \left\{ (D+F)^2 [(M_\pi^2 + M_K^2 - \omega^2) D_{4,\pi\pi}^p + D_{6,\pi\pi}^p] + \frac{1}{2}(D+F)^2 \right. \\ & [(M_\pi^2 + M_K^2 - \omega^2) D_{4,\pi\pi}^n + D_{6,\pi\pi}^n] + \frac{1}{2}(D-3F)^2 \left[ (M_K^2 - \frac{1}{3}M_\pi^2 - \omega^2) D_{4,\eta\eta}^n + D_{6,\eta\eta}^n \right] \\ & +(D+F)(D-3F) \left[ (\frac{1}{3}M_\pi^2 - \frac{1}{3}M_K^2 - \omega^2) D_{4,\pi\eta}^n + D_{6,\pi\eta}^n \right] \left. \right\} + \frac{\omega^2}{4f^4} (2\Sigma_K^n(\omega) + \Sigma_K^p(\omega)) \\ & + \frac{1}{12f^4} \left\{ (D^2 - F^2) G_\pi^n + (D+F)^2 G_\pi^p + (D-3F)(D-F) G_\eta^n \right\}, \end{aligned} \quad (5.4)$$

$$\begin{aligned} D_{\alpha,ij}^N = & \frac{1}{2\pi^2} \int_0^{k_F} d|\mathbf{k}| \frac{1}{|\mathbf{k}|^2 + M_i^2} \frac{1}{|\mathbf{k}|^2 + M_j^2} |\mathbf{k}|^\alpha, \quad \Sigma_i^N(\omega) = \frac{1}{2\pi^2} \int_0^{k_F} d|\mathbf{k}| \frac{|\mathbf{k}|^2}{\omega^2 - M_i^2 - |\mathbf{k}|^2}, \\ G_i^N = & \frac{1}{2\pi^2} \int_0^{k_F} d|\mathbf{k}| \frac{|\mathbf{k}|^2}{M_i^2 + |\mathbf{k}|^2} \frac{|\mathbf{k}|^2}{M_i^2}. \end{aligned} \quad (5.5)$$

Here the subscripts  $\pi$ ,  $\eta$  and  $K$  are the octet Goldstone bosons, the superscripts  $n$  and  $p$  stand for neutrons and protons.

The resulting correction denoted  $\delta T_{\rho_N}^{K^-n}$ , which satisfies the chiral symmetry constraint of Meißner et al. [66], is clearly nonlinear in density and repulsive as befits a Pauli exclusion effect.

### 5.1.2. Many-body correlations

For the second effect, the most important one is the correlation involving “particle-hole” excitations. This is of typically many-body nature. There are two classes of correlations one would have to consider. One involves non-strange particle-hole excitations and the other strange particle-nonstrange hole excitations. All these can be mediated by four-Fermi interactions described above.

We first consider the latter. These are depicted in Fig. 5.2. Since we are dealing with s-wave kaon interaction, the most important configuration that  $K^-$  can couple to is the  $\Lambda^*$  particle–nucleon hole (denoted as  $\Lambda^*N^{-1}$  with  $N$  either a proton ( $p$ ) or neutron ( $n$ )). Here we focus on the former type. Now for the s-wave in-medium kaon self-energy, the relevant four-Fermi interactions that involve a  $\Lambda^*$  can be reduced to a simple form involving two unknown constants

$$\mathcal{L}_{4-fermion} = C_{\Lambda^*}^S \bar{\Lambda}_v^* \Lambda_v^* \text{Tr } \bar{B}_v B_v + C_{\Lambda^*}^T \bar{\Lambda}_v^* \sigma^k \Lambda_v^* \text{Tr } \bar{B}_v \sigma^k B_v, \quad (5.6)$$

where  $C_{\Lambda^*}^{S,T}$  are the dimension  $-2$  ( $M^{-2}$ ) parameters to be fixed empirically and  $\sigma^k$  acts on baryon spinor.

Additional (in-medium) two-loop graphs that involve  $\Lambda^*N^{-1}$  excitations are given in Figs. 5.2c and 5.2d. They do not however involve contact four-Fermi interactions, so are calculable unambiguously.

We shall denote the sum of these contributions from Fig. 5.2 to the self-energy by  $\Pi_{\Lambda^*}$ . A simple calculation gives

$$\begin{aligned} \Pi_{\Lambda^*}(\omega) = & -\frac{g_{\Lambda^*}^2}{f^2} \left( \frac{\omega}{\omega + m_B - m_{\Lambda^*}} \right)^2 \left\{ C_{\Lambda^*}^S \rho_p (\rho_n + \frac{1}{2}\rho_p) - \frac{3}{2} C_{\Lambda^*}^T \rho_p^2 \right\} + \frac{g_{\Lambda^*}^2}{f^4} \rho_p \left( \frac{\omega}{\omega + m_B - m_{\Lambda^*}} \right) \\ & \times \omega^2 \left\{ (2\Sigma_K^p(\omega) + \Sigma_K^n(\omega)) - g_{\Lambda^*}^2 \left( \frac{\omega}{\omega + m_B - m_{\Lambda^*}} \right) (\Sigma_K^p(\omega) + \Sigma_K^n(\omega)) \right\}, \end{aligned} \quad (5.7)$$

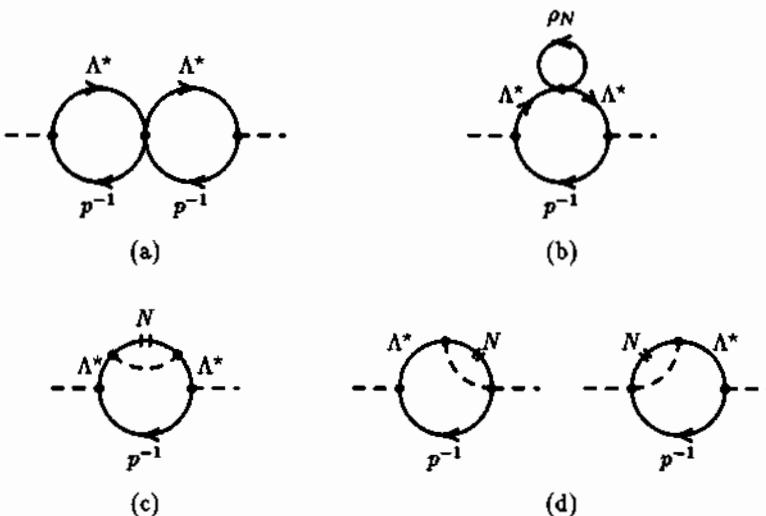


Fig. 5.2. Two-loop diagrams involving  $\Lambda^*$  contributing to the kaon self-energy. The diagrams (a) and (b) involve four-Fermi interactions, while the diagrams (c) and (d) do not involve four-Fermi interactions and hence can be unambiguously determined by on-shell parameters. Here the internal dotted line represents the kaon.

where  $g_{A^*}$  is the renormalized  $KN\Lambda^*$  coupling constant determined in [1] and  $\Sigma_K^N(\omega)$  is given by

$$\Sigma_K^N(\omega) = \frac{1}{2\pi^2} \int_0^{k_F} d|\mathbf{k}| \frac{|\mathbf{k}|^2}{\omega^2 - M_K^2 - |\mathbf{k}|^2}. \quad (5.8)$$

In Eq. (5.7), the first term comes from the diagrams of Figs. 5.2a and 5.2b and the second term from the diagrams of Fig. 5.2c and 5.2d. While the second term gives repulsion corresponding to a Pauli quenching, the first term can give either attraction or repulsion depending on the sign of  $(C_{A^*}^S[\rho_n + \frac{1}{2}\rho_p] - \frac{3}{2}C_{A^*}^T\rho_p)$  with the constants  $C_{A^*}^{S,T}$  being the only parameters that are not determined by on-shell data.

## 5.2. Kaonic atom

The complete self-energy to in-medium two-loop order is then

$$\Pi_K(\omega) = - \left( \rho_p T_{free}^{K-p}(\omega) + \rho_n T_{free}^{K-n}(\omega) \right) - \left( \rho_p \delta T_{\rho_N}^{K-p}(\omega) + \rho_n \delta T_{\rho_N}^{K-n}(\omega) \right) + \Pi_{A^*}(\omega). \quad (5.9)$$

The additional parameters  $C_{A^*}^{S,T}$  that are introduced at the level of four-Fermi interactions in the strange particle-hole sector require experimental data involving nuclei and nuclear matter. We shall now discuss how these constants can be fixed from kaonic atom data. In order to fix both of these constants, we would need data over a wide range of nuclei. One sees in (5.7) that for the symmetric matter, what matters is the combination  $(C_{A^*}^S - C_{A^*}^T)$ . At present, this is the only combination that we can hope to pin down from kaonic atom data. That leaves one parameter unfixed. We shall pick

**Table 5.1**

$K^-$  self-energies in symmetric nuclear matter ( $x = 0.5$ ) in unit of  $M_K^2$  at  $u = 0.97$  for  $g_{A^*}^2 = 0.25$ .  $\Delta V = M_K^* - M_K$  is the attraction (in unit of MeV) at a given  $(C_{A^*}^S - C_{A^*}^T)f^2$ .

$(C_{A^*}^S - C_{A^*}^T)f^2$	$M_K^*$	$\Delta V$	$-\rho T^{free}$	$-\rho \delta T^{free}$	$\Pi_{A^*}^1$	$\Pi_{A^*}^2$
1	402.3	-92.74	-0.2505	0.0467	-0.2489	0.1142
5	353.2	-141.8	-0.1820	0.0307	-0.3678	0.0286
10	327.2	-167.8	-0.1740	0.0258	-0.4297	0.0155
20	297.7	-197.3	-0.1749	0.0217	-0.4927	0.0082
50	254.0	-241.0	-0.1843	0.0176	-0.5731	0.0034
100	218.8	-276.2	-0.1929	0.0154	-0.6286	0.0016

$C_{A^*}^S$  for the reason to be explained later. We shall parametrize the proton and neutron densities by the proton fraction  $x$  and the nucleon density  $u = \rho/\rho_0$  as

$$\rho_p = x\rho, \quad \rho_n = (1-x)\rho, \quad \rho = u\rho_0. \quad (5.10)$$

Now what we know from the presently available kaonic atom data [18] is that the optical potential for the  $K^-$  in medium has an attraction of the order of

$$\Delta V \approx -(200 \pm 20) \text{ MeV} \quad \text{at } u = 0.97. \quad (5.11)$$

This implies approximately for  $x = 1/2$

$$(C_{A^*}^S - C_{A^*}^T)f^2 \approx 20. \quad (5.12)$$

Table 5.1 gives details of how this value is arrived at. It also lists the contributions of each chiral order to the self-energy (5.9),  $\Delta V = M_K^* - M_K$  and  $M_K^*$  which we shall loosely call “effective kaon mass”. This is not, strictly speaking, a mass but we shall refer to it as such in labeling the figures,

$$M_K^* \equiv \sqrt{M_K^2 + \Pi_K}. \quad (5.13)$$

Friedman et al. [18] note that their “nominal” optical potential gives an attraction of order of 800 MeV when extrapolated to three times the normal density. We show in Table 5.2 what our theory predicts at higher densities than normal. At  $u = 3$ , the net attraction is only about 1.7 times the one at  $u = 1$ .

Eq. (5.7) shows that for symmetric nuclear matter ( $x = 1/2$ ), the combination  $(C_{A^*}^S + C_{A^*}^T)$  does not enter into the self-energy formula. In order to extract it as needed for non-symmetric system as in compact star matter, we need information for nuclei with  $x \neq 1/2$ . This can be done from the results of Friedman et al. by noting that our self-energy is nonlinear in  $x$ , so

$$\frac{\partial \Delta V}{\partial x}(C_{A^*}^S, \rho \approx \rho_0)|_{x=1/2} \approx 400 b_1/b_0 \text{ MeV}, \quad (5.14)$$

where  $b_{0,1}$  are the constants given by Friedman et al. This relation determines the coefficient  $C_{A^*}^S$ . The result is shown in Table 5.3.

Table 5.2

$K^-$  self-energies in symmetric nuclear matter ( $x = 0.5$ ) in unit of  $M_K^2$  for  $g_{A^*}^2 = 0.25$  and  $(C_{A^*}^S - C_{A^*}^T)f^2 = 20$ .  $\Delta V \equiv M_K^* - M_K$  is the attraction (in unit of MeV) at given density.

$u$	$M_K^*$	$\Delta V$	$-\rho T^{free}$	$-\rho \delta T^{free}$	$\Pi_{A^*}^L$	$\Pi_{A^*}^2$
0.2	410.4	-84.57	-0.0565	0.0027	-0.2671	0.0081
0.4	369.9	-125.1	-0.0804	0.0067	-0.3663	0.0085
0.6	340.5	-154.5	-0.1093	0.0115	-0.4303	0.0086
0.8	316.2	-178.8	-0.1430	0.0168	-0.4725	0.0085
1.0	294.4	-200.6	-0.1808	0.0226	-0.4935	0.0081
1.2	275.0	-220.0	-0.2216	0.0286	-0.5041	0.0077
1.4	257.5	-237.5	-0.2647	0.0350	-0.5064	0.0072
1.6	240.8	-254.2	-0.3095	0.0416	-0.4975	0.0066
1.8	226.1	-268.9	-0.3549	0.0486	-0.4904	0.0061
2.0	211.7	-283.3	-0.4009	0.0557	-0.4739	0.0055
2.2	199.0	-296.0	-0.4465	0.0631	-0.4596	0.0050
2.4	186.7	-308.3	-0.4918	0.0707	-0.4406	0.0045
2.6	175.1	-319.9	-0.5364	0.0785	-0.4196	0.0041
2.8	164.1	-330.9	-0.5801	0.0865	-0.3977	0.0036
3.0	153.9	-341.1	-0.6226	0.0948	-0.3761	0.0032
3.2	144.5	-350.5	-0.6640	0.1034	-0.3552	0.0028
3.4	135.5	-359.5	-0.6971	0.1121	-0.3339	0.0025
3.6	128.6	-366.4	-0.7330	0.1216	-0.3238	0.0023
3.8	121.0	-374.0	-0.7689	0.1309	-0.3052	0.0021
4.0	113.2	-381.8	-0.8043	0.1403	-0.2832	0.0018

Table 5.3

Determination of  $C_{A^*}^S$  from the kaonic atom data[18] for  $(C_{A^*}^S - C_{A^*}^T)f^2 = 20$ .  $y = 0.41$  corresponds to no  $\Lambda(1405)$  mass shift in medium at the normal matter density.

$y = C_{A^*}^S f^2$	$\partial \Delta V / \partial x$	$b_1/b_0$
50	125.44 MeV	0.314
40	64.77 MeV	0.162
30	4.10 MeV	0.010
20	-56.58 MeV	-0.141
10	-117.25 MeV	-0.293
0.41	-175.43 MeV	-0.439
0	-177.92 MeV	-0.445
-10	-238.59 MeV	-0.596

Friedman et al. [18] find the acceptable value to be  $b_1/b_0 = -0.56 \pm 0.82$ . But there is one point which needs to be discussed in interpreting this number in the context of our theory. The constant  $C_{A^*}^S$  shifts linearly the effective in-medium mass of  $\Lambda(1405)$ , with the mass shift being given by

$$\delta m_{A^*} = \sum_{i=a,b} \delta \Sigma_{A^*}^{(i)} (\omega = m_{A^*} - m_B), \quad (5.15)$$

Table 5.4

Numerical values in MeV of  $r$  as function of  $x$  and  $u$ .

$r(u, x)$	$x = 0.0$	$x = 0.5$	$x = 1.0$
$u = 0.5$	30.78	38.25	30.78
$u = 1.0$	47.52	61.55	47.52
$u = 1.5$	60.20	79.75	60.20

Table 5.5

Self-energies for  $K^+$  in nuclear matter ( $x = 0.5$ ) in unit of  $M_K^2$  for  $g_{A^+}^2 = 0.25$  and  $(C_{A^+}^S - C_{A^+}^T)f^2 = 20$ .  $\Delta V \equiv M_K^* - M_K$  is the repulsion (in unit of MeV) at given density.

$u$	$M_K^*$	$\Delta V$	$-\rho T^{free}$	$-\rho\delta T^{free}$	$\Pi_{A^+}^1$	$\Pi_{A^+}^2$
0.2	505.8	10.8	0.0252	0.0202	-0.0019	-0.0024
0.4	514.1	19.1	0.0481	0.0430	-0.0076	-0.0052
0.6	521.7	26.7	0.0689	0.0638	-0.0174	-0.0077
0.8	527.9	32.9	0.0879	0.0895	-0.0312	-0.0110
1.0	533.4	38.4	0.1051	0.1183	-0.0493	-0.0146
1.2	538.4	43.4	0.1207	0.1505	-0.0716	-0.0187
1.4	542.9	47.9	0.1348	0.1868	-0.0982	-0.0234
1.6	547.1	52.1	0.1473	0.2265	-0.1291	-0.0286
1.8	551.0	56.0	0.1583	0.2698	-0.1645	-0.0343
2.0	554.1	59.1	0.1692	0.3281	-0.2041	-0.0418
2.2	557.5	62.5	0.1778	0.3816	-0.2483	-0.0488
2.4	560.8	65.8	0.1848	0.4372	-0.2971	-0.0562
2.6	563.2	68.2	0.1929	0.5129	-0.3501	-0.0661
2.8	566.2	71.2	0.1976	0.5771	-0.4079	-0.0747
3.0	568.2	73.2	0.2040	0.6638	-0.4698	-0.0860
3.2	571.0	76.0	0.2063	0.7356	-0.5369	-0.0957
3.4	572.8	77.8	0.2115	0.8341	-0.6078	-0.1086
3.6	574.5	79.5	0.2156	0.9352	-0.6833	-0.1220
3.8	576.9	81.9	0.2153	1.0210	-0.7642	-0.1336
4.0	578.4	83.4	0.2187	1.1350	-0.8486	-0.1486

where

$$\delta\Sigma_{A^+}^{(a)}(\omega) = -\frac{g_{A^+}^2}{f^2}\omega^2 (\Sigma_K^p(\omega) + \Sigma_K^s(\omega)), \quad \delta\Sigma_{A^+}^{(b)}(\omega) = -C_{A^+}^S(\rho_p + \rho_n). \quad (5.16)$$

The superscript (a, b) stands for the figures (a) and (b) of Fig. 5.4. The contribution from Fig. 5.4a is completely given with the known constants. The dependence on the unknown constant  $C_{A^+}^S$  appears linearly in the Fig. 5.4b. For the given values adopted here, the mass shift is numerically

$$\delta m_{A^+}(u, x, y) = [r(u, x) - 150.3 \times u \times y] \text{ MeV}, \quad (5.17)$$

where  $y = C_{A^+}^S f^2$  and  $r(u, x) \equiv \delta\Sigma_{A^+}^{(a)}$  with the numerical values given in Table 5.4.

For nuclear matter density  $u = 1$  and  $x = 1/2$ , the shift is

$$\delta m_{A^+}(u, x, y) \approx [62 - 150.3 \times y] \text{ MeV}, \quad (5.18)$$

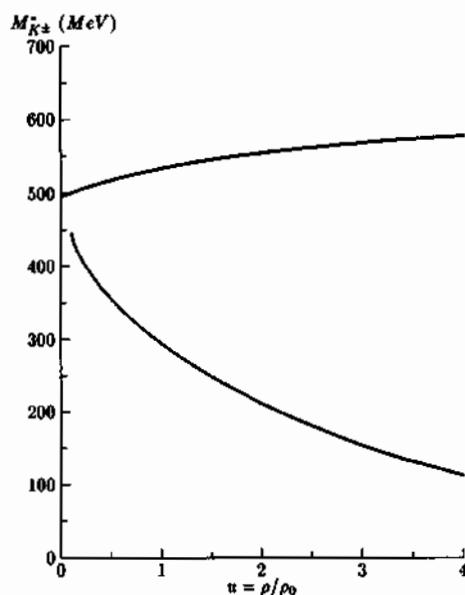


Fig. 5.3. Plot of the  $K^\pm$  effective potential in symmetric nuclear matter ( $x = 0.5$ ). The upper solid line corresponds to  $K^+$  “effective mass” and the lower solid line to  $K^-$  “effective mass” for  $(C_{A^*}^S - C_{A^*}^T)f^2 \approx 20$ .

with  $y = C_{A^*}^S f^2$ . It seems highly unlikely that the  $\Lambda(1405)$  will be shifted by hundreds of MeV in nuclear matter. This means that  $y$  must be of  $O(1)$ , and *not*  $O(10)$ . For  $y = 0.41$  which corresponds to  $b_1/b_0 \approx -0.4$ , there is no shift at normal matter density. We believe this is a reasonable value. In fact,  $y = 0$  is also acceptable. It would be interesting to measure the shift of  $\Lambda(1405)$  to fix the constant  $C_{A^*}^S$  more precisely although its precise magnitude seems to matter only a little for kaonic atoms and as it turns out, negligibly for kaon condensation.

Let us comment briefly on the role of multi-Fermion Lagrangians. This will eliminate another red herring in the literature. The Weinberg counting rule shows that the four-Fermi interactions are suppressed by  $O(Q^2)$  relative to the terms involving bilinears of Fermi fields. In general  $n$ -Fermi interactions will be suppressed by the same order relative to  $(n-1)$ -Fermi interactions. In considering kaon condensation, what this means in conjunction with the renormalization-group flow argument, is that  $n$ -Fermi interactions with  $n \geq 4$  are irrelevant in the RGE sense, and hence unimportant for condensation [22–24]. In Table 5.2, the situation with the kaonic atom data is a bit different. While the strength of the four-Fermi interaction,  $y$ , is not important its presence is essential for the attraction that seems to be required. This is in contrast to the kaon condensation which is driven by the “mass flow” with four-Fermi interactions being irrelevant in the RGE sense discussed in Section 3.3.

To exhibit the role of  $A^*$  in the kaon self-energy, we list each contribution of  $\Pi$ . Here  $\Pi_{free} = -\rho_N T_{free}^{K-N}$ ,  $\delta\Pi = -\rho_N \delta T^{K-N}$ ,  $\Pi_{A^*}^1$  corresponds to the first term of Eq. (5.7) which depends on  $C_{A^*}^{S,T}$  and  $\Pi_{A^*}^2$  to the second term independent of  $C_{A^*}^{S,T}$ . We observe that the  $C_{A^*}^{S,T}$ -dependent term plays a crucial role for attraction in kaonic atom. In Table 5.2 and Fig. 5.3, we list the predicted density dependence of the real part of the  $K^-$  optical potential for  $x = 0.5$  obtained for  $(C_{A^*}^S - C_{A^*}^T)f^2 \approx 20$ .

In Table 5.5 and Fig. 5.3 are given the properties of  $K^+$  in nuclear matter. The self-energy of  $K^+$

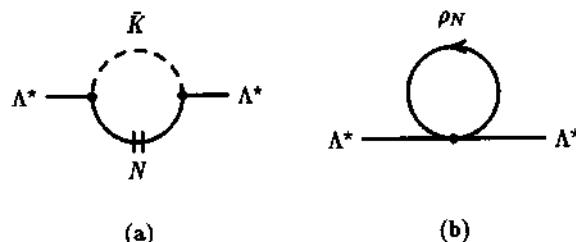


Fig. 5.4. One-loop diagrams contributing to the  $A^*$  mass shift in dense medium. The diagram (a) involves the intermediate states of  $K^- p$  and  $K^0 n$  and (b) the four-Fermi interaction  $C_F^S$ , with both protons and neutrons.

is simply obtained from that of  $K^-$  by crossing  $\omega \rightarrow -\omega$ . One can note here that the interaction of  $K^+$  with nuclear medium is quite weak as predicted by phenomenological models and supported by experiments. The  $M_K^*$  grows slowly as a function of density. This is a check of the consistency of the chiral expansion approach to kaon-nuclear interactions.

## 6. Kaon condensation

We have now all the ingredients needed to calculate the critical density for negatively charged kaon condensation in dense nuclear star matter. For this, we will follow the procedure given in [8]. As argued in [5], we need not consider pions when electrons with high chemical potential can trigger condensation through the process  $e^- \rightarrow K^-\nu_e$ . Thus we can focus on the spatially uniform condensate

$$\langle K^- \rangle = v_K e^{-i\mu t}, \quad (6.1)$$

where  $\mu$  is the chemical potential which is equal, by Baym's theorem [38], to the electron chemical potential.

### 6.1. Kaon effective mass and critical density

The energy density  $\tilde{\epsilon}$  – which is related to the effective potential in the standard way – is given by,

$$\begin{aligned}\tilde{\epsilon}(u, x, \mu, v_K) = & \frac{3}{5} E_F^{(0)} u^{5/3} \rho_0 + V(u) + u \rho_0 (1 - 2x)^2 S(u) - [\mu^2 - M_K^2 - H_K(\mu, u, x)] v_K^2 \\ & + \sum_{n \geq 2} a_n(\mu, u, x) v_K^n + \mu u \rho_0 x + \tilde{\epsilon}_e + \theta(|\mu| - m_\mu) \tilde{\epsilon}_\mu,\end{aligned}\quad (6.2)$$

where  $E_F^{(0)} = (p_F^{(0)})^2/2m_B$  and  $p_F^{(0)} = (3\pi^2\rho_0/2)^{1/3}$  are, respectively, Fermi energy and momentum at nuclear density. The  $V(u)$  is a potential for symmetric nuclear matter as described in [67] which is presumably subsumed in contact four-Fermi interactions (and one-pion-exchange – nonlocal – interaction) in the non-strange sector as mentioned above. It will affect the equation of state in the condensed phase but not the critical density, so we will drop it from now on. The nuclear symmetry energy  $S(u)$  – also subsumed in four-Fermi interactions in the non-strange sector – does play a role as we know from [67]: Protons enter to neutralize the charge of condensing  $K^-$ 's making the resulting

compact star “nuclear” rather than neutron star as one learns in standard astrophysics textbooks. We take the form advocated in [67]

$$S(u) = \left(2^{2/3} - 1\right) \frac{3}{5} E_F^{(0)} \left(u^{2/3} - F(u)\right) + S_0 F(u), \quad (6.3)$$

where  $F(u)$  is the potential contributions to the symmetry energy and  $S_0 \simeq 30 \text{ MeV}$  is the bulk symmetry energy parameter. We use three different forms of  $F(u)$  as in [67]

$$F(u) = u, \quad F(u) = \frac{2u^2}{1+u}, \quad F(u) = \sqrt{u}. \quad (6.4)$$

It will turn out that the choice of  $F(u)$  does not significantly affect the critical density. The contributions of the filled Fermi seas of electrons and muons are [8]

$$\begin{aligned} \bar{\epsilon}_e &= -\frac{\mu^4}{12\pi^2}, \\ \bar{\epsilon}_\mu &= \epsilon_\mu - \mu\rho_\mu = \frac{m_\mu^4}{8\pi^2} \left( (2t^2 + 1)t\sqrt{t^2 + 1} - \ln(t^2 + \sqrt{t^2 + 1}) \right) - \mu \frac{p_{F_\mu}^3}{3\pi^2}, \end{aligned} \quad (6.5)$$

where  $p_{F_\mu} = \sqrt{\mu^2 - m_\mu^2}$  is the Fermi momentum and  $t = p_{F_\mu}/m_\mu$ . Here we ignore hyperon Fermi seas in this calculation.

The ground-state energy prior to kaon condensation is obtained by extremizing the energy density  $\bar{\epsilon}$  with respect to  $x$ ,  $\mu$  and  $v_K$ ,

$$\frac{\partial \bar{\epsilon}}{\partial x}\Big|_{v_K=0} = 0, \quad \frac{\partial \bar{\epsilon}}{\partial \mu}\Big|_{v_K=0} = 0, \quad \frac{\partial \bar{\epsilon}}{\partial v_K^2}\Big|_{v_K=0} = 0, \quad (6.6)$$

from which we obtain three equations corresponding, respectively, to beta equilibrium, charge neutrality and dispersion relation,

$$\begin{aligned} \mu &= 4(1 - 2x)S(u), \quad 0 = -xu\rho_0 + \frac{\mu^3}{3\pi^2} + \theta(\mu - m_\mu) \frac{p_{F_\mu}^3}{3\pi^2}, \\ 0 &= D^{-1}(\mu, u, x) = \mu^2 - M_K^2 - \Pi_K(\mu, u, x) \equiv \mu^2 - M_K^{*2}(\mu, u, x). \end{aligned} \quad (6.7)$$

The proton fractions  $x(u)$  and the chemical potentials  $\mu$  prior to kaon condensation are plotted in Figs. 6.1 and 6.2 for various choices of the symmetry energy  $F(u)$ . We have solved these equations using for the kaon self-energy (a) the linear density approximation, Eq. (5.1) and (b) the full two-loop result, Eq. (5.9). Table 6.1a shows the case (a) for different symmetry energies Eq. (6.4). We see that the precise form of the symmetry energy does not matter quantitatively. The corresponding “effective kaon masses”  $M_K^*$  are plotted vs  $u$  in Figs. 6.3, 6.4 and 6.5 in solid line. Note that even in this linear density approximation kaon condensation *does* take place, *albeit* at a bit higher density than obtained before.

For the value that seems to be required by the kaonic atom data, (5.12), the critical density comes out to be about  $u_c \approx 3$ , rather close to the original Kaplan–Nelson value.

In Table 6.1b and Figs. 6.3, 6.4 and 6.5 are given the predictions for a wide range of values for  $C_A^S f^2$ . What is remarkable here is that while the  $C_A^{S,T}$ -dependent four-Fermi interactions are *essential*

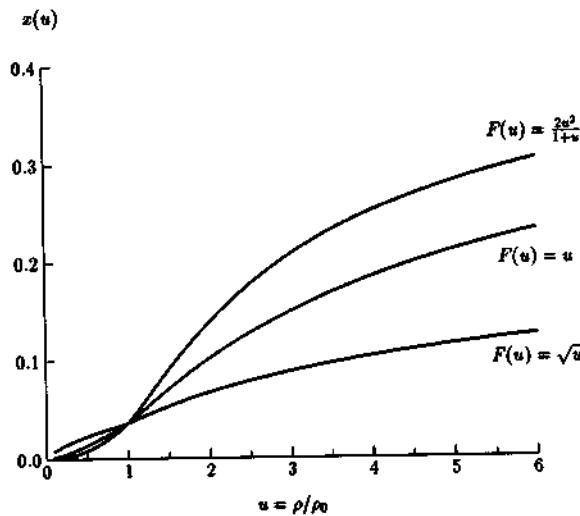


Fig. 6.1. Plot of the proton fraction  $x(u)$  prior to kaon condensation for different forms of  $F(u)$ .

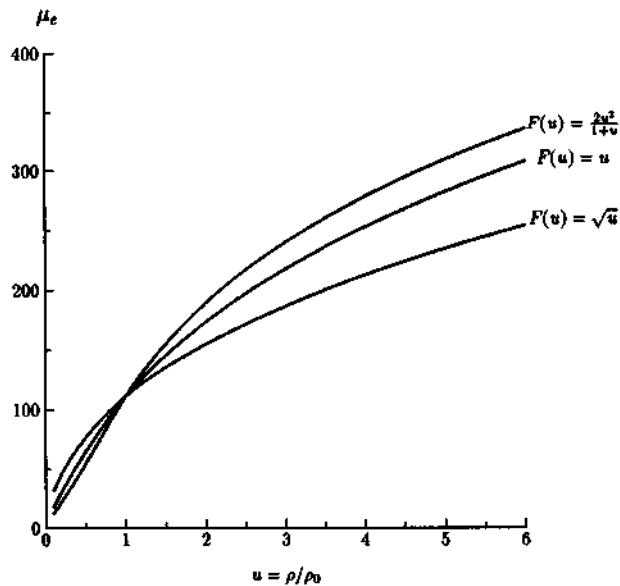


Fig. 6.2. Plot of the electron chemical potential  $\mu_e(u)$  prior to kaon condensation for different forms of  $F(u)$ .

for triggering kaon condensation, the critical density is quite insensitive to their strengths. In fact, as one can see in Table 6.2, reducing the constant  $(C_A^S - C_A^T)f^2$  that represents the kaonic atom attraction by an order of magnitude to 1 with  $C_A^S f^2 = 10$ , 0 modifies the critical density only to  $u_c \approx 3.3, 4.9$ , respectively.

As discussed in Section 4.2.2, there are ambiguities in the  $\mathcal{O}(Q^3)$  counter terms for off-shell.

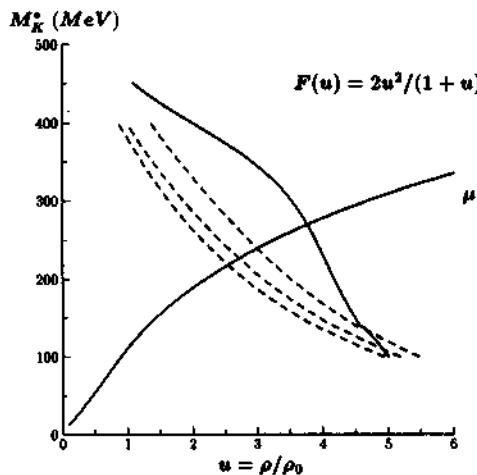


Fig. 6.3. Plot of the quantity  $M_K^*$  obtained from the dispersion formula  $D^{-1}(\mu, u) = 0$  vs the chemical potential  $\mu$  prior to kaon condensation for  $g_{A^*}^2 = 0.25$  and  $F(u) = 2u^2/(1+u)$ . The solid line corresponds to the linear density approximation and the dashed lines to the in-medium two-loop results for  $(C_{A^*}^S - C_{A^*}^T)f^2 = 20$  and  $C_{A^*}^S f^2 = 20, 10, 0$ , respectively, from the left. The point at which the chemical potential  $\mu$  intersects  $M_K^*$  corresponds to the critical point.

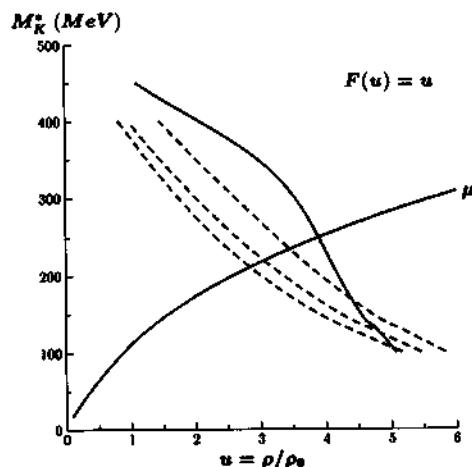


Fig. 6.4. The same as Fig. 6.3 for  $F(u) = u$ .



Fig. 6.5. The same as Fig. 6.3 for  $F(u) = \sqrt{u}$ .

$$\tilde{g}_{s,v}(\omega) = \alpha'_{s,v} \frac{\omega}{M_K} + \beta'_{s,v} \frac{\omega^3}{M_K^3} + L(\omega), \quad (6.8)$$

where  $L(\omega)$  is the loop contribution. In previous results, the critical densities are calculated under the assumption  $\alpha'_{s,v} = \beta'_{s,v}$ . To see the effect of these ambiguities, we will take simple parametrization. Because of the constraint in on-shell, we can parametrize  $\alpha'_{s,v}$  and  $\beta'_{s,v}$ ,

Table 6.1

Critical density  $u_c$  resulting from the in-medium two-loop chiral perturbation theory for  $g_{A^*}^2 = 0.25$ . (a) correspond to linear density approximation of Fig. 5.1a and (b) to the full two-loop result for  $(C_{A^*}^S - C_{A^*}^T)f^2 = 20$ .  $y = 0.41$  corresponds to no  $\Lambda(1405)$  mass shift in medium at the normal matter density. See Table 5.3.

	$y = C_{A^*}^S f^2$	$F(u) = \frac{2u^2}{1+u}$	$F(u) = u$	$F(u) = \sqrt{u}$
(a)		3.77	3.90	4.11
(b)	50	2.25	2.50	2.97
	40	2.33	2.58	3.08
	30	2.42	2.69	3.22
	20	2.54	2.84	3.41
	10	2.71	3.05	3.71
	0.41	2.98	3.43	4.28
	0	2.99	3.45	4.32
	-10	3.60	4.85	~6.41

Table 6.2

Critical density  $u_c$  resulting from the in-medium two-loop chiral perturbation theory for  $F(u) = u$ .

$g_{A^*}^2$	$(C_{A^*}^S - C_{A^*}^T)f^2$	$C_{A^*}^S f^2 = 100$	$C_{A^*}^S f^2 = 10$	$C_{A^*}^S f^2 = 0$
0.25	1	2.25	3.29	4.91
	10	2.25	3.16	3.76
	100	2.18	2.67	2.79
0.05	1	2.97	4.19	5.59
	10	2.96	4.01	4.64
	100	2.85	3.36	3.47

Table 6.3

$K^-$  self-energies in symmetric nuclear matter ( $x = 0.5$ ) in unit of  $M_K^2$  at  $u = 0.97$  for  $g_{A^*}^2 = 0.05$ .  $\Delta V = M_K^* - M_K$  is the attraction (in unit of MeV) at a given  $(C_{A^*}^S - C_{A^*}^T)f^2$ .

$(C_{A^*}^S - C_{A^*}^T)f^2$	$M_K^*$	$\Delta V$	$-\rho T^{free}$	$-\rho \delta T^{free}$	$H_{A^*}^1$	$H_{A^*}^2$
10	386.4	-108.6	-0.1142	0.0402	-0.3244	0.0092
20	364.6	-130.4	-0.1158	0.0335	-0.3805	0.0054
50	330.1	-164.9	-0.1279	0.0262	-0.4548	0.0026
70	315.9	-179.1	-0.1347	0.0240	-0.4831	0.0019
100	300.0	-195.0	-0.1429	0.0220	-0.5132	0.0014

$$\alpha'_{s,\nu} = sC_{s,\nu}, \quad \beta'_{s,\nu} = (1-s)C_{s,\nu}. \quad (6.9)$$

where  $C_{s,\nu}M_K^3 = (\bar{g}_{s,\nu}(M_K) - L(M_K))M_K^3 \approx -1000$  MeV. The resulting critical densities are summarized in Table 6.5; We see that the critical densities are not very sensitive to the  $s$  value.

To see how robust kaon condensation is with respect to the  $A^*KN$  coupling constant, let us take the extreme value  $g_{A^*}^2 \approx 0.05$  used in [51] which gives the wrong sign to the  $K^-p$  amplitude at

Table 6.4

Critical density  $u_c$  resulting from the in-medium two-loop ChPT for  $(C_{A^*}^S - C_{A^*}^T)f^2 = 100$ ,  $g_{A^*}^2 = 0.05$ 

$F(u)$	$C_{A^*}^S f^2 = 100$	$C_{A^*}^S f^2 = 50$	$C_{A^*}^S f^2 = 0$
$2u^2/1+u$	2.55	2.72	3.01
$u$	2.85	3.07	3.47
$\sqrt{u}$	3.42	3.72	4.35

Table 6.5

Critical density  $u_c$  for different choice of  $s$  under the constraint  $(C_{A^*}^S - C_{A^*}^T)f^2 = 20$ 

$F(u)$	$C_{A^*}^S f^2$	$s = 0.0$	$s = 0.5$	$s = 1.0$
$u$	0	3.79	3.45	3.10
	10	3.30	3.05	2.81
	20	3.03	2.84	2.65
$2u^2/1+u$	0	3.20	2.99	2.77
	10	2.87	2.71	2.54
	20	2.67	2.54	2.40
$\sqrt{u}$	0	5.21	4.32	3.60
	10	4.18	3.71	3.28
	20	3.76	3.41	3.09

threshold. In Table 6.3,  $\Delta V$  of kaonic atom ( $x = 0.5$ ) is given for  $u = 0.97$  and for various choices of  $(C_{A^*}^S - C_{A^*}^T)f^2$ . Constraining to the kaonic atom data implies approximately – within the range of error involved –  $(C_{A^*}^S - C_{A^*}^T)f^2 \approx 100$ . In Table 6.4, the resulting critical densities are given for  $(C_{A^*}^S - C_{A^*}^T)f^2 \approx 100$ , and those for other sets of  $C_{A^*}^S$  and  $C_{A^*}^T$  in Table 6.2. We see that given the constraint from kaonic atom data, the resulting critical densities are sensitive neither to the value of  $C_{A^*}^{ST}$  nor to  $g_{A^*}$ .

## 6.2. Essence of kaon condensation

To see which modes are involved in s-wave kaon condensation, we consider the dispersion formula at tree order,

$$D^{-1}(\omega) = \omega^2 - M_K^2 - \Pi(\omega). \quad (6.10)$$

As shown in Fig. 6.6 by solid line, the kaon “effective mass”  $M_K^*$  is reduced mainly by Weinberg–Tomozawa and the KN sigma terms when there are no  $A^*$  contribution. If we turn on the  $A^*$  coupling, there will be additional attractions. However since the effective mass  $M_K^*$  lies far from the  $A^*$ -pole contribution, the resulting magnitude of the attraction is small, i.e., the  $M_K^*$  remains nearly unmodified. Furthermore, since the  $A^*$ -pole is far outside of  $M_K^*$ , the condensed kaon mode remains the same independently of the  $A^*$ .

The results shown in Fig. 5.3 and Figs. 6.3, 6.4 and 6.5 correspond to the lowest branch which goes over to the  $K^-$  branch when the  $A^*KN$  coupling is turned off. Since the s-wave condensation is

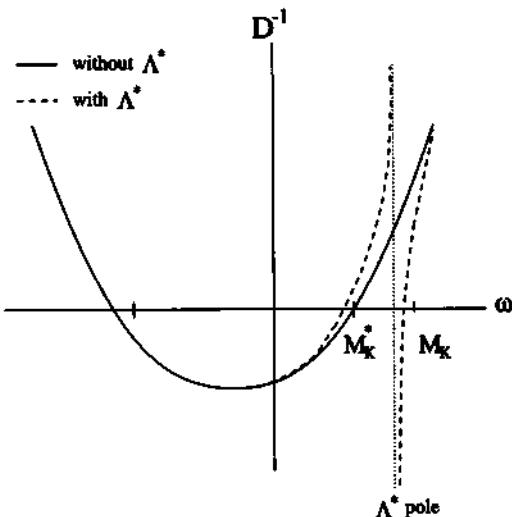


Fig. 6.6. Plot of  $D^{-1} = \mu^2 - M_K^2 - \Pi(\mu)$ .

driven by a mass flow, collective “phonons” do not seem to play an important role in contrast to the p-wave condensation (e.g., pion condensation) driven by the flow of a Yukawa coupling.

Summarizing the results,  $\Lambda^*$  may be crucial for understanding the KN scattering and kaonic atom data, but is irrelevant to determining the kaon condensation. The critical densities for wide ranges of  $\Lambda^*$  couplings in Table 6.2 confirm the unimportance of  $\Lambda^*$  contribution.

At the energies where kaon condensation sets in, we can also neglect one-loop contributions in addition to the  $\Lambda^*$  contributions. At this energy, the tree order contributions are dominant,

$$\begin{aligned}\mathcal{T}^{K^-p} &= \frac{1}{f^2} \left\{ \omega + \Sigma_{KN} \left( 1 - 0.37 \frac{\omega^2}{M_K^2} \right) + C \left( \frac{\omega^2}{M_K^2} + 0.083 \right) + \mathcal{O}(Q^3) \right\} + \mathcal{T}_{\Lambda^*}^{K^-p}, \\ \mathcal{T}^{K^-n} &= \frac{1}{f^2} \left\{ \frac{\omega}{2} + \Sigma_{KN} \left( 1 - 0.37 \frac{\omega^2}{M_K^2} \right) - C \left( \frac{\omega^2}{M_K^2} + 0.083 \right) + \mathcal{O}(Q^3) \right\},\end{aligned}\quad (6.11)$$

where  $\Sigma_{KN} \simeq 2.83 M_\pi$  and  $C = 15.0 \text{ MeV}$  which is negligible. Near  $\omega = M_K/2$  the approximate kaon self energy is given by

$$\Pi_{K^-}(\omega) \simeq -\rho_p \mathcal{T}^{K^-p} - \rho_n \mathcal{T}^{K^-n} \sim -\frac{3}{2f^2} \Sigma_{KN} \rho. \quad (6.12)$$

Then the optical potential is

$$\begin{aligned}\Delta V_{K^-} &\simeq \frac{1}{2} \frac{\Pi_{K^-}}{M_K(1 + M_K/m_p)} \simeq -\frac{3}{4f^2} \frac{\Sigma_{KN} \rho_0}{M_K(1 + M_K/m_p)} \left( \frac{\rho}{\rho_0} \right) \\ &\simeq -135 \text{ MeV} \left( \frac{\rho}{\rho_0} \right).\end{aligned}\quad (6.13)$$

This attraction is the essence of the kaon condensation in dense stellar matter. The comparison with the mean field approach [28] will be discussed later.

### 6.3. "Scaled" chiral Lagrangians

So far we have ignored four-Fermi interactions in the non-strange sector with the understanding that the effects not involving strangeness are to be taken from what we know from nuclear phenomenology. From the chiral Lagrangian point of view, this is not satisfactory. One would like to be able to describe both the non-strange and strange sectors *on the same footing* starting from a three-flavor chiral Lagrangian. While recent developments indicate that nuclear forces may be understood at low energies in terms of a chiral Lagrangian [11,12], it has not yet been possible to describe the ground-state property of nuclei including nuclear matter starting from a Lagrangian that has explicit chiral symmetry. So the natural question is: What about many-body correlations in the non-strange sector, not to mention those in the strange sector for which we are in total ignorance?

Montano, Politzer and Wise [20] addressed a related question in pion condensation in the chiral limit and found that four-Fermi interactions in the non-strange sector played an important role in inhibiting p-wave pion condensation in dense matter. It has of course been known since some time that the mechanism that quenches  $g_A$  in nuclear matter to a value close to unity banishes the pion condensation density beyond the relevant regime. As shown in [68], this mechanism can be incorporated by means of a four-Fermi interaction in a chiral Lagrangian in a channel corresponding to the spin-isospin mode (that is, the Landau–Migdal  $g'$  interaction).

#### 6.3.1. BR scaling

In this section we discuss a simple approach to including the main correlations in kaon condensation [1]. We cannot do so in full generality to all orders in density but we can select what we consider to be the dominant ones by resorting to the scaling argument (which we shall call "BR scaling") introduced by Brown and Rho [25,26]. How to implement the BR scaling in higher-order chiral expansion with the multiple scales that we are dealing with has not yet been completely worked out. It has up to date been formulated so as to be implemented *only at tree order* with the assumption that once the BR scaling is incorporated, higher-order terms are naturally suppressed. This assumption is checked *a posteriori*, [1] by taking the BR scaling into account at tree order, that is, at  $\mathcal{O}(Q^2)$  in our case, we will be including most of higher-order density dependences through the simple scaling.

As discussed in Ref. [1], we expect that the higher-density dependence modifies the Weinberg–Tomozawa term to

$$\sim \frac{1}{f^{*2}} K^\dagger \partial_0 K B_v^\dagger B_v, \quad (6.14)$$

where the asterisk denotes density dependence  $f^* = f(\rho)$ . We assume this to be given by the BR scaling,

$$\frac{f^*}{f} \approx 1 - 0.15u. \quad (6.15)$$

The self-energy for kaonic atoms with the BR scaling is given in Fig. 6.7 and tabulated in Tables 6.6 and 6.7. Comparing with Tables 5.1 and 5.2, we see that the BR scaling gives only a slightly more

Table 6.6

$K^-$  self-energies with BR scaling in symmetric nuclear matter ( $x = 0.5$ ) in unit of  $M_K^2$  at  $u = 0.97$  for  $g_{A^*}^2 = 0.25$ .  $\Delta V = M_K^* - M_K$  is the attraction (in unit of MeV) at a given  $(C_{A^*}^S - C_{A^*}^T)f^2$ .

$(C_{A^*}^S - C_{A^*}^T)f^2$	$M_K^*$	$\Delta V$	$-\rho T^{free}$	$-\rho \delta T^{free}$	$\Pi_{A^*}^1$	$\Pi_{A^*}^2$
1	392.2	-102.8	-0.3091	0.0424	-0.1888	0.0832
5	344.8	-150.2	-0.2598	0.0289	-0.3072	0.0233
10	319.1	-175.9	-0.2551	0.0245	-0.3669	0.0130
20	289.8	-205.2	-0.2578	0.0208	-0.4272	0.0070
50	246.5	-248.5	-0.2678	0.0171	-0.5038	0.0029
70	229.7	-265.3	-0.2720	0.0160	-0.5306	0.0021
100	211.7	-283.3	-0.2760	0.0151	-0.5570	0.0014

Table 6.7

$K^-$  self-energies with BR scaling in symmetric nuclear matter ( $x = 0.5$ ) in unit of  $M_K^2$  for  $g_{A^*}^2 = 0.25$  and  $(C_{A^*}^S - C_{A^*}^T)f^2 = 20$ .  $\Delta V \equiv M_K^* - M_K$  is the attraction (in unit of MeV) at given density.

$u$	$M_K^*$	$\Delta V$	$-\rho T^{free}$	$-\rho \delta T^{free}$	$\Pi_{A^*}^1$	$\Pi_{A^*}^2$
0.2	409.7	-85.26	-0.0590	0.0027	-0.2619	0.0079
0.4	369.8	-125.2	-0.0923	0.0067	-0.3647	0.0084
0.6	338.7	-156.3	-0.1373	0.0113	-0.4139	0.0082
0.8	311.4	-183.6	-0.1962	0.0164	-0.4316	0.0077
1.0	286.0	-209.0	-0.2698	0.0216	-0.4245	0.0068
1.2	261.1	-233.9	-0.3589	0.0268	-0.3962	0.0058
1.4	235.1	-259.9	-0.4649	0.0320	-0.3461	0.0046
1.6	206.5	-288.5	-0.5888	0.0368	-0.2772	0.0033
1.8	172.1	-322.9	-0.7311	0.0410	-0.1906	0.0019
2.0	127.8	-367.2	-0.8809	0.0446	-0.0981	0.0008

attractive potential at low densities. The attraction, however, increases significantly at higher densities. From Table 6.6, we find for kaonic atoms, approximately,

$$(C_{A^*}^S - C_{A^*}^T)f^2 \approx 20. \quad (6.16)$$

This is roughly the same as without the BR scaling since the scaling effect is not important at  $u = 0.97$ . The BR scaling becomes important in the region where condensation sets in. The corresponding critical density is given in Tables 6.8 and 6.9. The characteristic feature of the effect of the BR scaling is summarized in Figs. 6.8, 6.9 and 6.10. The remarkable thing to notice is that the critical density lowered to  $u_c \sim 2$  is completely insensitive to the parameters such as the form of the symmetry energy, the constants  $C_{A^*}^{ST}$  etc. in which possible uncertainties of the theory lie.

In order to verify the key hypothesis of the BR scaling – that the scaling subsumes higher order effects, thus *suppressing* higher chiral order effects of the scaled Lagrangian, we keep all the parameters *fixed* at the values determined at  $\mathcal{O}(Q^3)$ , BR-scale as described above, then ignore all terms of  $\mathcal{O}(Q^3)$  (i.e., ignore loop corrections except for the contribution of  $A^*$ ) and (A) set  $C_{A^*}^S = C_{A^*}^T = 0$  and (B) set  $(C_{A^*}^S - C_{A^*}^T)f^2 = 20$  and calculate the critical density. The results are given in Table 6.10. We see that the effects of loop corrections and four-Fermi interactions amount to less

Table 6.8

Critical density  $u_c$  with BR scaling resulting from the in-medium two-loop chiral perturbation theory for  $g_{A^*}^2 = 0.25$ . (a) correspond to linear density approximation of Fig. 5.1a and (b) to the full two-loop result for  $(C_{A^*}^S - C_{A^*}^T)f^2 = 20$ . See Table 5.3.

	$y = C_{A^*}^S f^2$	$F(u) = \frac{2u^2}{1+u}$	$F(u) = u$	$F(u) = \sqrt{u}$
(a)		2.11	2.16	2.22
(b)	20	2.03	2.11	2.22
	10	2.07	2.15	2.24
	0	2.13	2.20	2.27

Table 6.9

Critical density  $u_c$  with BR scaling resulting from the in-medium two-loop chiral perturbation theory for  $g_{A^*}^2 = 0.25$  and  $F(u) = u$ .

$(C_{A^*}^S - C_{A^*}^T)f^2$	$C_{A^*}^S f^2 = 100$	$C_{A^*}^S f^2 = 10$	$C_{A^*}^S f^2 = 0$
1	1.93	2.17	2.22
10	1.92	2.16	2.21
100	1.90	2.09	2.13

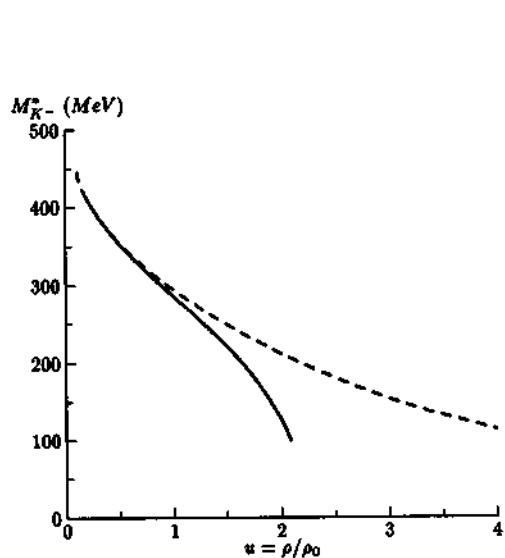


Fig. 6.7. Plot of the  $K^-$  effective potential in symmetric nuclear matter ( $x = 0.5$ ). The solid (dashed) line correspond to  $K^-$  "effective mass" with (without) BR scaling for  $(C_{A^*}^S - C_{A^*}^T)f^2 \approx 20$ .

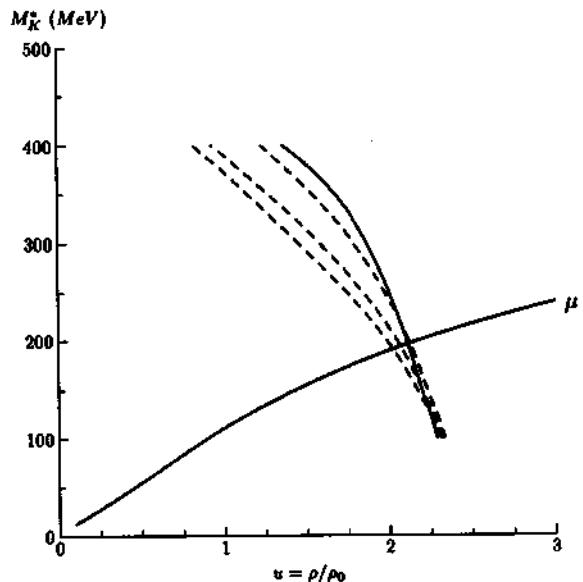
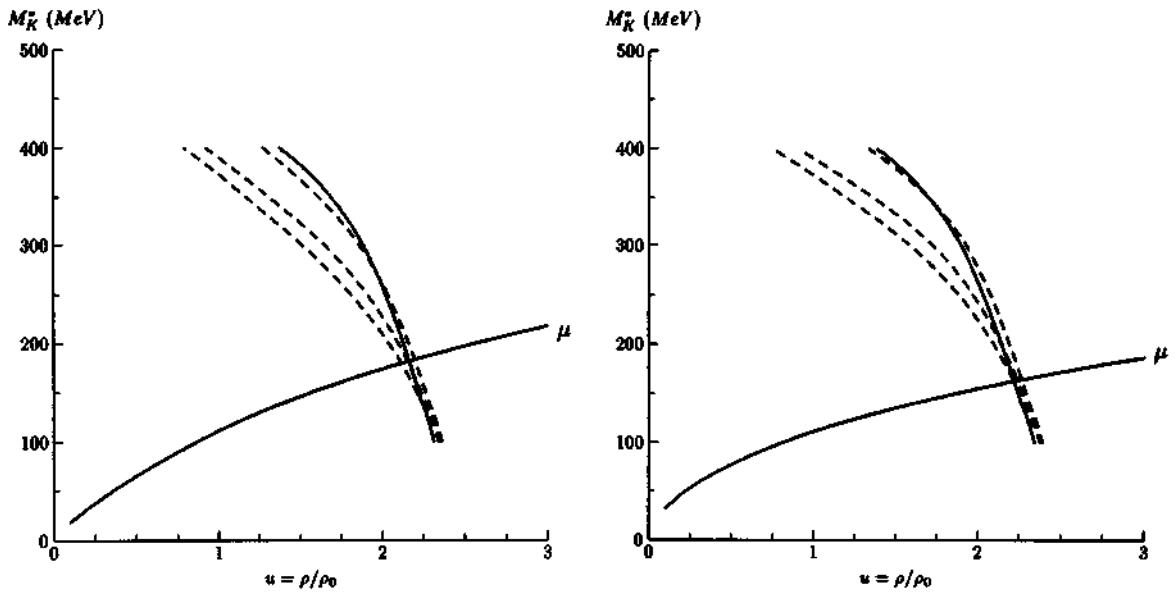


Fig. 6.8. The same as Fig. 6.3 with the BR scaling for  $F(u) = 2u^2/(1+u)$ .

Fig. 6.9. The same as Fig. 6.3 with the BR scaling for  $F(u) = u$ .Fig. 6.10. The same as Fig. 6.3 with the BR scaling for  $F(u) = \sqrt{u}$ .

than 10%. Recent mean-field approach of Brown and Rho [28] is consistent with our results. This will be discussed later.

To see the effect of BR scaling in the Weinberg-Tomozawa term more explicitly, ignore all terms of  $\mathcal{O}(Q^2)$  and  $\mathcal{O}(Q^3)$  except for the contribution of  $A^*$ , and (A) set  $C_{A^*}^S = C_{A^*}^T = 0$  and (B) set  $(C_{A^*}^S - C_{A^*}^T)f^2 = 20$  and calculate the critical density. The results are given in Table 6.11. In this table, once the BR scaling sets in, the Weinberg-Tomozawa term is more important than the scalar interaction. Even without the KN sigma term  $\Sigma_{KN}$ , we can have the kaon condensation with the critical density  $\rho_c < 3\rho_0$ . Recently Tatsumi et al. [27] obtained more or less the same results in the model-independent approach based on current algebra and PCAC. Their arguments will be presented later.

These results, Tables 6.10 and 6.11, may be taken as an a posteriori verification of the validity of the assumption made in deriving the scaling, although it is difficult to assess how reliable the absolute value of the predicted critical density (and the corresponding equation of state of the condensed phase) is.

### 6.3.2. Multi-fermion interactions

In this subsection, we consider the power counting when many nucleons are explicitly accounted for.

The momentum integral of the external nucleon (i.e., nucleon in matter),

$$\int_{k_f}^{k_i} d^3 k, \quad (6.17)$$

Table 6.10

Critical density  $u_c$  with BR scaling and without  $\mathcal{O}(Q^3)$  terms for  $g_{A^*}^2 = 0.25$  and (A)  $C_{A^*}^S = C_{A^*}^T = 0$  and (B)  $(C_{A^*}^S - C_{A^*}^T)f^2 = 20$ . This should be compared with the results of Table 6.9.

$y = C_{A^*}^S f^2$	$F(u) = \frac{2u^2}{1+u}$	$F(u) = u$	$F(u) = \sqrt{u}$
(A)	1.94	1.99	2.04
(B)	100	1.70	1.90
	10	1.88	2.02
	0	1.91	2.04

Table 6.11

Critical density  $u_c$  with BR scaling and with only  $\mathcal{O}(Q^1)$  terms for  $g_{A^*}^2 = 0.25$  and (A)  $C_{A^*}^S = C_{A^*}^T = 0$  and (B)  $(C_{A^*}^S - C_{A^*}^T)f^2 = 20$ . This should be compared with the results of Table 6.10.

$y = C_{A^*}^S f^2$	$F(u) = \frac{2u^2}{1+u}$	$F(u) = u$	$F(u) = \sqrt{u}$
(A)	2.47	2.59	2.66
(B)	100	1.97	2.13
	10	2.32	2.46
	0	2.40	2.52

implies that one nucleon carries the chiral power  $Q^3$ . Then we can define a new chiral power  $\nu'$  of the effective action as [69]

$$\nu' = \nu + 3(N_n - 1). \quad (6.18)$$

For one-nucleon process, our new chiral power is equal to  $\nu$ .

Starting from the power counting rule, we have

$$\nu = 4 - N_n - 2C + 2L + \sum_i \Delta_i. \quad (6.19)$$

The minimum power ( $\Delta_i = 0$  for all  $i$ ) of a 2-nucleon process ( $4 - N_n = 2$ ) corresponds to  $\nu = -2(\nu = 0)$  when  $C = 2(C = 1)$ . Then the total power of the effective action,  $C = 2$ , corresponds to  $\nu' = 1$ , which is reasonable because  $C = 2$  corresponds to the K-N scattering process with a non-interacting spectator nucleon. The genuine two-nucleon effect  $C = 1$  corresponds to  $\nu' = 3$ . Our result corresponding to Fig. 5.1 (b)~(f) is equivalent to a two-nucleon effect, as it involves two nucleon-momentum integrals ( $\nu' = 3$ ).

Now, consider the power of  $\Delta_i$  in each vertex,

$$\Delta_i = d_i + \frac{1}{2}n_i - 2. \quad (6.20)$$

By chiral symmetry, only derivative meson–nucleon coupling is allowed. This implies  $\Delta_i = 1$  for the following vertex:

$$K^\dagger \partial K (B_v^\dagger B_v)^2. \quad (6.21)$$

The chiral power of the effective action corresponding to this vertex is  $\nu' = 4$ . In the same way, the chiral powers corresponding to  $K^\dagger \partial K (B_v^\dagger B_v)^n$  with  $n \geq 2$  are higher than  $\nu' = 3$ . We do not need to consider these terms in our approach. The ONLY relevant multi-nucleon terms in our procedure are

$$C(B_v^\dagger B_v)^2, \quad C(B_v^\dagger \sigma B_v)^2. \quad (6.22)$$

These terms, however, are not considered explicitly in our calculation, since we cannot describe the ground state property of nuclei including nuclear matter starting from a Lagrangian that has explicit chiral symmetry. This problem is resolved in the mean-field approach of Brown and Rho [28]

#### 6.4. $\rho_s$ versus $\rho_B$

A criticism may be raised regarding the use in this review of the vector density  $B_v^\dagger B_v$  instead of the scalar density  $\bar{B}B$  (without the subscript  $v$ ) which would appear in relativistic Lagrangians [29,70]. However, in the heavy-fermion formalism that is used in this review – which is what is required for a consistent chiral expansion in the presence of baryon fields, the expansion is made in terms of  $\bar{B}_v(\cdots)B_v$  which is equivalent to  $B_v^\dagger(\cdots)B_v$  as one can easily verify by the projection operator that defines the subscripted baryon field. There is no approximation made in the calculation. Of course this quantity differs from the scalar density  $\bar{B}B$  (without the projection operator), particularly at high density. One might then wonder where the difference shows up in the HFF. It is easy to see where.

The difference between the scalar and vector density is  $\mathcal{O}(Q^2)$  in the chiral counting as we shall show. The scalar density in the relativistic approach is

$$\rho_s = \langle \bar{B}B \rangle = \left\langle \frac{m}{E} B^\dagger B \right\rangle. \quad (6.23)$$

This can be expanded as a power series of the nucleon momentum  $p$ ,

$$\rho_s = \left\langle \left( 1 - \frac{p^2}{m^2} + \mathcal{O}(p^4) \right) B^\dagger B \right\rangle \approx \rho_B - \left\langle \frac{p^2}{m^2} B^\dagger B \right\rangle. \quad (6.24)$$

Then the self-energy contribution coming from the leading order Lagrangian can be written as

$$H = A^{\nu=1} \rho_s \approx A^{\nu=1} \rho_B - A^{\nu=1} \left\langle \frac{p^2}{m^2} B^\dagger B \right\rangle. \quad (6.25)$$

Note that the first term of RHS is the leading order contribution in HBF, but the second term is higher order contribution by  $\mathcal{O}(Q^2)$  in HBF. This higher order term can be absorbed into  $c_1$ -type  $\nu = 3$  HBF-Lagrangian,

$$\begin{aligned} \mathcal{L}^{(3)} = & c_1 \text{Tr } \bar{B}_v(i\nu \cdot D)^3 B_v + c_2 \text{Tr } \bar{B}_v(i\nu \cdot D)(iD_\mu)^2 B_v \\ & + c_3 \text{Tr } \bar{B}_v(iD_\mu)(i\nu \cdot D)(iD^\mu) B_v + \dots \end{aligned} \quad (6.26)$$

If we consider the most general Lagrangian at  $\mathcal{O}(Q^3)$ , the effect of  $\rho_s$  applied to leading order Lagrangian is properly absorbed into the definition of the counter terms.

The effect of  $\rho_s$  to the next-to-leading order Lagrangian enters at  $\mathcal{O}(Q^4)$ , since the  $KN$  sigma term containing the quark mass matrix is of  $\mathcal{O}(Q^2)$ . Our calculation to the one-loop order is valid to  $\mathcal{O}(Q^3)$  and hence the difference does not figure in our theory to the order considered.

A consequence of the above observation is that it is inconsistent with the chiral expansion to use the scalar density in the Kaplan–Nelson term unless one includes at the same time all other counter terms of  $O(Q^4)$  (there are many with varying physical inputs, some known and some unknown) as well as all finite two-loop terms.

The one-loop corrections in our present paper evaluate *all* terms to the given order – i.e.,  $O(Q^3)$  – and all terms of  $O(Q^2)$  are correctly included in the Lagrangian  $\mathcal{L}^{(2)}$ .

## 7. Kaon condensation in other approaches

### 7.1. Mean-field approach of Brown and Rho

Recently, Brown and Rho [28] proposed a unified treatment of normal nuclear matter and kaon condensed matter using BR scaling. The former is described by Walecka's mean field theory which successfully describes the ground state properties of nuclear matter, while the latter by mean field chiral perturbation theory. In this section, we will summarize the mean field approach of Brown and Rho [28] to *s*-wave kaon condensation.

Consider the Weinberg–Tomozawa term and the sigma term that figure importantly in the  $K$ -nuclear interactions in symmetric nuclear matter,

$$\mathcal{L}_{KN} = -\frac{6i}{8f^2} (\bar{B}\gamma_0 B) \bar{K}\partial_\mu K + \frac{\Sigma_{KN}}{f^2} (\bar{B}B) \bar{K}K \equiv \mathcal{L}_\omega + \mathcal{L}_\sigma, \quad (7.1)$$

where  $K^T = (K^- \bar{K}^0)$  and  $f$  can be identified as free space pion decay constant  $f_\pi$ . In medium, this pion decay constant can be modified as  $f_\pi^*$ ,

$$\frac{f_\pi^*}{f_\pi} \approx 0.77, \quad (7.2)$$

as we shall see shortly.

One can interpret the first term of Eq. (7.1) as arising from integrating out the  $\omega$  meson. The resulting  $K^- N$  vector potential in medium is

$$V_{K^\pm} = \pm \frac{3}{8f_\pi^{*2}} \rho = \pm \frac{1}{3} V_N, \quad (7.3)$$

where  $V_N$  is the NN vector potential in Walecka mean field theory. The factor  $1/3$  arises because the  $\omega$  couples to a matter field kaon. The reason for this matter-field nature of the kaon is that all nonstrange hadrons become light in dense medium, so the kaon becomes in some sense heavy.

As for the second term of Eq. (7.1), we assume that the kaon behaves as a massive matter field. We therefore expect that it be coupled to the chiral scalar  $\chi$  as

$$\mathcal{L}_\sigma = \frac{1}{3} 2m_K g_\sigma^* \bar{K}K\chi, \quad (7.4)$$

where  $g_\sigma^* \approx 10$  [28] and the factor  $1/3$  accounts for one non-strange quark in the kaon as compared with three in the nucleon. Integrating out the scalar field, we have

$$\mathcal{L}_\sigma = 2m_K \frac{1}{3} \frac{g_\sigma^{*2}}{m_\sigma^{*2}} \bar{B} B \bar{K} K. \quad (7.5)$$

Comparing with the second term of Eq. (7.1), we have

$$\frac{\Sigma_{KN}}{f^2} \approx 2 \frac{m_K}{3} \frac{g_\sigma^{*2}}{m_\sigma^{*2}}. \quad (7.6)$$

Suppose we take the  $\Sigma_{KN}$  from the lattice calculations [61],  $\Sigma_{KN} \approx 3.2m_\pi$ . Then the left- and right-hand sides of Eq. (7.6) are equal if we set

$$f \approx f_\pi^*. \quad (7.7)$$

Then we have the scalar kaon–nucleon potential

$$S_{K^\pm} = \frac{1}{3} S_N, \quad (7.8)$$

where  $S_N$  is the NN scalar potential in Walecka mean field theory.

Given Walecka mean fields for nucleons, we can calculate the corresponding mean-field potential for  $K^-$ -nuclear interactions in symmetric nuclear matter. From the results obtained above, we have

$$S_{K^-} + V_{K^-} \approx \frac{1}{3}(S_N - V_N). \quad (7.9)$$

Phenomenology in Walecka mean-field theory gives  $(S_N - V_N) \leq -600$  MeV for  $\rho = \rho_0$  [71]. This leads to the prediction that at nuclear matter density

$$S_{K^-} + V_{K^-} \leq -200 \text{ MeV}. \quad (7.10)$$

This is consistent with the result of Friedman, Gal and Batty [18],

$$S_{K^-} + V_{K^-} = -200 \pm 20 \text{ MeV}. \quad (7.11)$$

In asymmetric nuclear matter, a simple way to extrapolate  $K^-$  energy into higher density ( $\rho_p \simeq \rho_n$ ) is

$$\epsilon_{K^-} = 494 \text{ MeV} - 200 \frac{\rho}{\rho_0} \text{ MeV}. \quad (7.12)$$

However neutron is the dominant element in neutron stars, and the  $K^-n$  interaction is smaller than the  $K^-p$  interaction. In our approach, for example, the leading order contribution – Weinberg–Tomozawa term (see Eq. (4.1).) – gives

$$T^{K^-n} = \frac{1}{2} T^{K^-p}. \quad (7.13)$$

Hence the net attraction is reduced, and the kaon effective mass in a neutron star becomes [28]

$$\epsilon_{K^-} = 494 \text{ MeV} - 171 \frac{\rho}{\rho_0} \text{ MeV}. \quad (7.14)$$

Considering the uncertainties of the simple approximation in Eq. (6.13), our result is consistent with that of mean field approach. In other words, the mean field approach is the key tool for describing kaon condensation. The resulting critical densities  $2\rho_0 < \rho_c < 4\rho_0$ , which are consistent with the results of original approach of kaon condensation [4], confirm the importance of mean field approach in kaon condensation.

## 7.2. Relativistic approach of Muto et al.

Following the model-independent approach based on current algebra and PCAC, Muto et al. discussed the interplay of scalar and vector interactions between kaons and nucleons in the context of kaon condensation [27,29].

The kaon-condensed state  $|K\rangle$  can be defined as

$$|K\rangle = U_K(\theta, \mu)|e\rangle, \quad U_K(\theta, \mu) = \exp(i\mu t\hat{Q}_{\text{em}}) \exp(i\theta F_4^5), \quad (7.15)$$

where  $\mu$  is the chemical potential needed for ensuring charge neutrality of the system,  $F_\alpha^5$  are the axial-vector charges which satisfy current algebra and  $\hat{Q}_{\text{em}}$  is the electromagnetic charge. The operator  $\exp(i\theta F_4^5)$  is essential for s-wave kaon condensation. Then the condensed field has the form

$$\langle K|K^\pm|K\rangle = (f/\sqrt{2}) \sin \theta \exp(\mp i\mu t). \quad (7.16)$$

The effective Hamiltonian in the kaon condensed state  $|K\rangle$  is

$$\mathcal{H}_{\text{eff}} = \mathcal{H}_{\text{chiral}} + \mu \hat{\rho}_{\text{em}}, \quad (7.17)$$

where  $\mathcal{H}_{\text{chiral}}$  denotes any chiral Hamiltonian with the structure  $\mathcal{H}_{\text{chiral}} = \mathcal{H}_{\text{symm}} + \mathcal{H}_{\text{SB}}$ .  $\mathcal{H}_{\text{symm}}$  is the chirally symmetric part and  $\mathcal{H}_{\text{SB}}$  the symmetry breaking one. Then the energy density in the condensed state is given by

$$\epsilon(\theta) \equiv \langle K|\mathcal{H}_{\text{eff}}|K\rangle = \langle e|\mathcal{H}_{\text{symm}}|e\rangle + \langle e|U_K^\dagger(\mathcal{H}_{\text{SB}} + \mu \hat{\rho}_{\text{em}})U_K|e\rangle, \quad (7.18)$$

By fully applying current algebra, they got

$$\begin{aligned} \Delta\epsilon(\theta) &\equiv \langle K|\mathcal{H}_{\text{eff}}|K\rangle - \langle e|\mathcal{H}_{\text{eff}}|e\rangle \\ &= \langle e|\mathcal{H}_{\text{SB}}(\theta)|e\rangle + \frac{1}{2}(\cos \theta - 1)\langle e|k_\alpha(V_3^\alpha + \sqrt{3}V_8^\alpha)|e\rangle - \frac{1}{2}k^2 f^2 \sin^2 \theta + \langle e|\mathcal{H}_1|e\rangle, \end{aligned} \quad (7.19)$$

for the energy difference with

$$\mathcal{H}_{\text{SB}}(\theta) \equiv U_K^\dagger \mathcal{H}_{\text{SB}} U_K = \mathcal{H}_{\text{SB}} - i \sin \theta [F_4^5, \mathcal{H}_{\text{SB}}] + (\cos \theta - 1) [F_4^5, [F_4^5, \mathcal{H}_{\text{SB}}]] \quad (7.20)$$

and  $\mathcal{H}_1 = k_\alpha \sin \theta A_\alpha^5$  being an “interaction” Hamiltonian which has no contribution in the lowest order. This simple form has essential features about low-energy kaon–hadron interactions, especially kaon–nucleon interaction; the scalar interaction is given by the  $KN$  sigma term,  $\Sigma_K = [F_4^5, [F_4^5, \mathcal{H}_{\text{SB}}]]$ , while the vector interaction comes from the Tomozawa–Weinberg term ( $\propto V_3^\alpha + \sqrt{3}V_8^\alpha$ ). Each vector current suggests the vector-meson exchange between kaon and nucleon;  $\rho$  meson and  $\omega$  and  $\phi$  mesons are responsible to  $V_3^\alpha$  and  $V_8^\alpha$ , respectively.

In order to reproduce the experimental data for low-energy  $KN$  scattering, it is not sufficient to treat Eq. (7.19) in the lowest order. Muto et al. proceeded to the second-order calculation with respect to  $\mathcal{H}_1$ . Using the pole-dominance hypothesis for the divergence of the axial current (PDDAC),  $\Delta\epsilon^{(2)}(\theta)$  can be written as

$$\Delta\epsilon^{(2)}(\theta) = -\frac{1}{2}i \int d^4x \langle e|T(k_\nu A_5^\nu(x) k_\lambda A_5^\lambda(0))|e\rangle \sin^2 \theta$$

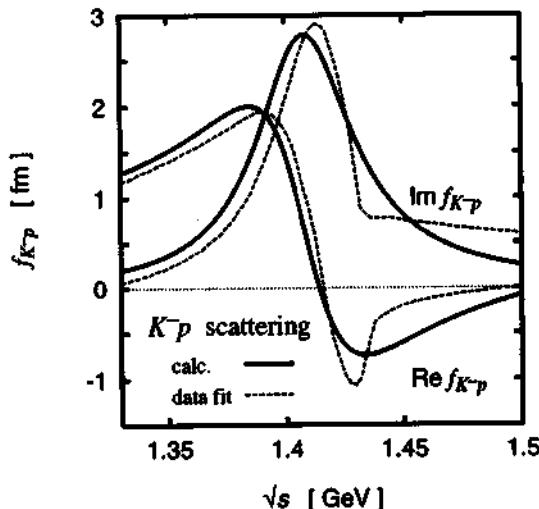


Fig. 7.1.  $K^- p$  scattering amplitude obtained in [27] (solid lines). Dashed lines are phenomenological fits of the scattering data in [72].

$$\begin{aligned}
 &= -\frac{1}{2}i \int d^4x \sum_{Y=A, \Sigma^0, A^*, \Sigma^{*0}} \sin^2 \theta \langle N | k_\nu k_\lambda T \{ A_3^\nu(x) Y \langle Y A_3^\lambda(0) \} | N \rangle \\
 &\quad - \frac{1}{2}i \sin^2 \theta k_\nu k_\lambda \alpha^{\nu\lambda} (\bar{N}N) + O(\rho_B^2).
 \end{aligned} \tag{7.21}$$

The last two terms in Eq. (7.21) represent the remaining smooth part and nuclear correlations of  $O(\rho_B^2)$ , respectively. Keeping only the  $s$ -wave resonance  $A^* = A(1405)$ ,  $\Delta\epsilon^{(2)}(\theta)$  is given by

$$\Delta\epsilon^{(2)}(\theta) = \Delta\epsilon_{\text{pole}}^{(2)}(\theta) + \Delta\epsilon_{\text{smooth}}^{(2)}(\theta) + (\text{correlations}) \tag{7.22}$$

with

$$\begin{aligned}
 \Delta\epsilon_{\text{pole}}^{(2)}(\theta) &\simeq -\frac{1}{4}\mu^2 \sin^2 \theta \frac{g_{A^*}^2}{\delta M - \mu} \rho_s(p), \\
 \Delta\epsilon_{\text{smooth}}^{(2)}(\theta) &\simeq -\frac{1}{2}\mu^2 f^2 \sin^2 \theta (d_p \rho_s(p) + d_n \rho_s(n)),
 \end{aligned} \tag{7.23}$$

where  $\rho_s(i)$ , ( $i = p, n$ ) denotes the scalar density of nucleon and  $\delta M = M_{A^*} - M \sim 465$  MeV the mass difference. Since the  $A^*$  lies just below the  $KN$  threshold, it plays an important role in the  $KN$  scattering physics, but as mentioned before, it has little effect on kaon condensation. They fitted the values of the coupling constant  $g_{A^*}$  and the range parameters  $d_{p,n}$  so as to fit the empirical data on the scattering lengths [72]:  $g_{A^*} = 0.7$ ,  $d_p = (0.37 - \Sigma_{KN}/m_K)/(m_K f^2)$  and  $d_n = (0.23 - \Sigma_{KN}/m_K)/(m_K f^2)$ . They allow some freedom in the value of  $\Sigma_{KN}$ . In Fig. 7.1 is plotted the  $K^- p$  scattering amplitude obtained by Muto et al. [27,29]. We can see that their result is consistent with our results Fig. 4.1.

For  $\theta \ll 1$  Eq. (7.19) can be expanded in  $\theta$ ,

$$\Delta\epsilon(\theta) = -D_K^{-1}(\mu, |\mathbf{k}| = 0; \rho_B) |K^-|^2 + O(\theta^4), \tag{7.24}$$

with the inverse propagator of kaons in a nuclear medium  $D_K^{-1}$ ,

$$D_K^{-1} = Z_\phi^{-1}(k_0^2 - \tilde{m}_K^{*2}) + \frac{\rho_B}{2f^2} k_0(1+x) + \frac{1}{2f^2} \frac{g_A^2 k_0^2}{\delta M - k_0} \cdot \rho_s(p), \quad (7.25)$$

where  $x$  is the proton fraction and

$$Z_\phi^{-1} = 1 + d_p \rho_s(p) + d_n \rho_s(n) \quad (7.26)$$

is the wave function renormalization factor. The effective mass of kaons can be defined as

$$\tilde{m}_K^{*2} = Z_\phi m_K^{*2}, \quad m_K^{*2} = m_K^2 - f^{-2} [\Sigma_{Kp} \rho_s(p) + \Sigma_{Kn} \rho_s(n)]. \quad (7.27)$$

Then the kaon mass is shifted by the scalar interaction and renormalized by the range effect. Then the excitation energy  $\omega(\rho_B) \equiv \omega_{k=0}$  in medium can be obtained by solving the equation;  $D_K^{-1} = 0$ . In order to incorporate the relativistic effects, they distinguished the scalar density,

$$\rho_s = 2 \int \frac{d^3 p}{(2\pi)^3} \frac{M}{E}, \quad (7.28)$$

from the baryon density  $\rho_B$  [29]. Furthermore they employed the relativistic mean-field theory ( $\sigma$ - $\omega$  model); then nucleon effective mass is given by  $M^* = M - g_\sigma \sigma$ . In Fig. 7.2, the kaon effective mass  $\omega(\rho_B)$  in symmetric matter is given as a function of density. Solid lines are results with the second-order effect and dashed lines those without it. Comparing them, one sees first that the  $K^-$  branch is quantitatively unmodified, decreasing with increasing density. On the other hand, by the second-order term, the  $K^+$  branch is strongly repelled from the dashed line reflecting the  $K^+N$  repulsion known experimentally. Comparing with our  $K^\pm$  effective in Fig. 5.3, the  $K^-$  branch is consistent in two approaches even though the repulsion of the  $K^+$  branch in our approach is rather small.

In Fig. 7.3 the effective masses and excitation energy of kaons in neutron star matter are given. We can see that the effective mass is not so reduced by the scalar interaction, whereas the excitation energy is considerably decreased by the vector interaction (the Tomozawa–Weinberg term). Even in the extreme case  $\Sigma_{KN} = 0$ , where the effective mass is almost unchanged, the kaon excitation energy is much reduced. They concluded that vector interaction due to the Tomozawa–Weinberg term is more important than the scalar interaction through the  $KN$  sigma term, and it brings about kaon condensation even in the absence of the sigma term. Their result is consistent with ours with the BR scaling as described in Subsection 6.3.1,

### 7.3. Kaons on the hypersphere

Recently, using the idea of Manton [73] for simulating density effects, Westerberg [30] explored S-wave kaon condensation in the bound-state approach to the Skyrme model on a 3-sphere.

A single hyperon may be modeled by the Skyrme Lagrangian in a flat space,

$$\mathcal{L} = \frac{f_\pi^2}{16} \text{Tr} (\partial_\mu U^\dagger \partial^\mu U) + \frac{1}{32e^2} \text{Tr} [\partial_\mu U U^\dagger, \partial_\nu U U^\dagger]^2 + \frac{f_\pi^2}{8} \text{Tr} \mathcal{M} (U + U^\dagger - 2), \quad (7.29)$$

where  $U(x, t) \in \text{SU}(3)$ .  $\mathcal{M}$  is proportional to the quark mass matrix given by (neglecting the  $u$  and  $d$  masses)

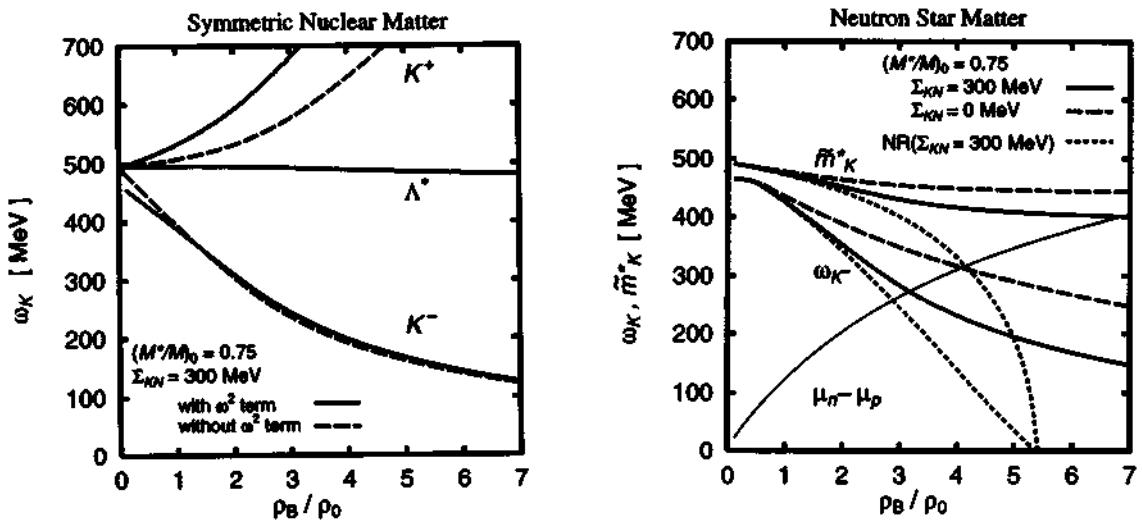


Fig. 7.2. Kaonic excitation mode in symmetric nuclear matter [27].  $(M^*/M)_0$  denotes the effective nucleon mass at the nuclear density  $\rho_0 \approx 0.17 \text{ fm}^{-3}$ .

Fig. 7.3.  $K^-$  excitation energy and effective kaon mass in neutron star matter[27].

$$\mathcal{M} = \left( \begin{array}{c|c} 0 & 0 \\ \hline 0^\dagger & m_K^2 \end{array} \right) \quad (7.30)$$

The standard fit to the nucleon and delta masses yields  $f_\pi = 129$  MeV and  $e = 5.45$ . The empirical value of the kaon mass is  $m_K = 495$  MeV. In addition to the local terms, Westerberg included the Wess-Zumino term,

$$S_{WZ} = -\frac{iN}{240\pi^2} \int_D d^5x \epsilon^{\mu\nu\alpha\beta\gamma} \text{Tr} (M_\mu M_\nu M_\alpha M_\beta M_\gamma), \quad (7.31)$$

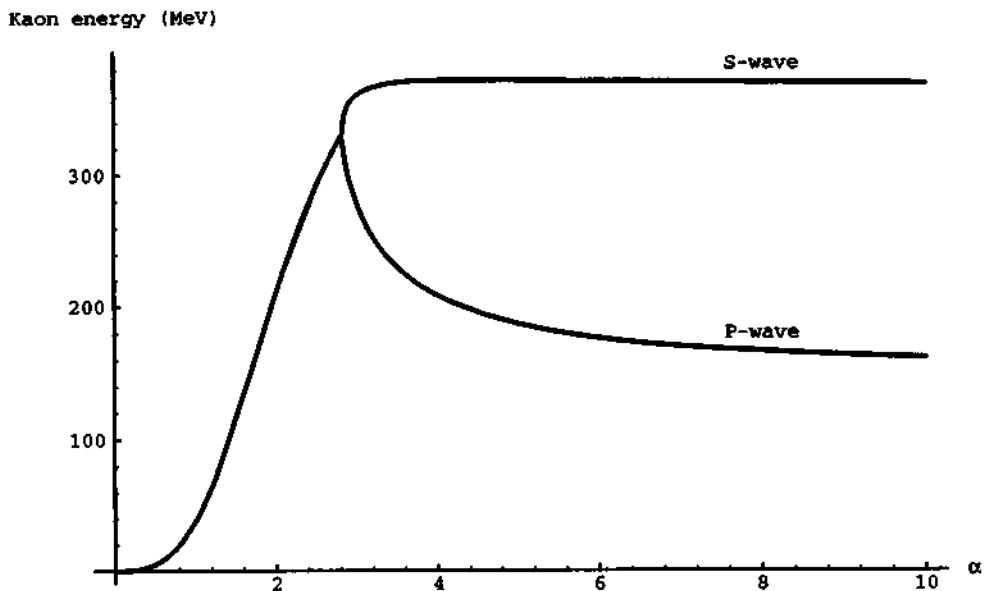
where  $M_\mu = \partial_\mu UU^\dagger$  and  $N$  is the number of colors. In the bound-state approach to the Skyrme model, the following ansatz is used:

$$\begin{aligned} U(x, t) &= \sqrt{U_\pi(x, t)} U_K(x, t) \sqrt{U_\pi(x, t)}, \\ U_K(x, t) &= \exp i \frac{2\sqrt{2}}{f_\pi} \left( \begin{array}{c|c} 0 & K(x, t) \\ \hline K(x, t)^\dagger & 0 \end{array} \right) \end{aligned} \quad (7.32)$$

where  $K$  is a complex spinor and  $U_\pi$  is realized by the hedgehog anzatz

$$U_\pi(x, t) = \left( \begin{array}{c|c} e^{iF(r)\hat{r}\cdot\hat{r}} & 0 \\ \hline 0^\dagger & 1 \end{array} \right) \quad (7.33)$$

with the boundary conditions  $F(0) = \pi$  and  $F(\infty) = 0$ . The eigenmodes for  $K(x, t)$  are written in terms of spin and isospin. The corresponding anzätze for S-wave and P-wave kaons are

Fig. 7.4. Kaon energy (MeV) vs  $\alpha$  [30].

$$K_S(\mathbf{x}, t) = k_S(r)\chi(t), \quad (7.34)$$

$$K_P(\mathbf{x}, t) = ik_P(r)\boldsymbol{\tau} \cdot \hat{\mathbf{r}}\chi(t). \quad (7.35)$$

In each case,  $\chi(t)$  are the dynamical variables arranged in a complex spinor. The field  $F(r)$  is determined classically using the non-kaon part of the Lagrangian. This baryon field is treated as a background field with which the kaons interact. The energy eigenstates form a Fock space for each kaon mode, where the energy required to create a kaon in a particular mode is given by its classical eigenfrequency (i.e. substitute  $\dot{\chi} = -i\omega\chi$  and minimize the action with respect to  $\omega$  and  $k(r)$ ). In flat space, the lowest energy modes for P-wave and S-wave are given respectively by  $\omega_P = 153$  MeV and  $\omega_S = 368$  MeV.

In [73], Manton suggested that a baryon crystal may be approximated by putting a single Skyrmion on a hypersphere of finite radius. This has the effect of allowing the baryon to interact with itself. Westerberg [30] continued with this idea by considering the bound state between a kaon and a baryon on the hypersphere. The spatial metric in a hypersphere of radius  $a$  is

$$ds^2 = a^2(d\rho^2 + \sin^2 \rho(d\theta^2 + \sin^2 \theta d\phi^2)), \quad (7.36)$$

where the possible ranges of three angular coordinates are

$$0 \leq \rho, \quad 0 \leq \theta \leq \pi, \quad 0 \leq \phi \leq 2\pi. \quad (7.37)$$

The baryon number density is given by the inverse volume of the hypersphere

$$\rho = \frac{1}{2\pi^2 a^3} \equiv \frac{e^3 F_\pi^3}{2\pi^2 a^3}, \quad (7.38)$$

where  $\alpha = aeF_\pi$  with  $F_\pi = 129$  MeV and  $e = 5.45$ . The kaon energy is shown in Fig. 7.4 which also shows that the chiral phase transition occurs at  $\alpha = \alpha_c = 2\sqrt{2}$ . For  $\alpha > \alpha_c$ , the P-wave and S-wave kaons have different masses, the difference in mass representing roughly the mass difference between  $A(1405)$  and  $A(1116)$ . For  $\alpha < \alpha_c$ , the S-wave and P-wave kaons become degenerate, with the kaon condensation occurring in the regime  $1 < \alpha < 2\sqrt{2}$ .

Solving the equation of state for the electron chemical potential, Westerberg [30] found the critical density to be at  $\alpha = 1.58$  corresponding to  $\rho \approx 3.7\rho_0$ , Eq. (7.38). This falls within our predicted range  $(2 \sim 4)\rho_0$ . However an unsatisfactory aspect of this result is that the kaon condensation sets in *after* – and *not before* as one expects – the chiral phase transition.

## 8. Discussion

In conclusion, we have shown that chiral perturbation theory at order  $N^2LO$  predicts kaon condensation in “nuclear star” matter at a density  $2 \lesssim u_c \lesssim 4$  with a large fraction of protons –  $x = 0.1 \sim 0.2$  at the critical point and rapidly increasing afterwards – neutralizing the negative charge of the condensed kaons. For this to occur, four-Fermi interactions involving  $A^*$  are found to play an important role in driving the condensation but the critical density is negligibly dependent on the strength of the four-Fermi interaction.

It is found that the BR scaling [25,26] favors a condensation at a density as low as twice the matter density and that when the BR scaling is operative, higher chiral corrections coming from the scaled Lagrangian are insignificant, justifying the basic assumption that goes into the derivation of the scaling relation. This suggest that at least for kaon condensation, the tree approximation with the scaled Lagrangian is consistent with the basic idea of the BR scaling.

Given the relatively low critical density obtained in the higher-order calculation of this paper, we consider it reasonable to assume that the compact-star properties will be qualitatively the same as in the tree-order calculation of Ref. [8].

Our treatment is still far from self-consistent as there are many nuclear correlation effects that are still to be taken into account. How to incorporate them in full consistency with chiral symmetry is not known. In particular the role of four-Fermi interactions in the non-strange channel, e.g., short-range nuclear correlations which involve both interaction terms in the Lagrangian and nuclear many-body effects, is still poorly understood. But Brown and Rho [28] are on the way to resolving this problem.

## Acknowledgements

My deepest gratitude goes to Professors Dong-Pil Min and Mannque Rho for their guidance and encouragement throughout my graduate years. Collaboration with Professor G.E. Brown was a great pleasure to me. This work was supported in part by the Korea Science and Engineering Foundation through the CTP of SNU and in part by the Korea Ministry of Education under Grant No. BSRI-95-2418.

## Appendix A. Neutron star in the core of supernova

In Section 2, we saw that the kaon condensation is a good candidate to explain the low maximum neutron star mass. Generally, neutron stars are believed to be formed in the cores of supernovae, like SN1987A. In this appendix, we review the formation of neutron star in the core of supernova, and the maximum neutron star mass.

### A.1. Neutron star story

*Discovery of neutron star.* Since the discovery of pulsar CP1919 (Cambridge Pulsar), a regularly pulsating radio source with period 1.337 sec at 81.5 MHz, by J. Bell and A. Hewish 1967, hundred of pulsars have been discovered by radio astronomers. Within a few months of the discovery of pulsars, it was believed that the pulsars were rapidly rotating neutron stars.

There have been three scenarios to explain such a regular radio sources. One of them is the eclipsing binary system, but it cannot explain the short period of pulses. Another scenario is a vibrating white dwarf, whose period of vibration can be slightly over a second ( $\geq 1\text{sec}$ ), just that of CP1919. But the subsequent discovery of pulsars that pulsed more rapidly than a second,  $1.6\text{ms} \sim 4.3\text{s}$ , made the vibrating white dwarf hypothesis untenable. The key observational facts are

- Pulsars have periods in the range 1.58 ms to 4.3 s. Some pulsar periods are as follows.
  - PSR 1937+214 : 1.558 ms
  - PSR 0531+21 (Crab) : 0.0331 s
  - PSR 1931+16 : 0.059 s
  - PSR 0833-45 (Vela) : 0.089 s
  - PSR 1845-19 : 4.308 s
- Pulsar periods increase very slowly as time goes on, and never decrease.
- Pulsars are remarkably good clocks. Some pulsar periods have been measured up to 13 significant digits.

The other possibility was a rapidly spinning star. If this were the cause of pulsars, the spinning object could be nothing but a neutron star, which would explains well the rapidity of their pulses and their impressive regularity.

*Which stars are neutron stars?* After the discovery of pulsars, the positions of pulsars were tested. In our galaxy, a rapid pulsar was found in the Vela supernova remnant and later in the heat of supernova in the Crab. Progress toward understanding of the relation between pulsars, neutron stars, and supernova continued.

If pulsars are spinning neutron stars, and if they are formed by supernova explosions, we might expect to find many pulsars located near supernova remnants. In nature, less than a dozen of the hundreds of known pulsars are in or near supernova remnants. Furthermore, although the pulsars are formed in a supernova explosion, the remnants should only be visible for a few tens of thousands of years. The pulsar, on the other hand, may be detectable for millions of years. The majority of the pulsars we see may simply have outlived the supernova remnants that surround them.

The absence of a pulsar in at least some supernova remnants is also understood because only about half the supernova, the Type II events, produce neutron stars [74]. There is no trace of neutron star at the center of Type I supernova.

**Table A.1**Distinguishing traits of compact objects.  $M_{\odot} = 1.989 \times 10^{33}$  g,  $R_{\odot} = 6.9599 \times 10^{10}$  cm.

Object	Mass ( $M$ )	Radius ( $R$ )	Mean Density ( $\text{g cm}^{-3}$ )
Sun	$M_{\odot}$	$R_{\odot}$	1
White dwarf	$\leq M_{\odot}$	$\sim 10^{-2} R_{\odot}$	$\leq 10^7$
Neutron star	$\sim 1 - 3 M_{\odot}$	$\sim 10^{-5} R_{\odot}$	$\leq 10^{15}$
Black hole	Arbitrary	$2GM/c^2$	$\sim M/R^3$

The most striking difference between Type I and Type II supernova is the presence of broad emission lines from hydrogen gas. Another striking difference is that they occurred in different places. Type I supernova, almost without exception, appeared in the elliptical galaxies, composed of ancient stars and formed early in the history of the universe. Type II supernova, on the other hand, appeared primarily in the arms of young spiral galaxies, among groups of bright, massive stars, glowing billows of hydrogen gas, and dark concentrations of dust.

*How supernovae give birth to neutron stars?* In the early 1930's, Chandrasekhar set a limit to the size of white dwarf. No carbon-rich white dwarf could support its weight if it were greater than about 1.4 times solar mass  $M_{\odot}$ ,

$$M_{CH} \approx 5.76 Y_e^2 M_{\odot}, \quad (\text{A.1})$$

with  $Y_e$  is electron fraction per baryon. In stellar collapse,  $Y_e$  is  $0.43 \sim 0.50$ . Stars more massive than  $M_{CH}$ , not the initial mass but its final mass when it exhausts its last bit of fuel, collapse, and that collapse turns into an explosion of epic proportions.

Let us follow the death of a star of  $16 M_{\odot}$ . As it passes into the final stages of its productive life, with heavier and heavier elements burning in its core and surrounding shells, the pace of change picks up to a frenzy. The last stages of energy generation, ending in the development of a core of iron, take only a matter of days, whereas its youth as a main sequence star, with hydrogen burning at the center, lasts millions of years. The star consists of a bloated ball of gas with a dense iron core slightly smaller than earth, but containing a little more mass than the sun, resting at the very center. Around this lies a thin layer of silicon, and around that shells of lighter elements as in Fig. A.1. The entire star may be somewhat between 50 and several hundreds times the diameter of the sun. No more energy can be produced by the iron core, yet it must support the immense weight that overlies it. The core has shrunk to a white dwarf, rigid and unyielding embedded in the still sputtering remnants of the star. But as the surrounding layer of silicon continues to burn, showering nuclei of iron onto the core, the mass of the core edges towards the Chandrasekhar limit. Collapse begins within minutes.

As iron core begins to fall inward under the pull of gravity, two things occur to make the collapse even more catastrophic. First, the temperature rises. Ordinarily this would raise the gas pressure and initiate a new round of nuclear burning. But the iron doesn't burn. Instead it begins to break apart by photo-disintegration,



The breakup drains energy from the star, cooling the material. But removing heat from the core of the star only lowers the pressure of the gas, hastening the collapse.

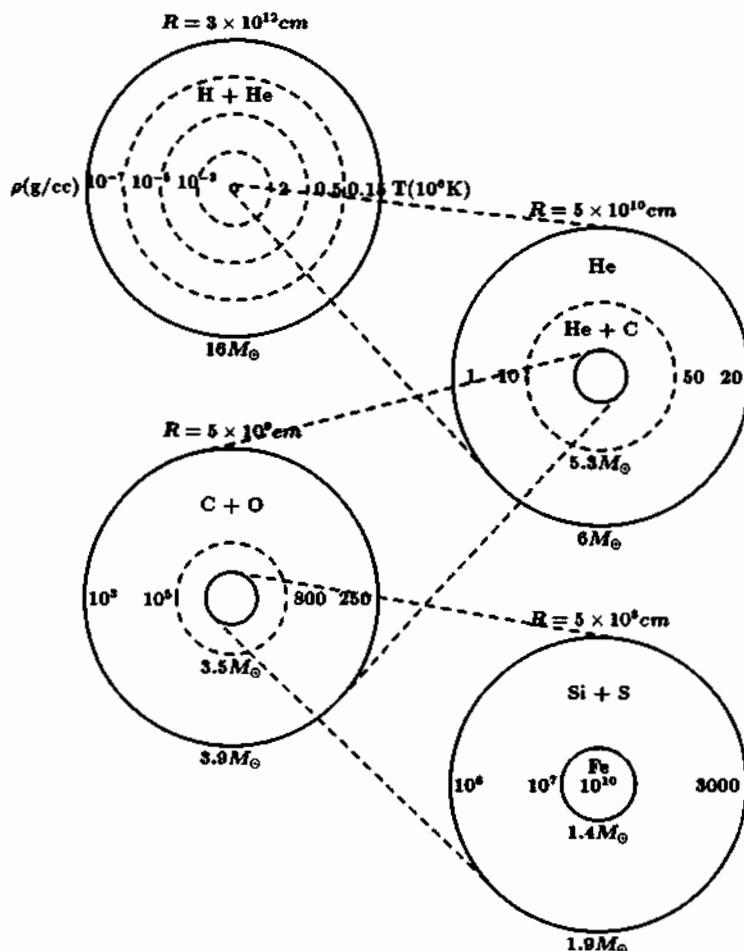


Fig. A.1. The interior of a star  $16 M_\odot$  just prior to a supernova: The radial dimension, given at the top of each diagram, is to scale and the central region of the star is enlarged from drawing to drawing by factors of  $\times 60$ ,  $\times 10$  and  $\times 10$ , respectively. The density ( $\text{g/cm}^3$ ) in the various zones of the star is labelled along the radius to the left of each diagram, the temperature (millions of kelvin) is labelled to the right, and the mass of each zone (solar masses) is labelled to the bottom. The iron core in the center of the last drawing is  $1/(8 \times 10^{15})$  of the volume of the blue supergiant star, although it contains about  $1/10$  of its mass [75,76].

One other event speeds up the catastrophe. As the star becomes more and more compressed, electrons are pushed so close to the protons in the nuclei of atoms that they begin to combine with the protons to become neutrons, releasing neutrinos in the process. The collapsing core rapidly changes from a dense ball of iron into an even denser glob of neutrinos, while the released neutrinos carry energy directly away from the center of the star with high efficiency. This accelerates the collapse even more. Once formed in the interior of the collapsing star, an initial burst of neutrinos travels out into space without difficulty.

But neutrinos carry energy, just as light does. Streaming out of the core of the collapsing star,

the neutrinos drain enormous amounts of energy almost instantaneously, in a burst that lasts only a fraction of a second. This rapidly lowers the pressure in the core. In less than a second, the burnt-out iron core tumbles inward, pell-mell toward its own center, where it is compressed to extraordinary densities. All this takes place so fast that the outer layers of the star don't even have time to respond. To distant observer, the star would seem unruffled for several hours at least, despite the violent events occurring deep inside. Unless one could see the initial burst of neutrinos coming out, the first catastrophic minutes of stellar distress would be completely hidden from view.

In most cases, the collapse turns into an equally rapid explosion. What causes this, ultimately, is the formation of a neutron star at the center of the star. When the density of the collapsing core reaches the density of an atomic nucleus, the material becomes exceedingly rigid once again. Neutron degeneracy, a situation similar to the electron degeneracy that supports white dwarfs, resists any further compression. The entire core has shrunk to a tiny ball.

All this happens so fast, however, that material continues to fall inward toward the neutron-rich ball at the center. When it reaches the rigid surface, it bounces. Almost immediately, layers of stellar material that were falling inward find themselves moving outward again, at speeds of about 10% the speed of light. This speed is far beyond the speed of sound in the gas of the star. Such a supersonic wave of material, shock wave, plows up an expanding sphere of compressed and heated material ahead of it.

What happens for the objects even more dense than neutron stars? This state of affairs should result if a star like the sun were compressed just a lot more than a neutron star. Pack the sun ( $M_{\odot} = 1.99 \times 10^{30}$  kg) into a ball whose radius is less than the Schwarzschild radius

$$R_{\text{Schwarz}} = \frac{2GM_{\odot}}{c^2} \approx 2.95 \text{ km}, \quad (\text{A.3})$$

it will shrink to black hole. A black hole is a body so compressed that its escape velocity exceeds the speed of light. Black holes can only be observed indirectly through the influence they exert on their environment. At least one good black candidate, Cygnus X-1, has been identified. In case of Cygnus X-1, the properties of the binary orbit – whose variability time scale  $\Delta t$  is  $\sim 1\text{--}10$  ms, and emission zone size is  $R \leq c\Delta t \leq 10^8$  cm – yield a rather firm lower mass limit for unseen companion of  $\sim 9M_{\odot}$  ( $9M_{\odot}\text{--}15M_{\odot}$ ). This is far above the upper mass limit for neutron stars and white dwarfs.

## A.2. Maximum neutron star mass

*Stellar death function.* That main sequence stars of mass  $\geq 25 - 30M_{\odot}$  must end up in black holes without producing nucleosynthesis, i.e., without returning matter to the galaxy, is required by the observed abundances of elements [77,78]. Maeder's argument is based on the measurement of  $\Delta Y/\Delta Z$ , the ratio of helium abundance to that of metals, in low-metallicity extragalactic  $H_{II}$  regions, especially irregular dwarf galaxies. The ratio can be measured with good accuracy [77],

$$\frac{\Delta Y}{\Delta Z} = 4 \pm 1.3. \quad (\text{A.4})$$

If all stable stars of mass up to  $\sim 100M_{\odot}$  were to explode, returning matter to the galaxy, this ratio would lie between 1 and 2. Helium is produced chiefly by relatively light stars, metals by heavy

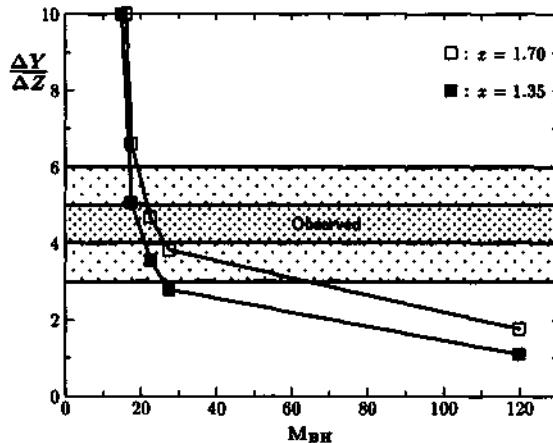


Fig. A.2. Values of the ratio  $\Delta Y/\Delta Z$  of the relative helium to metal enrichments as a function of  $M_{BH}$ . The data point correspond to initial metallicity  $Z=0.001$ . The observed range is indicated by shadings, the range 4 to 5 is the preferred one [78].

stars, so that cutting off the production by the heavy stars going directly into black holes without nucleosynthesis increases the  $\Delta Y/\Delta Z$ . Using the standard initial mass function for stars,

$$dN/dM = M^{-(1+x)}, \quad (\text{A.5})$$

with  $x = 1.35, 1.70$ , Maeder [78] found, that Pagel's [79] measurement on  $\Delta Y/\Delta Z$  was best reproduced by a cut off of nucleosynthesis at a main sequence stellar mass of  $\sim 22.5 M_\odot$  as shown in Fig. A.2.

There is considerable uncertainty in the initial mass function, as noted by Maeder [78], so that this limit could easily be  $\sim 30 M_\odot$  or even higher. Brown and Bethe estimated [3]  $30 M_\odot$  as the cut off for stars to drop directly into black holes without nucleosynthesis.

What about stars with mass  $20 M_\odot < M < 30 M_\odot$ ? Recent observation on SN1987A, whose progenitor mass is  $\sim 18 \pm 2 M_\odot$ , gives us an insight. Based on the empirical analysis, Brown and Bethe also argued that a large range of stars below this mass, down to  $\sim 18 M_\odot$ , can first accomplish nucleosynthesis and then collapse into black holes.

*SN1987A: neutron star or black hole?* SN1987A (February 23, 1987) in the Large Magellanic Cloud is the nearest and brightest supernova to be observed since SN1604AD (Kepler), and certainly the most important supernova since SN1054AD, the progenitor of Crab Nebula. Because of its brightness and proximity, it will be possible to observe SN1987A for many years as it expands to reveal its inner secrets. In contrast, typical supernovae, of which some 20–30 are observed each year, are some 1000 times further and  $10^5$  times fainter, and so become lost in their host galaxy within a year or two. Moreover, SN1987A has been observed at every wavelength band of the electromagnetic spectrum, from radio to  $\gamma$ -rays.

In supernova theory, the neutron star is followed by outburst of huge amount of neutrinos. In the case of SN1987A, neutrinos were detected by IMB and Kamiokande detector about three hours before the optical burst. The total energy emitted in  $\bar{\nu}_e$ 's is  $E_\nu \approx 3 \times 10^{52} \text{ erg/s}$ , while the decay time scale of burst is  $\Delta t \approx 10 \text{ sec}$ . These are quite consistent with the expected values. Hence, firstly, it is

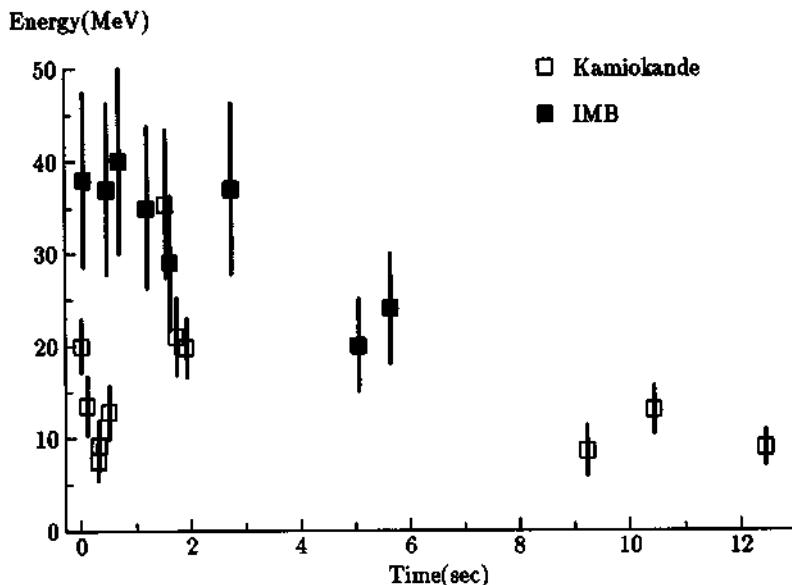
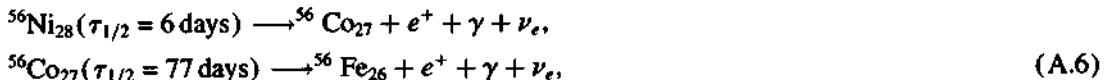


Fig. A.3. The energy of the neutrinos detected by the Kamiokande and IMB detectors is plotted as a single neutrino pulse. Most neutrinos arrived within the first second or two but there was a lower energy tail to the pulse which lasted more than 10 s.

believed that the neutron star was formed in SN1987A, even though there exist a time gap of about 7 seconds between eighth and ninth Kamiokande events, the gap being followed by another three events as shown in Fig. A.3 [76,80].

From the observation of radioactive decay of  $^{56}\text{Ni}$  and  $^{56}\text{Co}$



the mass of Ni ejecta of SN1987A is known to be  $0.075 M_\odot$ . Combining this result with the observed energy of SN1987A,  $E = 1.4 \pm 0.4 \text{ foe}$  (where  $\text{foe}$  stands for  $10^{51} \text{ erg}$ ), Brown and Bethe [31] obtain the range of the core mass of SN1987A,

$$M = 1.535 \pm 0.02 M_\odot. \quad (\text{A.7})$$

However, from the mass of the Hulse-Tayler pulsar, the lower limit of the compact core is known to be at least  $1.44 M_\odot$ . Consequently, the compact core mass of SN1987A must range

$$1.44 M_\odot < M_{\text{core}} < 1.56 M_\odot. \quad (\text{A.8})$$

Astronomers have been searching for a pulsar in the center of SN1987A remnants after the explosion. To see the explicit signal of a pulsar, one must wait until the remnants are transparent. If the pulsar is really formed in SN1987A, the X-ray signal should be detected within a few years after the explosion. The hypercritical accretion could hide the compact object for only  $\sim 1$  year, but after this time 1987A should be observed with a luminosity  $L = 4 \times 10^{38} \text{ ergs/s}$ , but the present light curve

is lower by two orders,  $L \sim 4 \times 10^{36} \text{ ergs/s}$  [81]. This compact object, after being a proto-neutron star for at least 10 seconds, appears to have collapsed into a black hole, low-mass black hole. The core of SN1987A may become a neutron star followed by neutrino emission, and later changed into low-mass blackhole with mass about  $1.5M_{\odot}$ .

A possible scenario was proposed by Bethe [82]. According to his arguments, a vigorous convection is produced in the supernova shock. But after 2 seconds, convection stops as heat is no longer supplied by neutrinos. As a result, a substantial fraction of the previously convecting material falls into the neutron star at the center. Bethe estimated this fraction to be about 10% of the mass in the shock wave, or about  $0.04M_{\odot}$ . If the neutron star were close to  $M_{\max}$ , the added  $0.04M_{\odot}$  could push the neutron star over the limit, and make it collapsing into a black hole. If this happened in SN1987A, it must have been more than 12 seconds after the first collapse, because at 12 sec, a neutrino was still observed at Kamiokande II. A softened equation of state might be associated with a delayed collapse of the young neutron star into a black hole.

From standard evolutionary analysis, it is believed that the neutron star is formed if the core mass lies between the maximum neutron star mass( $M_{\max}$ ) and the Chandrasekhar mass,

$$M_{CH} = 5.76Y_e^2 M_{\odot}, \quad (\text{A.9})$$

where  $Y_e$  is the electron fraction per baryon. In stellar collapse,  $Y_e$  is  $0.43 \sim 0.50$ . Because of the thermal pressure, the evolutionary lower bound of the neutron star is about  $(1.10 \sim 1.15) \times M_{CH} \approx (1.2 \sim 1.4) M_{\odot}$ .

According to Brown and Bethe [3], in the binary pulsar evolution, the accretion can proceed at the hypercritical rate

$$\dot{M} \geq 10^4 \dot{M}_{Edd}, \quad (\text{A.10})$$

where the Eddington limit is

$$\dot{M}_{Edd} = 1.5 \times 10^{-8} M_{\odot} \text{yr}^{-1}. \quad (\text{A.11})$$

Hence, if the neutron star mass were determined by the evolutionary scenario, massive neutron stars with masses exceeding  $1.5M_{\odot}$  should exist. But as shown below, none have been found.

*Observed supernova remnants* From radio, optical, x-ray and  $\gamma$ -ray surveys of supernova remnants, Helfand and Becker [74] came to the conclusion that nearly half of the supernova in the galaxy leave no observable remnant. This is understandable because half of the supernova are Type I, which leave no neutron stars. However, as discussed in detail by Van den Bergh et al. [83], the observed samples of supernova in our galaxy are at low galactic latitude, so their light is strongly absorbed on its way to Earth. They argued that Type II supernova outnumber Type I supernova by a factor of several. Thus the conclusions of Helfand and Becker [74] can be understood only if nearly half of the Type II supernova explosions do not form neutron stars. The possible candidate is, therefore, low mass black hole. This fact is consistent with the low masses of well measured neutron stars given in Fig. 2.1.

## Appendix B. Vertices and Feynman rules

In this appendix, the chiral Lagrangian used is given explicitly in terms of the component fields (meson octet, baryon octet and decuplet) we are interested in. The vertices of the Feynman graphs can be read off directly. The subscript  $n$  in  $\mathcal{L}_n$  denotes the number of lines attached to the given vertex. In Fig. 3.1, the three-point vertex  $\mathcal{L}_3$  enters in the diagrams (c,d,f), the four-point vertices  $\mathcal{L}_4^A$  (4-meson) and  $\mathcal{L}_4^B$  (2-meson + 2-baryon) in the diagrams (b,d) and (b,c,e), respectively and the higher-point vertices  $\mathcal{L}_5$  and  $\mathcal{L}_6$  figure in the diagrams (f) and (a), respectively.

$$\begin{aligned} \mathcal{L}_3 = & -\frac{1}{f} \left[ \bar{p}_v S_v^\mu p_v \left( (D+F)\partial_\mu \pi^0 - \frac{1}{\sqrt{3}}(D-3F)\partial_\mu \eta \right) + \bar{n}_v S_v^\mu n_v \left( -(D+F)\partial_\mu \pi^0 \right. \right. \\ & \left. \left. - \frac{1}{\sqrt{3}}(D-3F)\partial_\mu \eta \right) \right] - \frac{1}{f} \left[ \sqrt{2}(D+F)\bar{p}_v S_v^\mu n_v \partial_\mu \pi^+ \right. \\ & + \sqrt{2}(D-F)\bar{p}_v S_v^\mu \Sigma_v^+ \partial_\mu K^0 + \sqrt{2}(D-F)\bar{n}_v S_v^\mu \Sigma_v^- \partial_\mu K^+ - \frac{1}{\sqrt{3}}(D+3F)\bar{p}_v S_v^\mu \Lambda_v \partial_\mu K^+ \\ & - \frac{1}{\sqrt{3}}(D+3F)\bar{n}_v S_v^\mu \Lambda_v \partial_\mu K^0 + (D-F)\bar{p}_v S_v^\mu \Sigma_v^0 \partial_\mu K^+ - (D-F)\bar{n}_v S_v^\mu \Sigma_v^0 \partial_\mu K^0 + \text{h.c.} \left. \right] \\ & - \frac{C}{\sqrt{6}f} \left[ \sqrt{2}\bar{p}_v \partial^\mu \pi^0 \Delta_{v,\mu}^+ + \bar{p}_v \partial^\mu \pi^+ \Delta_{v,\mu}^0 + \frac{1}{\sqrt{2}}\bar{p}_v \partial^\mu K^+ \Sigma_{v,\mu}^{*0} \right. \\ & - \sqrt{3}\bar{p}_v \partial^\mu \pi^- \Delta_{v,\mu}^{++} - \bar{p}_v \partial^\mu K^0 \Sigma_{v,\mu}^{*+} + \sqrt{2}\bar{n}_v \partial^\mu \pi^0 \Delta_{v,\mu}^0 - \bar{n}_v \partial^\mu \pi^- \Delta_{v,\mu}^+ - \frac{1}{\sqrt{2}}\bar{n}_v \partial^\mu K^0 \Sigma_{v,\mu}^{*0} \\ & \left. + \sqrt{3}\bar{n}_v \partial^\mu \pi^+ \Delta_{v,\mu}^- + \bar{n}_v \partial^\mu K^+ \Sigma_{v,\mu}^{*-} + \text{h.c.} \right], \end{aligned} \quad (\text{B.1})$$

$$\begin{aligned} \mathcal{L}_4^A = & -\frac{1}{6f^2} \left[ \frac{1}{2}((\pi^0 + \sqrt{3}\eta) \vec{\partial}_\mu K^+) ((\pi^0 + \sqrt{3}\eta) \vec{\partial}_\mu K^-) + \left( -(K^+ \vec{\partial}_\mu K^-)(\pi^+ \vec{\partial}_\mu \pi^-) \right. \right. \\ & + (K^+ \vec{\partial}_\mu \pi^-)(K^- \vec{\partial}_\mu \pi^+) \left. \right) - (K^+ \vec{\partial}_\mu K^-)(K^+ \vec{\partial}_\mu K^-) + \left( (K^+ \vec{\partial}_\mu K^-)(\bar{K}^0 \vec{\partial}_\mu K^0) \right. \\ & \left. + (K^+ \vec{\partial}_\mu \bar{K}^0)(K^- \vec{\partial}_\mu K^0) \right) \right] + \frac{B_0}{6f^2} \left[ K^+ K^- \left( \frac{3m_u + m_s}{2}(\pi^0)^2 + \frac{m_u - m_s}{\sqrt{3}}\pi^0 \eta \right. \right. \\ & \left. \left. + \frac{m_u + 3m_s}{2}\eta^2 \right) + (2m_u + m_d + m_s)\pi^+ \pi^- K^+ K^- \right. \\ & \left. + (m_u + m_s)K^+ K^- K^+ K^- + (m_u + m_d + 2m_s)K^+ K^- \bar{K}^0 K^0 \right], \end{aligned} \quad (\text{B.2})$$

$$\begin{aligned} \mathcal{L}_4^B = & -\frac{i}{4f^2} \left[ (K^+ \vec{\partial}_\mu K^-) (2\bar{\Xi}_v^- v^\mu \Xi_v^- - \bar{\Sigma}_v^+ v^\mu \Sigma_v^+ + \bar{\Sigma}_v^- v^\mu \Sigma_v^- + \bar{\Xi}_v^0 v^\mu \Xi_v^0) \right. \\ & - (K^+ \vec{\partial}_\mu K^-) (2\bar{p}_v v^\mu p_v + \bar{n}_v v^\mu n_v) - \bar{p}_v v^\mu p_v \{ (\pi^+ \vec{\partial}_\mu \pi^-) - (\bar{K}^0 \vec{\partial}_\mu K^0) \} \\ & \left. + \bar{n}_v v^\mu n_v \{ (\pi^+ \vec{\partial}_\mu \pi^-) - 2(K^0 \vec{\partial}_\mu \bar{K}^0) \} \right] + \frac{i}{4f^2} \left[ \bar{p}_v v^\mu n_v (K^+ \vec{\partial}_\mu \bar{K}^0) \right. \end{aligned}$$

$$\begin{aligned}
& -\bar{\Sigma}_v^+ v^\mu p_v (K^- \vec{\partial}_\mu \pi^+) - \frac{1}{\sqrt{2}} \bar{n}_v v^\mu \Sigma_v^- ((\pi^0 + \sqrt{3}\eta) \vec{\partial}_\mu K^+) + \frac{1}{\sqrt{2}} \bar{n}_v v^\mu (\Sigma_v^0 - \sqrt{3}A_v) \\
& \times (\pi^- \vec{\partial}_\mu K^+) - \frac{1}{2} \bar{p}_v v^\mu (\Sigma_v^0 + \sqrt{3}A_v) ((\pi^0 + \sqrt{3}\eta) \vec{\partial}_\mu K^+) + \text{h.c.} \\
& -i \frac{3}{4f^2} \left[ (K^+ v \cdot \vec{\partial} K^-) \left( A_{v,v}^{++} A_v^{+,v} + \frac{2}{\sqrt{3}} \bar{A}_{v,v}^+ A_v^{+,v} + \frac{1}{\sqrt{3}} \bar{\Sigma}_{v,v}^{*+} \Sigma_v^{*+,v} \right. \right. \\
& \left. \left. + \frac{1}{\sqrt{3}} \bar{A}_{v,v}^0 A_v^{0,v} - \frac{1}{\sqrt{3}} \bar{\Xi}_{v,v}^{*0} \Xi_v^{*0,v} - \frac{1}{\sqrt{3}} \bar{\Sigma}_{v,v}^{*-} \Sigma_v^{*-,v} - \bar{\Omega}_{v,v}^- \Omega_v^{-,v} - \frac{2}{\sqrt{3}} \Xi_{v,v}^{*-} \Xi_v^{*-,v} \right) \right], \quad (B.3)
\end{aligned}$$

$$\begin{aligned}
\mathcal{L}_5 = & \frac{1}{6\sqrt{2}f^3} \left[ K^+ K^- \left\{ \sqrt{2} \bar{p}_v S_v^\mu p_v 2F(\partial_\mu \pi^0 + \sqrt{3}\partial_\mu \eta) - \sqrt{2} \bar{n}_v S_v^\mu n_v (D - F)(\partial_\mu \pi^0 + \sqrt{3}\partial_\mu \eta) \right\} \right. \\
& + K^+ K^- \left\{ (D + F) \bar{p}_v S_v^\mu n_v \partial_\mu \pi^+ + (D - F) \bar{p}_v S_v^\mu \Sigma_v^+ \partial_\mu K^0 + 2(D - F) \bar{n}_v S_v^\mu \Sigma_v^- \partial_\mu K^+ \right. \\
& - \frac{2}{\sqrt{6}} (D + 3F) \bar{p}_v S_v^\mu A_v \partial_\mu K^+ - \frac{1}{\sqrt{6}} (D + 3F) \bar{n}_v S_v^\mu A_v \partial_\mu K^0 \\
& \left. + \frac{2}{\sqrt{2}} (D - F) \bar{p}_v S_v^\mu \Sigma_v^0 \partial_\mu K^+ - \frac{1}{\sqrt{2}} (D - F) \bar{n}_v S_v^\mu \Sigma_v^0 \partial_\mu K^0 + \text{h.c.} \right\} \\
& + \frac{C}{12\sqrt{2}f^3} \left[ (K^+ K^-) \left( \sqrt{\frac{2}{3}} \bar{p}_v \partial^\mu \pi^0 A_{v,\mu}^+ + \sqrt{2} \bar{p}_v \partial^\mu \eta A_{v,\mu}^+ + \frac{1}{\sqrt{3}} \bar{p}_v \partial^\mu \pi^+ A_{v,\mu}^0 \right. \right. \\
& + \sqrt{\frac{2}{3}} \bar{p}_v \partial^\mu K^+ \Sigma_{v,\mu}^{*0} - \bar{p}_v \partial^\mu \pi^- A_{v,\mu}^{++} - \frac{1}{\sqrt{3}} \bar{p}_v \partial^\mu K^0 \Sigma_{v,\mu}^{*+} + \sqrt{\frac{2}{3}} \bar{n}_v \partial^\mu \pi^0 A_{v,\mu}^0 + \sqrt{2} \bar{n}_v \partial^\mu \eta A_{v,\mu}^0 \\
& \left. \left. + \bar{n}_v \partial^\mu \pi^+ A_{v,\mu}^- + \frac{2}{\sqrt{3}} \bar{n}_v \partial^\mu K^+ \Sigma_{v,\mu}^{*-} - \frac{1}{\sqrt{3}} \bar{n}_v \partial^\mu \pi^- A_{v,\mu}^+ - \frac{1}{\sqrt{6}} \bar{n}_v \partial^\mu K^0 \Sigma_{v,\mu}^{*0} \right) + \text{h.c.} \right], \quad (B.4)
\end{aligned}$$

$$\begin{aligned}
\mathcal{L}_6 = & -\frac{i}{96f^4} \left[ \bar{p}_v v^\mu p_v \left\{ (K^+ \vec{\partial}_\mu K^-) (5\pi^+ \pi^- + 5K^0 K^0) \right. \right. \\
& + 8K^+ K^- + (\pi^0)^2 + 2\sqrt{3}\pi^0\eta + 3\eta^2) + 7K^+ K^- (\pi^+ \vec{\partial}_\mu \pi^-) - 7(K^0 \vec{\partial}_\mu K^0) K^+ K^- \Big\} \\
& + \bar{n}_v v^\mu n_v \left\{ (K^+ \vec{\partial}_\mu K^-) (\frac{1}{2}(\pi^0)^2 + \sqrt{3}\pi^0\eta + \frac{3}{2}\eta^2 \right. \\
& \left. - 2\pi^+ \pi^- + 4K^+ K^- + 7K^0 K^0) - 5(K^0 \vec{\partial}_\mu K^0) K^+ K^- + 2K^+ K^- (\pi^+ \vec{\partial}_\mu \pi^-) \Big\} \Big], \quad (B.5)
\end{aligned}$$

where the quark masses are related to the meson masses as

$$M_\pi^2 = B_0(m_u + m_d) = 2B_0\hat{m}, \quad M_\eta^2 = \frac{2}{3}B_0(\hat{m} + 2m_s), \quad M_K^2 = B_0(\hat{m} + m_s). \quad (B.6)$$

The lowest-order Lagrangian involving baryon octet-meson-decuplet coupling is obtained from

$$\mathcal{L} = C(\bar{T}^\mu A_\mu B + \text{h.c.}) = C\epsilon^{ijk} \bar{T}_{ilm}^\mu A_{j,\mu}^l B_k^m + \text{h.c.} \quad (B.7)$$

Here the decuplet fields are given as

$T^{111}$	$T^{112}$	$T^{113}$	$T^{122}$	$T^{123}$	$T^{133}$	$T^{222}$	$T^{223}$	$T^{233}$	$T^{333}$
$\Delta^{++} \frac{1}{\sqrt{3}} \Delta^+ \frac{1}{\sqrt{3}} \Sigma^{*+} \frac{1}{\sqrt{3}} \Delta^0 \frac{1}{\sqrt{6}} \Sigma^{*0}$	$\frac{1}{\sqrt{3}} \Xi^{*0} \Delta^- \frac{1}{\sqrt{3}} \Sigma^- \frac{1}{\sqrt{3}} \Xi^{*-} \Omega^-$								

where  $T^{ijk}$  is symmetric with respect to i,j and k.

The propagator for the baryon octet in HFF is  $i/(v \cdot k)$ , where  $k^\mu$  is the residual momentum of the baryon as defined by  $p^\mu = m_b v^\mu + k^\mu$  [17]. The propagator for the decuplet is also simplified to the form  $[i/(v \cdot k - \delta_T)] \times (v^\mu v^\nu - g^{\mu\nu} - 4((d-3)/(d-1)) S_v^\mu S_v^\nu)$ . In the HFF, the spin operator takes the form  $S_v^\mu S_v^\nu = \frac{1}{4}(v^\mu v^\nu - g^{\mu\nu}) + (i/2)\epsilon^{\mu\nu\rho\beta} v_\alpha S_{v,\beta}$ .

Given the vertices and the propagators, one can immediately write down the the integral entering in the diagram (e) of Fig. 3.1,

$$I_i^E = \frac{1}{i} \int \frac{d^n k}{(2\pi)^n} \frac{v \cdot (2q - k)}{v \cdot k + i\epsilon} \frac{v \cdot (q + q' - k)}{(q - k)^2 - M_i^2 + i\epsilon} \\ = -(v \cdot q' + 2v \cdot q) \Delta_i + 2v \cdot q v \cdot (q + q') \Sigma_i(v \cdot q), \quad (\text{B.9})$$

where

$$\Delta_i = \frac{1}{i} \int \frac{d^n k}{(2\pi)^n} \mu^\epsilon \frac{1}{k^2 - M_i^2 + i\epsilon} = -\frac{M_i^2}{16\pi^2} \ln \frac{M_i^2}{\mu^2} - M_i^2 2L, \\ \Sigma_i(\omega) = \frac{1}{i} \int \frac{d^n k}{(2\pi)^n} \frac{1}{k^2 - M_i^2 + i\epsilon} \frac{1}{v \cdot k + \omega + i\epsilon} = \frac{1}{8\pi^2} \left[ \omega \left( 1 - \ln \frac{M_i^2}{\mu^2} \right) + \tilde{f}_i(\omega) \right] - 4\omega L. \quad (\text{B.10})$$

Here the divergent term  $L$  and the finite loop term  $\tilde{f}_i(\omega)$  are given by

$$L = \frac{1}{16\pi^2} \mu^{-\epsilon} \left( -\frac{2}{\epsilon} + \gamma - 1 - \ln 4\pi \right), \\ \tilde{f}_i(\omega) = \sqrt{\omega^2 - M_i^2} \left( \ln \left| \frac{\omega - \sqrt{\omega^2 - M_i^2}}{\omega + \sqrt{\omega^2 - M_i^2}} \right| + i2\pi\theta(\omega - M_i) \right) \quad \text{for } |\omega| > M_i, \\ = -\sqrt{M_i^2 - \omega^2} \left( \pi + 2\tan^{-1} \frac{\omega}{\sqrt{M_i^2 - \omega^2}} \right) \quad \text{for } |\omega| < M_i. \quad (\text{B.11})$$

Note that  $\tilde{f}_i(\omega)$  has a kink at  $\omega = M_i$ , which explains the nontrivial behavior of the  $KN$  scattering amplitude seen in Fig. 4.1. In order to get one-loop results, we also need the following well-known integrals

$$J_i(q^2) = \frac{1}{i} \int \frac{d^n k}{(2\pi)^n} \mu^\epsilon \frac{1}{k^2 - M_i^2 + i\epsilon} \frac{1}{(k - q)^2 - M_i^2 + i\epsilon} \\ = -\frac{1}{16\pi^2} \left( 1 + \ln \frac{M_i^2}{\mu^2} \right) - 2L + J_i(q^2), \quad (\text{B.12})$$

where

$$J_i(t) = \frac{1}{16\pi^2} \left[ 2 - \sqrt{1 - 4M_i^2/t} \ln \frac{\sqrt{1 - 4M_i^2/t + 1}}{\sqrt{1 - 4M_i^2/t - 1}} \right]. \quad (\text{B.13})$$

All the integrals needed for the diagrams in Fig. 3.1 can be obtained from  $\Delta_i$ ,  $J_i(q^2)$  and  $\Sigma_i(\omega)$ .

### Appendix C. Mass and wave function renormalizations

At one-loop order, we need one counter term,  $\mathcal{L}_c = -\delta m_B \text{Tr } \bar{B}B$ , to renormalize the nucleon mass. It does not affect  $KN$  scattering amplitude, but it is needed to absorb the divergences coming from one-loop self-energy graphs. The nucleon self-energy including the counter terms is

$$\begin{aligned} \Sigma_N(\omega) &= \delta m_B - 2a_1\hat{m} - 2a_2m_s - 2a_3(m_s + 2\hat{m}) - 2\omega\bar{h} - c_1\omega^3 \\ &\quad + \sum_i \frac{\lambda_i}{f^2} \mathcal{H}_i(\omega) + \sum_i \frac{4}{3} \frac{\lambda_{D,i}}{f^2} \mathcal{H}_i(\omega - \delta_T), \end{aligned} \quad (\text{C.1})$$

where  $\omega = v \cdot p - m_B v$  and

$$\begin{aligned} \bar{h} &= h_1\hat{m} + h_2m_s + h_3(2\hat{m} + m_s), \\ \mathcal{H}_i(\omega) &= -\frac{1}{4} [\omega\Delta_i + (M_i^2 - \omega^2)\Sigma_i(\omega)], \\ \Delta_i &= -\frac{M_i^2}{16\pi^2} \ln \frac{M_i^2}{\mu_{\text{d.r.}}^2} - M_i^2 2L, \quad \Sigma_i(\omega) = \frac{1}{8\pi^2} \left[ \omega \left( 1 - \ln \frac{M_i^2}{\mu_{\text{d.r.}}^2} \right) - f_i(\omega) \right] - 4\omega L, \\ f_i(\omega) &= \sqrt{M_i^2 - \omega^2} \left( \pi + 2 \sin^{-1} \frac{\omega}{M_i} \right), \quad L = \frac{1}{16\pi^2} \mu_{\text{d.r.}}^{-\epsilon} \left( -\frac{2}{\epsilon} + \gamma - 1 - \ln 4\pi \right), \end{aligned} \quad (\text{C.2})$$

and the coefficients  $\lambda_i$ 's are

$$\begin{aligned} \lambda_\pi &= 3(D + F)^2, \quad \lambda_K = \frac{1}{3} ((D + 3F)^2 + 9(D - F)^2), \\ \lambda_\eta &= \frac{1}{3}(D - 3F)^2, \quad \lambda_{D,\pi} = 4\lambda_{D,K} = 2C^2. \end{aligned} \quad (\text{C.3})$$

Here  $\mu_{\text{d.r.}}$  stands for the arbitrary mass scale  $\mu$  that arises in the dimensional regularization. In standard notation,  $Z_0$  ( $Z_3$ ) corresponds to the mass (wave function) renormalization. The counter term  $(Z_0 - 1)m_B \text{Tr } \bar{B}B$  corresponds to  $\delta m_B \text{Tr } \bar{B}B$ , and  $(Z_3 - 1)\text{Tr } \bar{B}iv \cdot \partial B$  to the  $h_i$  terms in the  $\nu = 3$  Lagrangian. One can compute  $Z_0$  and  $Z_3$  from  $\Sigma_N(\omega)$  or write down the corresponding counter terms as in renormalizable theory,  $(Z_0 - 1)m_B \text{Tr } \bar{B}B$  and  $(Z_3 - 1)\text{Tr } \bar{B}iv \cdot \partial B$ . However in a non-renormalizable theory such as ours, many counter terms enter at order by order, so it is more convenient to use the constraints on  $\Sigma_N(\omega)$  directly. Using a physical scheme, the convenient constraints are

$$\Sigma_N(\omega)|_{\omega=0} = 0, \quad \left. \frac{\partial \Sigma_N(\omega)}{\partial \omega} \right|_{\omega=0} = 0. \quad (\text{C.4})$$

In this case, the  $m_B$ , that figures in the leading-order Lagrangian, can be identified as the physical baryon mass. The constant  $\delta m_B$  is determined by the first constraint to absorb the divergence and  $\bar{h}$ , relevant for  $KN$  scattering, can be fixed by the second constraint. The latter is explicitly given by

$$\begin{aligned} \frac{\partial \Sigma_N(\omega)}{\partial \omega} \Big|_{\omega=0} = & -2\bar{h} + \sum_i \frac{\lambda_i}{f^2} \left( -\frac{1}{4}\Delta_i + M_i^2 \frac{\partial \Sigma_i(\omega)}{\partial \omega} \Big|_{\omega=0} \right) \\ & + \sum_i \frac{4}{3} \frac{\lambda_{D,i}}{f^2} \left( -\frac{1}{4}\Delta_i + 2\delta_T \Sigma(-\delta_T) + (M_i^2 - \delta_T^2) \frac{\partial \Sigma_i(\omega)}{\partial \omega} \Big|_{\omega=-\delta_T} \right). \end{aligned} \quad (\text{C.5})$$

The constraints

$$Z_N - 1 = \frac{\partial \Sigma_N(\omega)}{\partial \omega} \Big|_{\omega=0} = 0 \quad (\text{C.6})$$

completely determine  $\bar{h}$  as a function of  $\mu_{\text{d.r.}}$ . One convenient choice is  $\bar{h} = \bar{h}'(\mu_{\text{d.r.}}) + \sum_i \alpha_i L$  where  $L$  is the divergent piece, and  $\bar{h}'$  is the finite part that includes the chiral log terms  $\ln(M_i^2/\mu_{\text{d.r.}}^2)$ . The renormalized parameter  $\bar{h}'(\mu_{\text{d.r.}})$  is related with each other at different scales  $\mu_{\text{d.r.}}$  through the relation

$$\frac{\bar{h}'(\mu_{\text{d.r.}} = \mu_1)}{\bar{h}'(\mu_{\text{d.r.}} = \mu_2)} = \sum_i \beta_i(\delta_T) \ln \frac{\mu_1}{\mu_2}. \quad (\text{C.7})$$

Consider now kaon self-energy which we can write as

$$\Pi_K(q^2) = \sum_i \frac{\alpha_i}{12f^2} \Delta_i + \text{c.t.}, \quad (\text{C.8})$$

where  $\alpha_i$  are given by

$$\alpha_\pi = -3M_\pi^2, \quad \alpha_K = -6M_K^2, \quad \alpha_\eta = M_\eta^2. \quad (\text{C.9})$$

In the physical scheme, the constraints are

$$\Pi_K(q^2) \Big|_{q^2=M_K^2} = 0, \quad Z_K - 1 = \frac{\partial \Pi_K(q^2)}{\partial q^2} \Big|_{q^2=M_K^2} = 0 \quad (\text{C.10})$$

where  $M_K$  is the physical kaon mass. Here we pick the renormalization point at  $\mu_K = M_K$  for the kaon wave function renormalization. For  $\pi$  and  $\eta$ , we take  $M_\pi$  and  $M_\eta$  in the tree-order Lagrangian as physical masses. This physical scheme is independent of the arbitrary mass scale  $\mu_{\text{d.r.}}$  figuring in the dimensional regularization. Note however that the counter terms in kaon self-energy do not affect the  $KN$  scattering directly. So we shall not need the explicit magnitudes of these counter terms for our purpose.

Finally the scattering amplitudes can be obtained simply by calculating the six topologically distinct diagrams of Fig. 3.1 using physical masses

$$T^{KN} = \sum_{i=a,\dots,f} T_i^{KN}. \quad (\text{C.11})$$

This is because there is no contribution from the mass and wave function renormalizations,

$$T_{w,r}^{KN} = \alpha \frac{(1 - Z_N + 1 - Z_K)}{2f^2} v \cdot (q + q'), \quad (\text{C.12})$$

where  $\alpha = 2$  for  $K^+p$  and  $\alpha = 1$  for  $K^+n$ .

## Appendix D. Renormalization of $\mathcal{O}(Q^3)$ counter terms

The quantity  $\bar{g}_s(\bar{g}_v)$  is the crossing-odd t-channel isoscalar (isovector) contribution from one-loop plus counter terms which after the standard dimensional regularization, takes the form

$$\bar{g}_{s,v} = \alpha_{s,v} + \beta_{s,v} + \frac{1}{32\pi^2 f^2 M_K^2} \gamma_{s,v} + \sum_{i=\pi,K,\eta} \frac{\delta_{i,v}}{f^2 M_i^2} \left( L + \frac{1}{32\pi^2} \ln \frac{M_i^2}{\mu^2} \right), \quad (\text{D.1})$$

where  $L$  contains the divergence, and

$$\begin{aligned} \alpha_s &= -\frac{1}{M_K^2} \left( -\frac{3}{4}\bar{h} + (l_1 - 2l_2 + l_3)(\hat{m} + m_s) \right), \\ \beta_s &= -\frac{1}{2} ((g_2 - 2g_3 + g_6) + (g_4 - 2g_5 + g_7) + (g_8 - 2g_9)), \\ \alpha_v &= -\frac{1}{M_K^2} \left( -\frac{1}{4}\bar{h} + (l_1 + l_3)(\hat{m} + m_v) \right), \\ \beta_v &= -\frac{1}{2} ((g_2 + g_6) + (g_4 + g_7) + g_8), \end{aligned} \quad (\text{D.2})$$

with

$$\bar{h} = h_1 \hat{m} + h_2 m_s + h_3 (2\hat{m} + m_v). \quad (\text{D.3})$$

Here we have separated  $\alpha_{s,v}$  and  $\beta_{s,v}$  because in off-shell amplitudes, the constant  $\alpha_{s,v}$  is multiplied by  $\omega$  while the  $\beta_{s,v}$  is multiplied by  $\omega^3$ , thus behaving differently for  $\omega \neq M_K$ . The constants  $\gamma_{s,v}$  – coming from finite loop terms – and  $\delta_{i,v}$  – multiplying the divergence – are given by

$$\begin{aligned} \gamma_s &= -3M_K^2 - \frac{9}{4}M_K f_\pi (-M_K) - \frac{9}{4}M_K f_\eta (-M_K) - \frac{9}{4}(D+F)^2 M_\pi^2 \\ &\quad - \frac{1}{4}(D-3F)^2 M_\eta^2 + \left( \frac{1}{2}(D-F)^2 + \frac{1}{6}(D+3F)^2 \right) M_K^2 \\ &\quad + |C|^2 \left( (1+\sqrt{3}) C_\pi(-\delta_T) + \frac{1}{3\sqrt{3}} C_K(-\delta_T) - \frac{1}{4} F_K(-\delta_T) \right), \\ \gamma_v &= \frac{1}{3}M_\pi^2 - \frac{4}{3}M_K^2 + \frac{1}{4}M_K f_\pi (-M_K) - \frac{3}{4}M_K f_\eta (-M_K) + \frac{11}{12}(D+F)^2 M_\pi^2 \\ &\quad - \frac{1}{12}(D-3F)^2 M_\eta^2 + \left( -\frac{7}{6}(D-F)^2 + \frac{1}{18}(D+3F)^2 \right) M_K^2 \\ &\quad + |C|^2 \left( \left( 1 + \frac{1}{3\sqrt{3}} \right) C_\pi(-\delta_T) + \frac{1}{\sqrt{3}} C_K(-\delta_T) + \frac{2}{9} F_\pi(-\delta_T) + \frac{1}{36} F_K(-\delta_T) \right), \\ \delta_s^\pi &= \frac{9}{4}M_K^2 - \frac{3}{4}M_\pi^2 - \frac{27}{8}(D+F)^2 M_\pi^2 + |C|^2 (3 + 3\sqrt{3}) (2\delta_T^2 - M_\pi^2), \\ \delta_s^K &= -\frac{3}{2}M_K^2 - \left( \frac{27}{8}(D-F)^2 + \frac{3}{8}(D+3F)^2 \right) M_K^2 + |C|^2 \left( \frac{3}{4} + \frac{1}{2\sqrt{3}} \right) (2\delta_T^2 - M_K^2), \\ \delta_v^\pi &= \frac{9}{4}M_K^2 - \frac{3}{4}M_\eta^2 - \frac{3}{8}(D-3F)^2 M_\eta^2, \\ \delta_v^K &= -\frac{1}{4}M_K^2 - \frac{1}{4}M_\pi^2 - \frac{9}{8}(D+F)^2 M_\pi^2 + |C|^2 \left( \frac{5}{3} + \frac{1}{\sqrt{3}} \right) (2\delta_T^2 - M_\pi^2), \\ \delta_v^K &= \frac{1}{2}M_K^2 - \left( \frac{9}{8}(D-F)^2 + \frac{1}{8}(D+3F)^2 \right) M_K^2 + |C|^2 \left( -\frac{1}{12} + \frac{1}{2}\sqrt{3} \right) (2\delta_T^2 - M_K^2). \end{aligned}$$

$$\delta_v^0 = \frac{3}{4}M_K^2 - \frac{1}{4}M_\eta^2 - \frac{1}{8}(D - 3F)^2 M_\eta^2, \quad (D.4)$$

where  $C_i(-\delta_T)$ ,  $F_i(-\delta_T)$  and  $f_i(\omega)$  are

$$\begin{aligned} C_i(-\delta_T) &= -(M_i^2 + \delta_T^2) - 3\delta_T f_i(-\delta_T), \quad F_i(-\delta_T) = -\frac{4}{3}M_i^2 + 6\delta_T^2 + 6\delta_T f_i(-\delta_T), \\ f_i(\omega) &= \sqrt{\omega^2 - M_i^2} \ln \left| \frac{\omega + \sqrt{\omega^2 - M_i^2}}{\omega - \sqrt{\omega^2 - M_i^2}} \right| \quad \text{for } \omega^2 > M_i^2 \\ &= \sqrt{M_i^2 - \omega^2} \left( \pi + 2 \sin^{-1} \frac{\omega}{M_i} \right) \quad \text{for } \omega^2 < M_i^2. \end{aligned} \quad (D.5)$$

Note that  $f_i(\omega)$  result from diagram of Fig. 3.1e as explained in Appendix B. We write the counter terms as

$$\alpha_{s,v} = \alpha'_{s,v} + \alpha_{s,v}^{div}, \quad \beta_{s,v} = \beta'_{s,v} + \beta_{s,v}^{div}, \quad (D.6)$$

where the divergent parts  $\alpha_{s,v}^{div}$  and  $\beta_{s,v}^{div}$  are to cancel the divergence part  $L$  in Eq. (D.1) and the finite counter terms  $\alpha'$  and  $\beta'$  are to be fixed by experiments. After removing the divergences, the constant  $\bar{g}_{s,v}$  can be written as

$$\bar{g}_{s,v} = \alpha'_{s,v} + \beta'_{s,v} + \frac{1}{32\pi^2 f^2 M_K^2} \left( \gamma_{s,v} + \sum_{i=\pi, K, \eta} \delta_{s,v}^i \ln \frac{M_i^2}{\mu^2} \right). \quad (D.7)$$

Note that  $\bar{g}_{s,v}$  is  $\mu$ -independent and hence can be fixed from experiments independently of  $\mu$ . How to do this for off-shell amplitudes is described in the main text.

## Appendix E. $1/m$ corrections of baryon octet

Here we sketch briefly the derivation of the  $1/m_B$  corrections. Consider the lowest-order Lagrangian with baryon fields,

$$\mathcal{L} = \text{Tr } \bar{B} (i \not{D} B - m(1 - \not{\psi}) B + F \gamma^\mu \gamma_5 [A_\mu, B] + D \gamma^\mu \gamma_5 \{A_\mu, B\}). \quad (E.1)$$

Rewriting the fields as  $B = \frac{1}{2}\sqrt{2}B^a \lambda^a$ ,  $V_\mu = V_\mu^a \lambda^a$  and  $A_\mu = A_\mu^a \lambda^a$ , where  $\lambda^a$ ,  $a = 1, 8$  are the Gell-Mann matrices, the Lagrangian can be written

$$\mathcal{L} = \bar{B}^a i \not{D}^{ac} B^c - m \bar{B}^a \delta^{ac} (1 - \not{\psi}) B^c + \bar{B}^a \gamma^\mu \gamma_5 A_\mu^{ac} B^c. \quad (E.2)$$

Here

$$i \not{D}^{ac} = i \delta^{ac} \not{\partial} + i(2if^{abc}) \not{\gamma}^b, \quad A_\mu^{ac} = F(2if^{abc}) A_\mu^b + D(2d^{abc}) A_\mu^b, \quad (E.3)$$

and  $f^{abc}$ ,  $d^{abc}$  are  $SU(3)$  structure constants. The equation of motion for the heavy baryon field is

$$[g^{ac} - m(1 - \not{\psi}) \delta^{ac}] B^c = 0, \quad (E.4)$$

with  $g^{ac} = i \not{D}^{ac} + \gamma^\mu \gamma_5 A_\mu^{ac}$ . Now, decompose  $B^a$  using  $P_+ = \frac{1}{2}(1 + \not{\psi})$  and  $P_- = \frac{1}{2}(1 - \not{\psi})$ ,

$$B^a = P_+ B^a + P_- B^a = B_+^a + B_-^a. \quad (\text{E.5})$$

Applying  $P_-$  to Eq. (E.4) from the left, we have

$$B_-^c = \frac{1}{2m} P_- g^{ce} B_+^e + O\left(\frac{1}{m^2}\right). \quad (\text{E.6})$$

Now applying  $P_+$  to Eq. (E.4) and using Eq. (E.6), we get, modulo  $O(1/m^2)$ ,

$$P_+ \left( g^{ae} + \frac{1}{2m} g^{ac} P_- g^{ce} \right) B_+^e = 0. \quad (\text{E.7})$$

Thus the Lagrangian containing  $1/m_B$  correction term is

$$\mathcal{L} = \tilde{B}_+^a \left( g^{ae} + \frac{1}{2m} g^{ac} P_- g^{ce} \right) B_+^e. \quad (\text{E.8})$$

To write this explicitly, we note that the  $O(A^2)$  terms are given by

$$g^{ac} P_- g^{ce} = \gamma^\mu \gamma_5 A_\mu^{ac} P_- \gamma^\nu \gamma_5 A_\nu^{ce} = -A^{ac} P_+ A^{ce} = -(v \cdot A^{ac})(v \cdot A^{ce}), \quad (\text{E.9})$$

which leads to ( $m = m_B$ )

$$\mathcal{L}_{1/m} \longrightarrow -\frac{1}{2m_B} \tilde{B}^a (v \cdot A^{ac}) (v \cdot A^{ce}) B^e. \quad (\text{E.10})$$

From the identities

$$\begin{aligned} (2if^{abc})(2if^{cde}) &= \frac{1}{2} \text{Tr} (\lambda^a [\lambda^b, [\lambda^d, \lambda^e]]), \quad (2if^{abc})(2d^{cde}) = \frac{1}{2} \text{Tr} (\lambda^a [\lambda^b, \{\lambda^d, \lambda^e\}]), \\ (2d^{abc})(2if^{cde}) &= \frac{1}{2} \text{Tr} (\lambda^a \{\lambda^b, [\lambda^d, \lambda^e]\}), \\ (2d^{abc})(2d^{cde}) &= \frac{1}{2} \text{Tr} (\lambda^a \{\lambda^b, \{\lambda^d, \lambda^e\}\}) - \frac{2}{3} \text{Tr} (\lambda^a \lambda^b) \text{Tr} (\lambda^d \lambda^e), \end{aligned} \quad (\text{E.11})$$

we have

$$\begin{aligned} \tilde{B}^a (v \cdot A^{ac}) (v \cdot A^{ce}) B^e &= \text{Tr } \tilde{B} (\mathbf{D}^2 \{v \cdot A, \{v \cdot A, B\}\} + \mathbf{F} \mathbf{D} \{v \cdot A, \{v \cdot A, B\}\}) \\ &\quad + \mathbf{F}^2 [v \cdot A, [v \cdot A, B]] + \mathbf{F} \mathbf{D} [v \cdot A, \{v \cdot A, B\}] - \frac{4}{3} \mathbf{D}^2 \text{Tr} (\tilde{B} v \cdot A) \text{Tr} (v \cdot AB) \\ &= (\mathbf{D} + \mathbf{F})^2 \text{Tr } \tilde{B} (v \cdot A)^2 B + 2(\mathbf{D}^2 - \mathbf{F}^2) \text{Tr } \tilde{B} (v \cdot A) B (v \cdot A) \\ &\quad + (\mathbf{D} - \mathbf{F})^2 \text{Tr } \tilde{B} B (v \cdot A)^2 - \frac{4}{3} \mathbf{D}^2 \text{Tr} (\tilde{B} v \cdot A) \text{Tr} (v \cdot AB). \end{aligned} \quad (\text{E.12})$$

For s-wave scattering

$$\begin{aligned} \text{Tr } \tilde{B} A_0^2 B &\longrightarrow \frac{1}{4f^2} \left[ (\bar{N} \tau N) \cdot (\partial_t \bar{K} \tau \partial_t K) + (\bar{N} N) (\partial_t \bar{K} \partial_t K) \right], \\ \text{Tr } (\tilde{B} A_0) \text{Tr } (A_0 B) &\longrightarrow \frac{1}{4f^2} \left[ (\bar{N} \tau N) \cdot (\partial_t \bar{K} \tau \partial_t K) + (\bar{N} N) (\partial_t \bar{K} \partial_t K) \right], \\ \text{Tr } \tilde{B} B A_0^2 &\longrightarrow \frac{1}{2f^2} (\bar{N} N) (\partial_t \bar{K} \partial_t K), \quad \text{Tr } (\tilde{B} A_0 B A_0) \longrightarrow 0. \end{aligned} \quad (\text{E.13})$$

The  $1/m_B$  corrections are therefore

$$\mathcal{L}_{1/m} = \bar{d}_{s,\frac{1}{2}} \frac{1}{f^2} (\bar{N} N \partial_\mu \bar{K} \partial_\mu K) + \bar{d}_{v,\frac{1}{2}} \frac{1}{f^2} (\bar{N} \tau N) \cdot (\partial_\mu \bar{K} \tau \partial_\mu K), \quad (\text{E.14})$$

with

$$\begin{aligned} \bar{d}_{s,\frac{1}{2}} &= -\frac{1}{48m_B} [(D + 3F)^2 + 9(D - F)^2], \\ \bar{d}_{v,\frac{1}{2}} &= -\frac{1}{48m_B} [(D + 3F)^2 - 3(D - F)^2]. \end{aligned} \quad (\text{E.15})$$

## Appendix E. Decay width

In this review, the  $A^*$  coupling  $g_{A^*}^2$  was fixed by using a relativistic formula based on a relativistic chiral Lagrangian [1]. The  $A^*$  decay width in such formulation is

$$\Gamma_r = \frac{3g_{A^*}^2}{4\pi f^2} |p_\pi| \times \frac{2(E_\pi E_\Sigma + |p_\pi|^2)E_\pi - M_\pi^2(E_\Sigma - m_\Sigma)}{m_{A^*}}, \quad (\text{E.1})$$

with the pion momentum in c.m. frame given by

$$|p_\pi| = \sqrt{\frac{(m_{A^*}^2 - m_\Sigma^2 - M_\pi^2)^2 - 4M_\pi^2 m_\Sigma^2}{4m_{A^*}^2}} \approx 147 \text{ MeV}. \quad (\text{E.2})$$

Taking  $g^2 = 0.15$ , we have the decay width  $\Gamma_r \approx 50 \text{ MeV}$ . But in the heavy baryon approach, we have a larger momentum,

$$|p_\pi| \simeq \sqrt{(m_{A^*} - m_\Sigma)^2 - M_\pi^2} \approx 159 \text{ MeV}. \quad (\text{E.3})$$

As a result, the decay width is much larger than the relativistic approach,

$$\Gamma_{HBF} = \frac{6g_{A^*}^2}{4\pi f^2} \tilde{\omega}^2 \sqrt{\tilde{\omega}^2 - M_K^2} \approx 57 \text{ MeV} \quad (\text{E.4})$$

where  $\tilde{\omega} = (m_{A^*} - m_\Sigma)$  is the pion energy. This HBF decay width is about 20% larger than  $\Gamma_r \approx 50 \text{ MeV}$ . If we start from HBF decay width, we need a smaller  $g_{A^*}$  to reproduce the empirical decay width,

$$g_{A^*}^2 = 0.127. \quad (\text{E.5})$$

## Appendix G. $K^- p$ transition amplitudes

The transition matrix elements for various channels of threshold  $K^- p$  scattering are

$$\begin{aligned} T(K^- p \rightarrow \pi^- \Sigma^+) &= \frac{1}{4f^2} (2M_K + m_p - m_{\Sigma^+}) - \frac{g_{A^*}^2}{f^2} \frac{M_K(M_K + m_p - m_{\Sigma^+})}{m_p + M_K - m_{A^*}} \\ &\quad - \frac{1}{2f^2} a_2(2m_u + m_d + m_s) + \frac{1}{2f^2} (\bar{d}_3 + \bar{d}_7) M_K(M_K + m_p - m_{\Sigma^+}), \end{aligned}$$

$$\begin{aligned}
T(K^- p \rightarrow \pi^+ \Sigma^-) &= -\frac{g_A^2}{f^2} \frac{M_K(M_K + m_p - m_{\Sigma^-})}{m_p + M_K - m_{A^*}} + \frac{1}{2f^2} (\bar{d}_7 + \bar{d}_9) M_K (M_K + m_p - m_{\Sigma^-}), \\
T(K^- p \rightarrow K^- p) &= \frac{1}{f^2} M_K - \frac{g_A^2}{f^2} \frac{M_K^2}{m_p + M_K - m_{A^*}} \\
&\quad - \frac{1}{f^2} (m_u + m_s) (a_1 + a_2 + 2a_3) + \frac{1}{2f^2} (\bar{d}_1 + \bar{d}_3 + 2\bar{d}_5 + \bar{d}_7) M_K^2, \\
T(K^- p \rightarrow \bar{K}^0 n) &= \frac{1}{2f^2} M_K - \frac{g_A^2}{f^2} \frac{M_K(M_K + m_p - m_n)}{m_p + M_K - m_{A^*}} \\
&\quad - \frac{1}{2f^2} (m_u + m_d + 2m_s) a_1 + \frac{1}{2f^2} (\bar{d}_1 + \bar{d}_7) M_K (M_K + m_p - m_n), \\
T(K^- p \rightarrow \pi^0 \Sigma^0) &= \frac{1}{8f^2} (2M_K + m_p - m_{\Sigma^0}) - \frac{g_A^2}{f^2} \frac{M_K(M_K + m_p - m_{\Sigma^0})}{m_p + M_K - m_{A^*}} \\
&\quad - \frac{1}{4f^2} (3m_u + m_s) a_2 + \frac{1}{4f^2} (\bar{d}_3 + 2\bar{d}_7 + \bar{d}_9) M_K (M_K + m_p - m_{\Sigma^0}), \\
T(K^- p \rightarrow \pi^0 \Lambda) &= \frac{\sqrt{3}}{8f^2} (2M_K + m_p - m_A) - \frac{1}{4\sqrt{3}f^2} (3m_u + m_s) (-2a_1 + a_2) \\
&\quad + \frac{1}{4\sqrt{3}f^2} (-2\bar{d}_1 + \bar{d}_3 + \bar{d}_9) M_K (M_K + m_p - m_A), \\
T(K^- p \rightarrow \eta \Sigma^0) &= \frac{\sqrt{3}}{8f^2} (2M_K + m_p - m_{\Sigma^0}) - \frac{1}{4\sqrt{3}f^2} a_2 (m_u - 5m_s) \\
&\quad + \frac{1}{4\sqrt{3}f^2} (-\bar{d}_3 + \bar{d}_9) M_K (M_K + m_p - m_{\Sigma^0}), \\
T(K^- p \rightarrow \eta \Lambda) &= \frac{3}{8f^2} (2M_K + m_p - m_A) - \frac{g_A^2}{f^2} \frac{M_K(M_K + m_p - m_A)}{m_p + M_K - m_{A^*}} \\
&\quad - \frac{1}{12f^2} (2a_1 - a_2) (5m_s - m_u) + \frac{1}{12f^2} (2\bar{d}_1 - \bar{d}_3 + 6\bar{d}_7 + 5\bar{d}_9) M_K (M_K + m_p - m_A), \\
T(K^- p \rightarrow K^+ \Xi^-) &= -\frac{g_A^2}{f^2} \frac{M_K(M_K + m_p - m_{\Xi^-})}{m_p + M_K - m_{A^*}} + \frac{1}{2f^2} (\bar{d}_7 + \bar{d}_9) M_K (M_K + m_p - m_{\Xi^-}), \\
T(K^- p \rightarrow K^0 \Xi^0) &= -\frac{g_A^2}{f^2} \frac{M_K(M_K + m_p - m_{\Xi^0})}{m_p + M_K - m_{A^*}} + \frac{1}{2f^2} (\bar{d}_7 + \bar{d}_9) M_K (M_K + m_p - m_{\Xi^0}),
\end{aligned} \tag{G.1}$$

where the physical masses of baryon and meson are used. The processes involving  $\bar{K}$ ,  $\eta$  and  $\Xi$  do not figure in the branching ratios  $R_c$  and  $R_n$  [63].

## References

- [1] G.E. Brown, C.-H. Lee, M. Rho and V. Thorsson, Nucl. Phys. A 567 (1994) 937;  
C.-H. Lee, H. Jung, D.-P. Min and M. Rho, Phys. Lett. B 326 (1994) 14;  
C.-H. Lee, G.E. Brown and M. Rho, Phys. Lett. B 335 (1994) 266;  
C.-H. Lee, G.E. Brown, D.-P. Min and M. Rho, Nucl. Phys. A 585 (1995) 401;  
C.-H. Lee and M. Rho, in: Proceedings of International Symposium on Strangeness and Quark Matter, September 1-5, 1994, Krete;
- C.-H. Lee and M. Rho, in: Proceedings of International Workshop on Nuclear and Particle Physics, Chiral Dynamics in Hadrons and Nuclei, Feb. 6-10, 1995, Seoul, Korea, D.-P. Min and M. Rho, eds. (Seoul Nat'l Univ. Press, 1995).
- [2] C.-H. Lee, D.-P. Min and M. Rho, SNUTP-95-060(hep-ph/9505283), The role of  $\Lambda(1405)$  in kaon-proton interactions.
- [3] G.E. Brown and H.A. Bethe, Astrophys. J. 423 (1994) 659.
- [4] D.B. Kaplan and A.E. Nelson, Phys. Lett. B 175 (1986) 57.
- [5] G.E. Brown, K. Kubodera and M. Rho, Phys. Lett. B 192 (1987) 273.
- [6] H.D. Politzer and M.B. Wise, Phys. Lett. B 273 (1991) 156.
- [7] G.E. Brown, K. Kubodera, M. Rho and V. Thorsson, Phys. Lett. B 291 (1992) 355.
- [8] V. Thorsson, M. Prakash and J. Lattimer, Nucl. Phys. A 572 (1994) 693.
- [9] H. Yabu, S. Nakamura, F. Myhrer and K. Kubodera, Phys. Lett. B 315 (1993) 17.
- [10] M. Lutz, A. Steiner and W. Weise, Nucl. Phys. A 574 (1994) 755.
- [11] S. Weinberg, Phys. Lett. B 251 (1990) 288;  
S. Weinberg, Nucl. Phys. B 363 (1991) 3.
- [12] C. Ordóñez, L. Ray and U. van Kolck, Phys. Rev. Lett. 72 (1994) 1982.
- [13] M. Rho, Phys. Rev. Lett. 66 (1991) 1275.
- [14] T.-S. Park, D.-P. Min and M. Rho, Phys. Rep. 233 (1993) 341.
- [15] T.-S. Park, I.S. Towner and K. Kubodera, Nucl. Phys. A 579 (1994) 381.
- [16] T.-S. Park, D.-P. Min and M. Rho, Phys. Rev. Lett. 74 (1995) 4153.
- [17] E. Jenkins and A.V. Manohar, Phys. Lett. B 255 (1991) 558;  
E. Jenkins and A. Manohar, Phys. Lett. B 259 (1991) 353;  
E. Jenkins, Nucl. Phys. B 368 (1992) 190.
- [18] E. Friedman, A. Gal and C.J. Batty, Phys. Lett. B 308 (1993) 6;  
E. Friedman, A. Gal and C.J. Batty, Nucl. Phys. A 579 (1994) 518.
- [19] C.G. Callan and I. Klebanov, Nucl. Phys. B 262 (1985) 365;  
C.G. Callan, K. Hornbostel and I. Klebanov, Phys. Lett. B 202 (1988) 269.
- [20] D. Montano, H.D. Politzer and M.B. Wise, Nucl. Phys. B 375 (1992) 507.
- [21] B.R. Martin and M. Sakitt, Phys. Rev. 183 (1969) 1352.
- [22] R. Shankar, Rev. Mod. Phys. 66 (1994) 129.
- [23] J. Polchinski, in: Recent Directions in Particle Physics, J. Harvey and J. Polchinski, eds. (World Scientific, Singapore, 1993) p. 235.
- [24] H.K. Lee, M. Rho and S.J. Sin, Phys. Lett. B 348 (1995) 290.
- [25] G.E. Brown and M. Rho, Phys. Rev. Lett. 66 (1991) 2720.
- [26] G.E. Brown and M. Rho, hep-ph/9504250, Phys. Rep. to be published (1995), Chiral restoration in hot and/or dense matter.
- [27] T. Tatsumi, H. Fujii and T. Muto, Kyoto Univ. preprint KUNS-1355, Chiral symmetry and kaon condensation; H. Fujii, T. Maruyama, T. Muto and T. Tatsumi, Kyoto Univ. preprint KUNS-1348, Equation of state with kaon condensates and neutron stars.
- [28] G.E. Brown and M. Rho, nucl-th/9507028, From chiral mean field to Walecka mean field and kaon condensation.
- [29] T. Tatsumi, Prog. Theor. Phys. 80 (1988) 22;  
T. Muto and T. Tatsumi, Phys. Lett. B 283 (1992) 165;  
H. Fujii, T. Muto, T. Tatsumi and R. Tamagaki, Nucl. Phys. A 571 (1994) 758;  
T. Maruyama, H. Fujii, T. Muto and T. Tatsumi, Phys. Lett. B 337 (1994) 19.

- [30] K.M. Westerberg, Kaon Condensation in the Bound-State Approach to the Skyrme Model, PUPT-1521, hep-ph/9411430.
- [31] H.A. Bethe and G.E. Brown, *Astrophys. J.* 445 (1995) L129.
- [32] S.L. Shapiro and S.A. Teukolsky, *Black Holes, White Dwarfs and Neutron Stars: The physics of compact objects* (John Wiley, New York, 1983).
- [33] J.N. Bahcall and R.A. Wolf, *Phys. Rev. Lett.* 14 (1965) 343;  
J.N. Bahcall and R.A. Wolf, *Phys. Rev.* 140 (1965) B1445, B1452.
- [34] O. Maxwell, G.E. Brown, D.K. Campbell, R.F. Dashen and J.T. Manassah, *Astrophys. J.* 216 (1977) 77.
- [35] M. Rho, *Nucl. Phys. A* 231 (1974) 493.
- [36] K. Ohta and M. Wakamatsu, *Nucl. Phys. A* 234 (1974) 445.
- [37] D.H. Wilkinson, *Phys. Rev. C* 7 (1973) 930.
- [38] G. Baym, *Phys. Rev. Lett.* 30 (1973) 1340.
- [39] G. Ecker, UWThPh-1994-49, in: *Progress in Particle and Nuclear Physics*, Vol. 35 (Pergamon, Oxford, 1995), Chiral perturbation theory.
- [40] J. Bijnens, G. Ecker and J. Gasser, in: *The DAΦNE Physics Handbook*, second edition, L. Maiani, G. Pancheri and N. Paver, eds. (Frascati, 1995), Chiral perturbation theory.
- [41] V. Bernard, N. Kaiser and Ulf-G. Meißner, *Int. J. Mod. Phys. E* 4 (1995) 193.
- [42] H. Leutwyler, in: *Proceedings of the Workshop Hadrons 1994*, Gramado, RS, Brasil (1994), Principles of chiral perturbation theory.
- [43] S. Weinberg, *Physica A* 96 (1979) 327.
- [44] H. Georgi, *Weak interactions* (Benjamin Cummings, Menlo Park, CA, 1984).
- [45] H. Leutwyler, *Ann. Phys.* 235 (1994) 165.
- [46] E. Jenkins and A.V. Manohar, in: *Proceedings of the Workshop on Effective Field Theories of the Standard Model* Dobogókő, Hungary, Aug. 1991, U.-G. Meißner, ed. (World Scientific, Singapore, 1992).
- [47] J. Gasser and H. Leutwyler, *Nucl. Phys. B* 250 (1985) 465.
- [48] T. Barnes and E.S. Swanson, *Phys. Rev. C* 49 (1994) 1166.
- [49] O. Dumbrăjs, et al., *Nucl. Phys. B* 216 (1983) 277;  
A. Müller-Groeling, K. Holinde and J. Speth, *Nucl. Phys. A* 513 (1990) 557.
- [50] P.B. Siegel and W. Weise, *Phys. Rev. C* 38 (1988) 2221.
- [51] M.J. Savage, *Phys. Lett. B* 331 (1994) 411.
- [52] M.N. Butler, M.J. Savage and R.P. Springer, *Nucl. Phys. B* 399 (1993) 69.
- [53] C.L. Schat, N.N. Scoccola and C. Gobbi, *Nucl. Phys. A* 585 (1995) 627.
- [54] V. Bernard, N. Kaiser and U. -G. Meißner, *Phys. Lett. B* 309 (1993) 421.
- [55] S.L. Adler and R.F. Dashen, *Current Algebra and Applications to Particle Physics* (Benjamin, New York, 1968).
- [56] M. Lutz, A. Steiner and W. Weise, *Phys. Lett. B* 278 (1992) 29.
- [57] C. Callan, S. Coleman, J. Wess and B. Zumino, *Phys. Rev.* 177 (1969) 2247.
- [58] J. Gasser, M.E. Sainio and A. Svarc, *Nucl. Phys. B* 307 (1988) 779.
- [59] V. Thorsson and A. Wirzba, *Nucl. Phys. A* 589 (1995) 633.
- [60] S.-J. Dong and K.-F. Liu, in: *Proceedings of Lattice'94*, September 27-30, 1994, Bielefeld, *Nucl. Phys. B* (Supp.), to be published. (1995),  $\pi NN\sigma$  term and quark spin content of the nucleon.
- [61] M. Fukugita, Y. Kuramashi, M. Okawa and A. Ukawa, *Phys. Rev. D* 51 (1995) 5319.
- [62] N. Kaiser, P.B. Siegel and W. Weise, TUM/T39-95-5 (nucl-th/9505043), Chiral dynamics and the low energy kaon-nucleon interactions.
- [63] R.J. Nowak et al., *Nucl. Phys. B* 139 (1978) 61;  
D.J. Miller, R.J. Nowak and T. Tyminiecka, in: *Proceedings of the Workshop on Low and Intermediate Energy Kaon-Nucleon Physics*, Rome, March 24-28, 1980, E. Ferrari and G. Violini, eds. (Reidel, Dordrecht, 1981) p. 251.
- [64] W. Cameron et al., *Nucl. Phys. B* 78 (1974) 93.
- [65] R.A. Burnstein et al., *Phys. Rev. D* 10 (1974) 2767.
- [66] U. -G. Meißner, E. Oset and A. Pich, *Phys. Lett. B* 353 (1995) 161.
- [67] M. Prakash, T.L. Ainsworth and J.M. Lattimer, *Phys. Rev. Lett.* 61 (1988) 2518.
- [68] M. Rho, *Phys. Rep.* 240 (1994) 1.

- [69] H. Jung, in: Proceedings of the First International Symposium on Symmetries in Subatomic Physics, May 16-18, 1994, Taipei, Taiwan, R.O.C. (1995), Effective kaon mass in nuclear matter from chiral Lagrangians.
- [70] J. Schaffner, A. Gal, I.N. Mishustin, H. Stöcker and W. Greiner, Phys. Lett. B 344 (1994) 268.
- [71] B.D. Serot and J.D. Walecka, Adv. Nucl. Phys. 16 (1986) 1; B.D. Serot, Rep. Prog. Phys. 55 (1992) 1855.
- [72] A.D. Martin, Nucl. Phys. B 179 (1981) 33.
- [73] N.S. Manton, Comm. Math. Phys. 111 (1987) 469.
- [74] D.J. Helfand and R.H. Becker, Nature 307 (1984) 215.
- [75] S.E. Woosley, P.A. Pinto, P.G. Martin and T.A. Weaver, Astrophys. J. 318 (1987) 664.
- [76] P. Murdin, End in Fire : The supernova in the Large Magellanic Cloud (Cambridge, New York, 1990).
- [77] G.E. Brown, Nucl. Phys. A 574 (1994) 217c.
- [78] A. Maeder, Astron. Astrophys. 264 (1992) 105.
- [79] B.E.J. Pagel, E.A. Simonson, E.A. Terlevich and M.G. Edmunds, Mon. Not. R. Astron. Soc. 255 (1992) 325.
- [80] Kamales Kar, in: Proceedings of the School and Workshop on Supernovae and Stellar Evolution, Goa, India, March 10-17, 1989, A. Ray and T. Velusamy, eds. (World Scientific, Singapore, 1991) p. 222, Neutrinos from SN1987A - Observations and Expectations.
- [81] G.E. Brown and J.C. Weingartner, Astrophys. J. 436 (1994) 843.
- [82] H.A. Bethe, Astrophys. J. 419 (1993) 197.
- [83] S. Van den Bergh, R.D. McClure and R. Evans, Astrophys. J. 323 (1987) 44.

This page is intentionally left blank

---

## Appendix B

# Kaon Production in Heavy-Ion Collisions and Maximum Mass of Neutron Stars

G.Q. Li, C.-H. Lee and G.E. Brown

Reprinted with permission from *Physical Review Letters*, 79 (1997) 5214–5217

Copyright (1997) by the American Physical Society

This page is intentionally left blank

## Kaon Production in Heavy-Ion Collisions and Maximum Mass of Neutron Stars

G. Q. Li, C.-H. Lee, and G. E. Brown

*Department of Physics, State University of New York at Stony Brook, Stony Brook, New York 11794*

(Received 21 April 1997)

We determine an "empirical" kaon dispersion relation by analyzing and fitting recent experimental data on kaon production in heavy-ion collisions. We then investigate its effects on the hadronic equation of state at high densities and on neutron star properties. We find that the maximum mass of neutron stars can be lowered by about  $0.4M_{\odot}$  once kaon condensation as constrained by our empirical dispersion relation is introduced. We emphasize the growing interplay between hadron physics, relativistic heavy-ion physics, and the physics of compact objects in astrophysics. [S0031-9007(97)04917-X]

PACS numbers: 25.75.Dw, 24.10.Lx, 26.60.+c, 97.60.Jd

There is currently growing interplay between the physics of hadrons (especially the properties of hadrons in dense matter which might reflect spontaneous chiral symmetry breaking and its restoration), the physics of relativistic heavy-ion collisions (from which one might extract hadron properties in dense matter), and the physics of compact objects in astrophysics (which needs as inputs the information gained from the first two fields). A notable example is the kaon ( $K$  and  $\bar{K}$ ), which, being a Goldstone boson with strangeness, plays a special role in all of the three fields mentioned.

Ever since the pioneering work of Kaplan and Nelson [1] on the possibility of kaon condensation in nuclear matter, much theoretical effort has been devoted to the study of kaon properties in dense matter. Brown *et al.* [2] have carried out a detailed study of free-space and in-medium kaon-nucleon scattering using chiral perturbation theory. Yuba *et al.* [3] studied kaon in-medium properties based on phenomenological off-shell meson-nucleon interactions. Weise and collaborators [4] investigated this problem using the Nambu-Jona-Lasinio model, treating the kaon as a quark-antiquark excitation. Recently, they have extended the chiral perturbation calculation to include the coupled-channel effects which are important for the  $K^-$  meson [5]. Another type of study, which is based on the extension of the Walecka mean-field model from SU(2) to SU(3), was pursued by Schaffner *et al.* [6] and Knorren *et al.* [7]. Although, quantitatively, the results from these different models are not identical, qualitatively, a consistent picture has emerged; namely, in nuclear matter the  $K^+$  feels a weak repulsive potential, whereas the  $K^-$  feels a strong attractive potential.

Measurements of kaon spectra and flow have been carried out in heavy-ion collisions at the energies of the Darmstadt Schwerionen Synchrotron (SIS) [(1–2) $A$  GeV], the Brookhaven Alternating Gradient Synchrotron (AGS) (10 $A$  GeV), and the CERN Super Proton Synchrotron (SPS) (2004 GeV) [8]. By comparing transport model predictions with experimental data, one can study not only the global reaction dynamics but, more importantly, the kaon properties in dense matter. Of special interest is kaon production in heavy-ion collisions at SIS energies, as it

has been shown that particle production at subthreshold energies is sensitive to its properties in dense matter [9]. Recently, high quality data concerning  $K^+$  and  $K^-$  production in heavy-ion collisions at SIS energies have been published by the KaoS collaboration at Gesellschaft für Schwerionenforschung (GSI) [10]. The KaoS data show that the  $K^-$  yield at 1.84 GeV (projectile nucleus kinetic energy in the laboratory frame) agrees roughly with the  $K^+$  yield at 1.04 GeV. This is a nontrivial observation. These beam energies were purposely chosen such that the  $Q$  values for  $NN \rightarrow NK\bar{A}$  and  $NN \rightarrow NNK\bar{K}$  are both about –230 MeV. Near their respective production thresholds, the cross section for the  $K^-$  production in proton-proton interactions is 1–2 orders of magnitude smaller than that for  $K^+$  production [11]. In addition, antikaons are strongly absorbed in heavy-ion collisions, which should further reduce the  $K^-$  yield. The KaoS results of  $K^-/K^+ \sim 1$  indicate thus the importance of kaon medium effects which act oppositely on  $K^+$  and  $K^-$  production in nuclear medium.

Studies of neutron star properties also have a long history. A recent compilation by Thorsett, quoted by Brown [12], shows that well-measured neutron star masses are all less than  $1.5M_{\odot}$ . On the other hand, most of the theoretical calculations based on conventional nuclear equations of state (EOS) predict a maximum neutron star mass above  $2M_{\odot}$ . The EOS can, therefore, be substantially softened without running into contradiction with observation. Various scenarios have been proposed that can lead to a soft EOS, including the high-order self-interactions of the vector field [13], the possibility of kaon condensation [14], the existence of hyperons [7,15], and the transition to quark matter [16]. All of these possibilities need to be examined against the available empirical information from, e.g., relativistic heavy-ion collisions.

The chief aim of this paper is to determine, from the recent KaoS data on kaon production, together with the previous analysis of nucleon flow, kaon flow and dilepton spectra [17–20], an "empirical" kaon dispersion relation in dense matter. We will show that these data are consistent with the scenario that the  $K^+$  feels a weak repulsive potential and the  $K^-$  a strong attractive potential, as predicted by the chiral perturbation calculation. We then

study the effects of this empirical dispersion relation on the possibility of kaon condensation and on the neutron star properties. We find that  $K^-$  condensation happens at about  $3\rho_0$ , and the maximum mass of neutron stars is lowered by about  $0.4M_\odot$  once the kaon condensation is introduced. These values change by about 20% when different nuclear equations of state are used (see Ref. [21] for a detailed discussion).

We use the relativistic transport model for the description of heavy-ion collisions and for the calculation of kaon production [9]. The nucleon dynamics is governed by the

$$\begin{aligned} \epsilon_N = & \frac{2}{(2\pi)^3} \int_0^{K_F} dk \sqrt{k^2 + m_N^{*2}} + \frac{2}{(2\pi)^3} \int_0^{K_F} dk \sqrt{k^2 + m_N^{*2}} + W\rho + R \frac{1}{2} (\rho_p - \rho_n) - \frac{1}{2C_V^2} W^2 \\ & - \frac{1}{2C_\rho^2} R^2 + \frac{1}{2C_S^2} \Phi^2 + \frac{S'^2}{4C_S^2} d^2 \left[ \left( 1 - \frac{\Phi}{S'} \right)^{4/d} \left[ \frac{1}{d} \ln \left( 1 - \frac{\Phi}{S'} \right) - \frac{1}{4} \right] + \frac{1}{4} \right] - \frac{\xi}{24} W^4 - \frac{\eta}{2C_V^2 S'} \Phi W^2. \quad (1) \end{aligned}$$

The nucleon effective mass  $m_N^*$  is related to its scalar field  $\Phi$  by  $m_N^* = m_N - \Phi$ .  $W$  and  $R$  are the isospin-even and isospin-odd vector potentials, respectively. The last three terms give the self-interactions of the scalar field, the vector field, and the coupling between them. The meaning and values of various parameters in Eq. (1) can be found in [22]. In this paper, we use the parameter set T1 listed in Table 1 of [22].

From Eq. (1), we can derive a relativistic transport model for heavy-ion collisions. At SIS energies, the colliding system consists mainly of nucleons, delta resonances, and pions. While medium effects on pions are neglected, nucleons and delta resonances propagate in a common mean-field potential according to the Hamilton equation of motion,

$$\frac{dx}{dt} = \frac{p^*}{E^*}, \quad \frac{dp}{dt} = -\nabla_x(E^* + W), \quad (2)$$

where  $E^* = \sqrt{p^{*2} + m^*}$ . These particles also undergo stochastic two-body collisions [9]. In Fig. 1 we compare our results for proton and pion transverse mass spectra in central Ni + Ni collisions with experimental data from the FOPI collaboration [23]. The nice agreement with the data provides further support to the use of the chiral Lagrangian of [22] in the present analysis.

In heavy-ion collisions at SIS energies, kaons can be produced from pion-baryon and baryon-baryon collisions. For the former, we use cross sections obtained in the resonance model by Tsushima *et al.* [24]. For the latter, the cross sections obtained in the one-boson-exchange model of Ref. [25] are used. For antikaon production from pion-baryon collisions we use the parametrization proposed by Sibirtsev *et al.* [26]. For baryon-baryon collisions, we use a somewhat different parametrization, which describes the experimental data better, than used in Ref. [26]. In addition, the antikaon can also be produced from strangeness-exchange processes such as  $\pi Y \rightarrow \bar{K}N$ , where  $Y$  is a hyperon. The cross sections for these processes are obtained from the inverse ones,  $\bar{K}N \rightarrow \pi Y$ ,

chiral Lagrangian recently developed by Furnstahl, Tang, and Serot [22], which provides a very good description of nuclear matter and finite nuclei. Furthermore, a recent analysis [20] showed that this model reproduces nicely the nucleon flow [17] and dilepton spectra [19] in heavy-ion collisions, indicating that its extrapolation to  $(2-3)\rho_0$  is consistent with empirical information. We will thus use this model as our basis for the determination of the kaon dispersion relation and for the analysis of neutron star properties.

In the mean-field approximation, the energy density for the general case of asymmetric nuclear matter is given by

by the detailed-balance relation. The latter cross sections, together with the  $\bar{K}N$  elastic and absorption cross sections, are parametrized based on the available experimental data [11]. The details about elementary cross sections, the transport model, and the neutron star calculation will be given elsewhere [27].

From the chiral Lagrangian the kaon and antikaon in-medium energies can be written as [2]

$$\omega_K = [m_K^2 + k^2 - a_K \rho_S + (b_K \rho)^2]^{1/2} + b_K \rho, \quad (3)$$

$$\omega_{\bar{K}} = [m_{\bar{K}}^2 + k^2 - a_{\bar{K}} \rho_S + (b_{\bar{K}} \rho)^2]^{1/2} - b_{\bar{K}} \rho, \quad (4)$$

where  $b_K = 3/8f_\pi^2 \approx 0.333$  GeV fm<sup>3</sup>, and  $a_K$  and  $a_{\bar{K}}$  are two parameters that determine the strength of the attractive scalar potential for kaon and antikaon, respectively. If one considers only the Kaplan-Nelson term, then

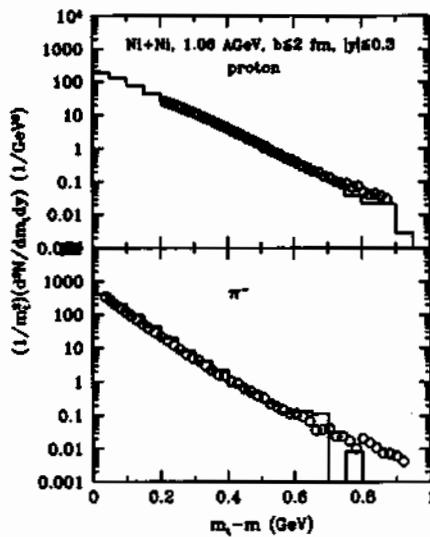


FIG. 1. Proton and  $\pi^-$  transverse mass spectra in central Ni + Ni collisions at 1.06 A GeV.

$a_K = a_{\bar{K}} = \Sigma_{KN}/f_\pi^2$ . In the same order, there is also the range term which acts differently on kaon and antikaon, and leads to different scalar attractions. Since the exact value of  $\Sigma_{KN}$  and the size of the higher-order corrections are still under intensive debate, we take the point of view that  $a_{K,\bar{K}}$  can be treated as free parameters and try to constrain them from the experimental observables. Since the  $KN$  interaction is relatively weak, impulse approximation should be reasonable at low densities. This provides some constraints on  $a_K$ . We find that  $a_K \approx 0.22 \text{ GeV}^2 \text{ fm}^3$ , corresponding to  $\Sigma_{KN} \approx 400 \text{ MeV}$ , gives a repulsive  $K^+$  potential of about 20 MeV at normal nuclear matter density, slightly smaller than the 25 MeV found in [28]. We will show later that this value also gives a good fit to the  $K^+$  spectra in heavy-ion collisions.

From the chiral Lagrangian we can also derive equations of motion for kaons [29],

$$\frac{dx}{dt} = \frac{\mathbf{p}^*}{\omega_{K,K} \mp b_K \rho}, \quad \frac{dp}{dt} = -\nabla_x \omega_{K,K}. \quad (5)$$

The minus (plus) sign corresponds to kaon (antikaon).

For  $K^+$  and  $K^-$  production in heavy-ion collisions, we consider two scenarios; namely, with and without kaon medium effects. As mentioned, we use  $a_K \approx 0.22 \text{ GeV}^2 \text{ fm}^3$  for  $K^+$ . For  $K^-$ , we adjust  $a_K$  such that we achieve a good fit to the experimental  $K^-$  spectra. We find  $a_K \approx 0.45 \text{ GeV}^2 \text{ fm}^3$ , which leads to a  $K^-$  potential of about  $-110 \text{ MeV}$  at normal nuclear matter density. This is somewhat smaller, in magnitude, than the "best" value of  $-200 \pm 20 \text{ MeV}$  extracted from kaonic atoms [30]. The latter value, however, depends sensitively on the extrapolation procedure from the surface of nuclei to their interiors [30]. On the other hand, kaon production in heavy-ion collisions and neutron star calculations are sensitive chiefly to higher densities  $[(2-3)\rho_0]$ , where our value should be more relevant.

The results for the  $K^+$  and  $K^-$  kinetic energy spectra are shown Fig. 2. The solid and dotted histograms give the results with and without kaon medium effects, respectively. The open circles are the experimental data from the KaoS collaboration [10]. For the  $K^+$ , it is seen that the results with kaon medium effects are in good agreement with the data, while those without kaon medium effects slightly overestimate the data. We note that the kaon feels a slightly repulsive potential; thus the inclusion of the kaon medium effects reduces the kaon yield. For the  $K^-$ , it is seen that without medium effects our results are about a factor 3–4 below the experimental data. With the inclusion of the medium effects, which reduces the antikaon production threshold, the  $K^-$  yield increases by about a factor of 3, and our results are in good agreement with the data. This is similar to the findings of Cassing *et al.* [31]. For both  $K^+$  and  $K^-$ , the differences between the two scenarios are most pronounced at low kinetic energies. The experimental data at these momenta will be very useful in discriminating the two scenarios.

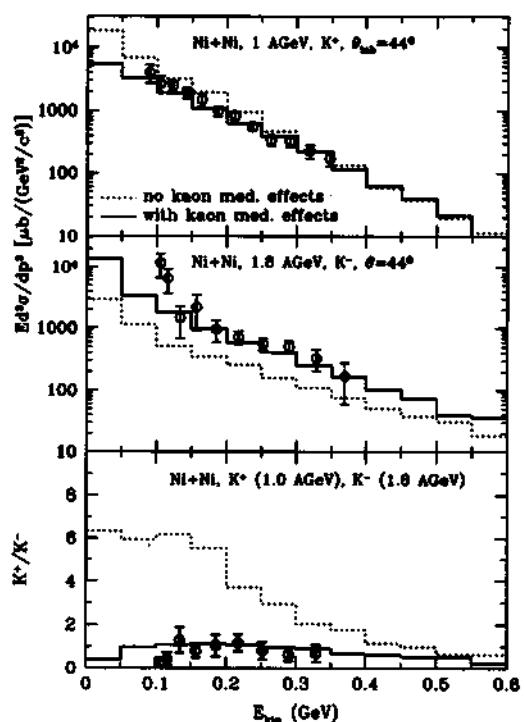


FIG. 2.  $K^+$  (upper window),  $K^-$  (middle window), and  $K^+/K^-$  (lower window) kinetic energy spectra in Ni + Ni collisions.

The effects of kaon and antikaon mean-field potentials can be more clearly seen by looking at their ratio as a function of the kinetic energy, which is shown in the lower window of Fig. 2. Without kaon medium effects, the  $K^+/K^-$  ratio decreases from about 7 at low kinetic energies to about 1 at high kinetic energies, which is in complete disagreement with the data. Since the antikaon absorption cross section by nucleons becomes large at low momentum, low-momentum antikaons are more strongly absorbed than high-momentum ones. This makes the  $K^+/K^-$  ratio increase with decreasing kinetic energies. When medium effects are included, we find that the  $K^+/K^-$  ratio is almost unity in the entire kinetic energy region, which agrees very well with the data. The shapes of the  $K^+$  and  $K^-$  spectra change in opposite ways in the presence of their mean-field potential. Kaons are "pushed" to high momenta by the repulsive potential, while antikaons are "pulled" to low momenta. The good description of the  $K^+/K^-$  ratio, together with the fit to the kaon flow in heavy-ion collisions [18,29], gives us confidence that our in-medium kaon and antikaon dispersion relations are reasonable.

We take the antikaon dispersion relation constrained by the heavy-ion data as empirical indication of an attractive antikaon potential in dense matter. We combine this with the energy density of Eq. (1) for nuclear matter,

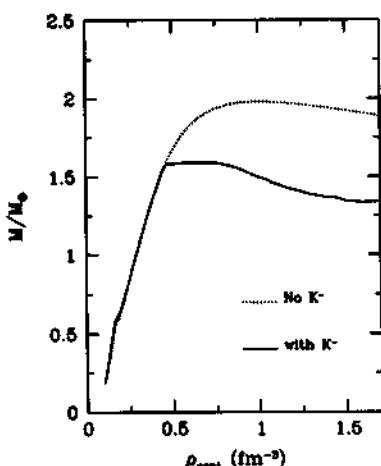


FIG. 3. Neutron star mass as a function of central density. The solid and dotted lines are obtained with and without  $K^-$  condensation.

and calculate neutron star properties. We find that, at about  $3\rho_0$ , the effective mass of the  $K^-$  drops below the chemical potential of the electron, indicating the onset of kaon condensation. After this the  $K^-$  density increases rapidly, leading to a large proton fraction in the neutron star. The results for neutron star mass as a function of central density are shown in Fig. 3, where the solid and dotted curves give the results with and without  $K^-$ , respectively. It is seen that the maximum mass of the neutron stars is reduced by about  $0.4M_\odot$  with the introduction of the kaon condensation. Both the critical density for kaon condensation and the amount of lowering in the maximum neutron star mass change by about 20% when different nuclear equations of state are used. The exact value of the maximum neutron star mass, on the other hand, depends more sensitively on the particular nuclear EOS used, as discussed in [13].

It should be mentioned that alternative explanations for lowering the maximum neutron star mass have been proposed. Of particular interest is the introduction of  $\Sigma^-$  hyperons in [15]. This is a complementary rather than competing scenario, since the  $\Sigma^-$ -particle-neutron-hole state has a  $p$ -wave coupling to the  $K^-$ . A unification of the scenarios can be achieved by introducing the "kaesobar," a linear combination of  $K^-$  and  $\Sigma^-$ -particle-neutron-hole [32], but the results of the present work will not be strongly modified.

In summary, we studied  $K^+$  and  $K^-$  production in Ni + Ni collisions at  $(1-2)A$  GeV, based on the relativistic transport model including the strangeness degrees of freedom. We found that the recent experimental data from the KaoS collaboration are consistent with the predictions of the chiral perturbation theory that the  $K^+$  feels a weak repulsive potential and the  $K^-$  feels a strong at-

tractive potential in the nuclear medium. Using the kaon in-medium properties constrained by heavy-ion data, we have studied the possibility of kaon condensation and its effects on neutron star properties. The critical density for kaon condensation was found to be about  $3\rho_0$ , and the maximum mass of neutron stars was found to be reduced by about  $0.4M_\odot$  once kaon condensation is introduced.

We are grateful to C. M. Ko, T. T. S. Kuo, M. Prakash, and M. Rho for useful discussions. We also thank N. Herrmann and P. Senger for useful communications. This work is supported in part by the Department of Energy under Grant No. DE-FG02-88ER40388.

- [1] D. B. Kaplan and A. E. Nelson, Phys. Lett. B **175**, 57 (1986).
- [2] G. E. Brown *et al.*, Nucl. Phys. **A567**, 937 (1994); C.-H. Lee *et al.*, Nucl. Phys. **A585**, 401 (1995).
- [3] H. Yabu *et al.*, Phys. Lett. B **315**, 17 (1993).
- [4] M. Lutz *et al.*, Nucl. Phys. **A574**, 755 (1994).
- [5] N. Kaiser *et al.*, Phys. Lett. B **365**, 12 (1996).
- [6] J. Schaffner *et al.*, Phys. Lett. B **334**, 268 (1994).
- [7] R. Knorren *et al.*, Phys. Rev. C **52**, 3470 (1995).
- [8] Special issue on QM'96, Nucl. Phys. **A610** (1996).
- [9] C. M. Ko and G. Q. Li, J. Phys. G **22**, 1673 (1996).
- [10] R. Barth *et al.*, Phys. Rev. Lett. **78**, 4027 (1997).
- [11] A. Baldini *et al.*, *Total Cross Sections of High Energy Particles* (Springer-Verlag, Heidelberg, 1988).
- [12] G. E. Brown, Proceedings of the Royal Dutch Academy Colloquium on Pulsar Timing, General Relativity and the Internal Structure of Neutron Stars, 1997 (to be published).
- [13] H. Müller and B. D. Serot, Nucl. Phys. **A606**, 508 (1996).
- [14] G. E. Brown *et al.*, Phys. Rev. D **37**, 2042 (1988).
- [15] N. K. Glendenning *et al.*, Phys. Rev. C **45**, 844 (1992).
- [16] N. K. Glendenning and S. Pei, Phys. Rev. C **52**, 2250 (1995).
- [17] M. D. Partland *et al.*, Phys. Rev. Lett. **75**, 2100 (1995).
- [18] J. Ritman *et al.*, Z. Phys. A **352**, 355 (1995).
- [19] G. Agakichiev *et al.*, Phys. Rev. Lett. **75**, 1272 (1995).
- [20] G. Q. Li *et al.*, nucl-th/9702023.
- [21] V. Thorsson *et al.*, Nucl. Phys. **A572**, 693 (1994).
- [22] R. J. Furnstahl *et al.*, Phys. Rev. C **52**, 1368 (1995).
- [23] N. Herrmann, Nucl. Phys. **A610**, 49c (1996).
- [24] K. Tsushima *et al.*, Phys. Lett. B **337**, 245 (1994); J. Phys. G **21**, 33 (1995).
- [25] G. Q. Li and C. M. Ko, Nucl. Phys. **A594**, 439 (1995); G. Q. Li, C. M. Ko, and W. S. Chung (to be published).
- [26] A. Sibirtsev *et al.*, Z. Phys. A **358**, 101 (1997).
- [27] G. Q. Li *et al.*, Nucl. Phys. **A625**, 372 (1997).
- [28] V. Koch, Phys. Lett. B **351**, 29 (1995).
- [29] G. Q. Li, C. M. Ko, and B.-A. Li, Phys. Rev. Lett. **74**, 235 (1995); G. Q. Li and C. M. Ko, Nucl. Phys. **A594**, 460 (1995); Phys. Rev. C **54**, R2159 (1996).
- [30] E. Friedman *et al.*, Nucl. Phys. **A579**, 518 (1994).
- [31] W. Cassing *et al.*, Nucl. Phys. **A614**, 415 (1997).
- [32] G. E. Brown, C.-H. Lee, and R. Rapp (to be published).

### Appendix C

## **$K^-/K^+$ Ratios in Relativistic Heavy-Ion Collisions**

G.Q. Li and G.E. Brown

Reprinted with permission from *Physical Review*, C58 (1998) 1698–1705

Copyright (1998) by the American Physical Society

This page is intentionally left blank

**$K^-/K^+$  ratios in relativistic heavy-ion collisions**

G. Q. Li and G. E. Brown

*Department of Physics and Astronomy, State University of New York at Stony Brook, Stony Brook, New York 11794*

(Received 2 April 1998)

We study  $K^-/K^+$  ratios as a function of centrality (participant nucleon number), transverse mass ( $m_T$ ), and rapidity, in heavy-ion collisions at beam energies between 1 and 2 A GeV. We use the relativistic transport model that includes explicitly the strangeness degrees of freedom and consider two scenarios for kaon properties in dense matter, one with and one without medium modifications of their properties. In both scenarios, the  $K^-/K^+$  ratio does not change very much with the centrality, while the  $K/\pi$  and  $R/\pi$  ratios increase with increasing centrality. Significant differences are predicted, both in magnitudes and shapes, for the  $m_T$  spectra and rapidity distributions of  $K^-/K^+$  ratio. Experimental measurement of these ratios, currently under investigation by the FOPI, KaoS, E866, and E895 Collaborations, will be useful in revealing the kaon in-medium properties. [S0556-2813(98)02809-X]

PACS number(s): 25.75.Dw, 26.60.+c, 24.10.Lx

**I. INTRODUCTION**

Whether and how hadronic properties, such as their masses, widths, and dispersion relations, are modified in hot and dense medium is a topic of great current interest. Of particular importance are the medium modifications of kaon properties, as they are related to both spontaneous and explicit chiral symmetry breaking, and they are useful inputs for the study of kaon condensation and neutron star properties [1,2]. Since the pioneering work of Kaplan and Nelson [3] on the possibility of kaon condensation in nuclear matter, a huge amount of theoretical effort has been devoted to the study of kaon properties in dense matter, using such diversified approaches as chiral perturbation theory [4–12], the Nambu-Jona-Lasinio model [13], and SU(3) Walecka-type mean-field model [14,15]. Although quantitatively results from these different models are not identical, qualitatively, a consistent picture has emerged; namely, in nuclear matter the  $K^+$  feels a weak repulsive potential, whereas the  $K^-$  feels a strong attractive potential.

Experimentally, in-medium properties of kaon and antikaon can be obtained from the analysis of kaon-nucleus scattering [16,17] and kaonic atom data [18]. The information so obtained is, unfortunately, restricted to low densities. For the study of kaon condensation, densities much higher than that accessible by kaonic atoms are involved. This can only be obtained by analyzing heavy-ion collision data on various observables involving kaon and antikaon. Measurements of kaon spectra and flow have been systematically carried out in heavy-ion collisions at SIS (1–2 A GeV), AGS (10 A GeV), and SPS (200 A GeV) energies [19]. The analysis of their yields, spectra, and in particular collective flow has indeed provided useful information about kaon properties in dense nuclear matter [20–28]. So far most of the experimental data from the FOPI [29–31] and KaoS [32–35] Collaborations at SIS/GSI seem to be consistent with the predictions from the chiral perturbation theory.

To put these conclusions on a firmer footing, additional observables and experimental data from independent Collaborations will certainly be useful. In this paper we study the centrality, transverse mass ( $m_T$ ), and rapidity dependence of the  $K^-/K^+$  ratio in heavy-ion collisions. Since the me-

dium effects act oppositely on kaon and antikaon, their ratio can reflect more precisely these effects, as was pointed out in Refs. [36,37]. We will consider Ni+Ni, Ru+Ru, and Au+Au collisions at beam energies between 1 and 2 A GeV. These systems are being investigated by the FOPI [38] and KaoS [39] Collaborations at SIS/GSI. Furthermore, the Au+Au collisions are also being analyzed by the E866 [36,37] and E895 [40,41] Collaborations at the AGS/BNL.

This paper is arranged as follows. In Sec. II, we briefly review the relativistic transport model, kaon in-medium properties, and elementary kaon production cross sections. The results are presented in Sec. III. The paper ends with a short summary in Sec. IV.

**II. THE RELATIVISTIC TRANSPORT MODEL AND KAON PRODUCTION**

Heavy-ion collisions involve very complicated nonequilibrium dynamics. One needs to use transport models in order to extract from experimental data the information about in-medium properties of hadrons. In this work we will use the relativistic transport model similar to that developed in Ref. [42]. Instead of the usual linear and nonlinear  $\sigma\text{-}\omega$  models, we base our model on the effective chiral Lagrangian recently developed by Furstahl, Tang, and Serot [43], which is derived using dimensional analysis, naturalness arguments, and provides a very good description of nuclear matter and finite nuclei. In the mean-field approximation, the energy density for the general case of asymmetric nuclear matter is given by

$$\begin{aligned} e_N = & \frac{2}{(2\pi)^3} \int_0^{K_F} dk \sqrt{k^2 + m_N^{*2}} + \frac{2}{(2\pi)^3} \int_0^{K_{F'}} dk \sqrt{k^2 + m_N^{*2}} \\ & + W\rho + R \frac{1}{2}(\rho_p - \rho_n) - \frac{1}{2C_V^2} W^2 - \frac{1}{2C_p^2} R^2 + \frac{1}{2C_S^2} \Phi^2 \\ & + \frac{S'^2}{4C_S^2} d^2 \left[ \left( 1 - \frac{\Phi}{S'} \right)^{4d} \left[ \frac{1}{d} \ln \left( 1 - \frac{\Phi}{S'} \right) - \frac{1}{4} \right] + \frac{1}{4} \right] \\ & - \frac{\xi}{24} W^4 - \frac{\eta}{2C_V^2} \frac{\Phi}{S'} W^2. \end{aligned} \quad (1)$$

The nucleon effective mass  $m_N^*$  is related to its scalar field  $\Phi$  by  $m_N^* = m_N - \Phi$ .  $W$  and  $R$  are the isospin-even and isospin-odd vector potentials, respectively. The last three terms give the self-interactions of the scalar field, the vector field, and the coupling between them. The meaning and values of various parameters in Eq. (1) can be found in Ref. [43].

From the energy density of Eq. (1), we can also derive a relativistic transport model for heavy-ion collisions. At SIS energies, the colliding system consists mainly of nucleons,  $\Delta$  resonances, and pions. While medium effects on pions are neglected, nucleons and  $\Delta$  resonances propagate in a common mean-field potential according to the Hamilton equation of motion

$$\frac{dx}{dt} = \frac{p^*}{E^*}, \quad \frac{dp}{dt} = -\nabla_x(E^* + W), \quad (2)$$

where  $E^* = \sqrt{p^{*2} + m^{*2}}$ . These particles also undergo stochastic two-body collisions, including both elastic and inelastic scattering.

In heavy-ion collisions at incident energies considered in this work, kaons can be produced from pion-baryon and baryon-baryon collisions. For the former we use cross sections obtained in the resonance model by Tsushima *et al.* [44]. For the latter the cross sections obtained in the one-boson-exchange model of Refs. [45,46] are used. Both models describe the available experimental data very well. For antikaon production from pion-baryon collisions we use the parametrization proposed by Sibirtsev *et al.* [47]. For baryon-baryon collisions, we use a somewhat different parametrization, which describes the experimental data better than Ref. [47]. In addition, the antikaon can also be produced from strangeness-exchange processes such as  $\pi Y \rightarrow \bar{K}N$  where  $Y$  is either a  $\Lambda$  or  $\Sigma$  hyperon. The cross sections for these processes are obtained from the reverse ones,  $\bar{K}N \rightarrow \pi Y$ , by the detailed-balance relation. All the parametrizations for the elementary cross sections and comparisons with experimental data can be found in our recent paper [28].

Particles produced in elementary hadron-hadron interactions in heavy-ion collisions cannot escape the environment freely. Instead, they are subjected to strong final-state interactions. For the kaon, because of strangeness conservation, its scattering with nucleons at low energies is dominated by elastic and pion production processes, which do not affect its final yield but changes its momentum spectra. The final-state interaction for the antikaon is much stronger. Antikaons can be destroyed in the strangeness-exchange processes. They also undergo elastic scattering. Both the elastic and absorption cross sections increase rapidly with decreasing antikaon momenta. This will have strong effects on the final  $K^-$  momentum spectra in heavy-ion collisions.

We will consider two scenarios for kaon properties in nuclear medium, one with and one without medium modification. From the chiral Lagrangian the kaon and antikaon in-medium energies can be written as [28]

$$\omega_K = [m_K^2 + k^2 - a_K \rho_S + (b_K \rho)^2]^{1/2} + b_K \rho, \quad (3)$$

$$\omega_{\bar{K}} = [m_K^2 + k^2 - a_{\bar{K}} \rho_S + (b_{\bar{K}} \rho)^2]^{1/2} - b_{\bar{K}} \rho, \quad (4)$$

where  $b_K = 3/(8f_\pi^2) \sim 0.333$  GeV fm<sup>3</sup>,  $a_K$  and  $a_{\bar{K}}$  are two parameters that determine the strength of the attractive scalar potential for kaon and antikaon, respectively. If one considers only the Kaplan-Nelson term, then  $a_K = a_{\bar{K}} = \Sigma_{KN}/f_\pi^2$ . In the same order, there is also the range term which acts differently on kaon and antikaon, and leads to different scalar attractions. Since the exact value of  $\Sigma_{KN}$  and the size of the higher-order corrections are still under intensive debate, we take the point of view that  $a_{K,\bar{K}}$  can be treated as free parameters and try to constrain them from the experimental observables in heavy-ion collisions. In Refs. [27,28] we found that  $a_K \sim 0.22$  GeV<sup>2</sup> fm<sup>3</sup> and  $a_{\bar{K}} \sim 0.45$  GeV<sup>2</sup> fm<sup>3</sup> provide a good description of kaon and antikaon spectra in Ni+Ni collisions at 1 and 1.8 A GeV. These values will be used in this work as well.

In nuclear medium, kaon and antikaon masses are modified, as are their production thresholds. We will then use  $\sqrt{s_0^*}$ , which are calculated with effective masses, in evaluating the in-medium production cross sections. This amounts to the change of threshold, or approximately, to the change of available phase space. In addition to the change in the production cross sections, the medium effects on kaon and antikaon also affect their momentum spectra, when they propagate in the mean-field potentials. The Hamilton equations of motion for kaon and antikaon are very similar to those for nucleons [21],

$$\frac{dr}{dt} = \frac{k}{\omega_{K,\bar{K}} \mp b_K \rho_N}, \quad \frac{dk}{dt} = -\nabla_x U_{K,\bar{K}}, \quad (5)$$

where the minus sign corresponds to kaon and the plus sign to antikaon. It is clearly that the  $K^+$  momentum increases and that of  $K^-$  decreases when they propagate in their respective mean field potentials. This affects significantly their momentum spectra, and especially the momentum spectra of their ratio.

### III. RESULTS AND DISCUSSIONS

In this section we present our results for the centrality, transverse mass, and rapidity dependences of  $K^-/K^+$  ratio. Before that, we first compare the proton and pion  $m$  spectra from our calculation with the available experimental data, taking central Ni+Ni collisions at 1.93A GeV as an example. In Fig. 1 we compare the proton transverse mass spectra in three rapidity bins with the experimental data from the FOPI Collaboration [48,49]. A similar comparison for the  $\pi^-$  transverse mass spectra is shown in Fig. 2. Our results are seen to be in good agreement with the data [48,49]. Similarly, in Ref. [28] the proton and  $\pi^-$  rapidity distributions obtained in our transport model were shown to be in good agreement with the FOPI data.

#### A. Centrality dependence

In this subsection, we discuss the centrality dependence of the  $K^-/K^+$  ratio. We use the participant nucleon number  $A_{\text{part}}$  as the measurement of the centrality. In Fig. 3 we show the  $A_{\text{part}}$  dependence of the pion multiplicity in Ni+Ni collisions at 1.93A GeV. The solid line gives 1/3 of the total pion number obtained in our calculation, while the circles

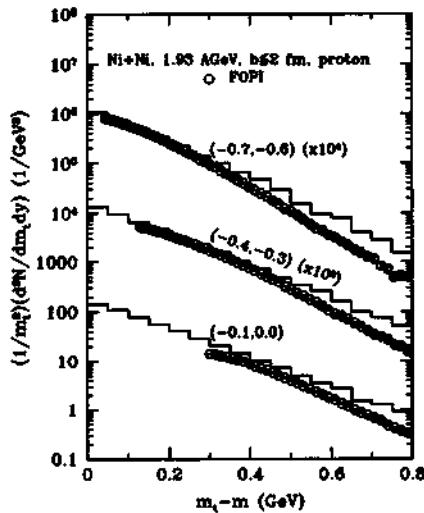


FIG. 1. Proton transverse mass spectra in central Ni+Ni collisions at 1.934 GeV in three rapidity bins. The open circles are experimental data from the FOPI Collaboration [48,49].

and squares are the FOPI data [50] for  $\pi^-$  and  $\pi^+$ , respectively. It is seen that our results are in good agreement with the data, and that the pion multiplicity increases almost linearly with the participant nucleon number. In other words, the  $N_\pi/A_{\text{part}}$  ratio is almost independent of the centrality.

On the other hand, it is well known that the  $K^+$  (usually shown in terms of  $K^+/\pi^+$  ratio) yield increases more than linearly with the  $A_{\text{part}}$  for beam energies ranging from 1A GeV [32–34,46] to 10A GeV [36,51]. Our results for  $K/\pi$  and  $\bar{K}/\pi$  ratios in Au+Au collisions are shown in Fig. 4. It is seen that either with or without kaon medium effects, these ratios increase more than linearly with the participant nucleon number. This is due to the increasing importance of

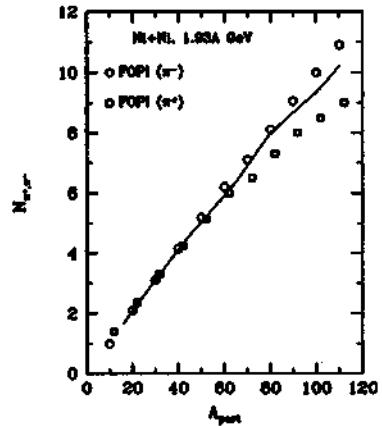


FIG. 3. Pion multiplicity as a function of the participant nucleon number in Ni+Ni collisions at 1.934 GeV.

the secondary processes involving baryon-resonances, pions, and hyperons in the case of antikaon production, when going from peripheral to central collisions. The E866 data for Au+Au collisions at 11.6A GeV/c indicate that  $K^-/\pi^+$  ratio also increases with increasing centrality [51].

The centrality dependence of the  $K^-/K^+$  ratios is shown in Figs. 5, 6, and 7, for Ni+Ni at 1.8A GeV, Ru+Ru at 1.69A GeV, and Au+Au at 2A GeV, respectively. For Ni+Ni collisions, we also show in Fig. 5 the experimental data from the KaoS Collaboration [35] by open circles. For all the systems considered, it is seen that either with or without medium effects, the  $K^-/K^+$  ratio does not depend very much on the centrality, since in both scenarios, the  $K/\pi$  and  $\bar{K}/\pi$  ratios increase at about the same rate with increasing centrality (see Fig. 4). The experimental data of the KaoS Collaboration for Ni+Ni collisions (Fig. 5) show little increase of the ratio towards central collisions within their sta-

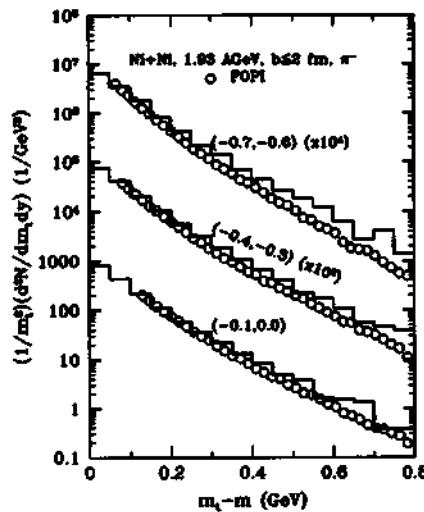


FIG. 2. Same as Fig. 1, for  $\pi^-$ .

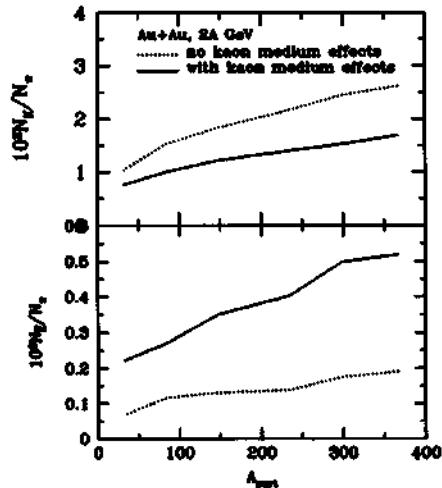


FIG. 4.  $K/\pi$  and  $\bar{K}/\pi$  ratios as a function of the participant nucleon number in Au+Au collisions at 2A GeV.

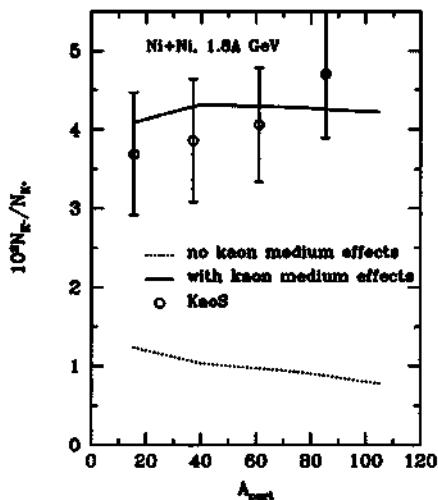


FIG. 5.  $K^-/K^+$  ratio as a function of the participant nucleon number in Ni+Ni collisions at 1.8A GeV. The open circles are the experimental data from the KaoS Collaboration [35].

tistical uncertainties. Preliminary data from the FOPI Collaboration for Ru+Ru at 1.69A GeV [52], and data from E866 Collaboration for Au+Au collisions at 10.6A GeV [36] both show weak centrality dependence for the  $K^-/K^+$  ratio.

Without kaon medium effects, the  $K^-/K^+$  ratio ranges from 0.006 in Ru+Ru at 1.69A GeV, to 0.01 in Ni+Ni collisions at 1.8A GeV, to 0.0075 in Au+Au collisions at 2A GeV. The increase of this ratio from Ru+Ru to Ni+Ni comes from the increase in beam energy, since at these energies the antikaon excitation function is steeper than that of kaon because of a higher threshold. On the other, the decrease of this ratio from Ni+Ni to Au+Au is due to the fact that in the large Au+Au system the antikaon absorption ef-

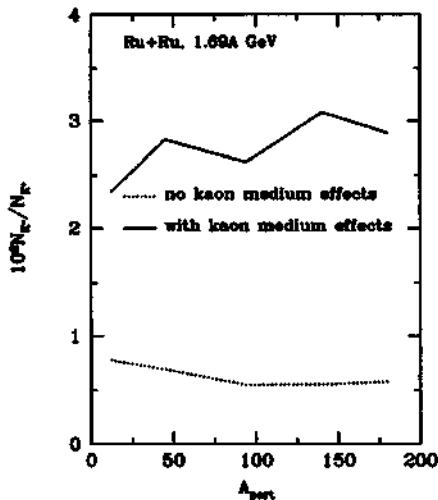


FIG. 6. Same as Fig. 5, for Ru+Ru collisions at 1.69A GeV.

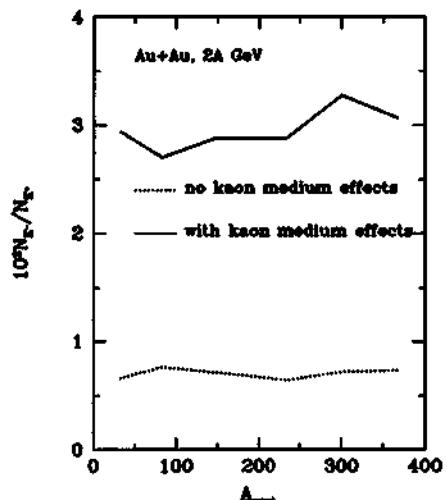


FIG. 7. Same as Fig. 5, for Au+Au collisions at 2A GeV.

fects become more significant. Qualitatively, the absorption probability is proportional to the product of the absorption cross section and the average size of the system. In this case the experimental data from the KaoS Collaboration for Ni+Ni collisions are significantly underestimated (see Fig. 5).

When kaon medium effects are included, the  $K^-/K^+$  ratio increase by about a factor of 4, for all the systems considered here. This increase results from a factor of 3 increase in antikaon yield and about 35% reduction in kaon yield. In this case, the  $K^-/K^+$  ratio from the KaoS Collaboration can be nicely explained. Furthermore, preliminary data from the FOPI Collaboration indicate a  $K^-/K^+$  ratio of about 0.02–0.03 in Ru+Ru collisions at 1.69A GeV. Our predictions that include kaon medium effects are seen to be in better agreement with these preliminary data than those without the kaon medium effects.

Naively, it is expected that when kaon medium effects are included, the  $K^-/K^+$  ratio should increase with the increasing centrality, since in central collisions reduction of the antikaon mass and the increase of kaon mass are the most significant. However, since the second processes  $\bar{K}\pi \rightarrow KN$  play an important role in antikaon production and hyperon yield, which are produced in association with kaons, is reduced most significantly in central collisions, the increase in the reduction of antikaon mass towards central collisions is largely compensated. The increase of the  $K^-/K^+$  ratio with increasing centrality when kaon medium effects are included is thus marginal (on the order of 10–20 %).

It is also of interest to show the beam energy dependence of the  $K^-/K^+$  ratio in central Au+Au collisions. This is done in Fig. 8 for impact parameter  $b=1$  fm. In both scenarios with and without kaon medium effects, the ratio is seen to increase as beam energy increases. This is understandable, as the antikaon production threshold is higher than that of the kaon, so at these energies, the antikaon production cross sections increases faster than that of the kaon.

Finally, we show in Fig. 9 the time evolution of central density, kaon yield, and antikaon yield in Au+Au collisions

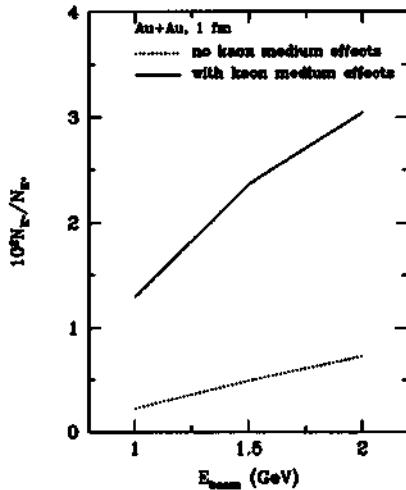


FIG. 8. Beam energy dependence of  $K^-/K^+$  ratio in central Au+Au collisions.

at  $1.5A$  GeV and  $b=3$  fm. It is seen that for a considerable duration of time the system is compressed to a density of around  $3\rho_0$ , and during this period of time, most of the kaons and antikaons are produced. As was shown in Ref. [28],  $K^-$  condensation is predicted to occur around three times normal nuclear matter based on current theoretical and empirical information. The experimental data on kaon and antikaon yields, spectra, ratios, and flow from Au+Au collisions at  $1.5A$  GeV will be very useful in pinning down the critical density for kaon condensation.

#### B. Transverse mass spectra

The effects of kaon and antikaon mean-field potentials can be more clearly seen by looking at their ratio as a func-

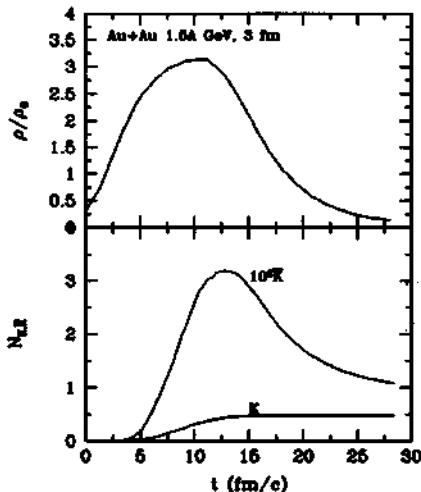


FIG. 9. Time evolution of central density as well as kaon and antikaon yields in Au+Au collisions at  $1.5A$  GeV and 3 fm.

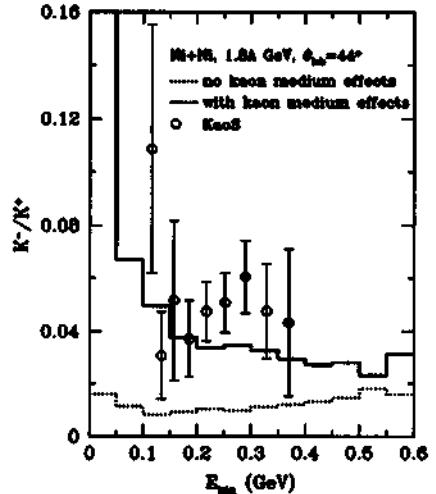


FIG. 10. Kinetic energy spectra of the  $K^-/K^+$  ratio in Ni+Ni collisions at  $1.8A$  GeV. The experimental data are from the KaoS Collaboration [35].

tion of their transverse mass (or kinetic energy). The results are shown in Figs. 10, 11, 12, and 13, for Ni+Ni at  $1.8A$  GeV, Ru+Ru at  $1.69A$  GeV, Au+Au at  $1.5A$  GeV, and Au+Au at  $2A$  GeV, respectively. The results for Ni+Ni are presented in terms of kinetic energy for minimum-biased collisions, in accordance with the experimental data from the KaoS Collaboration [35], shown in Fig. 10 with open circles.

When kaon medium effects are neglected, the  $K^-/K^+$  ratio is seen to increase slightly with increasing transverse momentum. Since the antikaon absorption cross section by nucleons becomes large at low momentum, low-momentum antikaons are more strongly absorbed than high-momentum ones. This makes the  $K^-/K^+$  ratio decrease at small trans-

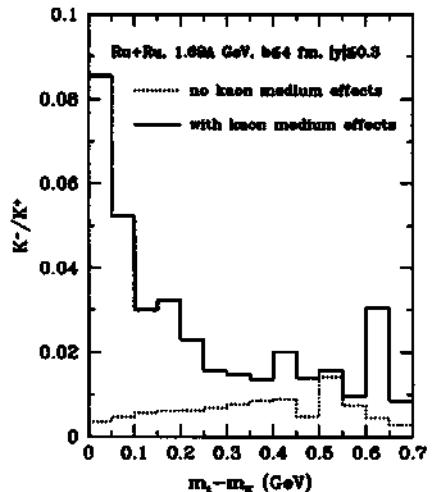


FIG. 11. Transverse mass spectra of  $K^-/K^+$  ratio in central Ru+Ru collisions at  $1.69A$  GeV.

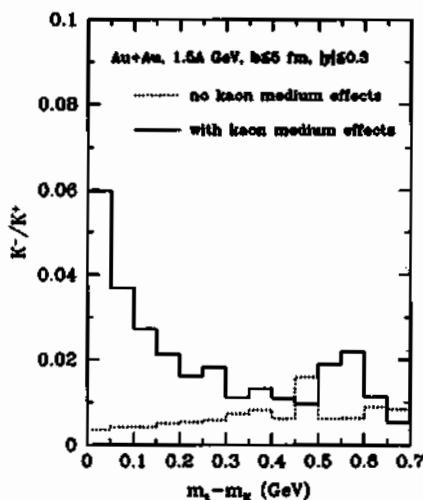


FIG. 12. Same as Fig. 11, for Au+Au collisions at 1.5A GeV.

verse mass. When medium effects are included, we find that the shape of the  $K^-/K^+$  ratio is completely different than that without kaon medium effects. The ratio is now seen to increase dramatically towards small transverse mass. For example, for Ru+Ru collisions, it increases from about 0.02 at  $m_t - m_K = 0.3$  GeV to about 0.08 when  $m_t - m_K$  approaches zero. The difference in the shape of the  $m_t$  spectra comes from the propagation of kaons and antikaons in their mean-field potential. Kaons are "pushed" to high momenta by the repulsive potential, while antikaons are "pulled" to low momenta, leading to an enhanced  $K^-/K^+$  ratio at small transverse masses. From Fig. 10, we also see that the effects of propagation in mean-field potentials are more pronounced in central collisions.

Experimental data from the KaoS Collaboration provided some indication for this dramatic increase of the  $K^-/K^+$

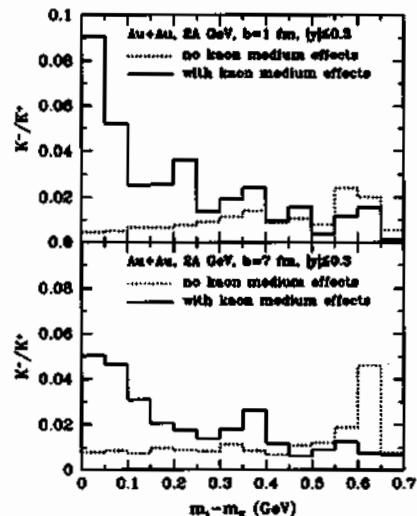
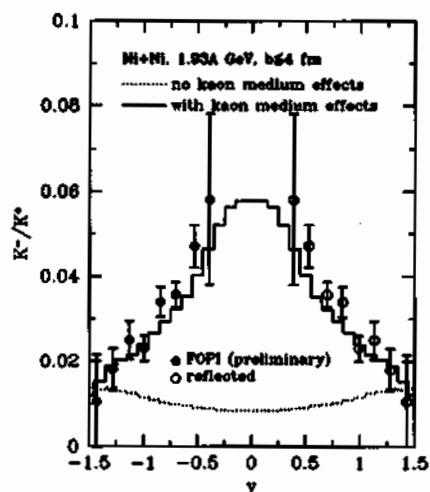


FIG. 13. Same as Fig. 11, for Au+Au collisions at 2A GeV.

FIG. 14. Rapidity distribution of  $K^-/K^+$  ratio in central Ni+Ni collisions at 1.93A GeV. The solid circles are preliminary experimental data from the FOPI Collaboration [52,54], while the open circles are obtained by reflecting the data with respect to the midrapidity.

ratio at small momenta [35]. Our results are compared with these data in Fig. 10. It is seen that our results including the kaon medium effects are in much better agreement with the data. Of course the statistical uncertainty of the data is still quite large (especially for the first datum point at 0.12 GeV). This will be improved in their recent analysis of Ni+Ni collisions at 1.93A GeV [39]. It will also be very useful if the ratio at kinetic energies less than 0.1 GeV can be measured experimentally.

It should be emphasized that in this work, as in our previous study of kaon production in heavy-ion collisions, the explicit momentum dependence of the kaon scalar and vector potential is not considered. This should be a reasonable approximation for heavy-ion collisions at 1–2 A GeV, as kaons (antikaon) produced in these reactions usually have small momenta. This approximation will break down at higher beam energies about 10A GeV. A recent theoretical calculation [53] indicated that the attractive antikaon potential becomes weaker as its momentum (relative to the medium) increases. Experimental data on  $K^-$ -nucleus scattering at high incident momenta also indicate a repulsive (rather than an attractive) antikaon optical-model potential. This momentum dependence of the kaon (antikaon) potential might explain the fact that so far the experimental data for heavy-ion collisions at 10A GeV do not show substantial enhancement of the  $K^-/K^+$  ratio approaching small transverse masses (momenta).

### C. Rapidity distribution

The rapidity distributions of the  $K^-/K^+$  ratio should provide quite similar information on kaon medium effects as its transverse mass spectra. The results for Ni+Ni at 1.93A GeV, Ru+Ru at 1.69A GeV, Au+Au at 1.5A GeV, and Au+Au at 2.0A GeV collisions are shown in Figs. 14,

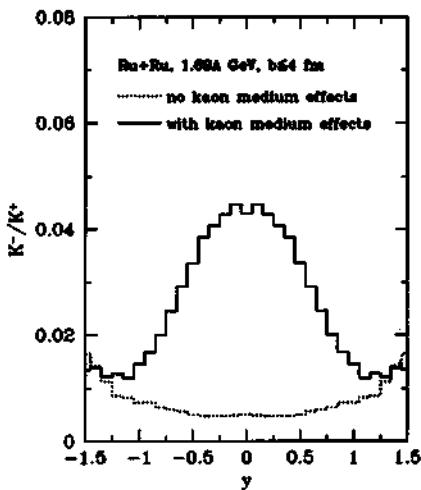


FIG. 15. Same as Fig. 14, for Ru+Ru collisions at 1.694 GeV.

15, 16, and 17, respectively. For Ni+Ni collisions we show also the preliminary data from the FOPI Collaboration [52,54] by circles.

Indeed, the magnitudes and the shapes of these rapidity distributions with and without kaon medium effects are very different. Without kaon medium effects, the rapidity distribution is seen to decrease slightly from target/projectile rapidities to midrapidity. This is again due to the large absorption cross section for slow-moving antikaons. When kaon medium effects are included, there appears not only an overall increase of the ratio, because of the increased production cross section of antikaons, but also a significant change of the shape of the rapidity distribution. The  $K^-/K^+$  ratio is seen to increase steadily from target/projectile rapidities to midrapidity. Our results with kaon medium effects are in much better agreement with the preliminary data from the FOPI Collaboration (Fig. 14).

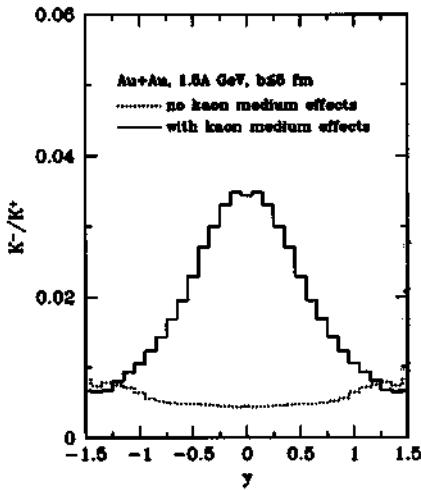


FIG. 16. Same as Fig. 14, for Au+Au collisions at 1.54 GeV.

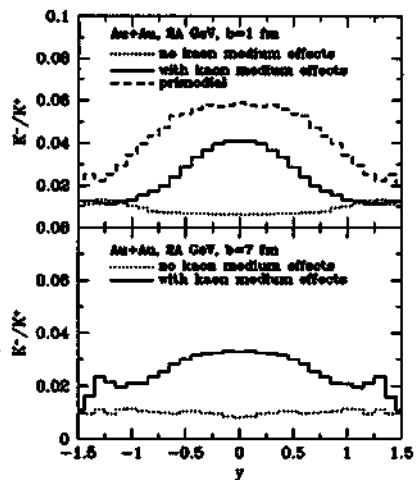


FIG. 17. Same as Fig. 14, for Au+Au collisions at 2A GeV.

For Au+Au collisions at 2A GeV and  $b=1$  fm, we also show in the figure the  $K^-/K^+$  ratio before any final-state interactions on kaons and antikaons are included, which means that even the antikaon absorption is turned off. We see that in this case, the ratio also decreases from midrapidity to projectile and target rapidities. The decrease of this ratio towards large rapidities (momenta) reflects the fact that antikaons are produced at a higher threshold than kaons, so that their momentum distributions are restricted more severely by the available energies. However, we need to emphasize that antikaon absorption and the fact that its absorption cross section increases at low momenta are well-known experimental facts and must be included in transport model calculations.

As mentioned in the Introduction, so far most of the experimental data from the FOPI and KaoS Collaborations at SIS/GSI are consistent with the chiral perturbation theory predictions for kaon in-medium properties. For an ultimate confirmation of these medium effects, it should be very useful if independent experimental data for Au+Au collisions at 2A GeV from the E866 and E895 Collaborations could become available.

#### IV. SUMMARY

In summary, we studied  $K^-/K^+$  ratios as a function of centrality (participant nucleon number), transverse mass ( $m_T$ ), and rapidity in Ni+Ni, Ru+Ru, and Au+Au collisions at beam energies between 1 and 2A GeV. We used the relativistic transport model that includes explicitly the strangeness degrees of freedom and considered two scenarios for kaon properties in dense matter, one with and one without medium modifications of their properties. In both scenarios, the  $K^-/K^+$  ratio does not change very much with the centrality, while the  $K/\pi$  and  $\bar{K}/\pi$  ratios increase with increasing centrality. Significant differences were predicted, both in magnitudes and shapes, for the  $m_T$  spectra and rapidity distributions of  $K^-/K^+$  ratio. We found that the experimental data from the KaoS Collaboration for the kinetic energy

spectra of the  $K^-/K^+$  ratio and those from the FOPI Collaboration for its rapidity distribution support the suggestion of kaon medium effects. We emphasize that the independent data from the E866 and E895 Collaborations for Au+Au collisions at 2A GeV will be useful in confirming or confronting these findings.

#### ACKNOWLEDGMENTS

We are grateful to N. Herrmann and P. Senger for sending us the data files, and to N. Herrmann, C. Ogilvie, and P. Senger for useful communications. This work was supported in part by the U.S. Department of Energy under Grant No. DE-FG02-88ER40388.

- 
- [1] G.E. Brown, K. Kubodera, D. Page, and P. Pizzecherro, Phys. Rev. D 37, 2042 (1988).
  - [2] V. Thorsson, M. Prakash, and J. M. Lattimer, Nucl. Phys. A572, 693 (1994).
  - [3] D. B. Kaplan and A. E. Nelson, Phys. Lett. B 175, 57 (1986).
  - [4] G. E. Brown, K. Kubodera, and M. Rho, Phys. Lett. B 192, 272 (1987).
  - [5] H. D. Politzer and M. B. Wise, Phys. Lett. B 273, 156 (1991).
  - [6] G. E. Brown, C.-H. Lee, M. Rho, and V. Thorsson, Nucl. Phys. A567, 937 (1994).
  - [7] C.-H. Lee, G. E. Brown, D. P. Min, and M. Rho, Nucl. Phys. A585, 401 (1995).
  - [8] N. Kaiser, P. B. Siegel, and W. Weise, Nucl. Phys. A594, 325 (1995).
  - [9] T. Waas, N. Kaiser, and W. Weise, Phys. Lett. B 379, 34 (1996).
  - [10] C.-H. Lee, Phys. Rep. 275, 255 (1996).
  - [11] N. Kaiser, T. Waas, and W. Weise, Nucl. Phys. A612, 297 (1997).
  - [12] T. Waas and W. Weise, Nucl. Phys. A625, 287 (1997).
  - [13] M. Lutz, A. Steiner, and W. Weise, Nucl. Phys. A574, 755 (1994).
  - [14] J. Schaffner, A. Gal, I. N. Mishustin, H. Stöcker, and W. Greiner, Phys. Lett. B 334, 268 (1994).
  - [15] R. Knorren, M. Prakash, and P. J. Ellis, Phys. Rev. C 52, 3470 (1995).
  - [16] G. E. Brown, C. B. Dover, P. B. Siegel, and W. Weise, Phys. Rev. Lett. 60, 2723 (1988).
  - [17] E. Friedman, A. Gal, and J. Mares, Nucl. Phys. A625, 272 (1997).
  - [18] E. Friedman, A. Gal, and C. J. Batty, Nucl. Phys. A579, 518 (1994).
  - [19] See e.g., *Proceedings of the Eleventh International Conference on Ultra-Relativistic Nucleus-Nucleus Collisions*, edited by A. M. Poskanzer, J. W. Harris, and L. S. Schroeder [Nucl. Phys. A590 (1995)]; *Proceedings of the Twelfth International Conference on Ultra-Relativistic Nucleus-Nucleus Collisions*, edited by P. Braun-Munzinger, H. J. Specht, R. Stock, and H. Stoecker [Nucl. Phys. A610 (1996)].
  - [20] G. Q. Li, C. M. Ko, and B. A. Li, Phys. Rev. Lett. 74, 235 (1995).
  - [21] G. Q. Li and C. M. Ko, Nucl. Phys. A594, 460 (1995).
  - [22] G. Q. Li and C. M. Ko, Phys. Rev. C 54, R2159 (1996).
  - [23] W. Cassing, E. L. Bratkovskaya, U. Mosel, S. Teis, and A. Sibirtsev, Nucl. Phys. A614, 415 (1997).
  - [24] E. L. Bratkovskaya, W. Cassing, and U. Mosel, Nucl. Phys. A622, 593 (1977).
  - [25] Z. S. Wang, A. Faessler, C. Fuchs, V. S. Uma Maheswari, and D. S. Kosov, Phys. Rev. Lett. 79, 4096 (1997).
  - [26] Z. S. Wang, A. Faessler, C. Fuchs, V. S. Uma Maheswari, and D. S. Kosov, Nucl. Phys. A628, 151 (1998).
  - [27] G. Q. Li, C.-H. Lee, and G. E. Brown, Phys. Rev. Lett. 79, 5214 (1997).
  - [28] G. Q. Li, C.-H. Lee, and G. E. Brown, Nucl. Phys. A625, 372 (1997).
  - [29] J. L. Ritman *et al.*, FOPI Collaboration, Z. Phys. A 352, 355 (1995).
  - [30] N. Herrmann for the FOPI Collaboration, Nucl. Phys. A610, 49c (1996).
  - [31] D. Best *et al.*, FOPI Collaboration, Nucl. Phys. A625, 307 (1997).
  - [32] D. Miskowiec *et al.*, KaoS Collaboration, Phys. Rev. Lett. 72, 3650 (1994).
  - [33] R. Elmer *et al.*, KaoS Collaboration, Phys. Rev. Lett. 77, 4886 (1996).
  - [34] P. Senger for the KaoS Collaboration, Heavy Ion Phys. 4, 317 (1996).
  - [35] R. Barth *et al.*, KaoS Collaboration, Phys. Rev. Lett. 78, 4027 (1997).
  - [36] C. Ogilvie, J. Phys. G 23, 1803 (1997).
  - [37] C. Ogilvie, Nucl. Phys. A630, 571c (1998).
  - [38] N. Herrmann (private communication).
  - [39] P. Senger (private communication).
  - [40] M. Justice *et al.*, EOS Collaboration, nucl-ex/9708005.
  - [41] D. Best for the E895 Collaboration, J. Phys. G 23, 1873 (1997).
  - [42] C. M. Ko, Q. Li, and R. Wang, Phys. Rev. Lett. 59, 1084 (1987).
  - [43] R. J. Furnstahl, H.-B. Tang, and B. D. Serot, Phys. Rev. C 52, 1368 (1995).
  - [44] K. Tsushima, S. W. Huang, and A. Faessler, Phys. Lett. B 337, 245 (1994); J. Phys. G 21, 33 (1995).
  - [45] G. Q. Li and C. M. Ko, Nucl. Phys. A594, 439 (1995).
  - [46] G. Q. Li, C. M. Ko, and W. S. Chung, Phys. Rev. C 57, 434 (1998).
  - [47] A. Sibirtsev, W. Cassing, and C. M. Ko, Z. Phys. A 358, 101 (1997).
  - [48] B. Hong *et al.*, FOPI Collaboration, Phys. Lett. B 407, 115 (1997).
  - [49] B. Hong *et al.*, FOPI Collaboration, Phys. Rev. C 57, 244 (1998).
  - [50] D. Pelte *et al.*, FOPI Collaboration, Z. Phys. A 359, 55 (1997).
  - [51] L. Ahle *et al.*, E866 Collaboration, Nucl. Phys. A610, 139c (1996).
  - [52] Y. Leifels, in *Proceedings of International Workshop on Hadrons in Dense Matter*, edited by W. Noerenberg (GSI, Darmstadt, 1997).
  - [53] M. Lutz, nucl-th/9709073.
  - [54] B. Hong, in *Proceedings of APCTP Workshop on Astro-Hadron Physics: Properties of Hadrons in Matter*, edited by G. E. Brown (World Scientific, Singapore, 1998).

---

## Appendix D

# Strangeness Equilibration at GSI Energies

G.E. Brown, M. Rho and C. Song

Reprinted with permission from *Nuclear Physics*, **A690** (2001) 184c-200c

Copyright (2001) by Elsevier Science B.V.

### **Commentary by G.E. Brown**

The work of Brown & Rho (2002) shows unambiguously that the chiral restoration transition in the chiral limit is accomplished by constituent quarks going massless. Strong indications of this had already been given by Brown, Buballa & Rho (1996) and Pirner & Wachs (1997). In the case of Paper 27 this means that with chiral restoration the factor of 9 enhancement of  $\omega$ -meson mean fields over the  $\rho$  disappears, and they become equal at chiral restoration. With finite temperature this was already seen from the quark number susceptibility by Gottlieb *et al.* (1987), where the isoscalar and isovector susceptibilities as measured in lattice calculations are equal. Finite density effects can now be connected with finite temperature effects calculated in lattice gauge simulations using the Nambu-Jona-Lasinio theory in mean field (G.E. Brown and Mannque Rho 2002).

Since, as is outlined in Paper 27, the presence of hyperons in neutron stars depends upon the strongly repulsive vector mean field acting on the neutrons in neutron-rich matter, it is clear that hyperons will not be present near the chiral restoration density.

Basically we are spelling out the scenario suggested in Paper 27, but now established by the work of Brown & Rho (2002), that kaon condensation happens before chiral restoration.

### **References**

- (1) Brown, G.E., Buballa, M., and Rho, M. (1996), *Nuclear Physics* **A596**, 503.
- (2) Brown, G.E., and Rho, M. (2002), nucl-th/0206021.
- (3) Gottlieb, S., Liu, W., Toussaint, D., Renken, R.C., and Sugar, R.I. (1987), *Physical Review Letters* **59**, 2247.
- (4) Pirner, H.-J., and Wachs, M. (1997), *Nuclear Physics* **A617**, 345.

This page is intentionally left blank



## Strangeness Equilibration at GSI Energies

G.E. Brown<sup>(a)</sup>, Mannque Rho<sup>(b)</sup> and Chaejun Song<sup>(a)</sup>

*(a) Department of Physics and Astronomy, State University of New York  
Stony Brook, NY 11794, USA*

*(b) Service de Physique Théorique, CE Saclay, 91191 Gif-sur-Yvette, France*

(November 30, 2000)

### Abstract

We develop the notion of “broad-band equilibration” in heavy-ion processes involving dense medium. Given density-dependent  $K^-$ -masses we show that the equilibration at GSI energies claimed to hold in previous treatments down to  $\sim \rho_0/4$ , can be replaced by a broad-band equilibration in which the  $K^-$ -meson and hyperons are produced in an essentially constant ratio independent of density. There are experimental indications that this also holds for AGS energies. We then proceed to argue that both  $K^+$  and  $K^-$  must get lighter in dense medium at some density  $\rho > \rho_0$  due to the decoupling of the vector mesons. As a consequence, kaon condensation in compact stars could take place before chiral restoration since the sum of bare quark masses in the kaon should lie below  $\mu_c$ . Another consequence of the decoupling vector interactions is that the quasi-particle picture involving (quasi)quarks, presumably ineffective at low densities, becomes more appropriate at higher densities as chiral restoration is approached.

## 1 Introduction

Following work by Hagedorn [1] on production of anti- ${}^3\text{He}$ , Cleymans et al [2] have shown that for low temperatures, such as found in systems produced at GSI, strangeness production is strongly suppressed. The abundance of  $K^+$  mesons, in systems assumed to be equilibrated, is given by [3],

$$n_{K^+} \sim e^{-E_{K^+}/T} V \left\{ g_K \int \frac{d^3 p}{(2\pi)^3} e^{-E_K/T} + g_\Lambda \int \frac{d^3 p}{(2\pi)^3} e^{-(E_\Lambda - \mu_B)/T} \right\}. \quad (1)$$

Here the  $g$ 's are the degeneracies. Because strangeness must be conserved in the interaction volume  $V$ , assumed to be that of the equilibrated system for each  $K^+$  which is produced, a particle of “negative strangeness”<sup>1</sup> containing  $s$ , say,  $\bar{K}$  or  $\Lambda$ , must also be produced, bringing in the  $\bar{K}$  or  $\Lambda$  phase space and Boltzmann factors. The  $K^+$  production is very small at GSI energies because of the low temperatures which give small Boltzmann factors for the  $\bar{K}$  and  $\Lambda$  in addition to the small Boltzmann factor for the  $K^+$ . Note the linear dependence on interaction volume which follows from the necessity to include  $\bar{K}$  or  $\Lambda$  phase space.

In an extensive and careful analysis, Cleymans, Oeschler and Redlich [3] show that measured particle multiplicity ratios  $\pi^+/p$ ,  $\pi^-/\pi^+$ ,  $d/p$ ,  $K^+/\pi^+$ , and  $K^+/K^-$  – but not  $\eta/\pi^0$  – in central Au-Au and Ni-Ni collisions at (0.8-2.0)A GeV are explained in terms of a thermal model with a common freeze-out temperature and chemical potential, if collective flow is included in the description. In other words, a scenario in which the kaons and anti-kaons are equilibrated appears to work well. This result is puzzling in view of a recent study by Bratkovskaya et al [4] that shows that the  $K^+$  mesons in the energy range considered would take a time of  $\sim 40$  fm/c to equilibrate. We remark that this is roughly consistent with the estimate for higher energies in the classic paper by Koch, Müller and Rafelski [5] that strangeness equilibration should take  $\sim 80$  fm/c. Such estimates have been applied at CERN energies and the fact that emergent particle abundances are described by Boltzmann factors with a common temperature  $\sim 165$  MeV [6] has been used as part of an argument that the quark/gluon plasma has been observed.

We interpret the result of [3] as follows. Since free-space masses are used for the hadrons involved, Cleymans et al [3] are forced to employ a  $\mu_B$  substantially less than the nucleon mass  $m_N$  in order to cut down  $\Lambda$  production as compared with  $K^-$  production, the sum of the two being equal to  $K^+$  production. This brings them to a diffuse system with density of only  $\sim \rho_0/4$  at chemical freeze-out. But this is much too low a density for equilibration.

We shall first show how this situation can be improved by replacing the  $K^-$  mass by the  $K^-$  energy at rest  $\omega_{K^-}^* \equiv \omega_-(k=0) < m_K$ . (The explicit formula for  $\omega_\pm$  is given later, see eq. (21).) In doing this, we first have to reproduce the  $K^+$  to  $K^-$  ratio found in the Ni + Ni experiments [7]:

$$n_{K^+}/n_{K^-} \simeq 30. \quad (2)$$

Cleymans et al reproduce the earlier smaller ratio of  $21 \pm 9$  with  $\mu_B = 750$  MeV and  $T = 70$  MeV. How this or rather (2) comes out is easy to see. The ratio of the second term on the RHS of eq.(1) to the first term is roughly the ratio of the exponential factors multiplied by the phase

<sup>1</sup>By “negative strangeness” we are referring to the negatively charged strange quark flavor. The positively charged anti-strange quark will be referred to as “positive strangeness.”

space volume

$$R = \frac{g_\Lambda}{g_K} \left( \frac{\bar{p}_\Lambda}{\bar{p}_K} \right)^3 \frac{e^{-(E_\Lambda - \mu_B)/T}}{e^{-E_{K^-}/T}} \approx \left( \frac{m_\Lambda}{m_K} \right)^{3/2} \frac{e^{-(m_\Lambda - \mu_B)/T}}{e^{-m_{K^-}/T}} \approx 21 \quad (3)$$

where we have used  $g_\Lambda \approx g_K \approx 2$ ,  $(\bar{p})^2/2m \simeq \frac{3}{2}T$ ,  $E_\Lambda = 1115$  MeV and  $E_K = 495$  MeV. We are able to convert  $E_\Lambda$  to  $m_\Lambda$  and  $E_{K^-}$  to  $m_{K^-}$  because  $E_\Lambda \simeq m_\Lambda + \frac{3}{2}T$  and  $E_{K^-} \simeq m_{K^-} + \frac{3}{2}T$  and the thermal energies cancel out in the ratio. This works as long as the masses are more than  $\sim 3T$ , where the nonrelativistic approximation is valid. Inclusion of the  $\Sigma$  and  $\Xi$  hyperons would roughly increase this number by 50% with the result that the ratio of  $K^-$  to  $\Lambda$ ,  $\Sigma$ ,  $\Xi$  production is <sup>2</sup>

$$\frac{n_{K^-}}{n_{\Lambda+\Sigma+\Xi}} \simeq 1/32. \quad (4)$$

Since a  $K^+$  must be produced to accompany each particle of one unit of strangeness (to conserve strangeness flavor), we then have

$$n_{K^+}/n_{K^-} \sim 33. \quad (5)$$

This is consistent with the empirical ratio (2). It should be noted that had we set  $\mu_B$  equal to  $m_N$ , we would have had the  $K^+$  to  $K^-$  ratio to be  $\sim 280$  because it costs so much less energy to make a  $\Lambda$  (or  $\Sigma$ ) rather than  $K^-$  in this case. In other words the chemical potential  $\mu_B$  is forced to lie well below  $m_N$  in order to penalize the hyperon production relative to that of the  $K^-$ 's.

One can see from fig.5 in Li and Brown [8] that without medium effects in the  $K^-$  mass, the  $K^+/K^-$  ratio is  $\sim 100$ , whereas the medium effect decreases the ratio to about 23. This suggests how to correctly redo the Cleymans et al's analysis, namely, by introducing the dropping  $K^-$  mass into it.

In Appendix A we show that positive strangeness production takes place chiefly at densities greater than  $2\rho_0$ . As the fireball expands to lower densities the amount of positive strangeness remains roughly constant. The number of  $K^+$ 's is such as to be in equilibrium ratio  $K^+/\pi^+$  with the equilibrated number density of pions at  $T = 70$  MeV,  $n_\pi \approx 0.37 T_{197}^3 \text{ fm}^{-3}$ . Only in this sense do the  $K^+$ 's equilibrate.

It is amusing to note that the "equilibrated ratio" of  $\sim 30$  for the  $n_{K^+}/n_{K^-}$  holds over a large range of densities for  $T = 70$  MeV, once density-dependent  $K^-$  masses are introduced, in that the ratio  $R$  of (3) is insensitive to density. (Remember that because of the small number of  $K^-$ 's, the number of  $K^+$ 's must be nearly equal to the number of hyperons,  $\Lambda$ ,  $\Sigma^-$  and  $\Xi$ , in order to conserve strangeness.) This insensitivity results because  $\omega_{K^-}^*$  decreases with density at roughly the same rate as  $\mu_B$  increases. We can write  $R$  of (3), neglecting possible changes in  $T$  and  $m_K$  in our lowest approximation, as <sup>3</sup>

$$R = \left( \frac{m_\Lambda}{m_K} \right)^{3/2} e^{(\mu_B + \omega_{K^-}^*)/T} e^{-m_\Lambda/T}. \quad (6)$$

<sup>2</sup>In order to reproduce this result with  $\mu_B = 750$  MeV and  $T = 70$  Mev within our approximation, we have assumed only the  $\Sigma^-$  and  $\Sigma^0$  hyperons to equilibrate with the  $\Lambda$ . This may be correct because the  $\Sigma^+$  and  $\Xi$  couple more weakly. Inclusion of the latter could change our result slightly. Probably they should be included in analysis of the AGS experiments at higher energies where they would be more copiously produced.

<sup>3</sup>To be fully consistent, we would also have to consider the medium modifications of the  $\Lambda$  and  $K^+$  properties. These and other improvements are left out for our rough calculation.

As will be further stressed later, the most important point in our arguments is that  $\mu_B + \omega_{K^-}^*$  is nearly constant with density. This is because whereas  $\mu_B$  increases from 860 MeV to 905 MeV as  $\rho$  goes from  $1.2\rho_0$  to  $2.1\rho_0$ ,  $\omega_{K^-}^*$  decreases from 380 MeV to 332 MeV, the sum  $\mu_B + \omega_{K^-}^*$  decreasing very slightly from 1240 MeV to 1237 MeV. Indeed, even at  $\rho = \rho_0/4$ ,  $\mu_B + \omega_{K^-}^* \sim 1218$  MeV, not much smaller.

We believe that the temperature will change only little in the region of dropping masses because in a consistent evolution (which we do not carry out here) the scalar field energy  $m_\sigma^2 \sigma^2/2$  in a mean field theory plays the role of an effective bag constant. In ref.[9] this is phrased in terms of a modified Walecka theory,

$$B_{eff} = \frac{1}{2} m_\sigma^2 \sigma^2 \Rightarrow \frac{1}{2} m_\sigma^2 (M_N/g_{\sigma NN})^2, \quad (7)$$

the  $\sigma$  going to  $M_N/g_{\sigma NN}$  as the nucleon effective mass goes to zero with chiral restoration. Most of the energy with compression to higher densities goes with this effective bag constant, estimated [9] to be  $\sim 280$  MeV/fm<sup>3</sup>, rather than heat, mocking up the behavior of a mixed phase with constant temperature. Moreover at  $\rho = 2\rho_0$  where  $m_N^*$  may be  $\sim 0.5m_N$ , only about 25% of the bag constant  $B$  may have been achieved, so there may be some increase in temperature. We shall, at the same level of accuracy, have to replace  $m_R$  in the prefactor of eq.(6) by  $\omega_{K^-}^*$ . We adjust the increase in temperature so that it exactly compensates for the decrease in prefactor so that the  $K^+/K^-$ -ratio is kept the same, as required by experiment [7]. We then find that the temperature at  $\rho = 2\rho_0$  must be increased from 70 to 95 MeV. This is roughly the change given by that in inverse slopes of  $K^-$  and  $K^+$  transverse momentum distributions found in going from low multiplicities to high multiplicities [7].

In any case, we see that  $R$  will depend but little on density. This near cancellation of changes in the factors is fortunate because the  $K^- + p \leftrightarrow \Lambda$  reaction, operating in the negative strangeness sector, is much stronger than the positive strangeness reactions, so the former should equilibrate to densities well below  $2\rho_0$  and we can see that the “apparent equilibration” might extend all the way down to  $\rho \sim \rho_0/4$ .

The near constancy of  $R$  with density also explains the fact the  $K^+/K^-$ -ratio does not vary with centrality<sup>4</sup>. Although  $R$  is the ratio of  $\Lambda$ 's to  $K^-$ 's, both of which are in the negative strangeness sector, nonetheless, the number of  $K^+$ 's must be equal to the sum of the two and since the  $\Lambda$ 's are much more abundant than the  $K^-$ 's,  $R$  essentially represents the  $K^+/K^-$ -ratio.

Detailed transport results of Bratkovskaya and Cassing (see Fig.1) show the last scattering of the detected  $K^-$  to be spread over all densities from  $\rho_0/2$  to  $3\rho_0$ , somewhat more of the last scatterings to come from the higher density. This seems difficult to reconcile with a scenario of the  $K^-$  numbers being decided at one definite density and temperature, but given our picture of dropping masses, one can see that the  $K^+/K^-$ -ratio depends little on density, that is, on  $\mu_B$  at which the  $K^-$  last interacts. In any case we understand from our above argument that the apparent density of equilibration can be chosen to be very low in a thermal description and still get more or less correct  $K^+/K^-$ -ratio.

<sup>4</sup>We are grateful to Helmut Oeschler for pointing this out to us.

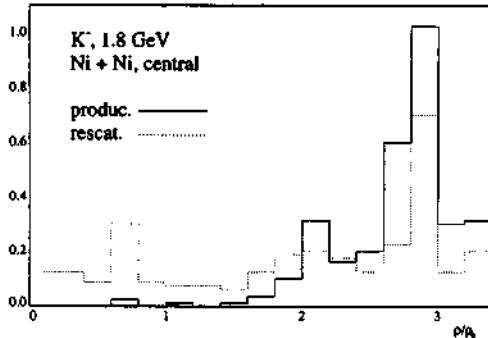


Figure 1: Calculations by Bratkovskaya and Cassing (private communication) which show the density of origin and that of the last interaction of the  $K^-$  mesons.

## 2 The Top-Down Scenario of $K^\pm$ Production

Brown and Rho [10] discussed fluctuations in the kaon sector in terms of a simple Lagrangian

$$\delta\mathcal{L}_{KN} = \frac{-6i}{8f^{*2}} (\bar{N}\gamma_0 N) \bar{K} \partial_t K + \frac{\Sigma_{KN}}{f^{*2}} (\bar{N}N) \bar{K} K + \dots \equiv \mathcal{L}_\omega + \mathcal{L}_\sigma + \dots \quad (8)$$

It was suggested there that at high densities, the constituent quark or quasi-quark description can be used with the  $\omega$ -meson coupling to the kaon viewed as a *matter field*<sup>5</sup> (rather than as a Goldstone boson). Such a description suggests that the  $\omega$  coupling to the kaon which has one non-strange quark is 1/3 of the  $\omega$  coupling to the nucleon which has three non-strange quarks. The  $\mathcal{L}_\omega$  in the Lagrangian was obtained by integrating out the  $\omega$ -meson as in the baryon sector. We may therefore replace it by the interaction

$$V_{K^\pm} \approx \pm \frac{1}{3} V_N. \quad (9)$$

In isospin asymmetric matter, we shall have to include also the  $\rho$ -meson exchange [10] with the vector-meson coupling treated in the top-down approach.

For the top-down scenario, we should replace the chiral Lagrangian (8) by one in which the “heavy” degrees of freedom figure explicitly. This means that  $\frac{1}{2f^{*2}}$  in the first term of (8) should be replaced by  $g^{*2}/m_\omega^{*2}$  and  $\frac{\Sigma_{KN}}{f^{*2}}$  in the second term by  $\frac{2}{3} m_K \frac{g_\omega^{*2}}{m_\omega^{*2}}$  assuming that both  $\omega$  and  $\sigma$  are still massive. We will argue in the next section that while the  $\omega$  mass drops, the ratio  $g^{*2}/m_\omega^{*2}$  should stay constant or more likely decrease with density and that beyond certain

<sup>5</sup>Briefly the “top-down” description of the kaon as used in [10] is as follows. As one approaches chiral restoration, the spontaneously generated (constituent) quark masses drop and the mass splitting between a constituent u (or d) quark and a constituent s quark relative to the sum becomes much larger than the same in the vacuum. In this case, it is better to treat, in the chiral quark picture, the s quark as a *matter field* rather than as a *chiral field* like the u or d quark. Even in zero-density environment, viewing the kaon as “massive” as in the Callan-Klebanov skyrmion for hypersons rather than as a Goldstone boson is a quantitatively reliable approximation. See [11] for extensive discussions on this matter.

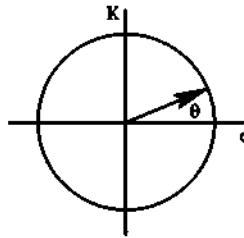


Figure 2: Projection onto the  $\sigma, K$  plane. The angular variable  $\theta$  represents fluctuation toward kaon mean field.

density above nuclear matter, the vector fields should decouple. On the other hand,  $g_\sigma$  is not scaled in the mean field that we are working with; the motivation for this is given in Brown, Buballa and Rho [9] who construct the chiral restoration transition in the mean field in the Nambu-Jona-Lasinio model. Thus

$$\frac{\Sigma_{KN}}{f^{*2}} \approx \frac{2}{3} m_K \frac{g_\sigma^2}{m_\sigma^{*2}}. \quad (10)$$

In this framework,  $m_\sigma^*$  is the order parameter for chiral restoration which drops à la BR scaling [12]:

$$\frac{m_\sigma^*}{m_\sigma} \equiv \Phi(\rho) \simeq \frac{1}{1 + y\rho/\rho_0} \quad (11)$$

with  $y \simeq 0.28$ , at least for  $\rho \lesssim \rho_0$ <sup>6</sup>.

Once the vector is decoupled, a simple way to calculate the in-medium kaon effective mass, equivalent to using the  $\mathcal{L}_\sigma$ , is to consider the kaon as fluctuation about the “ $\sigma$ ”-axis in the V-spin formalism [13] as depicted pictorially in Fig.2. The Hamiltonian for explicit chiral symmetry breaking is

$$\begin{aligned} H_{\chi SB} &= \Sigma_{KN}(\bar{N}N) \cos(\theta) + \frac{1}{2} m_K^2 f^{*2} \sin^2(\theta) \\ &\simeq \Sigma_{KN}(\bar{N}N) \left(1 - \frac{\theta^2}{2}\right) + \frac{1}{2} m_K^2 f^{*2} \theta^2 \end{aligned} \quad (12)$$

where the last expression is obtained for small fluctuation  $\theta$ . Dropping the term independent of  $\theta$ , we find

$$m_K^{*2} = m_K^2 \left(1 - \frac{\Sigma_{KN}(\bar{N}N)}{f^{*2} m_K^2}\right). \quad (13)$$

Using eq. (10) we obtain

$$m_K^{*2} = m_K^2 \left(1 - \frac{2 g_\sigma^2 (\bar{N}N)}{3 m_\sigma^{*2} m_K}\right). \quad (14)$$

<sup>6</sup> $y$  may well be different from this value for  $\rho > \rho_0$ . In fact the denominator of  $\Phi(\rho)$  could even be significantly modified from this linear form. At present there is no way to calculate this quantity from first principles.

In accord with Brown and Rho [10] we are proposing that eq.(13) should be used for low densities, in the Goldstone description of the  $K^\pm$ , and that we should switch over to eq.(14) for higher densities. It is possible that the  $m_K$  appearing in (14) should be replaced by  $m_K^*$  for self-consistency but the dropping of  $m_\sigma^{*2}$  makes the  $m_K^*$  of (14) decrease more rapidly than that of (13) so that eq.(13) with  $\langle \bar{N}N \rangle$  set equal to the vector density  $\rho$ <sup>7</sup>, a much used formula valid to linear order in density

$$m_K^{*2} \approx m_K^2 \left( 1 - \frac{\rho \Sigma_{KN}}{f^2 m_K^2} \right) \quad (15)$$

obviously gives too slow a decrease of  $m_K^*$  with density.

Although the above are our chief points, there are two further points to remark. One, even without scaling, our vector interaction on the kaon is still too large. Two, more importantly, there is reason to believe in the large  $\Sigma_{KN}$  term,

$$\Sigma_{KN} \sim 400 \text{ MeV.} \quad (16)$$

This comes from scaling of the pion sigma term

$$\Sigma_{KN} = \frac{(m_u + m_d)\langle N|\bar{u}u + \bar{s}s|N\rangle}{(m_u + m_d)\langle N|\bar{u}u + \bar{d}d|N\rangle} \Sigma_{\pi N}. \quad (17)$$

Taking  $m_s \sim 150$  MeV,  $m_u + m_d \sim 12$  MeV,  $\Sigma_{\pi N} = 46$  MeV and  $\langle N|\bar{s}s|N\rangle \sim \frac{1}{3}\langle N|\bar{d}d|N\rangle$  from lattice calculations [14], one arrives at (16).

Other authors, in adjusting the  $\Sigma$  term to fit the kaon-nucleon scattering amplitudes, have obtained a somewhat smaller  $\Sigma_{KN}$ . This can be understood in the chiral perturbation calculation of C.-H. Lee [15] where the only significant effect of higher chiral order terms can be summarized in the “range term”<sup>8</sup>; namely  $\Sigma_{KN}$  is to be replaced by an effective  $\Sigma$ ,

$$(\Sigma_{KN})_{eff} = (1 - 0.37\omega_K^{*2}/m_K^2)\Sigma_{KN}. \quad (18)$$

It should be pointed out that although the  $\Sigma_{KN}$  is important at low densities,  $\omega_K$  decreases with  $m_K^*$ , this “range-term” correction becomes less important at higher densities. This effect – which is easy to implement – is included in the realistic calculations.

### 3 Partial Decoupling of the Vector Interaction

#### 3.1 Evidence

There are both theoretical and empirical reasons why we believe that the vector interaction should decouple at high density.

<sup>7</sup>The correction to this approximation which may become important as the nucleon mass drops comes as a “1/m” correction in the heavy-fermion chiral perturbation theory (as for the “range term” mentioned below) and can be taken into account systematically. It can even be treated fully relativistically using a special regularization scheme being developed in the field. Our approximation does not warrant the full account of such terms, so we will not include this correction here.

<sup>8</sup>As mentioned, in the language of heavy-baryon chiral perturbation theory, this corresponds to a “1/m” correction term.

1. We first give the theoretical arguments. We know of three theoretical reasons why the vector coupling  $g_\omega^*$  should drop with density.

- The first is the observation by Song et al [16] that describing nuclear matter in terms of chiral Lagrangian in the mean field requires the ratio  $g_\omega^*/m_\omega^*$  to at least be roughly constant or even decreasing as a function of density. In fact to quantitatively account for non-linear terms in a mean-field effective Lagrangian, a dropping ratio is definitely favored<sup>9</sup>. For instance, as discussed in [16], the in-medium behavior of the  $\omega$ -meson field is encoded in the “FTS1” version of the non-linear theories of ref.[17]. In fact, because of the attractive quartic  $\omega$  term in the FTS1 theory, the authors of [17] have (for the parameter  $\eta = -1/2$  favored by experiments)  $g_\omega^{*2}/m_\omega^{*2} \simeq 0.8 g_\omega^2/m_\omega^2$  as modification of the quadratic term when rewritten in our notation. In other words, their vector mean field contains a partial decoupling already at  $\rho \approx \rho_0$  although they do not explicitly scale  $g_\omega$  as we do.

Historically, Walecka-type mean field theories with only quadratic interactions (i.e., linear Walecka model) gave compression moduli  $K \sim 500$  MeV, about double the empirical value. This is cured in nonlinear effective field theories like FTS1 by higher-dimension non-renormalizable terms which effectively decrease the growth in repulsion in density. As suggested in [16], an effective chiral Lagrangian with BR scaling can do the same (by the increase in magnitude of the effective scalar field with density) but more economically and efficiently.

- The second reason is perhaps more theoretical. Kim and Lee [18] have shown recently that in an effective QCD Lagrangian with baryons, pions and vector mesons put together in hidden gauge symmetric theory, the ratio  $g_V^*/m_V^*$  (where  $V$  stands for hidden gauge bosons) falls very rapidly with baryon chemical potential. The main agent for this behavior is found in [18] to be the pionic one-loop contribution linked to chiral symmetry which is lacking in the mean-field treatment for BR scaling [12]. One may argue on a more profound ground that the vector decoupling in approaching chiral phase transition is a flow to a fixed point. It has been argued recently by Harada and Yamawaki [19] that chiral symmetry restoration may correspond to a “vector manifestation” of chiral symmetry where the octet of Goldstone pions and the octet of longitudinal vector mesons belong to the representation  $(8, 1) \oplus (1, 8)$ . At this point, the vector coupling goes zero à la Georgi’s vector limit as does the vector meson mass [20].

The above arguments were made for the  $\rho$  which has a simple interpretation in terms of hidden gauge symmetry but it will apply to the  $\omega$  if the nonet symmetry continues to hold in nuclear medium. It is difficult to be quantitative as to how fast the ratio falls but it is clear that the drop is substantial already near normal nuclear matter density.

- Finally, close to chiral restoration in temperature, there is clear evidence from QCD for an equally rapid drop, specifically, from the quark number susceptibility that can be measured on the lattice [21]. The lattice calculation of the quark number susceptibility dealt with quarks and the large drop in the (isoscalar) vector mean field was

<sup>9</sup>Since the non-linear terms – though treated in the mean field – are fluctuation effects in the effective field theory approach, this represents a quantum correction to the BR scaling.

found to be due chiefly to the change-over from hadrons to quarks as the chiral restoration temperature is approached from below. The factor of 9 in the ratio  $g_{\omega NN}^2/g_{\omega QQ}^2$  (where  $Q$  is the constituent quark) should disappear in the change-over. Now since the electroweak properties of a constituent quark (quasiquark) are expected to be the same as those of a bare Dirac particle with  $g_A = 1$  and no anomalous magnetic moment (i.e., the QCD quark) [22] with possible corrections that are suppressed as  $1/N_c$  [23], there will be continuity between before and after the chiral transition. This is very much in accordance with the “Cheshire-Cat picture” developed elsewhere [11]. In fact, it is possible to give a dynamical (hadronic) interpretation of the above scenario. For instance in the picture of [24], this may be understood as the “elementary”  $\omega$  strength moving downwards into the “nuclear”  $\omega$ , the  $[N^*(1520)N^{-1}]^{J=1,I=0}$  isobar-hole state involving a single-quark spin flip [25]. The mechanism being intrinsic in the change-over of the degrees of freedom, we expect the same phenomenon to hold in density as well as in temperature. The upshot of this line of argument is that the suppression of the vector coupling is inevitable as density approaches the critical density for chiral transition.

We believe that the different behavior of vector and scalar mean fields, the latter to be discussed below, follows from their different roles in QCD. With the vector this is made clearer in the lattice calculations of the quark number susceptibility which involves the vector interactions. In Brown and Rho [21], it is shown that as the description changes from hadronic to quark/gluonic at  $T \sim T_c$ , the critical temperature for chiral restoration, the vector interaction drops by an order of magnitude, much faster than the logarithmic decrease due to asymptotic freedom. We expect a similar feature in density, somewhat like in the renormalization-group analysis for the isovector vector meson  $\rho$  of Kim and Lee [18]. The scalar interaction, on the other hand, brings about chiral restoration and must become more and more important with increasing density as the phase transition point is approached.

2. From the empirical side, the most direct indication of the decoupling of the vector interaction is from the baryon flow [26] which is particularly sensitive to the vector interaction. The authors of [26] find a form factor of the form

$$f_V(\mathbf{p}) = \frac{\Lambda_V^2 - \frac{1}{6}\mathbf{p}^2}{\Lambda_V^2 + \mathbf{p}^2} \quad (19)$$

with  $\Lambda_V = 0.9$  GeV is required to understand the baryon flow. Connecting momenta with distances, one finds that this represents a cutoff at

$$R_{cutoff} \sim \frac{\sqrt{6}}{\Lambda_V} \sim 0.5 \text{ fm.} \quad (20)$$

Furthermore it is well known that the vector mean field of the Walecka model must be modified, its increase as  $E/m_N$  removed, at a scale of  $\sim 1$  GeV. The reason for this is presumably that inside of  $R \sim 0.5$  fm, the finite size of the solitonic nucleon must be taken into account. A repulsion still results, but it is scalar in nature as found in [27] and for which there are direct physical indications [28].

### 3.2 Kaons at GSI

The  $K^+$  and  $K^-$  energies in the top-down scenario are given by

$$\omega_{\pm} = \pm \frac{\omega_K}{m_K} V_K + \sqrt{k^2 + m_K^{*2}} \quad (21)$$

with  $V_K$  given in (23) below. Although at high densities it will decouple, the term linear in  $V_K$  that figures in the range correction in  $(\Sigma_{KN})_{eff}$  will give a slightly different effective mass for  $K^+$  and  $K^-$  before decoupling. Although the large distance vector mean field must arise from vector meson exchange, this must be cut off at a reasonably large distance, say,  $\sim 0.5$  fm as indicated by the baryon-flow mentioned above.

For the GSI experiments with temperature  $\sim 70$  MeV, the nucleon and kaon momenta are  $|\mathbf{p}_N| \sim 444$  MeV and  $|\mathbf{p}_K| \sim 322$  MeV, respectively, and

$$f_V(p) \sim 0.82. \quad (22)$$

We therefore propose to use

$$\frac{V_K}{m_K} \approx \frac{1}{3} f_V^2 V_N(p=0)/m_K \sim 0.15 \quad (23)$$

This is small. Furthermore we assume it to be constant above  $\rho_0$ . This assumption amounts to taking the vector coupling to drop as  $\sim 1/\rho$ .

The above arguments could be quantified by a specific model. For example, as alluded to above, the low-lying  $\rho$ - and  $\omega$ -excitations in the bottom-up model can be built up as  $N^*$ -hole excitations [24]. At higher densities, these provide the low-mass strength. One might attempt to calculate the coupling constants to these excitations in the constituent-quark (or quasi-quark) model, which as we have suggested would be expected to be more applicable at densities near chiral restoration. Riska and Brown [25] find the quark model couplings to be a factor  $\sim 2$  lower than the hadronic ones [29].

## 4 Schematic Model

### 4.1 First try

On the basis of our above considerations, a first try in transport calculation might use the vector potential with the Song scaling [16] as  $g_\omega^*/m_\omega^* = \text{constant}$  and the effective mass

$$m_K^{*2} \approx m_K^2 \left( 1 - \frac{\rho(\Sigma_{KN})_{eff}}{f^2 m_K^2} \right) \quad (24)$$

with  $\Sigma_{KN} = 400$  MeV and  $(\Sigma_{KN})_{eff}$  given by eq. (18). While as argued above the vector coupling will decouple at very high densities, as  $\omega_K$  drops, the vector potential will become less important even at moderate densities since the factor  $\omega_K/m_K$  comes into the coupling of the vector potential to the kaon. The resulting  $\omega_K^*$  is plotted in Fig.3 vs. density.

Our schematic model (24) gives roughly the same mass as used by Li and Brown [8] to predict kaon and antikaon subthreshold production at GSI. For  $\rho \sim 3\rho_0$  it gives  $m_{K^-}^* \sim 230$  MeV, less than half the bare mass. We predict somewhat fewer  $K^-$ -mesons because the (attractive)

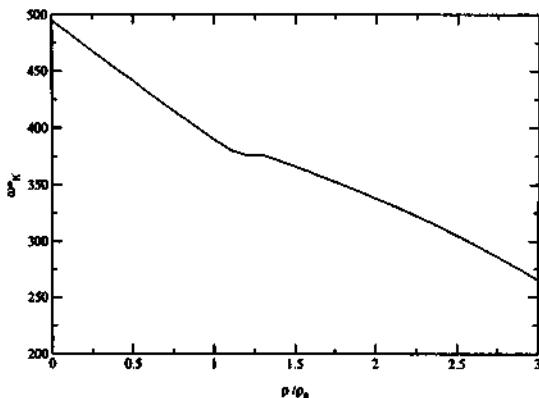


Figure 3:  $\omega_K^*$  vs.  $\rho/\rho_0$  for the schematic model (Section 4) with the vector decoupling eq.(23). The kink near  $\rho/\rho_0 \sim 1.2$  is an artifact of the oversimplified vector decoupling used for the calculation.

vector interaction is largely reduced if not decoupled. Cassing et al [30] have employed an  $m_K^*$  – somewhat lower than that given by eq.(24) to describe a lot of data.

The experimental data verify that our description is quite good up to the densities probed, i.e.,  $\sim 3 \rho_0$ . In order to go higher in density, we switch to our top-down description which through eq.(10) involves  $g_\sigma^2/m_\sigma^{*2}$ . Although the scalar interaction could have roughly the same form factor as the vector, cutting it off at  $\sim 0.5$  fm as mentioned above, we believe that this will be countered by the dropping scalar mass  $m_\sigma^*$  which must go to zero at chiral restoration (viewed as an order parameter). Treating the scalar interaction linearly as a fluctuation (as in (10)) cannot be expected to be valid all the way to chiral restoration but approaching the latter the  $\sigma$ -particle becomes the “dilaton” in the sense of Weinberg’s “mended symmetry” [31, 32] with mass going to zero (in the chiral limit) together with the pion.

At high densities at which the vector interaction decouples, the  $K^+$  and  $K^-$  will experience nearly the same very strong attractive interactions. This can be minimally expressed through the effective mass  $m_K^*$ . At low densities where the vector potential not only comes into play but slightly predominates over the scalar potential, the  $K^+$  will have a small repulsive interaction with nucleons. It is this interaction, extrapolated without medium effects by Bratkovskaya et al [4] which gives the long equilibration time of 40 fm/c. However, clearly the medium effects will change this by an order of magnitude.

#### 4.2 Implication on kaon condensation and maximum neutron-star mass

While in heavy-ion processes, we expect that taking  $m_\sigma^*$  to zero (or nearly zero in the real world) is a relevant limiting process, we do not have to take  $m_K^*$  to zero for kaon condensation in neutron stars, since the  $K^-$ -mass  $m_K^*$  must be brought down only to the electron chemical potential  $\mu_e \simeq E_F(e)$ , the approximate equality holding because the electrons are highly de-

generate. We should mention that it has been suggested that the electron chemical potential  $\mu_e$  could be kept low by replacing electrons plus neutrons by  $\Sigma^-$  hyperons (or more generally by exploiting Pauli exclusion principle with hyperon introduction) in neutron stars [33]. In this case, the  $\mu_e$  might never meet  $m_K^*$ .

Hyperon introduction may or may not take place, but even if it does, the scenario will be more subtle than considered presently. To see what can happen, let us consider what one could expect from a naive extrapolation to the relevant density, i.e.,  $\rho \sim 3\rho_0$ , based on the best available nuclear physics. The replacement of neutron plus electron will take place if the vector mean field felt by the neutron is still high at that density. The threshold for that would be

$$E_F^n + V_N + \mu_e \simeq M_{\Sigma^-} + \frac{2}{3}V_N + S_{\Sigma^-} \quad (25)$$

where  $E_F^n$  is the Fermi energy of the neutron,  $M_{\Sigma^-}$  the bare mass of the  $\Sigma^-$  and the  $S_{\Sigma^-}$  the scalar potential energy felt by the  $\Sigma^-$ . Here we are simply assuming that the two non-strange quarks of the  $\Sigma^-$  experience 2/3 of the vector mean field felt by the neutron. Extrapolating the FTS1 theory [17]<sup>10</sup> and taking into account in  $V_N$  the effect of the  $\rho$ -meson using vector dominance, we find  $E_F^n + V_N \sim 1064$  MeV at  $\rho \approx 3\rho_0$ . From the extended BPAL 32 equation of state with compression modulus 240 MeV [34], the electron chemical potential comes out to be  $\mu_e \simeq 214$  MeV. So the left-hand side of (25) is  $E_F^n + V_N + \mu_e \sim 1278$  MeV. For the right-hand side, we use the scalar potential energy for the  $\Sigma^-$  at  $\rho \approx 3\rho_0$  estimated by Brown, Lee and Rapp [35] to find that  $M_{\Sigma^-} + \frac{2}{3}V_N + S_{\Sigma^-} \sim 1240$  MeV. The replacement of neutron plus electron by  $\Sigma^-$  looks favored but only slightly.

What is the possible scenario on the maximum neutron star mass if we continue assuming that the calculation we made here can be trusted? A plausible scenario would be as follows.  $K^-$ -condensation supposedly occurs at about the same density and both the hyperonic excitation (in the form of  $\Sigma N^{-1}$  – where  $N^{-1}$  stands for the nucleon hole – component of the “kaesobar” [35]) and  $K^-$ -condensation would occur at  $T \sim 50$  MeV relevant to the neutron-star matter. Now if as is likely the temperature is greater than the difference in energies between the two possible phases, although the hyperons will be more important initially than the kaons, all of the phases will enter more or less equally in constructing the free energy of the system. In going to higher density the distribution between the different phases will change in order to minimize the free energy. Then it is clear that dropping from one minimum to another, the derivative of the free energy with density – which is just the pressure – will decrease as compared with the pressure from any single phase. This would imply that the maximum neutron star mass calculated with either hyperonic excitation or kaon condensation alone must be greater than the neutron star mass calculated with inclusion of both.

The story will be quite different if the vector field decouples. We showed in Section 3.1 that the isoscalar vector mean field must drop by a factor  $\gtrsim 9$  in the change-over from nucleons to quasiquarks as variables as one approaches the chiral restoration density. Hyperons will disappear during this drop. It is then highly likely that the kaon will condense before chiral restoration and that the kaon condensed phase will persist through the relevant range of densities which determine the maximum neutron star mass.

<sup>10</sup>There is nothing that would suggest that the effective Lagrangian valid up to  $\rho \sim \rho_0$  will continue to be valid at  $\rho \sim 3\rho_0$  without addition of higher mass-dimension operators, particularly if the chiral critical point is nearby. So this exercise can be taken only as indicative.

## 5 Concluding Remarks

By now there is a general consensus that the light-quark hadrons must behave differently in medium than in free space. This is understood in terms of a vacuum change induced by medium à la QCD. In this paper, we are re-confirming this property by arguing that not only the kaon mass but also its coupling to vector mesons should drop in matter with density. In particular, with the introduction of medium effects the apparent equilibration found in strangeness production at GSI can be increased from the baryon number density of  $\sim \frac{1}{4}\rho_0$  up to the much more reasonable  $\sim 2\rho_0$ .

From the baryon flow analysis we have direct indications that the vector interaction decouples from the nucleon at a three-momentum of  $|p| \sim 0.9$  GeV/c or at roughly  $0.5 m_N c$ . In colliding heavy ions this is reached at a kinetic energy per nucleon of  $\sim \frac{1}{8}m_N c^2$  which means a temperature of 78 MeV when equated to  $\frac{3}{2}T$ . This is just the temperature for chemical freeze-out at GSI energies. We have given several theoretical arguments why the vector coupling should drop rapidly with density.

Once the vector mean field, which acts with opposite signs on the  $K^+$  and  $K^-$  mesons is decoupled, these mesons will feel the same highly attractive scalar meson field. Their masses will fall down sharply; e.g., from eq.(24) with proposed parameters,

$$\frac{m_K^*}{m_K} \sim 0.5 \quad (26)$$

at  $\rho \approx 3\rho_0$  and possibly further because of the dropping  $m_\sigma^*$ . The differing slopes of  $K^+$  and  $K^-$  with kinetic energy will then develop after chemical freeze-out, as suggested by Li and Brown [8].

In this paper we have focused on the phenomenon at GSI energies. Here the chief role that the dropping  $K^-$ -mass played was to keep the combination  $\mu_B + \omega_{K^-}^*$  nearly constant so that low freeze-out density in the thermal equilibration scenario became irrelevant for the  $K^+/K^-$ -ratio. We suggest that the same scenario applies to AGS physics, where the freeze-out density in the thermal equilibration picture comes out to be  $\sim 0.35\rho_0$  [36]. In fact, there is no discernible dependence on centrality in the  $K^-/K^+$ -ratios measured at 4A GeV, 6.4 GeV, 8.4 GeV and 10.8A GeV [37]. From this it follows either that the ratio of produced  $K^-$  to hyperons is nearly independent of density or that the negative strangeness equilibrates down to a lower freeze-out density and then disperses. Given the relative weakness of the strange interactions, we believe the former to be the case. In fact, we suggest that near constancy with multiplicity of the  $K^-/K^+$ -ratio found experimentally be used to determine temperature dependence of  $\omega_{K^-}^*$  in the region of temperatures reached in AGS physics. As was done for GSI energies, the temperature can be obtained from the inverse slopes of the kaon and antikaon distributions of  $p_\perp$  and then the temperature dependence of  $\omega_{K^-}^*$  can be added to the density dependence as in [38], in such a way that  $\mu_B + \omega_{K^-}^*$  stays roughly constant as function of density. At least this can be done in the low-density regime considered in [36] when the approximation of a Boltzmann gas is accurate enough to calculate  $\mu_B$ . Our “broad-band equilibration”; i.e., the production of the same, apparently equilibrated ratio of  $K^-$ -mesons to hyperons over a broad band of densities, avoids complications in the way in which the  $K^-$  degree of freedom is mixed into other degrees of freedom at low density [39]. Most of the  $K^-$ -production will take place at the higher densities, as shown in Fig. 1, where the degrees of freedom other than  $K^-$  have been sent up to higher energies by the Pauli principle.

Unless the electron chemical potential in dense neutron star matter is prevented from

increasing with density (as might happen if the repulsive  $\Sigma$ -nuclear interaction turns to attraction at  $\rho > \rho_0$ ), kaon condensation will take place *before* chiral restoration. This has several implications at and beyond chiral restoration. For instance, its presence would have influence on the conjectured color superconductivity at high density, in particular regarding its possible coexistence with Overhauser effects, skyrmion crystal and other phases with interesting effects on neutron star cooling.

The phenomenon of vector decoupling, if confirmed to be correct, will have several important spin-off consequences. The first is that it will provide a refutation of the recent claim [40] that in the mean-field theory with a kaon-nuclear potential given by the vector-exchange (Weinberg-Tomozawa) term – both argued to be valid at high density – kaon condensation would be pushed up to a much higher density than that relevant in neutron-star matter. Our chief point against that argument is that the vector decoupling and the different role of scalar fields in QCD (e.g., BR scaling) described in this article cannot be accessed by the mean field reasoning used in [40] or by any other standard nuclear physics potential models. The second consequence that could be of a potential importance to the interpretation of heavy-ion experiments is that if the vector coupling<sup>11</sup> rapidly diminishes with density, the strong-coupling perturbation calculation of the vector response functions used in terms of “melting” vector mesons to explain [41], e.g., the CERES dilepton data must break down as one approaches the decoupling point that is out of reach of strong-coupling perturbative methods. It in turn provides yet one more justification (in addition to what has been already offered [42]) to the quasi-particle description in dense matter encoded in BR scaling and exploited in this paper.

Finally we should stress that the scenario presented in this paper – which is anchored on Brown-Rho scaling – should not be considered as an alternative to a possible quark-gluon scenario currently favored by the heavy-ion community. It is more likely a sort of “dual” description of the same physics along the line that the Cheshire-Cat Principle [11] embodies that would continue to apply at higher energies.

### Acknowledgments

The initial part of this work was done while we were visiting Korea Institute for Advanced Study whose hospitality is acknowledged. We are grateful for comments from Bengt Friman, Avraham Gal and Chang-Hwan Lee. One of the authors (GEB) is indebted to Elena Bratkovskaya, Peter Braun-Munzinger, Wolfgang Cassing, Che Ming Ko, Helmut Oeschler and Krzysztof Redlich for useful discussion and criticism. GEB and CS were supported by the US Department of Energy under Grant No. DE-FG02-88 ER40388. The work of CS was partly supported by Korea Science and Engineering Foundation.

### Appendix A

In this appendix we show how our argument that gives a correct  $K^+/K^-$  ratio can reproduce the  $K^+/\pi^+$  ratio. Let us leave  $T = 70$  MeV and choose  $\rho \sim 2\rho_0$  as educated guesses. We

<sup>11</sup>While the decoupling of the isoscalar vector interaction near the chiral restoration is highly plausible by the disappearance of a factor of 9 in the change-over from hadrons to quarks, the decoupling of the isovector vector interaction seems to be a lot more complex as indicated by the RG analysis of Kim and Lee [18]. What is not difficult to see is that their decoupling should occur at the same point as indicated in the quark number susceptibility [21].

are thereby increasing the equilibration density by a factor  $\sim 8$ . We then calculate the baryon density for this  $\mu_B$  and  $T = 70$  MeV and find  $\rho = 2\rho_0$  which checks the consistency.

According to Brown et al [38] the  $K^+$  production under these conditions will come chiefly from  $BB \rightarrow NAK$ , with excited baryon states giving most of the production. From the solid curve for  $\rho/\rho_0 = 2$  in fig.9 of [38] we find  $\langle\sigma v\rangle \sim 2 \times 10^{-3}$  mb =  $2 \times 10^{-4}$  fm $^2$ . The rate equation reads

$$\delta\Psi_K = \frac{1}{2}(\sigma_{BB}^{BYK}v_{BB})n_B^2 \simeq \frac{1}{2}(2 \times 10^{-4} \text{ fm}^2)\rho_B^2 = dn/dt = \frac{1}{9} \times 10^{-4} \text{ fm}^{-4} \quad (\text{A.1})$$

where  $B$ ,  $Y$  and  $K$  stand, respectively, for baryon, hyperon and kaon. For this, we have taken  $(\sigma_{BB}^{BYK}v_{BB})$  from fig.9 of [38] and  $\rho = 2\rho_0 = \frac{1}{3}$  fm $^{-3}$ . Choosing a time  $t = 10$  fm/c we obtain

$$n_{K^+} \simeq \delta\Psi_K t = \frac{2}{9} \times 10^{-3} \text{ fm}^{-3}. \quad (\text{A.2})$$

Now equilibrated pions have a density

$$n_\pi = 0.37(T/197\text{MeV})^3 \text{ fm}^{-3} = 0.016 \text{ fm}^{-3} \quad (\text{A.3})$$

for  $T = 70$  MeV. From (A.2) and (A.3) we get

$$n_{K^+}/n_{\pi^+} \simeq 0.0069 \quad (\text{A.4})$$

which is slightly below the “equilibrated value” of 0.0084 of Table 1 of Cleymans et al [3]. Production of  $K^+$  by pions may increase our number by  $\sim 25\%$  [43].

Our discussion of  $K^+$  production in this Appendix is in general agreement with earlier works by Ko and collaborators [43, 44]. In fact, if applied at the quark level, the vector mean field is conserved through the production process in heavy-ion collisions, so it affects only the strangeness condensation in which there is time for strangeness non-conservation. These earlier works establish that at the GSI energies the  $K^+$  content remains roughly constant once the fireball has expanded to  $\sim 2\rho_0$ , so that in this sense one can consider this as a chemical freeze-out density.

It should be noted that in the papers [30, 38, 43, 44] and others, the net potential – scalar plus vector – on the  $K^+$ -meson is slightly attractive at  $\rho \sim 2\rho_0$  even though the repulsive vector interaction is not decoupled. A hint for such change-over was noted already at nuclear matter density by Friedman, Gal and Mares [45]. Since in our top-down description the vector interaction can be thought of as applied to the quark (matter) field in the  $K^+$ , the total vector field on the initial components of a collision is then the same as on the final ones, so the vector mean fields have effectively no effect on the threshold energy of that process. Our proposal in this paper is that the vector mean field on the  $K^+$  should be below the values used by the workers in [30, 38, 43, 44] due to the decoupling. However, in comparison with [30], our total potential on the  $K^+$  at  $\rho = 2\rho_0$  is  $\sim -85$  MeV, as compared with their  $\sim -30$  MeV. We have not redone the calculation of [38] to take into account this difference.

## References

- [1] R. Hagedorn, CERN Yellow Report 71-12 (1971).

- [2] J. Cleymans, D. Elliot, A. Keränen and E. Suhonen, Phys. Rev. **C57** (1998) 3319.
- [3] J. Cleymans, H. Oeschler and K. Redlich, Phys. Rev. **C59** (1999) 1663; Phys. Lett. **485** (2000) 27.
- [4] E.L. Bratkovskaya, W. Cassing, C. Greiner, M. Effenberger, U. Mosel and A. Sibertsev, nucl-th/0001008, Nucl. Phys. **A675** (2000) 661.
- [5] P. Koch, B. Müller and J. Rafelski, Phys. Rep. **142** (1986) 167.
- [6] P. Braun-Munzinger, J. Stachel, J.P. Wessels, and N. Xu, Phys. Lett. **B365** (1996) 1.
- [7] M. Menzel et al., nucl-ex/0010013, to be published in Phys. Lett.; M. Menzel, University of Marburg PhD thesis, 2000; D. Best et al, Nucl. Phys. **A625** (1997) 307.
- [8] G.Q. Li and G.E. Brown, Phys. Rev. **C58** (1998) 1698.
- [9] G.E. Brown, Buballa and M. Rho, Nucl. Phys. **A609** (1996) 579.
- [10] G.E. Brown and M. Rho, Nucl. Phys. **A506** (1996) 503.
- [11] See, for an extensive treatment, the monograph by M.A. Nowak, M. Rho and I. Zahed, *Chiral Nuclear Dynamics* (World Scientific, Singapore, 1996).
- [12] G.E. Brown and M. Rho, Phys. Rev. Lett. **66** (1991) 2720.
- [13] G.E. Brown, K. Kubodera and M. Rho, Phys. Lett. **B192** (1987) 273.
- [14] K. F. Liu, S. J. Dong, T. Draper, D. Leinweber, J. Sloan, W. Wilcox and R. M. Woloshyn, Phys. Rev. **D59** (1999) 112001.
- [15] C.-H. Lee, Phys. Rep. **275** (1996) 256.
- [16] C. Song, Phys. Rep., to appear; nucl-th/0006030; C. Song, G. E. Brown, D.-P. Min, and M. Rho, Phys. Rev. C **56** (1997) 2244.
- [17] R.J. Furnsthal, H.-B. Tang and S.D. Serot, Phys. Rev. **C52** (1995) 1368.
- [18] Y. Kim and H.K. Lee, hep-ph/9905268; Phys. Rev. **C62** (2000) 037901.
- [19] M. Harada and K. Yamawaki, hep-ph/0010207.
- [20] H. Georgi, Phys. Rev. Lett. **63** (1989) 1917; Nucl. Phys. **B331** (1990) 217; M. Harada and K. Yamawaki, Phys. Rev. Lett. **83** (1999) 3374.
- [21] G.E. Brown and M. Rho, Phys. Rep. **269** (1996) 333.
- [22] S. Weinberg, Phys. Rev. Lett. **65** (1990) 1181.
- [23] S. Peris, Phys. Lett. **B268** (1991) 415.
- [24] Y. Kim, R. Rapp, G.E. Brown and M. Rho, Phys. Rev. **C62** (2000) 015202.
- [25] D.O. Riska and G.E. Brown, Nucl. Phys. A, in press, hep-ph/0005049.

- [26] P.K. Sahu, A. Hombach and W. Cassing, Nucl. Phys. **A640** (1998) 493; P.K. Sahu, W. Cassing, U. Mosel and A. Ohnishi, Nucl. Phys. **A672** (2000) 376.
- [27] V. Vento, M. Rho and G.E. Brown, Phys. Lett. **B103** (1981) 285.
- [28] K. Holinde, AIP Conference, Particles and Fields Series 43, Proc. LAMPF Workshop on  $(\pi, K)$  Physics, eds. B.F. Gibson, W.R. Gibbs and M.B. Johnson, Los Alamos, 1990.
- [29] B. Friman, M. Lutz and G. Wolf, nucl-th/0003012.
- [30] W. Cassing, E.L. Bratkovskaya, U. Mosel, S. Teis and A. Sibertsev, Nucl. Phys. **A614** (1997) 415.
- [31] S. Weinberg, in *Salamfestschrift: A Collection of Talks*, World Scientific Series in 20th Century Physics, eds. A. Ali, J. Ellis and S. Randjabar-Daemi (World Scientific, 1993).
- [32] S. Beane and U. van Kolck, Phys. Lett. **B328** (1994) 137.
- [33] N.K. Glendenning, Ap. J. **293** (1985) 470; *Compact stars* (Springer, New York, 1996); nucl-th/0009082.
- [34] M. Prakash, I. Bambaci, M. Prakash, P.J. Ellis, J.M. Lattimer and R. Knorren, Phys. Rep. **280** (1997) 1.
- [35] G.E. Brown, C.-H. Lee and R. Rapp, Nucl. Phys. **A639** (1998) 455c.
- [36] P. Braun-Munzinger, J. Stachel, J.P. Wessels and N. Xu, Phys. Lett. **B344** (1995) 43.
- [37] J. Dunlop, MIT PhD thesis, 1999; L. Ahle et al., Phys. Lett. **B490** (2000) 53.
- [38] G.E. Brown, C.M. Ko, Z.G. Wu and L.H. Xia, Phys. Rev. **C43** (1991) 1881.
- [39] A. Ramos and E. Oset, Nucl. Phys. **A671** (2000) 481.
- [40] J. Carlson, H. Heiselberg and V. Pandharipande, nucl-th/9912043.
- [41] See for review R. Rapp and J. Wambach, hep-ph/9909229.
- [42] M. Rho, “Effective field theories, Landau-Migdal Fermi-liquid theory and effective chiral Lagrangians for nuclear matter,” nucl-th/0007060.
- [43] W.S. Chung, G.Q. Li and C.M. Ko, Nucl. Phys. **A625** (1997) 347.
- [44] X.S. Fang, C.M. Ko, G.Q. Li and Y.M. Zheng, Nucl. Phys. **A575** (1994) 766.
- [45] E. Friedman, A. Gal and J. Mares, Nucl. Phys. **A625** (1997) 272.

This page is intentionally left blank

## Bibliography

1. Bethe, H.A., Brown, G.E., Applegate, J., and Lattimer, J.M. (1979) "Equation of state in the gravitational collapse of stars", *Nuclear Physics A* **324**, 487–533.
2. Bethe, H.A., and Brown, G.E. (1985) "How a supernova explodes", *Scientific American*, May 1985, **252**, 60–68.
3. Brown, G.E., and Weingartner, J.C. (1994), "Accretion onto and radiation from the compact object formed in SN 1987A", *Astrophysical Journal* **436**, 843–847.
4. Brown, G.E., and Bethe, H.A. (1994), "A scenario for a large number of low-mass black holes in the Galaxy", *Astrophysical Journal* **423**, 659–664.
5. Brown, G.E. (1995), "Neutron star accretion and binary pulsar formation", *Astrophysical Journal* **440**, 270–279.
6. Research News, "How collapsing stars might hide their tracks in black holes", *Science* **261** (1993), 831–832.
7. "Mystery of the missing star", *Discover*, December 1996, 111–115.
8. Bethe, H.A., and Brown, G.E. (1995), "Observational constraints on the maximum neutron star mass", *Astrophysical Journal* **445**, L129–L132.
9. Brown, G.E., Weingartner, J., and Wijers, R.A.M.J. (1996), "On the formation of low-mass black holes in massive binary stars", *Astrophysical Journal* **463**, 297–304.
10. Wettig, T., and Brown, G.E. (1996), "The evolution of relativistic binary pulsars", *New Astronomy* **1**, 17–34.
11. Brown, G.E. (1997), "Supernova explosions, black holes and nucleon stars", *Physikalische Blätter* **53**, 671–674.
12. Bethe, H.A., and Brown, G.E. (1998), "Evolution of binary compact objects that merge", *Astrophysical Journal* **506**, 780–789.
13. Bethe, H.A., and Brown, G.E. (1999), "Contribution of high-mass black holes to mergers of compact binaries", *Astrophysical Journal* **517**, 318–327.
14. Brown, G.E., Lee, C.-H., and Bethe, H.A. (1999), "The formation of high-mass black holes in low-mass X-ray binaries", *New Astronomy* **4**, 313–323.
15. Brown, G.E., Lee, C.-H., Wijers, R.A.M.J., and Bethe, H.A. (2000), "Evolution of black holes in the galaxy", *Physics Reports* **333–334**, 471–504.
16. Lee, H.K., Wijers, R.A.M.J., and Brown, G.E. (2000), "The Blandford-Znajek process as a central engine for a gamma-ray burst", *Physics Reports* **325**, 83–114.
17. Brown, G.E., Lee, C.-H., Wijers, R.A.M.J., Lee, H.K., Israeliyan, G., and Bethe, H.A. (2000), "A theory of gamma-ray bursts", *New Astronomy* **5**, 191–210.
18. Brown, G.E., Lee, C.-H., and Bethe, H.A. (2000), "Hypercritical advection-dominated accretion flow", *Astrophysical Journal* **541**, 918–923.
19. Brown, G.E., Lee, C.-H., Portegies Zwart, S.F., and Bethe, H.A. (2001), "Evolution of neutron star, carbon-oxygen white dwarf binaries", *Astrophysical Journal* **547**, 345–354.

20. Brown, G.E., Lee, C.-H., and Tauris, T. (2001), "Formation and evolution of black hole X-ray transient systems", *New Astronomy* **6**, 331–338.
21. Brown, G.E., Heger, A., Langer, N., Lee, C.-H., Wellstein, S., and Bethe, H.A. (2001), "Formation of high mass X-ray black hole binaries", *New Astronomy* **6**, 457–470.
22. McLaughlin, G.C., Wijers, R.A.M.J., Brown, G.E., and Bethe, H.A. (2002), "Broad and shifted iron-group emission lines in gamma-ray bursts as tests of the hypernova scenario", *Astrophysical Journal* **567**, 454–462.
23. Lee, C.-H., Brown, G.E., and Wijers, R.A.M.J. (2002), "Discovery of a black-hole mass-period correlation in soft X-ray transients and its implication for gamma-ray burst and hypernova mechanisms", *Astrophysical Journal* **575**, 996–1006.
24. [Appendix A] Lee, C.-H. (1996), "Kaon condensation in dense stellar matter", *Physics Reports* **275**, 255–341.
25. [Appendix B] Li, G.Q., Lee, C.-H., and Brown, G.E. (1997), "Kaon production in heavy-ion collisions and maximum mass of neutron stars", *Physical Review Letters* **79**, 5214–5217.
26. [Appendix C] Li, G.Q., and Brown, G.E. (1998), " $K^-/K^+$  ratios in relativistic heavy-ion collisions", *Physical Review C* **58** (1998) 1698–1705.
27. [Appendix D] Brown, G.E., Rho, M., and Song, C. (2001), "Strangeness equilibration at GSI energies", *Nuclear Physics A* **690** (2001) 184c–200c.

# FORMATION AND EVOLUTION OF BLACK HOLES IN THE GALAXY

In published papers H. A. Bethe and G. E. Brown worked out the collapse of large stars and supernova explosions. They went on to evolve binaries of compact stars, finding that in the standard scenario the first formed neutron star always went into a black hole in common envelope evolution. C.-H. Lee joined them in the study of black hole binaries and gamma ray bursts. They found the black holes to be the fossils of the gamma ray bursts. From their properties they could reconstruct features of the burst and of the accompanying hypernova explosions.

This invaluable book contains 23 papers on astrophysics, chiefly on compact objects, written over 23 years. The papers are accompanied by illuminating commentary. In addition there is an appendix on kaon condensation which the editors believe to be relevant to the equation of state in neutron stars, and to explain why black holes are formed at relatively low masses.

