A dying universe: the long-term fate and evolution of astrophysical objects

Fred C. Adams* and Gregory Laughlin†

Physics Department, University of Michigan, Ann Arbor, Michigan 48109

This paper outlines astrophysical issues related to the long-term fate of the universe. The authors consider the evolution of planets, stars, stellar populations, galaxies, and the universe itself over time scales that greatly exceed the current age of the universe. Their discussion starts with new stellar evolution calculations which follow the future evolution of the low-mass (M-type) stars that dominate the stellar mass function. They derive scaling relations that describe how the range of stellar masses and lifetimes depends on forthcoming increases in metallicity. They then proceed to determine the ultimate mass distribution of stellar remnants, i.e., the neutron stars, white dwarfs, and brown dwarfs remaining at the end of stellar evolution; this aggregate of remnants defines the "final stellar mass function." At times exceeding ~1–10 trillion years, the supply of interstellar gas will be exhausted, yet star formation will continue at a highly attenuated level via collisions between brown dwarfs. This process tails off as the galaxy gradually depletes its stars by ejecting the majority and driving a minority toward eventual accretion onto massive black holes. As the galaxy disperses, stellar remnants provide a mechanism for converting the halo dark matter into radiative energy. Posited weakly interacting massive particles are accreted by white dwarfs, where they subsequently annihilate with each other. Thermalization of the decay products keeps the old white dwarfs much warmer than they would otherwise be. After accounting for the destruction of the galaxy, the authors consider the fate of the expelled degenerate objects (planets, white dwarfs, and neutron stars) within the explicit assumption that proton decay is a viable process. The evolution and eventual sublimation of these objects is dictated by the decay of their constituent nucleons, and this evolutionary scenario is developed in some detail. After white dwarfs and neutron stars have disappeared, galactic black holes slowly lose their mass as they emit Hawking radiation. This review finishes with an evaluation of cosmological issues that arise in connection with the long-term evolution of the universe. Special attention is devoted to the relation between future density fluctuations and the prospects for continued large-scale expansion. The authors compute the evolution of the background radiation fields of the universe. After several trillion years, the current cosmic microwave background will have redshifted into insignificance; the dominant contribution to the radiation background will arise from other sources, including stars, dark-matter annihilation, proton decay, and black holes. Finally, the authors consider the dramatic possible effects of a nonzero vacuum energy density. [S0034-6861(97)00202-X]

CONTENTS

*fca@umich.edu
†gpl@boris.physics.lsa.umich.edu

I. INTRODUCTION

The long-term future of the universe and its contents is a topic of profound scientific and philosophical importance. With our current understanding of physics and astrophysics, many of the questions regarding the ultimate fate of the universe can now be quantitatively addressed. Our goal is to summarize and continue the development of a quantitative theory of the future.

Investigations of the early universe at both accessible and inaccessible energies have become commonplace, and a great deal of progress within this discipline has been made (see, for example, Weinberg, 1972, 1977; Kolb and Turner, 1990; Linde, 1990; Peebles, 1993; Zuckerman and Malkan, 1996). On the other hand, relatively little work has focused on the future of the universe. The details of the fiery denouement in store for a closed universe have been outlined by Rees (1969), whereas an overview of the seemingly more likely scenario in which the universe is either open or flat, and hence expands forever, was set forth in the seminal paper "Time Without End" (Dyson, 1979). The development of an open universe was also considered in detail by Islam (1977, 1979). The spirit of Rees, Islam, and Dyson's work inspired several followup studies (see also Rees, 1981). The forthcoming evolution of very-low-mass stars has been discussed in general terms by Salpeter (1982). The effects of matter annihilation in the late universe were studied (Page and McKee, 1981a, 1981b), and some aspects of proton decay have been explored (Dicus *et al.*, 1982; Turner, 1983). Finally, the possibility of self-reproducing inflationary domains has been proposed (Linde, 1988). In general, however, the future of the universe has not been extensively probed with rigorous calculations.

Because the future of the universe holds a great deal of intrinsic interest, a number of recent popular books have addressed the subject (e.g., Davies, 1994; Dyson, 1988; Barrow and Tipler, 1986; Poundstone, 1985). Authors have also grappled with the intriguing prospects for continued life, both human and otherwise, in the far future (e.g., Dyson, 1979; Frautschi, 1982; Barrow and Tipler, 1986; Linde, 1988, 1989; Tipler, 1992; Gott, 1993; Ellis and Coule, 1994). Our aim, however, is to proceed in as quantitative a manner as possible. We apply known physical principles to investigate the future of the universe on planetary, stellar, galactic, and cosmic scales. The issue of life, however alluring, is not considered here.

In standard big-bang cosmology, evolutionary epochs are usually expressed in terms of the redshift. When considering the far future, however, time itself is often the more relevant evolutionary measure. The immense dynamic range of time scales τ involved in the subject suggests a convenient logarithmic unit of time η, defined by

$$\eta \equiv \log_{10}\left[\frac{\tau}{(1\ \mathrm{yr})}\right]. \tag{1.1}$$

We refer to a particular integer value of η as a "cosmological decade." For example, the current age of the universe corresponds to $\eta \approx 10$.

The article of faith inherent in our discussion is that the laws of physics are constant in time, at least over the range of time scales $10 < \eta < 100$ under consideration. There is no general guarantee that this assumption holds. Nevertheless, modern cosmology suggests that physical laws have held constant from the Planck time to the present, i.e., over cosmological decades spanning the range $-50 \leqslant \eta \leqslant 10$, and there is little reason to expect that they will not continue to do so. We also implicitly assume that all of the relevant physics is known (with full awareness of the fact that our version of the future will be subject to revision as physical understanding improves).

This paper is organized in roughly chronological order, moving from events in the relatively near future to events in the far future. In Sec. II, we discuss physical processes that affect conventional stellar evolution; these processes will take place in the time range $10 < \eta < 15$. In Sec. III, we discuss events that lead to the disruption and death of the galaxy; these processes unfold over a time range $15 < \eta < 25$. Marching further into time, in Sec. IV, we discuss the fate of stellar objects in the face of very-long-term processes, including proton decay ($30 < \eta < 40$) and Hawking radiation ($60 < \eta < 100$). In Sec. V, we broaden our scope and focus on the long-term evolution of the universe as a whole. We conclude, in Sec. VI, with a general overview of our results. Since physical eschatology remains in its infancy, we emphasize the major unresolved issues and point out possible avenues for further research.

II. THE END OF CONVENTIONAL STELLAR EVOLUTION

At the present epoch, stars are the cornerstone of astrophysics. Stars mediate the appearance and evolution of galaxies, stars are responsible for evolving the chemical composition of matter, and stars provide us with much of the information we have regarding the current state of the universe.

For the next several thousand Hubble times, conventionally evolving stars will continue to play the central role. We thus consider the forthcoming aspects of our current epoch, which we term the *Stelliferous Era*. In particular, the fact that the majority of stars have barely begun to evolve motivates an extension of standard stellar evolution calculations of very-low-mass stars to time scales much longer than the current age of the universe. We also discuss continued star formation within the galaxy and the final mass distribution of stellar remnants.

A. Lifetimes of main-sequence stars

Low-mass stars are by far the most commonplace (see, for example, Henry, Kirkpatrick, and Simons,

1994), and they live for a long time. To a working approximation, the main-sequence (core-hydrogen-burning) lifetime of a star depends on its mass through the relation

$$\tau_* = 10^{10}\ \text{yr}\left[\frac{M_*}{1M_\odot}\right]^{-\alpha}, \tag{2.1a}$$

where the index $\alpha \approx 3$–4 for stars of low mass. In terms of cosmological decades η, we obtain

$$\eta_* = 10 - \alpha\ \log_{10}[M_*/1M_\odot]. \tag{2.1b}$$

Thus, for example, $\eta_* \approx 13$ for a small star with $M_* = 0.1\ M_\odot$. Indeed, at the present time, only stars with masses $M_* > 0.8\ M_\odot$ have had time to experience significant post-main-sequence evolution. Hence, a large fraction,

$$f \equiv \frac{\int_{M_{\min}}^{0.8}(dN/dm)dm}{\int_{M_{\min}}^{M_{\max}}(dN/dm)dm} \sim 80\%, \tag{2.2}$$

of all stars ever formed have yet to experience any significant evolution (here, dN/dm is the mass distribution—see Sec. II.E). We are effectively still in the midst of the transient initial phases of the stelliferous epoch.

Very little consideration has been given to the post-main-sequence development of stars that are small enough to outlive the current age of the universe. An essay by Salpeter (1982) contains a qualitative discussion regarding the evolution of M stars (especially with respect to ^{3}He production), but detailed stellar evolutionary sequences have not been presented in the literature. Nevertheless, there is a sizable collection of papers that discuss the pre-main-sequence and main-sequence properties of very-low-mass stars (e.g., Kumar, 1963; Copeland, Jensen, and Jorgensen, 1970; Grossman and Graboske, 1971; D'Antona and Mazzitelli, 1985; Dorman, Nelson, and Chau, 1989). The best comprehensive family of models spanning the M dwarfs and brown dwarfs is probably that of Burrows *et al.* (1993). Those authors devote attention to the formative contraction phases, as well as the exact mass of the minimum-mass star (which for their input physics occurs at $M_* = 0.0767\ M_\odot$). Evolution beyond 20 billion years was not considered (see also Burrows and Liebert, 1993).

The dearth of information regarding the fate of the M dwarfs has recently been addressed (Laughlin, Bodenheimer, and Adams, 1996). We have performed a detailed series of stellar evolution calculations, which follow the pre-main-sequence through post-main-sequence evolution of late M dwarfs, yielding the following picture of what lies in store for the low-mass stars.

Newly formed stars containing less mass than $M_* \sim 0.25\ M_\odot$ are fully convective throughout the bulk of their structure. The capacity of these stars to mix entirely their contents has several important consequences. First, these late M stars maintain access to their entire initial reserve of hydrogen, greatly extending their lifetimes in comparison to heavier stars like the sun, which see their fuel supply constricted by stratified radiative cores. Second, as recognized by Salpeter (1982), full convection precludes the buildup of composition gradients which are ultimately responsible (in part) for a star's ascent up the red giant branch. The lowest-mass stars burn all their hydrogen into helium over an $\eta = 13$ time scale and then quietly fade from prominence as helium white dwarfs. This general evolutionary scenario is detailed in Fig. 1 (adapted from Laughlin *et al.*, 1996), which charts the path in the Hertzsprung-Russell (H-R) diagram followed by low-mass stars of several different masses in the range $0.08\ M_\odot \leq M_* \leq 0.25\ M_\odot$.

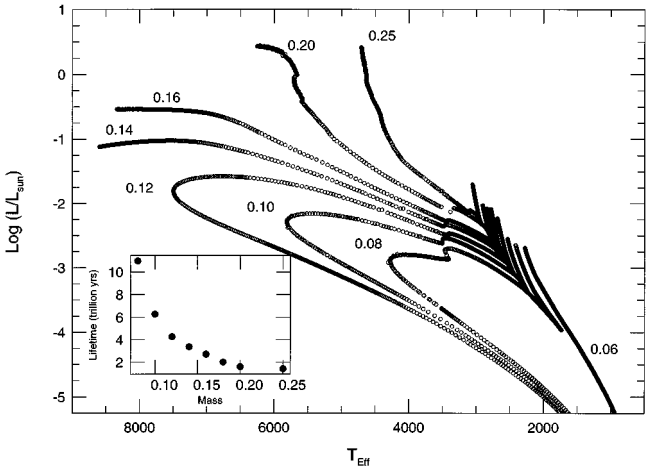

FIG. 1. The Hertzsprung-Russell diagram for low-mass stars for time scales much longer than the current age of the universe. The labeled curves show the evolutionary tracks for stars of varying masses, from 0.08 $M_\odot$ to 0.25 $M_\odot$, as well as the brown dwarf track for a substellar object with mass $M_* = 0.06\ M_\odot$. The inset figure shows the main-sequence lifetimes as a function of stellar mass.

Upon emerging from its parent cloud core, the lowest-mass star capable of burning hydrogen ($M_* \approx 0.08\ M_\odot$) descends the convective Hayashi track and arrives on the main sequence with a luminosity $L_* \sim 10^{-4}\ L_\odot$. The main-sequence phase is characterized by gradual prolonged increase in both luminosity and effective surface temperature as hydrogen is consumed. Due to the relatively low prevailing temperature in the stellar core ($T_c \approx 4\times10^6$ K), the proton-proton nuclear reaction chain is decoupled from statistical equilibrium, and the concentration of ^{3}He increases steadily until $\eta = 12.6$, at which time a maximum mass fraction of 16% ^{3}He has been attained. As the initial supply of hydrogen is depleted, the star heats up and contracts, burns the ^{3}He, increases in luminosity by a factor of 10, and more than doubles its effective temperature. After ~11 trillion years, when the star has become 90% ^{4}He by mass, a radiative core finally develops. The evolutionary time scale begins to accelerate, and hydrogen is exhausted relatively quickly in the center of the star. When nuclear burning within the modest resulting shell source can no longer provide the star's mounting energy requirements, the star begins to contract and cool and eventually becomes a helium white dwarf. Stars with masses up to $\sim 0.20\ M_\odot$ follow essentially this same evolutionary scenario. As stellar mass increases, radiative cores develop

sooner, and the stars perform increasingly dramatic blueward excursions in the H-R diagram.

A star with a slightly larger mass, $M_*=0.23\ M_\odot$, experiences the onset of a radiative core when the hydrogen mass fraction dips below 50%. The composition gradients that ensue are sufficient to briefly drive the star to lower effective temperature as the luminosity increases. In this sense, stars with mass $M_*=0.23\ M_\odot$ represent the lowest-mass objects that can become conventional "red giants." At these low masses, however, the full giant phase is not completed. Stars with initial mass $M_*<0.5\ M_\odot$ will be unable to generate the high central temperatures ($T_c\sim10^8$ K) required for the helium flash; these stars abort their ascent up the giant branch by veering to the left in the H-R diagram in the manner suggested by Fig. 1.

The steady increases in luminosity experienced by aging M dwarfs will have a considerable effect on the mass-to-light ratio of the galaxy. For example, as a $0.2\ M_\odot$ star evolves, there is a relatively fleeting epoch (at $\eta\approx12$) during which the star has approximately the same radius and luminosity as the present-day sun. Given that M dwarfs constitute the major fraction of all stars, the total luminosity of the galaxy will remain respectably large, $L_{\rm gal}\sim10^{10}\ L_\odot$, at this future date. This luminosity is roughly comparable to the characteristic luminosity $\mathcal{L}^*=3.4\times10^{10}\ L_\odot$ displayed by present-day galaxies (Mihalas and Binney, 1981).

B. Forthcoming metallicity effects

The foregoing evolutionary calculations assumed a solar abundance set. In the future, the metallicity of the galaxy will steadily increase as stars continue to process hydrogen and helium into heavy elements. It is thus useful to determine the effects of these metallicity increases.

1. Stellar lifetimes versus metallicity

First, it is possible to construct a simple scaling relation that clarifies how stellar lifetimes τ_* depend on the metallicity Z. The stellar lifetime is roughly given by amount of fuel available divided by the rate of fuel consumption, i.e.,

$$\tau_*\sim M_* X/L, \tag{2.3}$$

where M_* is the stellar mass and X is the hydrogen mass fraction. For relatively-low-mass stars, the luminosity L obeys the scaling relation

$$L\sim\kappa_0^{-1}\mu^{7.5}M_*^{5.5}, \tag{2.4}$$

where μ is the mean molecular weight of the star and where κ_0 is the constant of proportionality appearing in the usual opacity relation for stars (Clayton, 1983). Thus, for a given stellar mass M_*, the lifetime scales according to

$$\tau_*\sim\kappa_0 X\mu^{-7.5}. \tag{2.5}$$

To evaluate the stellar lifetime scaling relation, one needs to know how the parameters κ_0, X, and μ vary with metallicity. The opacity constant κ_0 is roughly linearly dependent on the metallicity, i.e.,

$$\kappa_0\sim Z. \tag{2.6}$$

The mean molecular weight μ can be approximately written in the form

$$\mu\approx\frac{2}{(1+3X+Y/2)}, \tag{2.7}$$

where Y is the helium mass fraction (e.g., see Clayton, 1983). By definition, the mass fractions obey the relation

$$X+Y+Z=1. \tag{2.8}$$

Finally, for this simple model, we write the helium abundance Y in the form

$$Y=Y_P+fZ, \tag{2.9}$$

where Y_P is the primordial abundance and the factor f accounts for the increase in helium abundance as the metallicity increases. Big-bang nucleosynthesis considerations indicate that $Y_P\approx1/4$ (Kolb and Turner, 1990), whereas $f\approx2$ based on the solar enrichment in Y and Z relative to the primordial values. Combining the above results, we obtain a scaling relation for the dependence of stellar lifetimes on metallicity,

$$\tau_*\sim Z(1-aZ)(1-bZ)^{7.5}, \tag{2.10}$$

where we have defined constants $a\equiv4(1+f)/3\approx4$ and $b\equiv8/9+20f/27\approx64/27$. This result implies that stellar lifetimes have a maximum value. In particular, we find that stars born with metallicity $Z\approx0.04$ live the longest. For larger values of Z, the reduction in nuclear fuel and the change in composition outweigh the lifetime-extending decrease in luminosity arising from the increased opacity.

A recent set of galactic chemical evolution calculations (Timmes, 1996) have probed far into the stelliferous epoch. The best indications suggest that the galactic abundance set will approach an asymptotically constant composition ($X\sim0.2$, $Y\sim0.6$, and $Z\sim0.2$) over a time scale $\eta\sim12$. As a consequence, any generations of stars formed after $\eta\sim12$ will suffer significantly shorter lifetimes than the theoretical maximum implied by Eq. (2.10).

2. Stellar masses versus metallicity

The maximum stable stellar mass decreases as metallicity increases. On the main sequence, the maximum possible mass is reached when the star's radiation pressure comes to dominate the thermal (gas) pressure within the star. Here, we introduce the usual ansatz that the total pressure at the center of the star can be written in the form $P_C=P_R+P_g$, where the thermal gas pressure is given by the fraction $P_g=\beta P_C$ and, similarly, $P_R=(1-\beta)P_C$. Using the ideal-gas law for the thermal pressure and the equation of state for a gas of photons, we can write the central pressure in the form

$$P_C=\left[\frac{3}{a}\frac{(1-\beta)}{\beta^4}\right]^{1/3}\left[\frac{k\rho_C}{\mu m_P}\right]^{4/3}, \tag{2.11}$$

where k is the Boltzmann constant and a is the radiation constant. The quantity μ is again the mean molecular weight and can be written in the form of Eq. (2.7). In hydrostatic equilibrium, the central pressure required to support a star of mass M_* can be expressed as

$$P_C \approx \left[\frac{\pi}{36}\right]^{1/3} G M_*^{2/3} \rho_C^{4/3}, \tag{2.12}$$

where ρ_C is the central density (see Phillips, 1994).

Equating the above two expressions (2.11) and (2.12), we can solve for the mass to find

$$M_* = \left[\frac{108}{\pi a}\frac{(1-\beta)}{\beta^4}\right]^{1/2}\left[\frac{k}{\mu m_P}\right]^2 G^{-3/2} \approx 40\ M_\odot \mu^{-2}, \tag{2.13}$$

where we have set β=1/2 to obtain the numerical value. The maximum stellar mass thus depends somewhat sensitively on the mean molecular weight μ, which in turn is a function of the metallicity. By applying the approximations (2.7), (2.8), and (2.9), one can write the maximum mass in the form

$$\begin{aligned} M_* &= 40\ M_\odot \{(2-5Y_P/4)-(3+5f/2)Z/2\}^2 \\ &\approx 114\ M_\odot (1-2.4Z)^2. \end{aligned} \tag{2.14}$$

Thus, for the expected asymptotic value of the metallicity, Z=0.2, the maximum-mass star is only $M_* \approx 30\ M_\odot$.

The continuously increasing metallicity of the interstellar medium will also have implications for low-mass stars. Higher metallicity leads to more effective cooling, which leads to lower temperatures, which in turn favors the formation of less massive stars (e.g., see the recent theory of the initial mass function by Adams and Fatuzzo, 1996). The initial mass function of the future should be skewed even more dramatically in favor of the faintest stars.

The forthcoming metallicity increases may also decrease the mass of the minimum-mass main-sequence star as a result of opacity effects (cf. the reviews of Stevenson, 1991; Burrows and Liebert, 1993). Other unexpected effects may also occur. For example, when the metallicity reaches several times the solar value, objects with mass $M_* = 0.04\ M_\odot$ may quite possibly halt their cooling and contraction and land on the main sequence when thick ice clouds form in their atmospheres. Such "frozen stars" would have an effective temperature of $T_* \approx 273$ K, far cooler than the current minimum-mass main-sequence stars. The luminosity of these frugal objects would be more than a thousand times smaller than the dimmest stars of today, with commensurate increases in longevity.

C. The fate of the Earth and the Sun

A popular and frequently quoted scenario for the demise of the Earth involves destruction through evaporation during the Sun's asymptotic giant branch (AGB) phase. As the Sun leaves the horizontal branch and expands to become an AGB star, its outer radius may swell to such an extent that the photospheric radius overtakes the current orbital radius of the Earth. If this state of affairs comes to pass, then two important processes will affect the Earth: (1) Evaporation of material due to the extreme heat, and (2) Orbital decay through frictional drag. This second process drives the Earth inexorably into the giant Sun, thereby increasing the efficacy of the evaporation process. Once the Earth finds itself inside the Sun, the time scale for orbital decay is roughly given by the time required for the expiring Earth to sweep through its mass, M_E, in solar material. This short time interval is given by

$$\tau = \frac{M_E}{\rho_\odot (\pi R_E^2) v_{\rm orbit}} \approx 50\ \text{yr}, \tag{2.15}$$

where $\rho_\odot \sim 10^{-6}$ g/cm^3 is the mass density of solar material at the photosphere, $R_E \approx 6370$ km is the radius of the Earth, and $v_{\rm orbit} \approx 30$ km/s is the orbital speed. Hence the demise of the Earth will befall it swiftly, even in comparison to the accelerated stellar evolution time scale inherent to the asymptotic giant branch. The Earth will be efficiently dragged far inside the Sun and vaporized in the fierce heat of the stellar plasma, its sole legacy being a small (0.01%) increase in the metallicity of the Sun's distended outer envelope.

Recent work suggests, however, that this dramatic scene can be avoided. When the Sun reaches a luminosity of $\sim 100 L_\odot$ on its ascent of the red giant branch, it will experience heavy mass loss through the action of strong stellar winds. Mass loss results in an increase in the orbital radii of the planets and can help the Earth avoid destruction. However, the actual amount of mass loss remains uncertain; estimates are based largely on empirical measurements (see Reimers, 1975), but it seems reasonable that the Sun will diminish to $\sim 0.70\ M_\odot$ when it reaches the tip of the red giant branch, and will end its AGB phase as a carbon white dwarf with mass $\sim 0.5\ M_\odot$. Detailed stellar evolution calculations for the Sun have been made by Sackmann, Boothroyd, and Kraemer (1993). In their best-guess mass-loss scenario, they find that the orbital radii for both the Earth and Venus increase sufficiently to avoid being engulfed during the AGB phase. Only with a more conservative mass-loss assumption, in which the Sun retains $0.83\ M_\odot$ upon arrival on the horizontal branch, does the solar radius eventually overtake the Earth's orbit.

D. Continued star formation in the galaxy

Galaxies can live only as long as their stars. Hence it is useful to estimate how long a galaxy can sustain normal star formation (see, for example, Shu, Adams, and Lizano, 1987) before it runs out of raw material. One would particularly like to know when the last star forms.

There have been many studies of the star formation history in both our galaxy and other disk galaxies (e.g., Roberts, 1963; Larson and Tinsley, 1978; Rana, 1991; Kennicutt, Tamblyn, and Congdon, 1994). Although many uncertainties arise in these investigations, the results can be roughly summarized as follows. The gas

depletion time τ_R for a disk galaxy is defined to be the current mass in gas, $M_{\rm gas}$, divided by the star formation rate SFR, i.e.,

$$\tau_R \equiv \frac{M_{\rm gas}}{SFR}. \tag{2.16}$$

For typical disk galaxies, this time scale is comparable to the current age of the universe; Kennicutt *et al.* (1995) cite a range $\tau_R \approx 5$–15 Gyr. The actual time scale for (total) gas depletion will be longer because the star formation rate is expected to decrease as the mass in gas decreases. For example, if we assume that the star formation rate is proportional to the current mass in gas, we derive a total depletion time of the form

$$\tau = \tau_R \ln[M_0/M_F], \tag{2.17}$$

where M_0 is the initial mass in gas and M_F is the final mass. For typical disk galaxies, the initial gas mass is $M_0 \sim 10^{10}\ M_\odot$ (see Table 5 of Kennicutt *et al.*, 1995). Thus, if we take the extreme case of $M_F = 1\ M_\odot$, the total gas depletion time is only $\tau \approx 23\tau_R \approx 120$–350 Gyr. In terms of cosmological decades, the gas depletion time becomes $\eta_D = 11.1$–11.5.

Several effects tend to extend the gas depletion time scale beyond this simple estimate. When stars die, they return a fraction of their mass back to the interstellar medium. This gas recycling effect can prolong the gas depletion time scale by a factor of 3 or 4 (Kennicutt *et al.*, 1995). Additional gas can be added to the galaxy through infall onto the galactic disk, but this effect should be relatively small (see the review of Rana, 1991); the total mass added to the disk should not increase the time scale by more than a factor of 2. Finally, if the star formation rate decreases more quickly with decreasing gas mass than the simple linear law used above, then the depletion time scale becomes correspondingly larger. Given these complications, we expect the actual gas depletion time will fall in the range

$$\eta_D = 12\text{–}14. \tag{2.18}$$

Thus, by the cosmological decade $\eta \approx 14$, essentially all normal star formation in galaxies will have ceased. Coincidentally, low-mass M dwarfs have life expectancies that are comparable to this time scale. In other words, both star formation and stellar evolution come to an end at approximately the same cosmological decade.

There are some indications that star formation may turn off even more dramatically than outlined above. Once the gas density drops below a critical surface density, star formation may turn off completely (as in elliptical and SO galaxies). The gas may be heated entirely by its slow accretion onto a central black hole.

These results indicate that stellar evolution is confined to a reasonably narrow range of cosmological decades. It is presumably impossible for stars to form and burn hydrogen before the epoch of recombination in the universe (at a redshift $z \sim 1000$ and hence $\eta \sim 5.5$). Thus significant numbers of stars will exist only within the range

$$5.5 < \eta < 14. \tag{2.19}$$

The current epoch ($\eta \sim 10$) lies near the center of this range of (logarithmic) time scales. On the other hand, if we use a linear time scale, the current epoch lies very near the beginning of the stelliferous era.

E. The final mass function

When ordinary star formation and conventional stellar evolution have ceased, all of the remaining stellar objects will be in the form of brown dwarfs, white dwarfs, neutron stars, and black holes. One way to characterize the stellar content of the universe at this epoch is by the mass distribution of these objects; we refer to this distribution as the "final mass function" or FMF. Technically, the final mass function is not final in the sense that degenerate objects can also evolve and thereby change their masses, albeit on vastly longer time scales. The subsequent evolution of degenerate objects is discussed in detail in Sec. IV.

Two factors act to determine the FMF: (1) The initial distribution of stellar masses [the initial mass function (IMF) for the progenitor stars], and (2) The transformation between initial stellar mass and the mass of the final degenerate object. Both of these components can depend on cosmological time. In particular, one expects that metallicity effects will tend to shift the IMF toward lower masses as time progresses.

The initial mass function can be specified in terms of a general log-normal form for the mass distribution $\psi = dN/d\ \ln m$,

$$\ln\psi(\ln m) = A - \frac{1}{2\langle\sigma\rangle^2}\{\ln[m/m_C]\}^2, \tag{2.20}$$

where A, m_C, and $\langle\sigma\rangle$ are constants. Throughout this discussion, stellar masses are written in solar units, i.e., $m \equiv M_*/(1\ M_\odot)$. This general form for the IMF is motivated both by the current theory of star formation and by general statistical considerations (Larson, 1973; Elmegreen and Mathieu, 1983; Zinnecker, 1984; Adams and Fatuzzo, 1996). In addition, this form is (roughly) consistent with observations (Miller and Scalo, 1979), which suggest that the shape parameters have the values $\langle\sigma\rangle \approx 1.57$ and $m_C \approx 0.1$ for the present-day IMF (see also Salpeter, 1955; Scalo, 1986; Rana, 1991). The constant A sets the overall normalization of the distribution and is not of interest here.

For a given initial mass function, we must find the final masses m_F of the degenerate objects resulting from the progenitor stars with a given mass m. For the brown dwarf range of progenitor masses, $m < m_H$, stellar objects do not evolve through nuclear processes and hence $m_F = m$. Here, the scale $m_H \approx 0.08$ is the minimum stellar mass required for hydrogen burning to take place.

Progenitor stars in the mass range $m_H \leq m \leq m_{SN}$ eventually become white dwarfs, where the mass scale $m_{SN} \approx 8$ is the minimum stellar mass required for the star to explode in a supernova (note that the mass scale m_{SN} can depend on the metallicity—see Jura, 1986). Thus, for the white dwarf portion of the population, we must specify the transformation between progenitor mass m

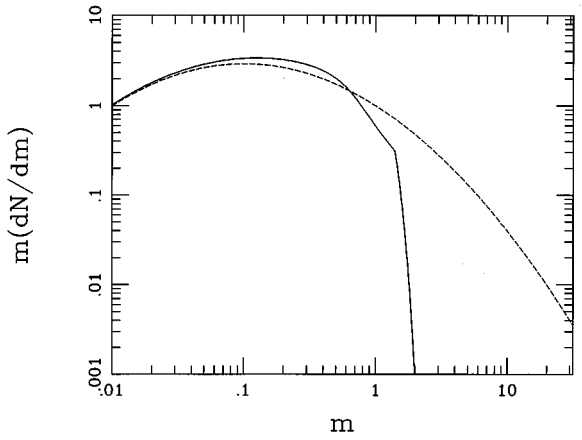

FIG. 2. The final mass function (FMF) for stars. Solid curve shows the predicted distribution $m(dN/dm)$ for the masses of the degenerate stellar objects (brown dwarfs, white dwarfs, and neutron stars) remaining at the cosmological epoch when conventional star formation has ceased. The dashed curve shows the mass distribution of the initial progenitor population (the initial mass function).

and white dwarf mass m_{WD}. The results of Laughlin *et al.* (1996) indicate that stars with main-sequence masses $m<0.4$ will undergo negligible mass loss in becoming helium white dwarfs. Unfortunately, this relationship remains somewhat ill defined at higher masses, mostly due to uncertainties in red giant mass-loss rates (e.g., see Wood, 1992). For the sake of definiteness, we adopt the following transformation between progenitor mass and white dwarf mass,

$$m_{WD}=\frac{m}{1+\alpha m}\exp[\beta m], \tag{2.21}$$

with $\alpha=1.4$ and $\beta=1/15$. This formula is consistent with the models of Wood (1992) over the appropriate mass range and approaches the expected form $m_{WD}=m$ in the low-mass limit.

Stars with large initial masses, $m>m_{SN}$, end their lives in supernova explosions and leave behind a neutron star (although black holes can also, in principle, be produced). The mass of the remnant neutron star is expected to be near the Chandrasekhar limit $m_{Ch}\approx 1.4$, as confirmed in the case of the binary pulsar (Manchester and Taylor, 1977).

To compute the FMF, one convolves the initial mass function with the transformations from progenitor stars to white dwarfs and neutron stars. The final mass function that results is shown in Fig. 2. For comparison, the initial mass function is also shown (as the dashed curve). Notice that the two distributions are similar for masses less than the Chandrasekhar mass ($\sim 1.4\ M_\odot$) and completely different for larger masses.

Once the FMF has been determined, one can estimate the number and mass fractions of the various FMF constituents. We define $\mathcal{N}_{BD}$ to be the fraction of brown dwarfs by number and $\mathcal{M}_{BD}$ to be the fraction of brown dwarfs by mass, with analogous fractions for white dwarfs ($\mathcal{N}_{WD}$ and $\mathcal{M}_{WD}$) and neutron stars ($\mathcal{N}_{NS}$ and $\mathcal{M}_{NS}$). For an IMF of the form (2.20) with present-day values for the shape parameters, we obtain the following number fractions:

$$\mathcal{N}_{BD}=0.45,\quad \mathcal{N}_{WD}=0.55,\quad \mathcal{N}_{NS}=0.0026. \tag{2.22}$$

Similarly, for the mass fractions one finds

$$\mathcal{M}_{BD}=0.097,\quad \mathcal{M}_{WD}=0.88,\quad \mathcal{M}_{NS}=0.024. \tag{2.23}$$

Thus brown dwarfs are expected to the present in substantial numbers, but most of the mass will reside in the form of white dwarfs. Neutron stars will make a relatively small contribution to the total stellar population. The above values for $\mathcal{N}_{NS}$ and $\mathcal{M}_{NS}$ were obtained under the assumption that all stars $m>m_{SN}\sim 8$ produce neutron stars. In reality, a portion of these high-mass stars may collapse to form black holes instead, but this complication does not materially affect the basic picture described above.

III. DEATH OF THE GALAXY

We have argued that over the long term, the galaxy will incorporate a large fraction of the available baryonic matter into stars. By the cosmological decade $\eta=14$–15, the stellar component of the galaxy will be in the form of seemingly inert degenerate remnants. Further galactic activity will involve these remnants in phenomena that unfold over time scales ranging from $\eta\sim 15$ to $\eta\sim 30$. This time period is part of what we term the *Degenerate Era.*

The course of this long-term galactic dynamical evolution is dictated by two generalized competing processes. First, in an isolated physical system containing any type of dissipative mechanism (for example, gravitational radiation, or extremely close inelastic encounters between individual stars), the system must evolve toward a state of lower energy while simultaneously conserving angular momentum. The net outgrowth of this process is a configuration in which most of the mass is concentrated in the center and most of the angular momentum is carried by small parcels at large radii. (The present-day solar system presents a good example of this process at work.) Alternatively, a second competing trend occurs when collisionless relaxation processes are viable. In a galaxy, distant encounters between individual stars are effectively collisionless. Over time, stars tend to be evaporated from the system. The end product of this process is a tightly bound agglomeration (perhaps a massive black hole) in the center, containing only a fairly small fraction of the total mass. Hence one must estimate the relative efficiencies of both collisionless and dissipative processes in order to predict the final state of the galaxy. This same competition occurs for physical systems on both larger scales (e.g., galaxy clusters) and smaller scales (e.g., globular clusters).

In addition to gravitational radiation and dynamical relaxation, occasional collisions between substellar objects—brown dwarfs—provide a channel for continued star formation at a very slow rate. Collisions and mergers involving two white dwarfs will lead to an occa-

sional type-I supernova, whereas rare impacts involving neutron stars will engender even more exotic bursts of energy. Such events are impressive today. They will be truly spectacular within the cold and impoverished environment of an evolved galaxy.

A. Dynamical relaxation of the galaxy

A stellar system such as a galaxy relaxes dynamically because of stellar encounters. The characteristic time scale associated with this process in the case of purely stellar systems is well known and can be written as

$$\tau_{\rm relax} = \frac{R}{v} \frac{N}{12 \ln(N/2)}, \tag{3.1}$$

where R is the size of the system, v is the typical random velocity, and N is the total number of stars (for further discussion, see Lightman and Shapiro, 1978; Shu, 1982; Binney and Tremaine, 1987). The logarithmic factor appearing in the denominator takes into account the effects of many small-angle deflections of stars through distant encounters. The time scale for stars to evaporate out of the system is roughly given by

$$\tau_{\rm evap} = 100 \tau_{\rm relax} \sim 10^{19} \text{ yr}, \tag{3.2}$$

where we have used R=10 kpc, v=40 km/s, and $N=10^{11}$ to obtain the numerical result. We thus obtain the corresponding estimate

$$\eta_{\rm evap} = 19 + \log_{10}[R/10 \text{ kpc}] + \log_{10}[N/10^{11}]. \tag{3.3}$$

Thus stars escape from the galaxy with a characteristic time scale $\eta \approx 19$–20 (see also Islam, 1977; Dyson, 1979).

The stellar dynamical evolution of the galaxy is more complicated than the simple picture outlined above. First, the galaxy is likely to have an extended halo of dark matter, much of which may be in nonbaryonic form. Since this dark halo does not fully participate in the dynamical relaxation process, the halo tends to stabilize the system and makes the stellar evaporation time scale somewhat longer than the simple estimate given above.

Other dynamical issues can also be important. In globular clusters, for example, mass segregation occurs long before stellar evaporation, and binary star heating plays an important (actually dominant) role in the long-term evolution. On the other hand, Eq. (3.1) is formally valid only if the stars are not bound into binary or triple systems. Binary interaction effects can be important for the long-term evolution of the stellar component of the galaxy. In particular, the presence of binaries can increase the effective interaction cross section and can lead to a variety of additional types of interactions. Both three-body interactions and binary-binary interactions are possible. As a general rule, interactions lead to hard binaries becoming harder and wide binaries becoming softer or even disrupted (''ionized''). Binaries that become sufficiently hard (close) can spiral inwards, become mass-transfer systems, and eventually explode as supernovae. These effects are just now becoming understood in the context of globular cluster evolution (for further discussion of these dynamical issues, see, for example, Chernoff and Weinberg, 1990; Hut *et al.*, 1992).

Galaxies in general, and our galaxy in particular, live in groups or clusters. These larger-scale systems will also undergo dynamical relaxation processes analogous to those discussed above. However, a more immediate issue that can affect our galaxy in the relatively near future is the possibility of merging with other galaxies in the local group, in particular Andromeda (M31). The orbits of nearby galaxies have been the subject of much study (e.g., Peebles, 1994), but large uncertainties remain. For the current separation between the Milky Way and M31 (d=0.75 Mpc) and radial velocity (v_r=120 km/s), the two galaxies will experience a close encounter at a time $\Delta t = 6 \times 10^9$ yr in the future (i.e., at η=10.2). Whether this encounter will lead to a collision/merger or simply a distant passage depends on the tangential velocity component, which is not well determined. The models of Peebles (1994) suggest that the distance of closest approach will lie in the range 20–416 kpc, with more models predicting values near the upper end of this range. Thus more work is necessary to determine whether or not the Milky Way is destined to collide with M31 in the relatively near future.

However, even if our galaxy does not collide with M31 on the first pass, the two galaxies are clearly a bound binary pair. The orbits of binary galaxy pairs decay relatively rapidly through dynamical friction (Binney and Tremaine, 1987; Weinberg, 1989). Thus, even if a collision does not occur on the first passing, M31 and the Milky Way will not survive very long as individual spiral galaxies. On a time scale of approximately η=11–12, the entire local group will coalesce into one large stellar system.

B. Gravitational radiation and the decay of orbits

Gravitational radiation acts in the opposite direction: it causes orbits to lose energy and decay so that the stars move inward. We first consider the case of a galaxy and its constituent stars. As a given star moves through the potential well of a galaxy, its orbit decays through gravitational radiation (Weinberg, 1972; Misner, Thorne, and Wheeler, 1973). The rate of energy loss is proportional to the square of the quadrupole moment of the whole system (see also Ohanian and Ruffini, 1994). For the case in which the galaxy has a large-scale quadrupole moment (e.g., a bar), the rate of energy loss from gravitational radiation can be written in the simple form

$$\frac{\dot{E}}{E} = \left(\frac{v}{c} \right)^5 \tau^{-1}, \tag{3.4}$$

where $\tau = 2\pi R/v$ is the orbit time. For a galaxy, the rotation curve is almost flat with a nearly constant velocity $v \sim 200$ km/s. The time scale τ_{GR} for gravitational radiation is thus given by

$$\tau_{GR} = \frac{2\pi R}{v} \left(\frac{v}{c} \right)^{-5} \approx 10^{24} \text{ yr} \left(\frac{R}{R_0} \right), \tag{3.5}$$

where R_0=10 kpc is a reference length scale for the galaxy. We thus obtain the estimate

$$\eta_{GR}=24+\log_{10}[R/10\ \text{kpc}]. \tag{3.6}$$

This time scale corresponds to $\sim 10^{16}$ orbits around the galactic center. Notice that if the stars are radiating incoherently in a nearly smooth potential, the time scale becomes longer by a factor of $M_{\rm gal}/M_*$, where M_* is the mass of the star and $M_{\rm gal}$ is the effective galactic mass. Notice also that gravitational orbital decay takes substantially longer than stellar evaporation from the galaxy (see the previous section). Thus the evolution of the galaxy will be dominated by the collisionless process, and hence the majority of stellar remnants will be ejected into intergalactic space rather than winding up in the galactic core (see also Islam, 1977; Dyson, 1979; Rees, 1984).

Gravitational radiation also causes the orbits of binary stars to lose energy and decay. Of particular importance is the decay of binary brown dwarf stars. The eventual coalescence of these systems can lead to the formation of a new hydrogen-burning star, provided that the mass of the entire system is larger than the hydrogen-burning limit $M_H \sim 0.08\ M_\odot$. The time scale τ_{OD} for orbital decay can be written

$$\tau_{OD}=\frac{\pi}{2}\frac{c^5 R_0^4}{G^3 M_*^3}, \tag{3.7}$$

where M_* is the mass of the stars and R_0 is the initial orbital separation. Inserting numerical values and writing the result in terms of cosmological decades, we obtain the result

$$\eta_{OD}=19.4+4\log_{10}[R_0/(1\ \text{AU})] -3\log_{10}[M_*/(1\ M_\odot)]. \tag{3.8}$$

This result also applies to planetary orbits (see Sec. III.F below).

C. Star formation through brown dwarf collisions

Once all of the interstellar material has been used up, one viable way to produce additional stars is through the collisions of brown dwarfs. These objects have masses too small for ordinary hydrogen burning to take place and hence their supply of nuclear fuel will remain essentially untapped. Collisions between these substellar objects can produce stellar objects with masses greater than the hydrogen-burning limit, i.e., stars of low mass. We note that the search for brown dwarfs has been the focus of much observational work (see, for example, Tinney, 1995) and the existence of these objects is now on firm ground (e.g., Golimowski *et al.*, 1995; Oppenheimer *et al.*, 1995).

1. Collision time scales

After conventional star formation in the galaxy has ceased, the total number of brown dwarfs in the galaxy will be N_0. Although the value of N_0 is uncertain and is currently the subject of much current research (e.g., Alcock *et al.*, 1993; Aubourg *et al.*, 1993; Tinney, 1995), we expect that N_0 is roughly comparable to the number of ordinary stars in the galaxy today, $N_0 \sim 10^{11}$ (see Sec. II.C). The rate Γ at which these brown dwarfs collide is given by

$$\Gamma=\frac{N\sigma v}{V}=-\frac{1}{N}\frac{dN}{dt}, \tag{3.9}$$

where N is the number of brown dwarfs in a galaxy with volume V, σ is the collision cross section (see below), and v is the typical relative velocity. This equation can be integrated to obtain

$$N(t)=\frac{N_0}{1+t/\tau_C}, \tag{3.10}$$

where τ_C is the characteristic time scale

$$\tau_C=\Gamma^{-1}\sim 10^{22}\ \text{yr}, \tag{3.11}$$

or, equivalently,

$$\eta_C=22+\log_{10}[V/(20\ \text{kpc})^3]-\log_{10}[v/(200\ \text{km/s})]. \tag{3.12}$$

To obtain this numerical value for the time scale, we have assumed that the collision cross section is given by the geometrical cross section of the brown dwarfs; this assumption is justified below. We have also used the numerical values $V\sim(20\ \text{kpc})^3$ and $v\sim 200$ km/s which are characteristic of the galactic halo.

The estimate of collision rates given here is somewhat conservative. Nearby stellar encounters can lead to the formation of binaries through tidal excitation of modes on the stars (see Press and Teukolsky, 1977; Lee and Ostriker, 1986). These binaries can eventually decay and thereby lead to additional stellar collisions.

The time scale (3.12) is the time required for the halo population of brown dwarfs to change. Notice that this time scale is larger than the evaporation time scale calculated in Sec. III.A. This ordering makes sense because distant encounters (which lead to evaporation) must be much more frequent than true collisions. For $\eta<\eta_C$, the collision rate of brown dwarfs for the entire galaxy is given by $\Gamma_{\rm tot}=N/\tau_C\sim 10^{-11}\ \text{yr}^{-1}$. The typical outcome of a brown dwarf collision will be the production of a stellar object with mass $M_*\sim 0.1\ M_\odot$, large enough to burn hydrogen. The stellar (main-sequence) lifetime of such a star is roughly 1×10^{13} yr. This stellar evolutionary time scale is longer than the time scale on which stars are forming. As a result, the galaxy will produce many stars through this process and will contain $\sim$50 hydrogen burning stars at a time for the cosmological decades $\eta>14$.

Notice that the time scale for producing stars through brown dwarf collisions is generally much shorter than the orbit decay time for brown dwarf binaries. For orbital decay, Eq. (3.8) implies that $\eta\sim 22.5+4\log_{10}(R/1\ \text{AU})$. Thus brown dwarf collisions provide the dominant mechanism for continued star formation while the galaxy remains intact.

2. Collision cross sections

To complete this argument, we must estimate the cross section for colliding brown dwarfs. Consider two

brown dwarfs with a relative velocity $v_{\rm rel}$. For simplicity, we consider the case of equal-mass brown dwarfs with mass m. The orbital angular momentum of the system is given by

$$J = m v_{\rm rel} b, \tag{3.13}$$

where b is the impact parameter. When the two dwarfs collide and form a composite star of mass $\sim 2m$, the angular momentum can be written

$$I\Omega = f(2m) R^2 \Omega, \tag{3.14}$$

where R is the stellar radius, Ω is the rotation rate, and f is a numerical constant of order unity which depends on the internal structure of the star. We next invoke the constraint that the rotation rate of the final state must be less than the breakup speed, i.e.,

$$\Omega^2 R^2 < \frac{G(2m)}{R}. \tag{3.15}$$

Combining the above results, we obtain a bound on the impact parameter b that can lead to a bound final system. We thus obtain

$$b^2 < \frac{8 f^2 G m R}{v_{\rm rel}^2}, \tag{3.16}$$

which can be used to estimate the cross section,

$$\sigma \approx \pi b^2 = \frac{8 \pi f^2 G m R}{v_{\rm rel}^2}. \tag{3.17}$$

Using typical numerical values, we find that $b \sim R \sim 10^{10}$ cm, which is roughly comparable to the radius of the brown dwarf (Burrows *et al.*, 1993).

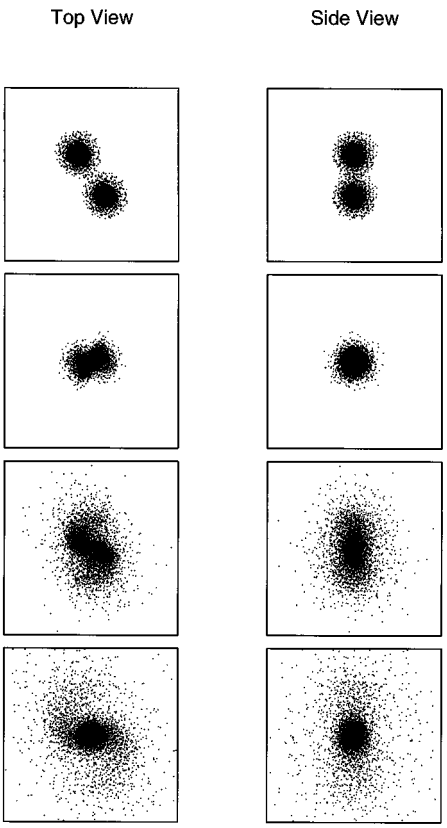

FIG. 3. Numerical simulation of a collision between two brown dwarfs. The two initial objects have masses less than that required for hydrogen burning; the final product of the collision is a true star and is capable of sustained hydrogen fusion. The two stars collide with a relative velocity of 200 km/s and an impact parameter of ~ 1 stellar radius. The top series of panels shows the collision from a side view; the bottom series of panels shows the top view.

3. Numerical simulations and other results

In order to illustrate the viability of this collision process, we have done a set of numerical simulations using smooth-particle hydrodynamics (SPH). We find that collisions between substellar objects can indeed form final products with masses greater than the minimum mass required to burn hydrogen. Examples of such collisions are shown in Fig. 3. In these simulations, density structures from theoretical brown dwarf models (Laughlin and Bodenheimer, 1993) are delivered into impact with relative velocity 200 km/s. The hydrodynamic evolutionary sequences shown are adiabatic. One expects that the emergent stellar mass object will contract toward the main sequence on a Kelvin-Helmholtz time scale and then initiate hydrogen burning.

Finally, we note that white dwarfs will also collide in the galactic halo. As outlined in Sec. II.E, we expect roughly comparable numbers of white dwarfs and brown dwarfs at the end of the stelliferous era. Although the white dwarfs are actually smaller in radial size, they are more massive and hence have a larger gravitational enhancement to their interaction cross section. As a result, the net cross section and hence the net interaction rate of white dwarfs should be roughly comparable to that of brown dwarfs (Sec. III.C.1). When white dwarfs collide with each other, several different final states are possible, as we discuss below.

If the two white dwarfs are sufficiently massive, it is possible that the collision product will have a final mass that exceeds the Chandrasekhar limit ($M_{Ch} \approx 1.4\ M_\odot$) and hence can explode in a supernova. Using the final mass function (see Sec. II.E and Fig. 2), we estimate that roughly one-third of the white dwarfs will have masses greater than $0.7\ M_\odot$ and hence only about one-tenth of the collisions can possibly result in an object exceeding the Chandrasekhar mass limit. The supernova rate from these collisions can thus be as large as $\Gamma_{SN} \sim 10^{-12}$ yr^{-1}, although it will be somewhat smaller in practice due to inefficiencies.

The most common type of collision is between two low-mass white dwarfs—the final mass function peaks at the mass scale $M_* \approx 0.13\ M_\odot$. These low-mass objects will have an almost pure helium composition. If the final product of the collision has a mass larger than the minimum mass required for helium burning ($M_{\rm He} \approx 0.3\ M_\odot$), then the product star could land on the helium main sequence (see, for example, Kippenhahn and Weigert,

1990). In order for the star to burn helium, the collision must be sufficiently energetic to impart enough thermal energy into the star; otherwise, the star will become just another helium white dwarf. Another possibility exists for collisions between white dwarfs of slightly larger masses. If the product of the collision has a mass smaller than the Chandrasekhar mass and larger than the minimum mass to burn carbon ($0.9\ M_\odot \leq M_* \leq 1.4\ M_\odot$), the product star could land on the carbon main sequence. Thus this mode of late-time star formation can lead to an interesting variety of stellar objects.

D. The black-hole accretion time

Large black holes tend to accrete stars and gas and thereby increase their mass. The black-hole accretion time is the characteristic time scale for a black hole in the center of a galaxy to swallow the rest of the galaxy. If we consider collisions of the black hole with stars, and ignore the other processes discussed above (gravitational radiation and stellar evaporation), the time for the black hole to absorb the stars in the galaxy is given by

$$\tau = \frac{V}{\sigma v}, \tag{3.18}$$

where $V = R^3$ is the volume of the galaxy, v is the typical speed of objects in the galaxy ($v \sim 200$ km/s), and σ is the effective cross section of the black hole. As a starting point, we write the cross section in the form

$$\sigma = \Lambda \pi R_S^2, \tag{3.19}$$

where Λ is a dimensionless enhancement factor due to gravitational focusing, and where R_S is the Schwarzschild radius R_S given by

$$R_S = GM/c^2. \tag{3.20}$$

We thus obtain the time scale

$$\tau = 10^{30}\ \text{yr}[M/10^6\ M_\odot]^{-2}[R/10\ \text{kpc}]^3 \Lambda^{-1}, \tag{3.21a}$$

or, equivalently,

$$\eta_{\text{accrete}} = 30 - 2\log_{10}[M/10^6\ M_\odot] + 3\log_{10}[R/10\ \text{kpc}] - \log_{10}[\Lambda]. \tag{3.21b}$$

The time scale $\eta_{\text{accrete}} \sim 30$ is much longer than the time scale for both stellar evaporation and gravitational radiation (see also the following section). As a consequence, at these late times, all the stars in a galaxy will either have evaporated into intergalactic space or will have fallen into the central black hole via gravitational radiation decay of their orbits. Of course, as the black hole mass grows, the accretion time scale decreases. Very roughly, we expect ~1–10% of the stars to fall to the central black hole and the remainder to be evaporated; the final mass of the central black hole will thus be $M_{BH} \sim 10^9$–$10^{10}\ M_\odot$.

One can also consider this process on the size scale of superclusters. When $\eta \sim 30$, supercluster-sized cosmological density perturbations of length R will have long since evolved to nonlinearity and will be fully decoupled from the Hubble flow. One can imagine an ensemble of $\sim 10^9$–$10^{10}\ M_\odot$ black holes that have descended from dead galaxies and are now roaming freely and hoovering up an occasional remaining star in the volume R^3. The characteristic time scale for this process is

$$\eta_{\text{accrete}} = 33 - 2\log_{10}[M/10^9\ M_\odot] + 3\log_{10}[R/10\ \text{Mpc}]. \tag{3.22}$$

As for the case of the galaxy, however, this straightforward scenario is compromised by additional effects. Gravitational radiation will continuously cause the orbits of the black holes to decay, and some of them may eventually merge. Stellar encounters both with other stars and with other black holes will lead to stellar evaporation from the supercluster-sized system. Over the long term, one expects that the supercluster will consist of a very large central black hole with a majority of the stars and many of the original $\sim 10^9$–$10^{10}\ M_\odot$ galactic black holes escaping to large distances. In other words, the supercluster-sized system will behave somewhat analogously to the galaxy, except that it will contain a larger size scale, a longer time scale, and two widely disparate mass scales (namely, a stellar mass scale $M_* \sim 1\ M_\odot$, and a black-hole mass scale $M_{BH} \sim 10^9$–$10^{10}\ M_\odot$). Equipartition effects between the two mass scales will come into play and will drive the galactic black holes toward the center while preferentially ejecting the stellar remnants. In principle, this hierarchy can extend up to larger and larger perturbation length scales, although the relevant time scales and detailed dynamics become more uncertain as one proceeds with the extrapolation.

E. Annihilation and capture of halo dark matter

Galactic halos consist largely of dark matter, much of which may reside in nonbaryonic form. Although the nature and composition of this dark matter remains an important open question, one of the leading candidates is weakly interacting massive particles, usually denoted as WIMPs. These particles are expected to have masses in the range M_W=10–100 GeV and to interact through the weak force and gravity only (see the reviews of Diehl *et al.*, 1995; Jungman, Kamionkowski, and Griest, 1996; see also the recent proposal of Kane and Wells, 1996). Many authors have studied the signatures of WIMP annihilation, usually with the hope of finding a detectable signal. One can apply the results of these studies to an estimate of the time scale for the depletion of WIMPs from a galactic halo.

We first consider the case of direct particle-particle annihilation. Following the usual conventions, the rate Γ_W for WIMP annihilation in the halo can be written in the form

$$\Gamma_W = n_W \langle \sigma v \rangle, \tag{3.23}$$

where n_W is the number density of WIMPs in the halo and $\langle \sigma v \rangle$ is the average value of the annihilation cross section times velocity. If WIMPs make up a substantial

mass fraction of the galactic halo, their number density is expected to be roughly $n_W \sim 1$ cm^{-3}. The typical velocity of particles in the galactic halo is ~ 200 km/s. Using the most naive dimensional argument, we can estimate the interaction cross section as

$$\sigma \sim M_W^2 G_F^2 \sim 5 \times 10^{-38} \text{ cm}^2 \left[\frac{M_W}{1 \text{ GeV}} \right]^2, \tag{3.24}$$

where M_W is the mass of the particle and G_F is the Fermi constant. The true cross section has additional factors which take into account spin dependences, mixing angles, and other model-dependent quantities (see Diehl *et al.*, 1995; Jungman *et al.*, 1996); the form (3.24) is thus highly approximate, but adequate for our purposes. We also note that the relic abundance of dark-matter particles is determined by the interaction cross section; in order for the abundance to be cosmologically significant, the interaction cross section must be of order $\sigma \sim 10^{-37}$ cm^2 (see Kolb and Turner, 1990).

Putting all of the above results together, we can estimate the time scale τ_W for the population of WIMPs to change,

$$\tau_W = \Gamma^{-1} = \frac{1}{n_W \langle \sigma v \rangle} \sim 3 \times 10^{22} \text{ yr}. \tag{3.25}$$

Thus, in terms of cosmological decades, we obtain the annihilation time scale in the form

$$\eta_W = 22.5 - \log_{10} \left[\frac{\langle \sigma v \rangle}{10^{-30} \text{ cm}^3 \text{ s}^{-1}} \right] - \log_{10} \left[\frac{n_W}{1 \text{ cm}^{-3}} \right]. \tag{3.26}$$

It takes a relatively long time for WIMPs to annihilate via direct collisions. In particular, the annihilation time scale is much longer than the stellar evaporation time scale (Sec. III.A).

Another important related effect is the capture of WIMPs by astrophysical objects. The process of WIMP capture has been studied for both the Sun (Faulkner and Gilliland, 1985, Press and Spergel, 1985) and the Earth (Freese, 1986) as a means of helping to detect the dark matter in the halo (see also Krauss, Srednicki, and Wilczek, 1986; Gould, 1987, 1991). Although WIMP capture by the Sun and the Earth can be important for dark-matter detection, the lifetimes of both (main-sequence) stars and planets are generally too small for WIMP capture to significantly affect the total population of particles in the galactic halo. On the other hand, stellar remnants, in particular white dwarfs, can be sufficiently long lived to have important effects.

In astrophysical objects, WIMPs are captured by scattering off of nuclei. When the scattering event leads to a final velocity of the WIMP that is less than the escape speed of the object, then the WIMP has been successfully captured. For the case of white dwarfs, we can make the following simple estimate of the capture process. The mean free path of a WIMP in matter with white dwarf densities is generally less than the radius of the star. In addition, the escape speed from a white dwarf is large, roughly ~ 3000 km/s, which is much larger than the velocity dispersion of WIMPs in the halo. As a result, to first approximation, most WIMPs that pass through a white dwarf will be captured. The WIMP capture rate Γ_{W*} by a white dwarf is thus given by

$$\Gamma_{W*} = n_W \sigma_{WD} v_{\text{rel}}, \tag{3.27}$$

where $\sigma_{WD} \sim 10^{18}$ cm^2 is the cross-sectional area of the white dwarf and $v_{\text{rel}} \sim 200$ km/s is the relative velocity. The capture rate is thus

$$\Gamma_{W*} \sim 10^{25} \text{ s}^{-1}. \tag{3.28}$$

With this capture rate, a white dwarf star can consume its weight in WIMPs on a time scale of $\sim 10^{24}$ yr. The total mass in WIMPs in the halo is expected to be a factor of 1–10 times the mass of stars, which will be mostly in the form of white dwarfs at these late times (See II.E). As a result, the time scale for white dwarfs to deplete the entire halo population of WIMPs via capture is roughly given by

$$\tau \sim 10^{25} \text{ yr} \quad \text{or} \quad \eta \sim 25. \tag{3.29}$$

The actual time scales will depend on the fraction of the galactic halo in nonbaryonic form and on the properties (e.g., mass) of the particles; these quantities remain unknown at this time.

The annihilation of halo WIMPs has important consequences for both the galaxy itself and for the white dwarfs. Basically, the galaxy as a whole loses mass while the white dwarfs are kept hotter than they would be otherwise. The population of captured WIMPs inside the star will build up to a critical density at which the WIMP annihilation rate is in equilibrium with the WIMP capture rate (see, for example, Jungman *et al.*, 1996). Furthermore, most of the annihilation products will be absorbed by the star, and the energy is eventually radiated away (ultimately in photons). The net result of this process (along with direct annihilation) is thus to radiate away the mass of the galactic halo on the time scales given by Eqs. (3.26) and (3.29). This process competes with the evaporation of stars through dynamic relaxation (Sec. III.A) and the decay of stellar orbits through gravitational radiation (Sec. III.B).

Since the time scale for WIMP evaporation is much longer than the dynamical time scale, the galaxy will adiabatically expand as the halo radiates away. In the outer galaxy, the dark matter in the halo dominates the gravitational potential well, and hence the stars in the outer galaxy will become unbound as the halo mass is radiated away. Since WIMPs do not dominate the potential inside the solar circle, the corresponding effects on the inner galaxy are relatively weak.

The white dwarf stars themselves will be kept hot by this WIMP capture process with a roughly constant luminosity given by

$$L_{WD} = \mathcal{F} m_W \Gamma_{W*} = \mathcal{F} m_W n_W \sigma_{WD} v_{\text{rel}} \sim 4 \times 10^{-12} \ L_\odot, \tag{3.30}$$

where $\mathcal{F}$ is an efficiency factor (expected to be of order unity) which takes into account the loss of energy from the star in the form of neutrinos. With this luminosity,

the white dwarf has a surface temperature $T_* \approx 63$ K, where we have assumed a typical white dwarf mass $M_* = 0.5\ M_\odot$. As a reference point, we note that an entire galaxy of such stars has a total luminosity comparable to that of the Sun, $L_{\rm gal} \sim 1\ L_\odot$. However, most of the radiation will be emitted at infrared wavelengths, $\lambda \sim 50\ \mu$m.

For completeness, we note that axions provide another viable candidate for the dark matter in the galactic halo (see Chapter 10 of Kolb and Turner, 1990). These particles arise from solutions to the strong CP problem in quantum chromodynamics (see Peccei and Quinn, 1977a, 1977b; Weinberg, 1978; Wilczek, 1978). The coupling of the axion to the photon allows the axion to decay to a pair of photons with a lifetime τ_a given by

$$\tau_a \approx 2 \times 10^{17}\ \text{yr} (m_a / 1\ \text{eV})^{-5}, \tag{3.31}$$

where m_a is the mass of the axion; we have assumed representative values for the remaining particle physics parameters. Relic axions with sufficient numbers to contribute to the dark-matter budget of the universe have masses in the range $10^{-6}\ \text{eV} < m_a < 10^{-3}\ \text{eV}$, where the value depends on the production mechanism. Using these mass values, we obtain an allowed range of axion decay time scales,

$$32 \leqslant \eta_a \leqslant 47. \tag{3.32}$$

F. The fate of planets during galactic death

Planets can be loosely defined as objects that are small enough (in mass) to be supported by ordinary Coulomb forces rather than by degeneracy pressure. Over the long term, planets suffer from several deleterious processes. They can be vaporized by their evolving parent stars, and their orbits can either decay or be disrupted. Barring these more imminent catastrophes, planets will evaporate as their protons decay (see Sec. IV.H).

The theory of general relativity indicates that planetary orbits slowly decay via emission of gravitational radiation (see Sec. III.B). To fix ideas, consider a planet orbiting a star of mass M_* at an initial orbital radius R. Gravitational radiation drives orbital decay on a time scale given by

$$\tau = \frac{2\pi R}{v} \left(\frac{v}{c} \right)^{-5} = 2.6 \times 10^{19}\ \text{yr} \left(\frac{R}{1\ \text{AU}} \right)^4 \left(\frac{M_*}{1\ M_\odot} \right)^{-3}, \tag{3.33}$$

or, in terms of cosmological decades,

$$\eta = 19.4 + 4 \log_{10}[R/1\ \text{AU}] - 3 \log_{10}[M_*/1\ M_\odot]. \tag{3.34}$$

In the interim, planets can be dislodged from their parent stars during encounters and collisions with interloping stars. The time scale for these dislocations is given by the time interval required to produce a sufficiently close encounter with another star. Very roughly, if a perturbing star intrudes within a given planet's orbit, then the planet is likely to be entirely ejected from the system. This time scale is given by

$$\tau = \frac{1}{n \sigma v}, \tag{3.35}$$

where n is the number density of stars ($\sim 0.1\ \text{pc}^{-3}$ in our galaxy today), v is the relative velocity (~ 100 km/s), and the cross section σ is determined by the orbital radius of the planet ($\sigma \approx \pi R^2$). Inserting these values, one finds

$$\tau = 1.3 \times 10^{15}\ \text{yr} \left(\frac{R}{1\ \text{AU}} \right)^{-2}, \tag{3.36}$$

$$\eta = 15.1 - 2 \log_{10}[R/1\ \text{AU}], \tag{3.37}$$

where R is the radius of the planetary orbit.

Comparing Eq. (3.33) with Eq. (3.36), we find that the time scale for gravitational radiation is equal to that of stellar encounters for planetary orbits of radius $R = 0.2$ AU, which is about half the radius of the orbit of Mercury in our own solar system. One might guess then, that very close planets, such as the recently discovered companion to 51 Pegasus (Mayor and Queloz, 1995; Marcy, Butler, and Williams, 1996), will eventually merge with their parent stars as a result of radiative orbital decay, while planets with larger initial orbits (e.g., the giant planets in our solar system) will be stripped away from their parent stars as a consequence of stellar encounters. However, since the time scale for stellar evolution ($\eta_* < 14$) is much shorter than the time scale for orbital decay, close-in planets around solar-type stars will be destroyed during the red giant phases long before their orbits are constricted by general relativity. Only the inner planets of low-mass M dwarfs (which experience no giant phases) will find their fate sealed by gravitational radiation.

IV. LONG-TERM FATE OF DEGENERATE STELLAR OBJECTS

Brown dwarfs, white dwarfs, neutron stars, and black holes have lifetimes that not only are much longer than the current age of the universe ($\eta = 10$), but also greatly exceed the expected lifetime of the galaxy ($\eta = 20$–25). Due to a general lack of urgency, the ultimate fate of these objects has not yet been extensively considered. Nevertheless, these objects will not live forever. If the proton is unstable, then proton decay will drive the long-term evolution of degenerate stellar objects. Black holes are essentially unaffected by proton decay, but they gradually dissipate via the emission of Hawking radiation. Both proton decay and Hawking radiation yield many interesting astrophysical consequences. In the following discussion, we work out the details of these processes (see also Feinberg, 1981; Dicus *et al.*, 1982).

A. Proton decay

In grand unified theories (GUTs), the proton is unstable and has a finite, albeit quite long, lifetime. For example, the proton can decay through the process

$$p \rightarrow e^+ + \pi^0, \tag{4.1}$$

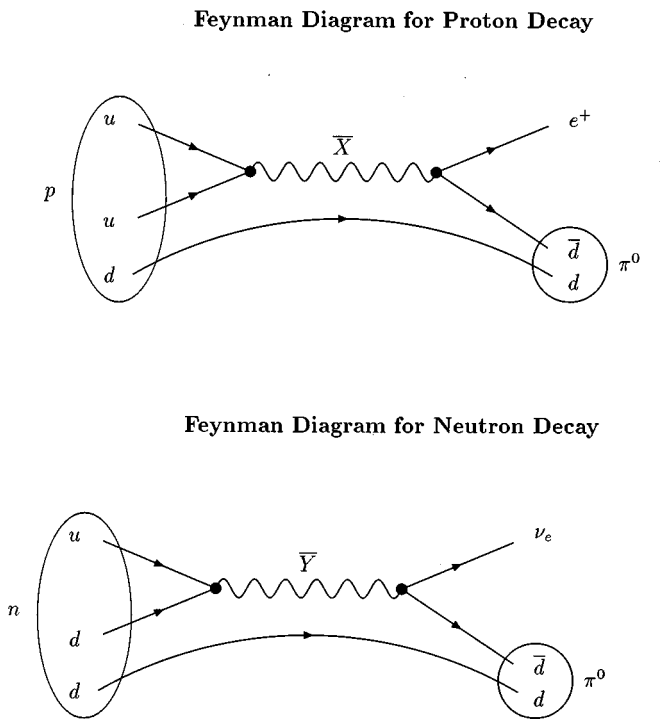

FIG. 4. Representative Feynman diagrams for proton decay (top diagram) and neutron decay (bottom diagram) shown in terms of the constituent quarks (u,d,$\bar{d}$). These processes are the form expected for the simplest grand unified theories (GUTs). The particles X and Y are the intermediate vector bosons that mediate the baryon-number-violating process and are expected to have masses comparable to the GUT scale $\sim 10^{16}$ GeV.

and the Feynman diagram for this decay process is shown in Fig. 4. Many different additional decay channels are possible and the details ultimately depend on the particular theory (e.g., see the reviews of Langacker, 1981; Perkins, 1984). In particular, we note that many other decay products are possible, including neutrinos. If protons are unstable, then neutrons will also be unstable over a commensurate time scale. Free neutrons are of course unstable to β decay on a very short time scale ($\sim$10 minutes); however, bound neutrons will be unstable through processes analogous to the proton-decay modes (e.g., see Fig. 4). In the present context, the protons and neutrons of interest are bound in "heavy" nuclei (mostly carbon and helium) within white dwarfs.

For the simplest class of GUTs, as illustrated by the decay modes shown in Fig. 4, the rate of nucleon decay Γ_P is roughly given by

$$\Gamma_P = \alpha_5^2 \frac{m_P^5}{M_X^4}, \tag{4.2}$$

where m_P is the proton mass and α_5 is a dimensionless coupling parameter (see Langacker, 1981; Perkins, 1984; Kane, 1993). The mass scale M_X is the mass of the particle, which mediates the baryon-number-violating process. The decay rate should also include an extra numerical factor which takes into account the probability that the interacting quarks (which participate in the decay) are in the same place at the same time; this numerical factor is less than unity, so that the proton lifetime is larger by a corresponding factor. To a first approximation, the time scale for proton decay is thus given by

$$\tau_P \approx 10^{37}\ \text{yr} \left[\frac{M_X}{10^{16}\ \text{GeV}} \right]^4, \tag{4.3}$$

where we have taken into account the aforementioned numerical probability factor. The corresponding cosmological time scale is

$$\eta_P = 37 + 4 \log_{10}[M_X / 10^{16}\ \text{GeV}]. \tag{4.4}$$

Notice that this time scale has a very sensitive dependence on the mass scale M_X of the mediating boson.

We want to find the allowed range for the proton lifetime. This time scale is constrained from below by current experimental limits on the lifetime of the proton (e.g., Perkins, 1984). The proton lifetime must be greater than $\eta \sim 32$ (10^{32} yr), where the exact limit depends on the particular mode of proton decay (Particle Data Group, 1994). Finding an upper bound is more difficult. If we restrict our attention to the class of proton-decay processes for which Eq. (4.4) is valid, then we must find an upper bound on the mass scale M_X. Following cosmological tradition, we expect the scale M_X to be smaller than the Planck scale $M_{\rm Pl} \approx 10^{19}$ GeV, which implies a range for the proton lifetime

$$32 < \eta_P < 49. \tag{4.5}$$

The lower bound is set by experimental data; the upper bound is more suggestive than definitive (see also Sec. IV.F).

We can find a more restrictive range for the proton lifetime for the special case in which the decay mode from some GUT is responsible for baryogenesis in the early universe. (Note that some baryon-number-violating process is necessary for baryogenesis to take place—see Sakharov, 1967). Let us suppose that the decay mode from some GUT is valid and that baryogenesis takes place at an energy scale in the early universe $E_B \sim M_X$. This energy scale must be less than the energy scale E_I of the inflationary epoch (Guth, 1981). The inflationary energy scale is constrained to be less than $\sim 10^{-2}\ M_{\rm Pl}$ in order to avoid overproducing scalar density perturbations and gravitational radiation perturbations (Lyth, 1984; Hawking, 1985; Krauss and White, 1992; Adams and Freese, 1995). Combining these two constraints, we obtain the following suggestive range for the time scale for proton decay:

$$32 < \eta_P < 41. \tag{4.6}$$

Although a range of nine orders of magnitude in the relevant time scale seems rather severe, the general tenor of the following discussion does not depend critically on the exact value. For the sake of definiteness, we adopt $\eta_P = 37$ as a representative time scale.

B. White dwarfs powered by proton decay

On a sufficiently long time scale, the evolution of a white dwarf is driven by proton decay. When a proton

decays inside a star, most of the primary decay products (e.g., pions and positrons) quickly interact and/or decay themselves to produce photons. For example, the neutral pion π^0 decays into a pair of photons with a lifetime of $\sim 10^{-16}$ sec; positrons, e^+, last only $\sim 10^{-15}$ sec before annihilating with an electron and producing gamma rays. Therefore one common net result of proton decay in a star is the eventual production of four photons through the effective reaction

$$p + e^- \rightarrow \gamma + \gamma + \gamma + \gamma, \tag{4.7}$$

where the typical energy of the photons is given by $E_\gamma \sim m_P/4 \sim 235$ MeV. These photons have a relatively short mean free path within the star and will thermalize and diffuse outwards through a random-walk process with a characteristic time scale of $\sim 10^5$ yr, much shorter than the evolutionary time scale of the system. Additionally, some fraction of the decay products are in the form of neutrinos, which immediately leave the system.

When proton decay is a white dwarf's primary energy source, the luminosity is

$$L_*(t) = \mathcal{F} N_0 E \Gamma_P e^{-\Gamma_P t} \approx \mathcal{F} M(t) \Gamma_P, \tag{4.8}$$

where $N_0 \sim 10^{57}$ is the initial number of protons in the star, $E \sim 1$ GeV is the net energy produced per decay, and Γ_P is the decay rate. The factor $\mathcal{F}$ is an efficiency parameter that takes into account the fraction of energy lost in the form of neutrinos. Very roughly, we expect $\sim 1/3$ of the energy in the decay products to be in neutrinos and hence $\mathcal{F} \sim 2/3$ (e.g., Dicus *et al.*, 1982). The exact value of the fraction $\mathcal{F}$ depends on the branching ratios for a particular GUT and hence is model dependent. For a typical decay rate of $\Gamma_P \sim 10^{-37}$ yr^{-1}, the luminosity in solar units becomes

$$L_* \sim 10^{-24}\ L_\odot. \tag{4.9}$$

It is perhaps more illuminating to express this stellar luminosity in ordinary terrestrial units. A white dwarf fueled by proton decay generates approximately 400 watts, enough power to run a few light bulbs, or, alternately, about 1/2 horsepower. An entire galaxy of such stars has a total luminosity of $L_{\rm gal} \sim 10^{-13}\ L_\odot$, which is much smaller than that of a single hydrogen-burning star.

The total possible lifetime for a star powered by proton decay is given by

$$\tau = \frac{1}{\Gamma_P} \ln[N_0 / N_{\rm min}], \tag{4.10}$$

where $N_0 \sim 10^{57}$ is the initial number of nucleons in the star and $N_{\rm min}$ is the minimum number of nucleons required to consider the object a star. If, for example, one takes the extreme case of $N_{\rm min}=1$, the time required for the star to completely disappear is $t \approx 130/\Gamma_P$; in general we obtain

$$\eta_* = \eta_P + \log_{10}[\ln(N_0 / N_{\rm min})]. \tag{4.11}$$

As we show in Sec. IV.D, the object ceases to be a star when $N_{\rm min} \sim 10^{48}$ and hence $\eta_* \approx \eta_P + 1.3$.

During the proton-decay phase, the stellar surface temperature is given by

$$T_*^4 = \frac{\mathcal{F} N_0 E \Gamma_P}{4 \pi \sigma_B R_*^2} e^{-\Gamma_P t}, \tag{4.12}$$

where we have assumed that the spectral energy distribution is simply a blackbody (σ_B is the Stefan-Boltzmann constant). For a 1-$M_\odot$ star and the typical decay rate Γ_P, the effective stellar temperature is $T_* \sim 0.06$ K. This temperature will be enormously hotter than the temperature of the universe's background radiation at the cosmological decade $\eta = 37$.

As a white dwarf loses mass via proton decay, the star expands according to the usual mass/radius relation

$$R_* M_*^{1/3} = 0.114 \frac{h^2}{G m_e m_P^{5/3}} (Z/A)^{5/3}, \tag{4.13}$$

where Z and A are the atomic number and atomic weight of the white dwarf material (see, for example, Chandrasekhar, 1939; Shu, 1982; Shapiro and Teukolsky, 1983). For simplicity, we shall take typical values and use $A = 2Z$. If we also rewrite the white dwarf mass/radius relation in terms of natural units, we obtain the relation

$$R_* = 1.42 \left(\frac{M_{\rm Pl}}{m_e} \right) \left(\frac{M_{\rm Pl}}{m_P} \right) \left(\frac{M_*}{m_P} \right)^{-1/3} m_P^{-1}. \tag{4.14}$$

While the white dwarf is in the proton-decay phase of its evolution, the star follows a well defined track in the H-R diagram, i.e.,

$$L_* = L_0 (T_* / T_0)^{12/5}, \tag{4.15}$$

or, in terms of numerical values,

$$L_* = 10^{-24}\ L_\odot \left[\frac{T_*}{0.06\ \text{K}} \right]^{12/5}. \tag{4.16}$$

We note that the white dwarf mass/radius relation depends on the star's chemical composition, which changes as the nucleons decay (see the following section). This effect will cause the evolutionary tracks to depart slightly from the 12/5 power law derived above. However, this modification is small and will not be considered here.

C. Chemical evolution in white dwarfs

Over the duration of the proton-decay phase, the chemical composition of a white dwarf is entirely altered. Several different effects contribute to the change in chemical composition. The nucleon-decay process itself directly alters the types of nuclei in the star and drives the chemical composition toward nuclei of increasingly lower atomic numbers. However, pycnonuclear reactions can occur on the relevant (long) time scales and build nuclei back up to higher atomic numbers. In addition, spallation interactions remove protons and neutrons from nuclei; these free nucleons then interact with other nuclei and lead to further changes in composition.

In the absence of pycnonuclear reactions and spallation, the chemical evolution of a white dwarf is a simple cascade toward lower atomic numbers. As protons and neutrons decay, the remaining nuclei become correspondingly smaller. Some of the nuclear products are radioactive and will subsequently decay. Given the long time scale for proton decay, these radioactive nuclei are extremely short lived. As a result, only the stable isotopes remain. At relatively late times, when the total mass of the star has decreased by a substantial factor (roughly a factor of ten as we show below), almost all of the nuclei left in the star will be in the form of hydrogen.

At high densities and low temperatures, nuclear reactions can still take place, although at a slow rate. The quantum-mechanical zero-point energy of the nuclei allows them to overcome the Coulomb repulsion and fuse. In natural units, the nuclear reaction rate can be written in the form

$$W = 4\left(\frac{2}{\pi^3}\right)^{1/2} S(Z^2\alpha\mu)^{3/4} R_0^{-5/4} \times \exp[-4Z(\alpha\mu R_0)^{1/2}], \tag{4.17}$$

where μ is the reduced mass of the nucleus, R_0 is the average spacing between nuclei, and α is the fine-structure constant (see Shapiro and Teukolsky, 1983). A slightly different form for this reaction rate can be derived by including anisotropic and electron screening effects (Salpeter and Van Horn, 1969), but the basic form is similar. The parameter $S(E)$ is a slowly varying function of energy which takes into account the probability of two nuclei interacting, given that tunneling has occurred. Specifically, the parameter S is related to the cross section $\sigma(E)$ through the relation

$$\sigma(E) = \frac{S(E)}{E}\,\mathcal{T}, \tag{4.18}$$

where $\mathcal{T}$ is the tunneling transition probability. The parameter S can be determined either from direct experiments or from theoretical calculations (see Shapiro and Teukolsky, 1983; Bahcall, 1989).

In order to evaluate the time scale for pycnonuclear reactions to occur, one needs to determine the spacing R_0 of the nuclei, or, equivalently, the number density of particles. Using the white dwarf mass/radius relation, we obtain the result

$$\mu R_0 = 2.29A\left(\frac{M_{\rm Pl}}{m_e}\right)\left(\frac{M_{\rm Pl}}{m_P}\right)\left(\frac{M_*}{m_P}\right)^{-2/3} \approx 4060A\,m_*^{-2/3}, \tag{4.19}$$

where A is average the atomic weight of the nuclei and where we have defined $m_* \equiv M_*/M_\odot$.

We can now obtain a rough estimate for the efficiency of pycnonuclear reactions' building larger nuclei within white dwarfs. As a reference point, we note that for a density of $\rho \sim 10^6$ g cm^{-3}, the time scale for hydrogen to fuse into helium is $\sim 10^5$ yr (Salpeter and van Horn, 1969; Shapiro and Teukolsky, 1983), which is much shorter than the proton-decay time scale. However, the form of Eq. (4.17) shows that the rate of nuclear reactions becomes highly suppressed as the reacting nuclei become larger. The exponential suppression factor roughly has the form $\sim\exp[-\beta Z A^{1/2}]$, where the numerical factor $\beta \approx 22$. Thus, as the quantity $ZA^{1/2}$ increases, the rate of nuclear reactions decreases exponentially. For example, if $Z=6$ and $A=12$ (for carbon), this exponential term is a factor of $\sim 10^{-190}$ smaller than that for hydrogen. Because of this large exponential suppression, fusion reactions will generally not proceed beyond helium during the late-time chemical evolution considered here. Thus the net effect of pycnonuclear reactions is to maintain the decaying dwarf with a predominantly helium composition down to a lower mass scale.

Spallation is another important process that affects the chemical evolution of white dwarf stars during the epoch of proton decay. The high-energy photons produced through proton decay can interact with nuclei in the star. The most common result of such an interaction is the emission of a single free neutron, but charged particles (protons), additional neutrons, and gamma rays can also result (Hubbell, Grimm, and Overbo, 1980). The free neutrons will be promptly captured by other nuclei in a type of late-time *s* process (the *r* process is of course dramatically irrelevant). The free protons can produce heavier nuclei through pycnonuclear reactions, as described above. Both of these mechanisms thus allow heavier elements to build up in the star, albeit at a very slow rate and a very low abundance. Thus the process of spallation initially produces free neutrons and protons; but these nucleons are incorporated into other nuclei. As a result, the net effect of spallation is to remove nucleons from some nuclei and then give them back to other nuclei within the star. The result of this redistribution process is to widen the distribution of the atomic numbers (and atomic weights) for the nuclei in the star.

In order to assess the importance of spallation processes, we must consider the interaction cross section. To leading order, the cross section for nuclear absorption of photons is a single ''giant resonance'' with a peak at about 24 MeV for light nuclei and a width in the range Γ=3–9 MeV. The relative magnitude of this resonance feature is $\sim$20 mb (see, for example, Brune and Schmidt, 1974; Hubbell, Gimm, and Overbo, 1980), roughly a factor of 30 smaller than the total interaction cross section (which is dominated by scattering and pair production). For each proton-decay event, $\sim$940 MeV of matter is converted into photons, with some neutrino losses. When these photons cascade downward in energy through the resonance regime (at $\sim$24 MeV), there will be 20–40 photons and about one in 30 will produce a spallation event. Hence, on average, each proton-decay event leads to approximately one spallation event.

Spallation products allow the interesting possibility that a CNO cycle can be set up within the star. The time scale for pycnonuclear reactions between protons (produced by spallation) and carbon nuclei is short compared to the proton-decay time scale. The time scale for pycnonuclear reactions between protons and nitrogen nuclei is comparable to the proton-decay time scale.

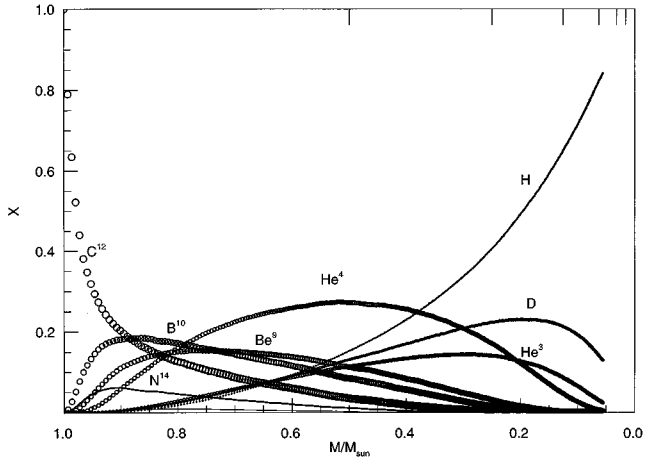

FIG. 5. Chemical evolution of a white dwarf star during proton decay. The curves show the mass fractions of the major component nuclei in the star as a function of time, which is measured here in terms of the stellar mass. The initial state is a 1.0 $M_\odot$ white dwarf made of pure ^{12}C. This simulation includes the effects of spallation and radioactivity (see text).

Thus, in principle, the white dwarf can set up a CNO cycle analogous to that operating in upper-main-sequence stars (see Shu, 1982; Clayton, 1983; Kippenhahn and Weigert, 1990). The energy produced by this cycle will be small compared to that produced by proton decay and hence this process does not actually affect the luminosity of the star. However, this cycle will affect the chemical composition and evolution of the star. As usual, the net effect of the CNO cycle is to build four free protons into a helium nucleus and to maintain an equilibrium abundance of the intermediate nitrogen and oxygen nuclei.

In order to obtain some understanding of the chemical evolution of white dwarfs, we have performed a simple numerical simulation of the process. Figure 5 shows the results of this calculation for a 1-$M_\odot$ white dwarf with an initial chemical composition of pure carbon ^{12}C. The simulation assumes that radioactive isotopes decay immediately as they are formed through the preferred decay modes. For each proton-decay event, a spallation event also occurs (see above) and leads to the removal of a nucleon from a random nucleus; the spallation products are then assumed to fuse immediately and randomly with other nuclei through the *s* process and pycnonuclear reactions. The spallation process builds up a small abundance of nuclei heavier than the original ^{12}C, particularly ^{13}C, which has a substantial mass fraction at "early" times. The white dwarf evolves through successive phases in which smaller and smaller nuclei are the dominant elements by mass fraction. The star never builds up a significant lithium fraction due to the immediate fission of newly formed ^{8}Be into α particles. The star has a broad phase during which ^{4}He dominates the composition. When the white dwarf has lost about 60% of its original mass, the hydrogen mass fraction begins to predominate.

D. Final phases of white dwarf evolution

In the final phases in the life of a white dwarf, the star has lost most of its mass through proton decay. When the mass of the star becomes sufficiently small, two important effects emerge: First, degeneracy is lifted and the star ceases to be a white dwarf. And second, the object becomes optically thin to its internal radiation produced by proton decay and thus ceases to be a star. In the following discussion, we present simple estimates of the mass scales at which these events occur.

When the star has lost enough of its initial mass to become nondegenerate, most of the nucleons in the star will be in the form of hydrogen (see the previous section). A cold star composed of pure hydrogen will generally have a thick envelope of molecular hydrogen surrounding a degenerate core of atomic hydrogen. As the stellar mass continues to decline through the process of proton decay, the degenerate core becomes increasingly smaller and finally disappears altogether. This transition occurs when the degeneracy energy, the Coulomb energy, and the self-gravitational energy of the star are all comparable in magnitude; this event, in turn, occurs when the central pressure P_C drops below a critical value of roughly a few Megabars ($P_C \sim 10^{12}$ dyne/cm^2). The central pressure in a star can be written in the form

$$P_C = \beta \frac{G M_*^2}{R_*^4}, \tag{4.20}$$

where β is a dimensionless number of order unity. Using the white dwarf mass/radius relation in the form of Eq. (4.13) and setting $Z = A = 1$, we find the central pressure as a function of stellar mass,

$$P_C \approx \frac{\beta}{410} M_{\rm Pl}{}^{-10} m_e^4 m_P^{20/3} M_*^{10/3}, \tag{4.21}$$

or, equivalently (in cgs units),

$$P_C \approx 2 \times 10^{21} \ \text{dyne/cm}^2 \left(\frac{M_*}{1 \ M_\odot} \right)^{10/3}. \tag{4.22}$$

Combining these results, we find that the mass scale $M_{*\rm nd}$ at which the star becomes nondegenerate is given by

$$M_{*\rm nd} \approx 10^{-3} \ M_\odot. \tag{4.23}$$

This mass scale is roughly the mass of a giant planet such as Jupiter (for more detailed discussion of this issue, see also Hamada and Salpeter, 1961; Shu, 1982). At this point in its evolution, the star has a radius $R_* \sim 0.1 \ R_\odot \sim 7 \times 10^9$ cm and a mean density of roughly $\rho \sim 1$ g/cm^3; these properties are also comparable to those of Jupiter. As a reference point, notice also that neutral hydrogen atoms packed into a cubic array with sides equal to one Bohr radius would give a density of 1.4 g/cm^3. At this transition, a star powered by proton decay has luminosity $L_* \approx 10^{-27} \ L_\odot$ and effective surface temperature $T_* \approx 0.0034$ K.

Once the star becomes nondegenerate, it follows a new track in the H-R diagram. The expressions for the

luminosity and surface temperature [see Eqs. (4.8) and (4.12)] remain valid, but the mass/radius relation changes. Since the density of matter is determined by Coulomb forces for the small mass scales of interest, the density is roughly constant with a value $\rho_0 \sim 1$ g/cm^3. We can thus use the simple relationship $M_* = 4\pi\rho_0 R_*^3/3$. Combining these results, we obtain the relation

$$L_* = \frac{36\pi}{\mathcal{F}^2} \frac{\sigma_B^3}{\Gamma_P^2 \rho_0^2} T_*^{12}, \tag{4.24}$$

or, in terms of numerical values,

$$L_* \approx 10^{-27} \ L_\odot \left[\frac{T_*}{0.0034 \ \mathrm{K}} \right]^{12}. \tag{4.25}$$

This steep power law implies that the effective temperature of the star does not change very much during the final phases of evolution (the mass has to decrease by 12 orders of magnitude in order for the temperature to change by a factor of 10). As a result, effective surface temperatures of order $T_* \sim 10^{-3}$ K characterize the final phases of stellar evolution.

As the star loses mass, it also becomes increasingly optically thin to radiation. As an object becomes transparent, it becomes difficult to meaningfully consider the remnant as a star. An object becomes optically thin when

$$R_* n \sigma < 1, \tag{4.26}$$

where n is the number density of targets and σ is the cross section of interaction between the radiation field and the stellar material. In this present context, we must consider whether the star is optically thin both to the gamma rays produced by proton decay and to the internal radiation at longer wavelengths characteristic of its bolometric surface temperature. This latter condition is required for the radiation field to be thermalized.

We first consider the conditions for which the star becomes optically thin to the gamma rays (with energies $E_\gamma \sim 250$ MeV) produced by proton decay. Since we are considering the interaction of gamma rays with matter, we can write the cross section in the form

$$\sigma = C \sigma_T = C \frac{8\pi}{3} \frac{\alpha^2}{m_e^2}, \tag{4.27}$$

where C is a dimensionless number (of order unity) and σ_T is the Thompson cross section. To a rough approximation, the density will be $\rho \sim 1$ g/cm^3 and hence the number density will have a roughly constant value $n \sim 10^{24}$ cm^{-3}. Using these values, we find that the "star" will be safely optically thick to gamma rays provided its characteristic size is larger than about one meter. In other words, the object must be as big as a large rock. These rocks will not, however, look very much like stars. At the extremely low bolometric temperatures characteristic of the stellar photospheres at these late times, the wavelength of the photospheric photons will be macroscopic and hence will interact much less strongly than the gamma rays. As a result, the spectral energy distribution of these objects will suffer severe departures from blackbody spectral shapes.

In order to consider the optical depth of the star to its internal radiation field, we rewrite the condition (4.26) using the relation $n\sigma = \rho\kappa$, where κ is the opacity. As derived above [Eq. (4.24)], the surface temperature is a slowly varying function in this final phase of evolution; as a result, the wavelength of photons in the stellar photosphere will be of order $\lambda \sim 100$ cm. The interaction of this radiation with the star depends on the chemical purity and the crystal-grain structure of the stellar material. We can obtain a very rough estimate of the opacity by scaling from known astrophysical quantities. For interstellar graphite, for example, the opacity at $\lambda = 100$ μm is roughly $\kappa \sim 1$ cm^2/g and scales with wavelength according to $\kappa \propto \lambda^{-2}$ (see Draine and Lee, 1984). We thus estimate that the opacity in the outer layers of the star/rock will be $\kappa \sim 10^{-8}$ cm^2/g. Thus, in order for the star to be optically thick to its internal radiation, its radius must be $R_* > 10^8$ cm, which corresponds to a mass scale of

$$M_{* \mathrm{thin}} \sim 10^{24} \ \mathrm{g}. \tag{4.28}$$

All of these values should be regarded as highly approximate.

From these results, the ultimate future of white dwarfs, and indeed our own Sun, becomes clear: A white dwarf emerges from degeneracy as a pure sphere of hydrogen when the mass drops below $M_* \sim 10^{-3} \ M_\odot$. Finally, the remaining object becomes transparent to its own internal radiation when its mass dwindles to $M_* \sim 10^{24}$ g, and at this point it is no longer a star. Stellar evolution thus effectively comes to an end.

Just prior to the conclusion of stellar evolution, the white dwarf experiences about 2000 proton-decay events per second and hence has a luminosity of $L_* \sim 10^{-33} \ L_\odot \sim 4$ erg/s and a temperature $T_* \sim 10^{-3}$ K. The time at which this transition occurs is given by $\tau \sim 21 \Gamma_P^{-1}$.

Given these results, we can now describe the complete evolution of a 1.0-$M_\odot$ star (e.g., the Sun), from its birth to its death. The entire evolution of such a star in the Hertzsprung-Russell diagram is plotted in Fig. 6. The star first appears on the stellar birthline (Stahler, 1988) and then follows a pre-main-sequence track onto the main sequence. After exhausting its available hydrogen, the star follows conventional post-main-sequence evolution, including red giant, horizontal branch, red supergiant, and planetary nebula phases. The star then becomes a white dwarf with mass $M_* \approx 0.5 \ M_\odot$ and cools along a constant radius track. The white dwarf spends many cosmological decades ($\eta = 11$–25) near the center of the diagram ($L_* = 10^{14}$ W; $T_* = 63$ K), where the star is powered by annihilation of WIMPs accreted from the galactic halo. When the supply of WIMPs is exhausted, the star cools relatively quickly and obtains its luminosity from proton decay ($L_* \approx 400$ W). The star then follows the evolutionary track in the lower right part of the diagram (with $L_* \sim T_*^{12/5}$) until mass loss from proton decay causes the star to become nondegenerate. The star then becomes a rocklike object supported by Coulomb forces and follows a steeper track (with $L_* \sim T_*^{12}$)

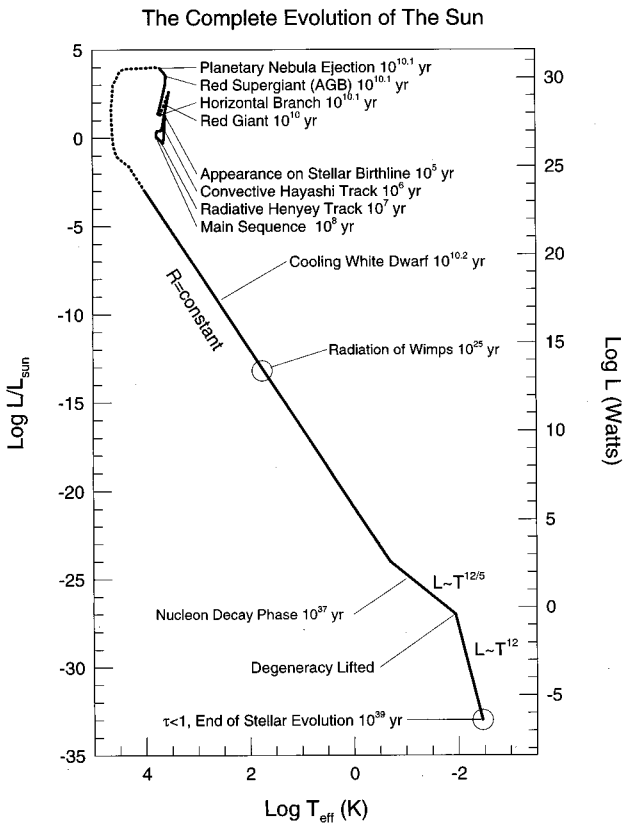

FIG. 6. The complete evolution of the Sun (or any 1-$M_\odot$ star) in the H-R diagram. The track shows the overall evolution of a star, from birth to final death. The star first appears in the H-R diagram on the stellar birthline and then follows a pre-main-sequence track onto the main sequence. After its post-main-sequence evolution (red giant, horizontal branch, red supergiant, and planetary nebula phases), the star becomes a white dwarf and cools along a constant-radius track. The star spends many cosmological decades η=11–25 at a point near the center of the diagram ($L_*=10^{14}$ W; T_*=63 K), where the star is powered by annihilation of WIMPs accreted from the galactic halo. When the supply of WIMPs is exhausted, the star cools relatively quickly and obtains its luminosity from proton decay ($L_*\approx$400 W). The star then follows the evolutionary track in the lower right part of the diagram (with $L_*\sim T_*^{12/5}$) until mass loss from proton decay causes degeneracy to lift. At η=39, the object becomes optically thin and stellar evolution comes to an end.

in the H-R diagram until it becomes optically thin. At this point, the object ceases to be a star and stellar evolution effectively comes to an end. During its entire lifetime, the Sun will span roughly 33 orders of magnitude in luminosity, 9 orders of magnitude in mass, and 8 orders of magnitude in surface temperature.

E. Neutron stars powered by proton decay

The evolution of neutron stars powered by proton decay is qualitatively similar to that of white dwarfs. Since neutron stars are (roughly) the same mass as white dwarfs, and since proton decay occurs on the size scale of an individual nucleon, the luminosity of the neutron star is given by Eqs. (4.8) and (4.9). To leading order, the mass/radius relation for a neutron star is the same as that of white dwarfs with the electron mass m_e replaced by the neutron mass [see Eqs. (4.13) and (4.14)]. Neutron stars are thus ~2000 times smaller than white dwarfs of the same mass and have appropriately warmer surface temperatures. Neutron stars undergoing nucleon decay follow a track in the H-R diagram given by

$$L_* = 10^{-24}\ L_\odot \left[\frac{T_*}{3\ \mathrm{K}}\right]^{12/5}. \tag{4.29}$$

The final phases of the life of a neutron star will differ from those of a white dwarf. In particular, the neutrons in a neutron star come out of degeneracy in a somewhat different manner than do the electrons in a white dwarf. Within a neutron star, the neutrons exist and do not β decay (into protons, electrons, and antineutrinos) because of the extremely high densities, which are close to nuclear densities in the stellar interior. On the exterior, however, every neutron star has a solid crust composed of ordinary matter. As a neutron star squanders its mass through nucleon decay, the radius swells and the density decreases. The outer layers of the star are less dense than the central regions and hence the outer region will experience β decay first. Thus, as the mass decreases, neutrons in the outer portion of the star begin to β decay into their constituent particles and the star must readjust itself accordingly; the net effect is that the crust of ordinary matter thickens steadily and moves inwards towards the center. Once the stellar mass decreases below a critical value M_{C*}, the crust reaches the center of the star and the transition becomes complete. At this point, the star will resemble a white dwarf more than a neutron star.

This process defines a minimum-mass neutron star (see Shapiro and Teukolsky, 1983), which is roughly characterized by the parameters

$$M_{C*} = 0.0925\ M_\odot, \quad \rho_C = 1.55\times 10^{14}\ \mathrm{g\ cm^{-3}},$$
$$R_* = 164\ \mathrm{km}, \tag{4.30}$$

where ρ_C is the central density of the star. It is hard to imagine current-day astrophysical processes which produce stellar objects near this limit. The transformation from a neutron star to a white dwarf occurs with a time scale given by

$$\tau = \frac{1}{\Gamma_P}\ln[M_0/M_{C*}] \approx \frac{2.7}{\Gamma_P}, \tag{4.31}$$

where $M_0\approx 1.4\ M_\odot$ is the initial mass of the neutron star. Notice that neutron stars have a possible mass range of only a factor of ~15, considerably smaller than the mass range available to white dwarfs.

F. Higher-order proton decay

Not all particle physics theories predict proton decay through the process described above with decay rate Γ_P [Eq. (4.2) and Fig. 4]. In theories that do not allow pro-

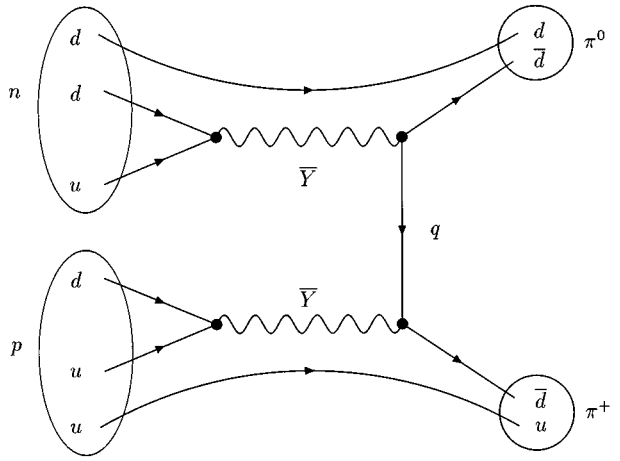

FIG. 7. Representative Feynman diagram for nucleon decay for a $\Delta B=2$ process, i.e., a decay involving two nucleons. The net result of this interaction (shown here in terms of the constituent quarks) is the decay of a neutron and a proton into two pions, $n+p\rightarrow\pi^0+\pi^+$. The Y particle mediates the baryon-number-violating process. Similar diagrams for neutron-neutron decay and for proton-proton decay can be obtained by changing the type of spectator quarks.

ton decay through this first-order process, the proton can often decay through second-order processes and/or through gravitational effects. By a second-order process, we mean an interaction involving two protons and/or neutrons, i.e., $\Delta B=2$, where B is the baryon number. The decay rate for these alternate decay channels is typically much smaller than that discussed above. In this section, we discuss the decay rates and time scales for these higher-order processes (see also Feinberg, Goldhaber, and Steigman, 1978; Wilczek and Zee, 1979; Mohapatra and Marshak, 1980; Weinberg, 1980).

We first consider a class of theories which allow baryon number violation, but do not have the proper vertices for direct proton decay ($\Delta B=1$). In such theories, proton decay can sometimes take place through higher-order processes ($\Delta B>1$). For example, if the quarks in two nucleons interact as shown in Fig. 7, the decay rate is roughly given by

$$\Gamma_2\sim\alpha_5^4\frac{m_P^9}{M_X^8}. \tag{4.32}$$

Even for this higher-order example, the theory must have the proper vertices for this process to occur. We note that some theories forbid this class of decay channels and require $\Delta B=3$ reactions in order for nucleon decay to take place (e.g., Castano and Martin, 1994; Goity and Sher, 1995). For the example shown in Fig. 7, the decay rate is suppressed by a factor of $(m_P/M_X)^4\sim10^{64}$ relative to the simplest GUT decay channel. As a result, the time scale for proton decay through this second-order process is roughly given by

$$\tau_{P2}\approx10^{101}\ \text{yr}\left[\frac{M_X}{10^{16}\ \text{GeV}}\right]^8, \tag{4.33}$$

and the corresponding cosmological time scale is

$$\eta_{P2}=101+8\log_{10}[M_X/10^{16}\ \text{GeV}]. \tag{4.34}$$

In order for this decay process to take place, the protons involved must be near each other. For the case of interest, the protons in white dwarfs are (mostly) in carbon nuclei and hence meet this requirement. Similarly, the neutrons in a neutron star are all essentially at nuclear densities. Notice, however, that free protons in interstellar or intergalactic space will generally not decay through this channel.

The proton can also decay through virtual black-hole processes in quantum gravity theories (e.g., Zel'dovich, 1976; Hawking, Page, and Pope, 1979; Page, 1980; Hawking, 1987). Unfortunately, the time scale associated with this process is not very well determined, but it is estimated to lie in the range

$$10^{46}\ \text{yr}<\tau_{PBH}<10^{169}\ \text{yr}, \tag{4.35}$$

with the corresponding range of cosmological decades

$$46<\eta_{PBH}<169. \tag{4.36}$$

Thus, within the (very large) uncertainty, this time scale for proton decay is commensurate with the second-order GUT processes discussed above.

We note that many other possible modes of nucleon decay exist. For example, supersymmetric theories can give rise to a double neutron decay process of the form shown in Fig. 8(a) (see Goity and Sher, 1995). In this case, two neutrons decay into two neutral kaons. Within the context of standard GUTs, decay channels involving higher-order diagrams can also occur. As another example, the process shown in Fig. 8(b) involves three intermediate vector bosons and thus leads to a proton lifetime approximately given by

$$\eta_{P3}=165+12\log_{10}[M_X/10^{16}\ \text{GeV}]. \tag{4.37}$$

Other final states are possible (e.g., three pions), although the time scales should be comparable. This process [Fig. 8(b)] involves only the most elementary baryon-number-violating processes, which allow interactions of the general form $qq\rightarrow q\bar{q}$. As a result, this decay mode is likely to occur even when the lower-order channels are not allowed.

Finally, we mention the case of sphalerons, which provide yet another mechanism that can lead to baryon number violation and hence proton decay. The vacuum structure of the electroweak theory allows for the nonconservation of baryon number; tunneling events between the different vacuum states in the theory give rise to a change in baryon number (for further details, see Rajaraman, 1987; Kolb and Turner, 1990). Because these events require quantum tunneling, the rate for this process is exponentially suppressed at zero temperature by the large factor $f=\exp[4\pi/\alpha_W]\sim10^{172}$, where α_W is the fine-structure constant for weak interactions. In terms of cosmological decades, the time scale for proton decay through this process has the form $\eta_P=\eta_0+172$, where η_0 is the natural time scale (for no suppression). Using the light crossing time of the proton to determine the natural time scale (i.e., we optimistically take $\eta_0=-31$), we

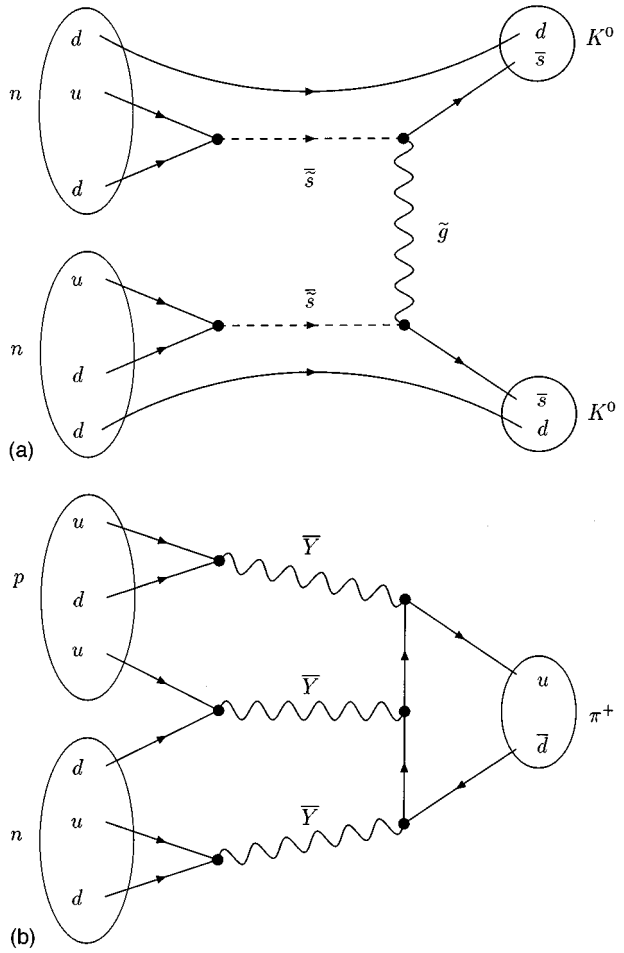

FIG. 8. Representative Feynman diagram for higher-order nucleon decay processes, shown here in terms of the constituent quarks. (a) Double neutron decay for a supersymmetric theory. The net reaction converts two neutrons n into two neutral kaons K^0. The tildes denote the supersymmetric partners of the particles. (b) Double nucleon decay involving three intermediate vector bosons Y. Other final states are possible (e.g., three pions), but the overall decay rate is comparable and implies a decay time scale $\eta_P \sim 165+12 \log_{10}[M_Y/10^{16}\ \mathrm{GeV}]$.

obtain the crude estimate $\eta_P \approx 141$. Since this time scale is much longer than the current age of the universe, this mode of proton decay has not been fully explored. In addition, this process has associated selection rules ('t Hooft, 1976) that place further limits on the possible events which exhibit nonconservation of baryon number. However, this mode of baryon number violation could play a role in the far future of the universe.

To summarize this discussion, we stress that many different mechanisms for baryon number violation and proton decay can be realized within modern theories of particle physics. As a result, it seems likely that the proton must eventually decay with a lifetime somewhere in the range

$$32 < \eta_P < 200, \tag{4.38}$$

where the upper bound was obtained by using $M_X \sim M_{\mathrm{Pl}}$ in Eq. (4.37).

To put these very long time scales in perspective, we note that the total number N_N of nucleons in the observable universe (at the present epoch) is roughly $N_N \sim 10^{78}$. Thus, for a decay time of $\eta = 100$, the expected number N_D of nucleons that have decayed within our observable universe during its entire history is far less than unity, $N_D \sim 10^{-12}$. The experimental difficulties involved in detecting higher-order proton-decay processes thus become clear.

If the proton decays with a lifetime corresponding to $\eta \sim 100$–200, the evolution of white dwarfs will be qualitatively the same as the scenario outlined above, but with a few differences. Since the evolutionary time scale is much longer, pycnonuclear reactions will be much more effective at building the chemical composition of the stars back up to nuclei of high atomic number. Thus stars with a given mass will have higher atomic numbers for their constituent nuclei. However, the nuclear reaction rate [Eq. (4.17)] has an exponential sensitivity to the density. As the star loses mass and becomes less dense [according to the white dwarf mass/radius relation (4.13), (4.14)], pycnonuclear reactions will shut down rather abruptly. If these nuclear reactions stop entirely, the star will quickly become pure hydrogen, and proton decay through a two-body process will be highly suppressed. However, hydrogen tends to form molecules at these extremely low temperatures. The pycnonuclear reaction between the two protons in a hydrogen molecule proceeds at a fixed rate independent of the ambient conditions and has a time scale of roughly $\eta \approx 60$ (see Dyson, 1979, Shapiro and Teukolsky, 1983, and Sec. III.C for simple estimates of pycnonuclear reaction rates). This reaction will thus convert the star into deuterium and helium on a time scale significantly shorter than that of higher-order proton decay. The resulting larger nuclei can then still decay through a second- or third-order process. We also note that this same mechanism allows for hydrogen molecules in intergalactic space to undergo proton decay through a two-body process.

G. Hawking radiation and the decay of black holes

Black holes cannot live forever; they evaporate on long time scales through a quantum-mechanical tunneling process that produces photons and other products (Hawking, 1975). In particular, black holes radiate a thermal spectrum of particles with an effective temperature given by

$$T_{BH} = \frac{1}{8\pi G M_{BH}}, \tag{4.39}$$

where M_{BH} is the mass of the black hole. The total lifetime of the black hole thus becomes

$$\tau_{BH} = \frac{2560\pi}{g_*} G^2 M_{BH}^3, \tag{4.40}$$

where g_* determines the total number of effective degrees of freedom in the radiation field. Inserting numeri-

cal values and scaling to a reference black hole mass of $10^6\ M_\odot$, we find the time scale

$$\tau_{BH} = 10^{83}\ \text{yr}[M_{BH}/10^6\ \ M_\odot]^3, \tag{4.41}$$

or, equivalently,

$$\eta_{BH} = 83 + 3\log_{10}[M_{BH}/10^6\ \ M_\odot]. \tag{4.42}$$

Thus even a black hole with a mass comparable to a galaxy ($M_{BH} \sim 10^{11}\ M_\odot$) will evaporate through this process on the time scale $\eta_{BH} \sim 98$. One important consequence of this result is that for $\eta > 100$, a large fraction of the universe will be in the form of radiation, electrons, positrons, and other decay products.

H. Proton decay in planets

Planets will also eventually disintegrate through the process of proton decay. Since nuclear reactions have a time scale ($\eta \sim 1500$) much longer than that of proton decay and hence are unimportant (see Dyson, 1979), the chemical evolution of the planet is well described by a simple proton-decay cascade scenario (see Sec. IV.C). In particular, this cascade will convert a planet initially composed of iron into a hydrogen lattice in ~6 proton half-lives, or equivalently, on a time scale given by

$$\tau_{\text{planet}} \approx \frac{6\ln 2}{\Gamma_P} \approx 10^{38}\ \text{yr}; \quad \eta_{\text{planet}} \approx 38. \tag{4.43}$$

This time scale also represents the time at which the planet is effectively destroyed.

During the epoch of proton decay, planets radiate energy with an effective luminosity given by

$$L_{\text{planet}} = \mathcal{F} M_{\text{planet}}(t)\Gamma_P \approx 10^{-30}\ \ L_\odot \left[\frac{M_{\text{planet}}}{M_E}\right], \tag{4.44}$$

where M_E is the mass of the Earth and where we have used a proton-decay lifetime of 10^{37} yr. The efficiency factor $\mathcal{F}$ is expected to be of order unity. Thus the luminosity corresponds to ~0.4 mW.

V. LONG-TERM EVOLUTION OF THE UNIVERSE

In spite of the wealth of recent progress in our understanding of cosmology, the future evolution of the universe cannot be unambiguously predicted. In particular, the geometry of the universe as a whole remains unspecified. The universe can be closed ($k=+1$; $\Omega>1$), flat ($k=0$; $\Omega=1$), or open ($k=-1$; $\Omega<1$). In addition, the contribution of vacuum energy density remains uncertain and can have important implications for the long-term evolution of the universe.

A. Future expansion of a closed universe

If the universe is closed, then the total lifetime of the universe, from big bang to big crunch, can be relatively short in comparison with the characteristic time scales of many of the physical processes considered in this paper. For a closed universe with density parameter $\Omega_0 > 1$, the total lifetime τ_U of the universe can be written in the form

$$\tau_U = \Omega_0(\Omega_0 - 1)^{-3/2}\pi H_0^{-1}, \tag{5.1}$$

where H_0 is the present value of the Hubble parameter (see, for example, Peebles, 1993). Notice that, by definition, the age $\tau_U \to \infty$ as $\Omega_0 \to 1$. Current cosmological observations suggest that the Hubble constant is restricted to lie in the range 50–100 km s^{-1} Mpc^{-1} (e.g., Riess, Press, and Kirshner, 1995), and hence the time scale H_0^{-1} is restricted to be greater than ~10 Gyr. Additional observations (e.g., those of Loh and Spillar, 1986) suggest that $\Omega_0 < 2$. Using these results, we thus obtain a lower bound on the total lifetime of the universe,

$$\tau_U > 20\pi\ \ \text{Gyr}. \tag{5.2}$$

In terms of the time variable η, this limit takes the form

$$\eta_U > 10.8. \tag{5.3}$$

This limit is not very strong—if the universe is indeed closed, then there will be insufficient time to allow for many of the processes we describe in this paper.

We also note that a closed-universe model can in principle be generalized to give rise to an oscillating universe. In this case, the big crunch occurring at the end of the universe is really a "big bounce" and produces a new universe of the next generation. This idea originated with Lemaître (1933) and has been subsequently considered in many different contexts (from Tolman, 1934 to Peebles, 1993).

B. Density fluctuations and the expansion of a flat or open universe

The universe will either continue expanding forever or will collapse back in on itself, but it is not commonly acknowledged that observations are unable to provide a definitive answer to this important question. The goal of many present-day astronomical observations is to measure the density parameter Ω, which is the ratio of the density of the universe to that required to close the universe. However, measurements of Ω do not necessarily determine the long-term fate of the universe.

Suppose, for example, that we can ultimately measure Ω to be some value Ω_0 (either less than or greater than unity). This value of Ω_0 means that the density within the current horizon volume has a given ratio to the critical density. If we could view the universe (today) on a much larger size scale (we can't because of causality), then the mean density of the universe of that larger size scale need not be the same as that which we measure within our horizon today. Let Ω_{big} denote the ratio of the density of the universe to the critical density on the aforementioned larger size scale. In particular, we could measure a value $\Omega_0 < 1$ and have $\Omega_{\text{big}} > 1$, or, alternately, we could measure $\Omega_0 > 1$ and have $\Omega_{\text{big}} < 1$. This possibility has been discussed at some length by Linde (1988, 1989, 1990).

To fix ideas, consider the case in which the local value of the density parameter is $\Omega_0 \approx 1$ and the larger scale value is $\Omega_{\rm big}=2>1$. (Note that Ω is not constant in time and hence this value refers to the time when the larger scale enters the horizon.) In other words, we live in an apparently flat universe, which is actually closed on a larger scale. This state of affairs requires that our currently observable universe lies within a large-scale density fluctuation of amplitude

$$\frac{\Delta\rho}{\rho} = \frac{\Omega_0 - \Omega_{\rm big}}{\Omega_{\rm big}} = -\frac{1}{2}, \tag{5.4}$$

where the minus sign indicates that we live in a locally underdense region. Thus a density perturbation with amplitude of order unity is required; furthermore, as we discuss below, the size scale of the perturbation must greatly exceed the current horizon size.

On size scales comparable to that of our current horizon, density fluctuations are constrained to be quite small ($\Delta\rho/\rho \sim 10^{-5}$) because of measurements of temperature fluctuations in the cosmic microwave background radiation (Smoot *et al.*, 1992; Wright *et al.*, 1992). On smaller size scales, additional measurements indicate that density fluctuations are similarly small in amplitude (Meyer, Cheng, and Page, 1991; Gaier *et al.*, 1992; Schuster *et al.*, 1993). The microwave background also constrains density fluctuations on scales *larger than the horizon* (see Grischuk and Zel'dovich, 1978), although the sensitivity of the constraint decreases with increasing size scale λ according to the relation $\sim(\lambda_{\rm hor}/\lambda)^2$, where $\lambda_{\rm hor}$ is the horizon size. Given that density fluctuations have amplitudes of roughly $\sim 10^{-5}$ on the size scale of the horizon today, the smallest size scale λ_1 for which fluctuations can be of order unity is estimated to be

$$\lambda_1 \sim 300\lambda_{\rm hor} \approx 10^6 \ \ {\rm Mpc}. \tag{5.5}$$

For a locally flat universe ($\Omega_0 \approx 1$), density fluctuations with this size scale will enter the horizon at a time $t_1 \approx 3\times 10^7 t_0 \approx 3\times 10^{17}$ yr, or, equivalently, at the cosmological decade

$$\eta_1 \approx 17.5. \tag{5.6}$$

This time scale represents a lower bound on the (final) age of the universe if the present geometry is spatially flat. In practice, the newly closed universe will require some additional time to recollapse [see Eq. (5.1)] and hence the lower bound on the total age becomes approximately $\eta > 18$.

The situation is somewhat different for the case of an open universe with $\Omega_0 < 1$. If the universe is open, then the expansion velocity will (relatively) quickly approach the speed of light, i.e., the scale factor will expand according to $R \propto t$ (for this discussion, we do not include the possibility that $\Omega_0 = 1-\epsilon$, where $\epsilon \ll 1$, i.e., we consider only manifestly open cases). In this limit, the (comoving) particle horizon expands logarithmically with time and hence continues to grow. However, the speed of light sphere—the distance out to which particles in the universe are receding at the speed of light—approaches a constant in comoving coordinates. As a result, density perturbations on very large scales will remain effectively "frozen out" and are thus prevented from further growth as long as the universe remains open. Because the comoving horizon continues to grow, albeit quite slowly, the possibility remains for the universe to become closed at some future time. The logarithmic growth of the horizon implies that the time scale for the universe to become closed depends exponentially on the size scale λ_1 for which density perturbations are of order unity. The resulting time scale is quite long ($\eta \gg 100$), even compared to the time scales considered in this paper.

To summarize, if the universe currently has a nearly flat spatial geometry, then microwave background constraints imply a lower bound on the total age of universe, $\eta > 18$. The evolution of the universe at later times depends on the spectrum of density perturbations. If large-amplitude perturbations ($\Delta\rho/\rho > 1$) enter the horizon at late times, then the universe could end in a big crunch at some time $\eta > \eta_1 = 17.5$. On the other hand, if the very-large-scale density perturbations have small amplitude ($\Delta\rho/\rho \ll 1$), then the universe can continue to expand for much longer time scales. If the universe is currently open, then large-scale density perturbations are essentially frozen out.

C. Inflation and the future of the universe

The inflationary universe scenario was originally invented (Guth, 1981) to solve the horizon problem and the flatness problem faced by standard big-bang cosmology (see also Albrecht and Steinhardt, 1982; Linde, 1982). The problem of magnetic monopoles was also a motivation, but will not be discussed here. In addition, inflationary models which utilize "slowly rolling" scalar fields can produce density fluctuations that later grow into the galaxies, clusters, and superclusters that we see today (Guth and Pi, 1982; Hawking, 1982; Starobinsky, 1982; Bardeen, Steinhardt, and Turner, 1983).

During the inflationary epoch, the scale factor of the universe grows superluminally (usually exponentially with time). During this period of rapid expansion, a small causally connected region of the universe inflates to become large enough to contain the presently observable universe. As a result, the observed homogeneity and isotropy of the universe can be explained, as well as the observed flatness. In order to achieve this resolution of the horizon and flatness problems, the scale factor of the universe must inflate by a factor of e^{N_I}, where the number of e-foldings $N_I \sim 60$. At the end of this period of rapid expansion, the universe must be rethermalized in order to become radiation dominated and recover the successes of standard big-bang theory.

Since the conception of inflation, many models have been produced and many treatments of the requirements for sufficient inflation have been given (e.g., Steinhardt and Turner, 1984; Kolb and Turner, 1990; Linde, 1990). These constraints are generally written in terms of explaining the flatness and causality of the universe at the present epoch. However, it is possible, or

even quite likely, that inflation will solve the horizon and flatness problems far into the future. In this discussion, we find the number N_I of inflationary e-foldings required to solve the horizon and flatness problems until a future cosmological decade η.

Since the number of e-foldings required to solve the flatness problem is (usually) almost the same as that required to solve the horizon problem, it is sufficient to consider only the latter (for further discussion of this issue, see, Kolb and Turner, 1990; Linde, 1990). The condition for sufficient inflation can be written in the form

$$\frac{1}{(HR)_\eta} < \frac{1}{(HR)_B}, \tag{5.7}$$

where the left-hand side of the inequality refers to the inverse of the product of the Hubble parameter and the scale factor evaluated at the future cosmological decade η and the right-hand side refers to the same quantity evaluated at the beginning of the inflationary epoch.

The Hubble parameter at the beginning of inflation takes the form

$$H_B^2 = \frac{8\pi}{3}\frac{M_I^4}{M_{\rm Pl}^2}, \tag{5.8}$$

where M_I is the energy scale at the start of inflation (typically, the energy scale $M_I \sim 10^{16}$ GeV, which corresponds to cosmological decade $\eta_I \sim -44.5$). Similarly, the Hubble parameter at some future time η can be written in the form

$$H_\eta^2 = \frac{8\pi}{3}\frac{M_\eta^4}{M_{\rm Pl}^2}, \tag{5.9}$$

where the energy scale M_η is defined by

$$\rho(\eta) \equiv M_\eta^4 = \rho_0 R_\eta^{-3}. \tag{5.10}$$

In the second equality, we have written the energy density in terms of its value ρ_0 at the present epoch and we assume that the universe remains matter dominated. We also assume that the evolution of the universe is essentially adiabatic from the end of inflation (scale factor $R_{\rm end}$) until the future epoch of interest (scale factor R_η), i.e.,

$$\frac{R_{\rm end}}{R_\eta} = \frac{T_\eta}{fM_I}, \tag{5.11}$$

where $T_\eta = T_0/R_\eta$ is the cosmic microwave background (CMB) temperature at time η and $T_0 \approx 2.7$ K is the CMB temperature today. The quantity fM_I is the CMB temperature at the end of inflation, after thermalization, and we have introduced the dimensionless factor $f<1$.

Combining all of the above results, we obtain the constraint for sufficient inflation,

$$e^{N_I} = \frac{R_{\rm end}}{R_B} > \frac{M_I T_0 R_\eta^{1/2}}{f\sqrt{\rho_0}}. \tag{5.12}$$

Next, we write the present-day energy density ρ_0 in terms of the present-day CMB temperature T_0,

$$\rho_0 = \beta^2 T_0^4, \tag{5.13}$$

where $\beta \approx 100$. The number of e-foldings is thus given by

$$N_I = \ln[R_{\rm end}/R_B] = \ln[M_I/\beta T_0] + \tfrac{1}{2}\ln R_\eta - \ln f. \tag{5.14a}$$

Inserting numerical values and using the definition (1.1) of cosmological decades, we can write this constraint in the form

$$N_I \approx 61 + \ln[M_I/(10^{16}\text{GeV})] + \tfrac{1}{3}(\eta - 10)\ln 10. \tag{5.14b}$$

For example, in order to have enough inflation for the universe to be smooth and flat up to the cosmological decade $\eta = 100$, we require $N_I \approx 130$ e-foldings of inflation. This value is not unreasonable in that $N_I = 130$ is just as natural from the point of view of particle physics as the $N_I = 61$ value required by standard inflation.

We must also consider the density perturbations produced by inflation. All known models of inflation produce density fluctuations, and most models predict that the amplitudes are given by

$$\frac{\Delta\rho}{\rho} \approx \frac{1}{10}\frac{H^2}{\dot{\Phi}}, \tag{5.15}$$

where H is the Hubble parameter and Φ is the scalar field responsible for inflation (Guth and Pi, 1982; Hawking, 1982; Starobinsky, 1982; Bardeen, Steinhardt, and Turner, 1983). In models of inflation with more than one scalar field (e.g., La and Steinhardt, 1989; Adams and Freese, 1991), the additional fields can also produce density fluctuations in accordance with Eq. (5.15).

In order for these density fluctuations to be sufficiently small, as required by measurements of the cosmic microwave background, the potential $V(\Phi)$ for the inflation field must be very flat. This statement can be quantified by defining a "fine-tuning parameter" λ_{FT} through the relation

$$\lambda_{FT} \equiv \frac{\Delta V}{(\Delta\Phi)^4}, \tag{5.16}$$

where ΔV is the change in the potential during a given portion of the inflationary epoch and $\Delta\Phi$ is the change in the scalar field over the same period (Adams, Freese, and Guth, 1991). The parameter λ_{FT} is constrained to less than $\sim 10^{-8}$ for all models of inflation of this class and is typically much smaller, $\lambda_{FT} \sim 10^{-12}$, for specific models. The required smallness of this parameter places tight constraints on models of inflation.

The aforementioned constraints were derived by demanding that the density fluctuations [Eq. (5.15)] be sufficiently small in amplitude over the size scales of current cosmological interest, i.e., from the horizon size (today) down to the size scale of galaxies. These density perturbations are generated over $N_\delta \approx 8$ e-foldings during the inflationary epoch. However, as discussed in Sec. V.B, large-amplitude density fluctuations can come across the horizon in the future and effectively close the universe (see also Linde, 1988, 1989, 1990). In order for the universe to survive (not become closed) up until

some future cosmological decade η, density fluctuations must be small in amplitude for all size scales up to the horizon size at time η [within an order of magnitude—see Eq. (5.1)]. As a result, inflation must produce small-amplitude density fluctuations over many more e-foldings of the inflationary epoch, namely,

$$N_\delta \approx 8 + \tfrac{1}{3}(\eta - 10)\ln 10, \tag{5.17}$$

where η is the future cosmological decade of interest. For example, for η=100 we would require $N_\delta \approx 77$. Although this larger value of N_δ places a tighter bound on the fine-tuning parameter λ_{FT}, and hence a tighter constraint on the inflationary potential, such bounds can be accommodated by inflationary models (see Adams, Freese, and Guth, 1991 for further discussion). Loosely speaking, once the potential is flat over the usual N_δ=8 e-foldings required for standard inflationary models, it is not that difficult to make it flat for N_δ=80.

D. Background radiation fields

Many of the processes discussed in this paper will produce background radiation fields, which can be important components of the universe (see, Bond, Carr, and Hogan, 1991 for a discussion of present-day backgrounds). Stars produce radiation fields and low-mass stars will continue to shine for several more cosmological decades (Sec. II). The net effect of WIMP capture and annihilation in white dwarfs (Sec. III.E) will be to convert a substantial portion of the mass energy of galactic halos into radiation. Similarly, the net effect of proton decay (Sec. IV) will convert the mass energy of the baryons in the universe into radiation. Finally, black holes will evaporate as well (Sec. IV.H), ultimately converting their rest mass into radiation fields. As we show below, each of these radiation fields will dominate the radiation background of the universe for a range of cosmological decades, before being successively redshifted to insignificance.

The overall evolution of a radiation field in an expanding universe can be described by the simple differential equation

$$\frac{d\rho_{\rm rad}}{dt} + 4\frac{\dot{R}}{R}\rho_{\rm rad} = S(t), \tag{5.18}$$

where $\rho_{\rm rad}$ is the energy density of the radiation field and $S(t)$ is a source term (see, for example, Kolb and Turner, 1990).

Low-mass stars will continue to shine far into the future. The source term for this stellar radiation can be written in the form

$$S_*(t) = n_* L_* = \epsilon_* \Omega_* \rho_0 R^{-3} \frac{1}{t_*}, \tag{5.19}$$

where L_* and n_* are the luminosity and number density of the low-mass stars. In the second equality, we have introduced the present-day mass fraction of low-mass stars Ω_*, the nuclear burning efficiency $\epsilon_* \sim 0.007$, the effective stellar lifetime t_*, and the present-day energy density of the universe ρ_0. For this example, we have written these expressions for a population of stars with only a single mass; in general, one should of course consider a distribution of stellar masses and then integrate over the distribution. As a further refinement, one could also include the time dependence of the stellar luminosity L_* (see Sec. II).

For a given geometry of the universe, we find the solution for the background radiation field from low-mass stars,

$$\rho_{{\rm rad}*} = \epsilon_* \Omega_* \rho(R) f \frac{t}{t_*}, \tag{5.20}$$

where the dimensionless factor f=1/2 for an open universe and f=3/5 for a flat universe. This form is valid until the stars burn out at time $t = t_*$. After that time, the radiation field simply redshifts in the usual manner, $\rho_{{\rm rad}*} \sim R^{-4}$.

For the case of WIMP annihilation in white dwarfs, the source term is given by

$$S_W(t) = L_* n_* = \Omega_W \rho_0 R^{-3} \Gamma, \tag{5.21}$$

where L_* and n_* are the luminosity and number density of the white dwarfs. In the second equality, we have written the source in terms of the energy density in WIMPs, where Ω_W is the present-day mass fraction of WIMPs and Γ is the effective annihilation rate. The solution for the background radiation field from WIMP annihilation can be found,

$$\rho_{\rm wrb}(t) = f \Omega_W \rho(R) \Gamma t, \tag{5.22}$$

where the dimensionless factor f is defined above. This form is valid until the galactic halos begin to run out of WIMP dark matter at time $t \sim \Gamma^{-1} \sim 10^{25}$ yr, or until the galactic halo ejects nearly all of its white dwarfs. We note that direct annihilation of dark matter will also contribute to the background radiation field of the universe. However, this radiation will be highly nonthermal; the annihilation products will include gamma rays with characteristic energy $E_\gamma \sim 1$ GeV.

For the case of proton decay, the effective source term for the resulting radiation field can be written

$$S_P(t) = \mathcal{F} \Omega_B \rho_0 R^{-3} \Gamma_P \, e^{-\Gamma_P t}, \tag{5.23}$$

where Ω_B is the present-day contribution of baryons to the total energy density ρ_0, Γ_P is the proton-decay rate, and $\mathcal{F}$ is an efficiency factor of order unity. For a given geometry of the universe, we obtain the solution for the background radiation field from proton decay,

$$\rho_{\rm prb}(t) = \mathcal{F} \, \Omega_B \rho(R) F(\xi), \tag{5.24}$$

where $F(\xi)$ is a dimensionless function of the dimensionless time variable $\xi \equiv \Gamma_P t$. For an open universe,

$$F(\xi) = \frac{1 - (1+\xi)e^{-\xi}}{\xi}, \tag{5.25}$$

whereas for a flat universe,

$$F(\xi) = \xi^{-2/3} \int_0^\xi x^{2/3} e^{-x} dx = \xi^{-2/3} \gamma(5/3, \xi), \tag{5.26}$$

where $\gamma(5/3,\xi)$ is the incomplete gamma function (Abramowitz and Stegun, 1972).

For black-hole evaporation, the calculation of the radiation field is more complicated because the result depends on the mass distribution of black holes in the universe. For simplicity, we shall consider a population of black holes with a single mass M and mass fraction Ω_{BH} (scaled to the present epoch). The source term for black-hole evaporation can be written in the form

$$S_{BH}(t)=\Omega_{BH}\rho_0 R^{-3}\frac{1}{3\tau_{BH}}\frac{1}{1-t/\tau_{BH}}, \tag{5.27}$$

where $\tau_{\rm BH}$ is the total lifetime of a black hole of the given mass M [see Eq. (4.37)]. For an open universe, we obtain the solution for the background radiation field from black-hole evaporation

$$\rho_{\rm bhr}(t)=\Omega_{BH}\rho(R)F(\xi), \tag{5.28}$$

where the dimensionless time variable $\xi=t/\tau_{BH}$. For an open universe, the dimensionless function $F(\xi)$ is given by

$$F(\xi)=\frac{1}{3\xi}\left\{\ln\left[\frac{1}{1-\xi}\right]-\xi\right\}, \tag{5.29}$$

whereas for a flat universe,

$$F(\xi)=\frac{1}{3\xi^{2/3}}\int_0^{\xi}\frac{x^{2/3}dx}{1-x}. \tag{5.30}$$

Each of the four radiation fields discussed here has the same general time dependence. For times short compared to the depletion times, the radiation fields have the form

$$\rho(t)\approx\Omega_X\rho(R)\Gamma_X t, \tag{5.31}$$

where Ω_X is the present-day abundance of the raw material and Γ_X is the effective decay rate (notice that we have neglected dimensionless factors of order unity). After the sources (stars, WIMPs, protons, black holes) have been successively exhausted, the remaining radiation fields simply redshift away, i.e.,

$$\rho(t)=\rho(t_{\rm end})(R/R_{\rm end})^{-4}, \tag{5.32}$$

where the subscript refers to the end of the time period during which the ambient radiation was produced.

Due to the gross mismatch in the characteristic time scales, each of the radiation fields will provide the dominate contribution to the radiation content of the universe over a given time period. This trend is illustrated in Fig. 9, which shows the relative contribution of each radiation field as a function of cosmological time η. For purposes of illustration, we have assumed an open universe and the following source abundances: low-mass stars $\Omega_*=10^{-3}$, weakly interacting massive particles $\Omega_W=0.2$, baryons $\Omega_B=0.05$, and black holes $\Omega_{BH}=0.1$. At present, the cosmic microwave background (left over from the big bang itself) provides the dominant radiation component. The radiation field from starlight will dominate the background for the next several cosmological decades. At cosmological decade $\eta\sim16$, the radiation field resulting from WIMP annihilation will overtake the starlight background and become the dominant component. At the cosmological decade $\eta\sim30$, the WIMP annihilation radiation field will have redshifted away and the radiation field from proton decay will begin to dominate. At much longer time scales, $\eta\sim60$, the radiation field from black-hole evaporation provides the dominant contribution (where we have used 10^6-$M_\odot$ black holes for this example).

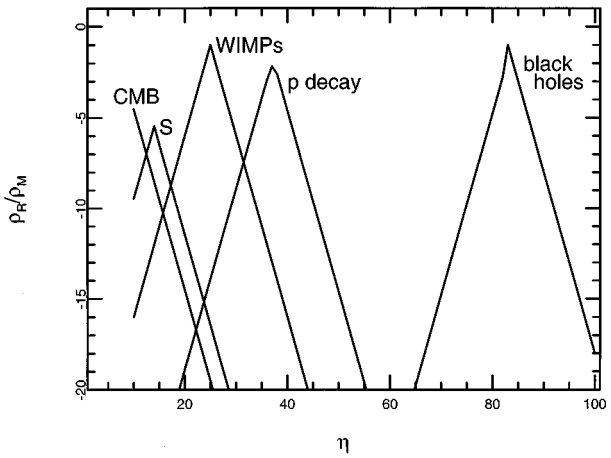

FIG. 9. Background radiation fields in the universe. The vertical axis represents the ratio of the energy density in radiation to the total energy density (assuming the universe remains matter dominated). The horizontal axis is given in terms of cosmological decades η. The various curves represent the radiation fields from the cosmic microwave background (CMB), light from low-mass stars (S), radiation from WIMP annihilation in white dwarfs (WIMPs), radiation from proton decay (p decay), and black-hole evaporation (black holes).

The discussion thus far has focused on the total energy density $\rho_{\rm rad}$ of the background radiation fields. One can also determine the spectrum of the background fields as a function of cosmological time, i.e., one could follow the time evolution of the radiation energy density per unit frequency. In general, the spectra of the background radiation fields will be nonthermal for two reasons:

(1) The source terms are not necessarily perfect blackbodies. The stars and black holes themselves produce nearly thermal spectra, but objects of different masses will radiate like blackbodies of different temperatures. One must therefore integrate over the mass distribution of the source population. It is interesting that this statement applies to all of the above sources. For the first three sources (low-mass stars, white dwarfs radiating WIMP annihilation products, and white dwarfs powered by proton decay), the mass distribution is not very wide and the resulting composite spectrum is close to that of a blackbody. For the case of black holes, the spectrum is potentially much wider, but the mass distribution is far more uncertain.

(2) The expansion of the universe redshifts the radiation field as it is produced and thereby makes the resultant spectrum wider than a thermal distribution. How-

ever, due to the linear time dependence of the emission [Eq. (5.31)], most of the radiation is emitted in the final cosmological decade of the source's life. The redshift effect is thus not as large as one might naively think.

To summarize, the radiation fields will experience departures from a purely thermal distribution. However, we expect that the departures are not overly severe.

The above results, taken in conjunction with our current cosmological understanding, imply that it is unlikely that the universe will become radiation dominated in the far future. The majority of the energy density at the present epoch is (most likely) in the form of nonbaryonic dark matter of some kind. A substantial fraction of this dark matter resides in galactic halos, and some fraction of these halos can be annihilated and hence converted into radiation through the white dwarf capture process outlined in Sec. III.E. However, an equal or larger fraction of this dark matter resides outside of galaxies and/or can escape destruction through evaporation from galactic halos. Thus, unless the dark matter particles themselves decay into radiation, it seems that enough nonbaryonic dark matter should survive to keep the universe matter dominated at all future epochs; in addition, the leftover electrons and positrons will help prevent the universe from becoming radiation dominated (see also Page and Mckee, 1981a, 1981b).

E. Possible effects of vacuum energy density

If the universe contains a nonvanishing contribution of vacuum energy to the total energy density, then two interesting long-term effects can arise. The universe can enter a second inflationary phase, in which the universe expands superluminally (Guth, 1981; see also Linde, 1982; Albrecht and Steinhardt 1983). Alternately, the vacuum can, in principle, be unstable and the universe can tunnel into an entirely new state (e.g., Coleman, 1977, 1985). Unfortunately, the contribution of the vacuum to the energy density of the universe remains unknown. In fact, the "natural value" of the vacuum energy density appears to be larger than the cosmologically allowed value by many orders of magnitude. This discrepancy is generally known as the "cosmological-constant problem" and has no currently accepted resolution (see the reviews of Weinberg, 1989; Carroll, Press, and Turner, 1992).

1. Future inflationary epochs

We first consider the possibility of a future inflationary epoch. The evolution equation for the universe can be written in the form

$$\left(\frac{\dot{R}}{R}\right)^2 = \frac{8\pi G}{3}(\rho_M + \rho_{\rm vac}), \tag{5.33}$$

where R is the scale factor, ρ_M is the energy density in matter, and $\rho_{\rm vac}$ is the vacuum energy density. We have assumed a spatially flat universe for simplicity. The matter density varies with the scale factor according to $\rho_M \sim R^{-3}$, whereas the vacuum energy density is constant. We can define the ratio

$$\nu \equiv \rho_{\rm vac}/\rho_0, \tag{5.34}$$

i.e., the ratio of the vacuum energy density to that of the matter density ρ_0 at the present epoch. We can then integrate Eq. (5.6) into the future and solve for the time $t_{\rm vac}$ at which the universe becomes vacuum dominated. We find the result

$$t_{\rm vac} = t_0 + \tau \frac{\sinh^{-1}[1] - \sinh^{-1}[\nu^{1/2}]}{\nu^{1/2}}, \tag{5.35}$$

where t_0 is the present age of the universe and we have defined $\tau \equiv (6\pi G\rho_0)^{-1/2}$; both time scales t_0 and τ are approximately 10^{10} yr.

Several results are immediately apparent from Eq. (5.35). If the vacuum energy density provides any appreciable fraction of the total energy density at the present epoch (in other words, if ν is not too small), then the universe will enter an inflationary phase in the very near future. Furthermore, almost any nonvanishing value of the present-day vacuum energy will lead the universe into an inflationary phase on the long time scales considered in this paper. For small values of the ratio ν, the future inflationary epoch occurs at the cosmological decade given by

$$\eta_{\rm inflate} \approx 10 + \frac{1}{2}\log_{10}\left[\frac{1}{\nu}\right]. \tag{5.36}$$

For example, even for a present-day vacuum contribution as small as $\nu \sim 10^{-40}$, the universe will enter an inflationary phase at the cosmological decade $\eta_{\rm inflate} \approx 30$, long before protons begin to decay. In other words, the traditional cosmological-constant problem becomes even more severe when we consider future cosmological decades.

If the universe enters into a future inflationary epoch, several interesting consequences arise. After a transition time comparable to the age of the universe at the epoch (5.36), the scale factor of the universe will begin to grow superluminally. Because of this rapid expansion, all of the astrophysical objects in the universe become isolated and eventually become out of causal contact. In other words, every given comoving observer will see an effectively shrinking horizon (the particle horizon does not actually get smaller, but this language has become common in cosmology—see Ellis and Rothman, 1993 for further discussion of horizons in this context). In particular, astrophysical objects, such as galaxies and stars, will cross outside the speed-of-light sphere and hence disappear from view. For these same astrophysical objects, the velocity relative to the observer becomes larger than the speed of light and their emitted photons are redshifted to infinity.

2. Tunneling processes

We next consider the possibility that the universe is currently in a false vacuum state. In other words, a lower-energy vacuum state exists and the universe can

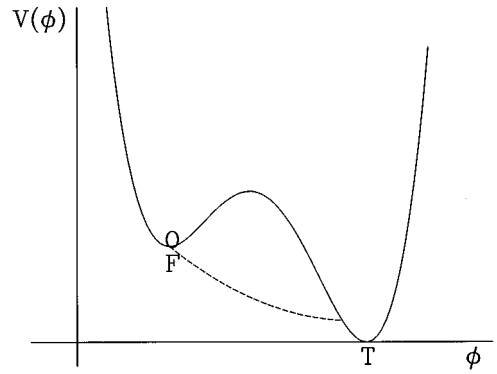

FIG. 10. Potential $V(\Phi)$ of a scalar field which determines the vacuum state of the universe. This potential has both a false vacuum state (labeled F) and a true vacuum state (labeled T). As illustrated by the dashed curve, the universe can tunnel from the false vacuum state into the true vacuum state at some future time.

someday tunnel to that lower-energy state. This problem, the fate of the false vacuum, was first explored quantitatively by Voloshin *et al.* (1975) and by Coleman (1977). Additional effects have been studied subsequently, including gravity (Coleman and De Luccia, 1980) and finite-temperature effects (e.g., Linde, 1983).

To obtain quantitative results, we consider an illustrative example in which the vacuum energy density of the universe can be described by the dynamics of a single scalar field. Once a field configuration becomes trapped in a metastable state (the false vacuum), bubbles of the true vacuum state nucleate in the sea of false vacuum and begin growing spherically. The speed of the bubble walls quickly approaches the speed of light. The basic problem is to calculate the tunneling rate (the decay probability) from the false vacuum state to the true vacuum state, i.e., the bubble nucleation rate $\mathcal{P}$ per unit time per unit volume. For tunneling of scalar fields at zero temperature (generally called quantum tunneling), the four-dimensional Euclidean action S_4 of the theory largely determines this tunneling rate. The decay probability $\mathcal{P}$ can be written in the form

$$\mathcal{P}=Ke^{-S_4}, \tag{5.37}$$

where K is a determinental factor (see Coleman, 1977, 1985). For purposes of illustration, we assume a generic quartic potential of the form

$$V(\Phi)=\lambda\Phi^4-a\Phi^3+b\Phi^2+c\Phi+d. \tag{5.38}$$

We can then write the action S_4 in the form

$$S_4=\frac{\pi^2}{3\lambda}(2-\delta)^{-3}\mathcal{R}(\delta), \tag{5.39}$$

where $\delta\equiv 8\lambda b/a^2$ and where $\mathcal{R}$ is a slowly varying function that has a value near unity for most of the range of possible quartic potentials (Adams, 1993). The composite shape parameter δ varies from 0 to 2 as the potential $V(\Phi)$ varies from having no barrier height to having nearly degenerate vacua (see Fig. 10).

Even though Eqs. (5.37)–(5.39) describe the tunneling rate, we unfortunately do not know what potential (if any) describes our universe and hence it is difficult to obtain a precise numerical estimate for this time scale. To get some quantitative feeling for this problem, we consider the following example. For the case of no tunneling barrier (i.e., for $S_4=0$), the characteristic decay probability is given by $\mathcal{P}_0\sim K\sim M_V^4$, where M_V is the characteristic energy scale for the scalar field. For $M_V=10^{16}$ GeV (roughly the GUT scale), $\mathcal{P}_0\sim 10^{129}$ $\mathrm{s}^{-1}\,\mathrm{cm}^{-3}$. With this decay rate, the universe within a characteristic volume M_V^{-3} would convert from false vacuum to true vacuum on a time scale of $\sim 10^{-24}$ s. Clearly, however, the actual decay time scale must be long enough that the universe has not decayed by the present epoch. In order to ensure that the universe has survived, we require that no nucleation events have occurred within the present horizon volume [$\sim$(3000 Mpc)^3] during the current age of the universe ($\sim 10^{10}$ yr). This constraint implies that the action S_4 must be sufficiently large to suppress nucleation, in particular,

$$S_4>231\ln 10\approx 532. \tag{5.40}$$

The question then becomes: is this value for S_4 reasonable? For the parameter λ, a reasonable range of values is $0.1<\lambda<1$; similarly, for δ, we take the range $0.1<\delta<1.9$. Using the form (5.39) for the action and setting $\mathcal{R}=1$, we find the approximate range

$$0.5<S_4<3\times 10^4. \tag{5.41}$$

Thus the value required for the universe to survive to the present epoch [Eq. (5.40)] can be easily realized within this simple model. In the future, however, the universe could tunnel into its false vacuum state at virtually any time, as soon as tomorrow or as late as $\eta=10^4$. If and when this tunneling effect occurs, the universe will change its character almost completely. The physical laws of the universe, or at least the values of all of the physical constants, would change as the phase transition completes (see Sher, 1989 and Crone and Sher, 1990 for a discussion of changing laws of physics during a future phase transition). The universe, as we know it, would simply cease to exist.

Vacuum tunneling of the entire universe is certainly one of the more speculative topics considered in this paper. Nevertheless, its inclusion is appropriate since the act of tunneling from a false vacuum into a true vacuum would change the nature of the universe more dramatically than just about any other physical process.

It is also possible for the universe to spontaneously create "child universes" through a quantum tunneling process roughly analogous to that considered above (Sato *et al.*, 1982; Blau, Guendelman, and Guth, 1987; Hawking, 1987). In this situation, a bubble of false vacuum energy nucleates in an otherwise empty space time. If this bubble is sufficiently large, it will grow exponentially and will eventually become causally disconnected from the original space time. In this sense, the newly created bubble becomes a separate "child universe." The newly created universe appears quite differ-

ent to observers inside and outside the bubble. Observers inside the bubble see the local universe in a state of exponential expansion. Observers outside the bubble, in the empty space-time background, see the newly created universe as a black hole that collapses and becomes causally disconnected. As a result, these child universes will not greatly affect the future evolution of our universe because they (relatively) quickly become out of causal contact.

One potentially interesting effect of these child universes is that they can, in principle, receive information from our universe. Before the newly created universe grows out of causal contact with our own universe, it is connected through a relativistic wormhole, which can provide a conduit for information transfer and perhaps even the transfer of matter (see Visser, 1995 for further discussion of wormholes and transferability). The implications of this possibility are the subject of current debate (for varying points of view, see, for example, Linde, 1988, 1989; Tipler, 1992; Davies, 1994).

F. Speculations about energy and entropy production in the far future

Thus far in this paper, we have shown that entropy can be generated (and hence work can be done) up to cosmological decades $\eta \sim 100$. For very long time scales $\eta \gg 100$, the future evolution of the universe becomes highly uncertain, but the possibility of continued entropy production is very important (see Sec. VI.D). Here, we briefly assess some of the possible ways for energy and entropy to be generated in the far future.

1. Continued formation and decay of black holes

For the case of a flat spatial geometry for the universe, future density perturbations can provide a mechanism to produce entropy. These density perturbations create large structures which can eventually collapse to form black holes. The resulting black holes, in turn, evaporate by emitting Hawking radiation and thus represent entropy (and energy) sources (see also Page and McKee, 1981a; Frautschi, 1982). Density perturbations of increasingly larger size scale λ will enter the horizon as the universe continues to expand. The corresponding mass scale M_λ of these perturbations is given by

$$M_\lambda = M_0 \left(\frac{t_\lambda}{t_0} \right), \tag{5.42}$$

where t_λ is the time at which the perturbation enters the horizon and $M_0 \approx 10^{22}\ M_\odot$ is the total mass within the present-day horizon (at time t_0).

The time t_λ represents the time at which a given perturbation enters the horizon and begins to grow; a large structure (such as a black hole) can only form at some later time after the perturbation becomes nonlinear. Suppose that a density perturbation has an initial amplitude δ_λ when it enters the horizon. In the linear regime, the perturbation will grow according to the usual relation

$$\delta = \delta_\lambda \left(\frac{t}{t_\lambda} \right)^{2/3}, \tag{5.43}$$

where $\delta \equiv \Delta\rho/\rho$ and $t > t_\lambda$ (see Peebles, 1993). Using this growth law, the epoch $\eta_{\rm nl}$ at which the perturbation becomes nonlinear can be written in the form

$$\eta_{\rm nl} = \eta_\lambda - \tfrac{3}{2} \log_{10} \delta_\lambda . \tag{5.44}$$

For example, if the perturbation has an amplitude $\delta_\lambda = 10^{-4}$, then it becomes nonlinear at time $\eta_{\rm nl} = \eta_\lambda + 6$. Since we are interested in very long time scales $\eta > 100$, the difference between the horizon crossing time η_λ and the time $\eta_{\rm nl}$ of nonlinearity is not overly large.

One possible result of this process is the production of a large black hole with a mass $M_{BH} \sim M_\lambda$. The time scale for such a black hole to evaporate through the Hawking process is given by

$$\eta_{BH} = 101 + 3\,\eta_\lambda , \tag{5.45}$$

where we have combined Eqs. (4.42) and (5.42). Since $\eta_{BH} \gg \eta_\lambda \sim \eta_{nl}$, the universe can form black holes faster than they can evaporate. Thus, for the case of a geometrically flat universe, future density perturbations can, in principle, continue to produce black holes of increasingly larger mass. In this case, the universe will always have a source of entropy—the Hawking radiation from these black holes.

We note that these bound perturbations need not necessarily form black holes. The material is (most likely) almost entirely nondissipative and collisionless, and will thus have a tendency to form virialized clumps with binding energy per unit mass of order $\sim \delta c^2$. Thus, unless the perturbation spectrum is tilted so that δ is of order unity on these much larger scales, the ensuing dynamics is probably roughly analogous to that of a cluster-mass clump of cold dark matter in our present universe. However, even if the mass of the entire perturbation does not form a single large black hole, smaller-scale structures can in principle form black holes, in analogy to those currently in the centers of present-day galaxies. In addition, it is possible that the existing black holes can merge faster than they evaporate through the Hawking process (see also Sec. III.D). Thus the possibility remains for the continued existence of black holes in the universe.

The process outlined here, the formation of larger and larger black holes, can continue as long as the universe remains spatially flat and the density perturbations that enter the horizon are not overly large. The inflationary universe scenario provides a mechanism to achieve this state of affairs, at least up to some future epoch [see Sec. V.C and in particular Eq. (5.14)]. Thus the nature of the universe in the far future $\eta \gg 100$ may be determined by the physics of the early universe (in particular, inflation) at the cosmological decade $\eta \sim -45$.

Notice that at these very late times, $\eta \gg 100$, the matter entering the horizon will already be "processed" by the physical mechanisms described earlier in the paper. The nucleons will have (most likely) already decayed and the matter content of the universe will be mostly electrons,

positrons, and nonbaryonic dark-matter particles. Annihilation of both e^+-e^- pairs and dark matter will occur simultaneously with perturbation growth and hence the final mass of the black hole will be less than M_λ. This issue must be studied in further depth.

2. Particle annihilation in an open universe

If the universe is open, however, then future density perturbations are effectively frozen out (see Sec. V.B) and the hierarchy of black holes described above cannot be produced. For an open universe, continued energy and entropy production is more difficult to achieve. One process that can continue far into the future, albeit at a very low level, is the continued annihilation of particles. Electrons and positrons represent one type of particle that can annihilate (see also Page and McKee, 1981a, 1981b), but the discussion given below applies to a general population of particles.

Consider a collection of particles with number density n. The time evolution of the particle population is governed by the simple differential equation

$$\frac{dn}{dt}+3Hn=-\langle\sigma v\rangle n^2, \tag{5.46}$$

where $H=\dot{R}/R$ is the Hubble parameter and $\langle\sigma v\rangle$ is the appropriate average of interaction cross section times the speed (e.g., see Kolb and Turner, 1990). Since we are interested in the case for which the expansion rate is much larger than the interaction rate, the particles are very far from thermal equilibrium and we can neglect any back reactions that produce particles. For this example, we consider the universe to be open, independent of the activity of this particle population. As a result, we can write $R\propto t$ and hence $H=1/t$. We also take the quantity $\langle\sigma v\rangle$ to be a constant in time (corresponding to s-wave annihilation). With these approximations, the differential equation (5.46) can be integrated to obtain the solution

$$n(t)=n_1\left(\frac{t_1}{t}\right)^3\{1+\Delta_\infty[1-(t_1/t)^2]\}^{-1}, \tag{5.47}$$

where we have defined the quantity

$$\Delta_\infty\equiv\tfrac{1}{2}n_1t_1\langle\sigma v\rangle \tag{5.48}$$

and where we have invoked the boundary condition

$$n(t_1)=n_1=\text{const}. \tag{5.49}$$

Analogous solutions for particle annihilation can be found for the case of a flat universe ($H=2/3t$) and an inflating universe (H=const).

The difference between the solution (5.47) and the simple adiabatic scaling solution $n(t)=n_1(t_1/t)^3$ is due to particle annihilation, which is extremely small but nonzero. This statement can be quantified by defining the fractional difference Δ between the solution (5.47) and the adiabatic solution, i.e.,

$$\Delta(t)\equiv\frac{\Delta n}{n}(t)=\Delta_\infty[1-(t_1/t)^2]. \tag{5.50}$$

Over the entire (future) lifetime of the universe, the comoving fraction of particles that annihilate is given by the quantity Δ_∞, which is both finite and typically much less than unity. For example, if we consider the largest possible values at the present epoch ($\sigma\approx\sigma_T\approx10^{-24}$ cm^2, $n_1\approx10^{-6}$ cm^{-3}, $t_1\approx3\times10^{17}$ s, and $v=c$), then $\Delta_\infty\approx10^{-2}$. The fraction Δ_∞ will generally be much smaller than this example. The fact that the fraction Δ_∞ is finite implies that the process of particle annihilation can provide only a finite amount of energy over the infinite time interval $\eta_1<\eta<\infty$.

3. Formation and decay of positronium

Another related process that will occur on long time scales is the formation and eventual decay of positronium. This process has been studied in some detail by Page and McKee (1981a, 1981b; see also the discussion of Barrow and Tipler, 1986); here we briefly summarize their results. The time scale for the formation of positronium in a flat universe is given by

$$\eta_{\text{form}}\approx85+2(\eta_P-37)-\tfrac{2}{3}\log_{10}[\Omega_e], \tag{5.51}$$

where η_P is the proton lifetime (see Sec. IV) and where Ω_e is the mass fraction of $e^\pm$ after proton decay. For a flat or nearly flat universe, most of the electrons and positrons become bound into positronium. In an open universe, some positronium formation occurs, but most electrons and positrons remain unattached.

At the time of formation, the positronium atoms are generally in states of very high quantum number (and have radii larger than the current horizon size). The atoms emit a cascade of low-energy photons until they reach their ground state; once this occurs, the positronium rapidly annihilates. The relevant time scale for this decay process is estimated to be

$$\eta_{\text{decay}}\approx141+4(\eta_P-37)-\tfrac{8}{3}\log_{10}[\Omega_e]. \tag{5.52}$$

VI. SUMMARY AND DISCUSSION

Our goal has been to present a plausible and quantitative description of the future of the universe. Table I outlines the most important events in the overall flow of time, as well as the cosmological decades at which they occur [see Eq. (1.1)]. In constructing this table, representative values for the (often uncertain) parameters have been assumed; the stated time scales must therefore be viewed as approximate. Furthermore, as a general rule, both the overall future of the universe and the time line suggested in Table I become more and more uncertain in the face of successively deeper extrapolations into time. Some of the effects we have described will compete with one another, and hence not all the relevant physical processes can proceed to completion. Almost certainly, parts of our current time line will undergo dramatic revision as physical understanding improves. We have been struck by the remarkable natural utility of the logarithmic "clock," η, in organizing the passage of time. Global processes that can characterize the entire universe rarely span more than a few cosmo-

TABLE I. Important events in the history and future of the universe.

The big bang	$\eta=-\infty$
Planck Epoch	−50.5
GUT Epoch	−44.5
Electroweak phase transition	−17.5
Quarks become confined into hadrons	−12.5
Nucleosynthesis	−6
Matter domination	4
Recombination	5.5
First possible stellar generation	6
Formation of the Galaxy	9
Formation of the Solar System	9.5
Today: The Present Epoch	10
Our Sun dies	10.2
Close encounter of Milky Way with Andromeda (M31)	10.2
Lower bound on the age of closed universe	10.8
Lifetime of main-sequence stars with lowest mass	13
End of conventional star formation	14
Planets become detached from stars	15
Star formation via brown dwarf collisions	16
Lower bound on age of flat universe (with future $\Delta\rho/\rho>1$)	18
Stars evaporate from the Galaxy	19
Planetary orbits decay via gravitational radiation	20
WIMPs in the galactic halo annihilate	22.5
Star formation via orbital decay of brown dwarf binaries	23
Stellar orbits in the galaxy decay via gravitational radiation	24
White dwarfs deplete WIMPs from the galactic halo	25
Black holes accrete stars on galactic size scale	30
Black holes accrete stars on cluster size scale	33
Protons decay	37
Neutron stars β-decay	38
Planets destroyed by proton decay	38
White dwarfs destroyed by proton decay	39
Axions decay into photons	42
Hydrogen molecules experience pycnonuclear reactions	60
Stellar-sized black holes evaporate	65
Black holes with $M=10^6\ M_\odot$ evaporate	83
Positronium formation in a flat universe	85
Galaxy-sized black holes evaporate	98
Black hole with mass of current horizon scale evaporates	131
Positronium decay in a flat universe	141
Higher-order proton-decay processes	~100–200

logical decades, and the ebb and flow of events is dispersed quite evenly across a hundred and fifty orders of magnitude in time, i.e., $-50<\eta<100$.

A. Summary of results

Our specific contributions to physical eschatology can be summarized as follows:

(1) We have presented new stellar evolution calculations which show the long-term behavior of very-low-mass stars (see Fig. 1). Stars with very small mass ($\sim 0.1\ M_\odot$) do not experience any red giant phases. As they evolve, these stars become steadily brighter and bluer, reaching first a maximum luminosity and second a maximum temperature, prior to fading away as helium white dwarfs.

(2) Both stellar evolution and conventional star formation come to an end at the cosmological decade $\eta\sim 14$. This time scale only slightly exceeds the longest evolution time for a low-mass star. It also corresponds to the time at which the galaxy runs out of raw material (gas) for producing new stars. The era of conventional stars in the universe is confined to the range $6<\eta<14$.

(3) We have introduced the final mass function (FMF), i.e., the distribution of masses for the degenerate stellar objects left over from stellar evolution (see Fig. 2). Roughly half of these objects will be white dwarfs, with most of the remainder being brown dwarfs. Most of the mass, however, will be in the form of white dwarfs [see Eqs. (2.22) and (2.23)].

(4) We have explored a new mode of continued star formation through the collisions of substellar objects (see Fig. 3). Although the time scale for this process is quite long, this mode of star formation will be the leading source of new stars for cosmological decades in the range $15<\eta<23$.

(5) We have presented a scenario for the future evolution of the galaxy. The galaxy lives in its present state until a time of $\eta\sim 14$ when conventional star formation ceases and the smallest ordinary stars leave the main sequence. For times $\eta>14$, the principle mode of additional star formation is through the collisions and mergers of brown dwarfs (substellar objects). The galaxy itself evolves through the competing processes of orbital decay via gravitational radiation and the evaporation of stars into the intergalactic medium via stellar encounters. Stellar evaporation is the dominant process, and most of the stars will leave the system at a time $\eta\sim 19$. Some fraction (we roughly estimate ~0.01–0.10) of the galaxy is left behind in its central black hole.

(6) We have considered the annihilation and capture of weakly interacting massive particles (WIMPs) in the galactic halo. In the absence of other evolutionary processes, the WIMPs in the halo annihilate on the time scale $\eta\sim 23$. On the other hand, white dwarfs can capture WIMPs and thereby deplete the halo on the somewhat longer time scale $\eta\sim 25$. The phenom-

enon of WIMP capture indicates that white dwarf cooling will be arrested rather shortly at a luminosity $L_* \sim 10^{-12}\ L_\odot$.

(7) Depending on the amount of mass loss suffered by the Sun when it becomes a red giant, the Earth may be vaporized by the Sun during its asymptotic giant phase of evolution; in this case, the Earth will be converted to a small (0.01%) increase in the solar metallicity. In general, however, planets can end their lives in a variety of ways. They can be vaporized by their parent stars, ejected into interstellar space through close stellar encounters, merge with their parent stars through gravitational radiation, and eventually disappear as their protons decay.

(8) We have discussed the allowed range for the proton lifetime. A firm lower bound on the lifetime arises from current experimental searches. Although no definitive upper limit exists, we can obtain a suggestive upper "bound" on the proton lifetime by using decay rates suggested by GUTs and by invoking the constraint of the mass of the mediating boson, $M_X < M_{\rm Pl} \sim 10^{19}$ GeV. We thus obtain an expected range for the proton lifetime

$$32 < \eta_P < 49 + 76(N-1), \tag{6.1}$$

where the integer N is order of the process, i.e., the number of mediating bosons required for the decay to take place. Even for the third-order case, we have $\eta_P < 201$. Quantum gravity effects also lead to proton decay with time scales in the range $46 < \eta_P < 169$. Finally, sphalerons imply $\eta_P \sim 140$.

(9) We have presented a scenario for the future evolution of sunlike stars (see Fig. 6). In this case, stars evolve into white dwarf configurations as in conventional stellar evolution. On sufficiently long time scales, however, proton decay becomes important. For cosmological decades in the range $20 < \eta < 35$, the mass of the star does not change appreciably, but the luminosity is dominated by the energy generated by proton decay. In the following cosmological decades, $\eta = 35$–37, mass loss plays a large role in determining the stellar structure. The star expands as it loses mass and follows the usual mass/radius relation for white dwarfs. The chemical composition changes as well (see Fig. 5). Proton decay by itself quickly reduces the star to a state of pure hydrogen. However, pycnonuclear reactions will be sufficient to maintain substantial amounts of helium (^{3}He and ^{4}He) until the mass of the star decreases below $\sim 0.01\ M_\odot$. During the proton-decay phase of evolution, a white dwarf follows a well-defined track in the H-R diagram given by $L_* \propto T_*^{12/5}$. After the stellar mass decreases to $M_* \approx 10^{-3} M_\odot$, the star is lifted out of degeneracy and follows a steeper track $L_* \propto T_*^{12}$ in the H-R diagram.

(10) If proton decay does not take place through the first-order process assumed above, then white dwarfs and other degenerate objects will still evolve, but on a much longer time scale. The relevant physical process is likely to be proton decay through higher-order effects. The time scales for the destruction and decay of degenerate stars obey the ordering

$$\eta_P \ll \eta_{BH} \ll \eta_{P2}, \tag{6.2}$$

where $\eta_P \sim 37$ is the time scale for first-order proton decay, $\eta_{BH} \sim 65$ is the time scale for a stellar-sized black hole to evaporate, and $\eta_{P2} \sim 100$–200 is the time scale for proton decay through higher-order processes.

(11) In the future, the universe as a whole can evolve in a variety of different possible ways. Future density perturbations can come across the horizon and close the universe; this effect would ultimately lead (locally) to a big crunch. Alternately, the universe could contain a small amount of vacuum energy (a cosmological-constant term) and could enter a late-time inflationary epoch. Finally, the universe could be currently in a false vacuum state and hence kevorking on the brink of instability. In this case, when the universe eventually tunnels into the true vacuum state, the laws of physics and hence the universe as we know it would change completely.

(12) As the cosmic microwave background redshifts away, several different radiation fields will dominate the background. In the near term, stellar radiation will overtake the cosmic background. Later on, the radiation produced by dark-matter annihilation (both direct and in white dwarfs) will provide the dominant contribution. This radiation field will be replaced by that arising from proton decay, and then, eventually, by the radiation field arising from evaporation of black holes (see Fig. 9).

B. Eras of the future universe

Our current understanding of the universe suggests that we can organize the future into distinct eras, somewhat analogous to geological eras:

(a) *The Radiation-Dominated Era.* $-\infty < \eta < 4$. This era corresponds to the usual time period in which most of the energy density of the universe is in the form of radiation.

(b) *The Stelliferous Era.* $6 < \eta < 14$. Most of the energy generated in the universe arises from nuclear processes in conventional stellar evolution.

(c) *The Degenerate Era.* $15 < \eta < 37$. Most of the (baryonic) mass in the universe is locked up in degenerate stellar objects: brown dwarfs, white dwarfs, and neutron stars. Energy is generated through proton decay and particle annihilation.

(d) *The Black-Hole Era.* $38 < \eta < 100$. After the epoch of proton decay, the only stellarlike objects remaining are black holes of widely disparate masses, which are actively evaporating during this era.

(e) *The Dark Era.* $\eta > 100$. At this late time, protons have decayed and black holes have evaporated. Only the waste products from these processes remain: mostly photons of colossal wavelength, neu-

trinos, electrons, and positrons. The seeming poverty of this distant epoch is perhaps more due to the difficulties inherent in extrapolating far enough into the future, rather than an actual dearth of physical processes.

C. Experimental and theoretical implications

Almost by definition, direct experiments that test theoretical predictions of the very-long-term fate of the universe cannot be made in our lifetimes. However, this topic in general and this paper in particular have interesting implications for present-day experimental and theoretical work. If we want to gain more certainty regarding the future of the universe and the astrophysical objects within it, then several issues must be resolved. The most important of these are as follows:

(1) Does the proton decay? What is the lifetime? This issue largely determines the fate of stellar objects in the universe for time scales longer than $\eta \sim 35$. If the proton is stable to first-order decay processes, then stellar objects in general and white dwarfs in particular can live in the range of cosmological decades $\eta < 100$. If the proton is also stable to second-order decay processes, then degenerate stellar objects can live for a much longer time. On the other hand, if the proton does decay, a large fraction of the universe will be in the form of proton decay products (neutrinos, photons, positrons, etc.) for times $\eta > 35$.

(2) What is the vacuum state of the universe? This issue plays an important role in determining the ultimate fate of the universe itself. If the vacuum energy density of the universe is nonzero, then the universe might ultimately experience a future epoch of inflation. On the other hand, if the vacuum energy density is strictly zero, then future (large) density perturbations can, in principle, enter our horizon and lead (locally) to a closed universe and hence a big crunch.

(3) What is the nature of the dark matter? Of particular importance is the nature of the dark matter that makes up galactic halos. The lifetime of the dark-matter particles is also of great interest.

(4) What fraction of the stars in a galaxy are evaporated out of the system and what fraction are accreted by the central black hole (or black holes)? This issue is important because black holes dominate the energy and entropy production in the universe in the time range $36 < \eta < 100$ and the mass of a black hole determines its lifetime.

(5) Does new physics occur at extremely low temperatures? As the universe evolves and continues to expand, the relevant temperatures become increasingly small. In the scenario outlined here, photons from the cosmic microwave background and other radiation fields, which permeate all of space, can redshift indefinitely in accordance with the classical theory of radiation. It seems possible that classical theory will break down at some point. For example, in an open universe, the CMB photons will have a wavelength longer than the current horizon size ($\sim$3000 Mpc) at a time $\eta \sim 40$, just after proton decay. Some preliminary models for future phase transitions have been proposed (Primack and Sher, 1980; Suzuki, 1988; Sher, 1989), but this issue calls out for further exploration.

D. Entropy and heat death

The concept of the heat death of the universe has troubled many philosophers and scientists since the mid-nineteenth century when the second law of thermodynamics was first understood (e.g., Helmholz, 1854; Clausius, 1865, 1868). Very roughly, classical heat death occurs when the universe as a whole reaches thermodynamic equilibrium; in such a state, the entire universe has a constant temperature at all points in space and hence no heat engine can operate. Without the ability to do physical work, the universe ''runs down'' and becomes a rather lifeless place. Within the context of modern big-bang cosmology, however, the temperature of the universe is continually changing and the issue shifts substantially; many authors have grappled with this problem, from the inception of big-bang theory (e.g., Eddington, 1931) to more recent times (Barrow and Tipler, 1978, 1986; Frautschi, 1982). A continually expanding universe never reaches true thermodynamic equilibrium and hence never reaches a constant temperature. Classical heat death is thus manifestly avoided. However, the expansion can, in principle, become purely adiabatic so that the entropy in a given comoving volume of the universe approaches (or attains) a constant value. In this case, the universe can still become a dull and lifeless place with no ability to do physical work. We denote this latter possibility as cosmological heat death.

Long-term entropy production in the universe is constrained in fairly general terms for a given class of systems (Bekenstein, 1981). For a spatially bounded physical system with effective radius R, the entropy S of the system has a well-defined maximum value. This upper bound is given by

$$S \leqslant \frac{2\pi R E}{\hbar c}, \tag{6.3}$$

where E is the total energy of the system. Thus, for a bounded system (with finite size R), the ratio S/E of entropy to energy has a firm upper bound. Furthermore, this bound can be actually attained for black holes (see Bekenstein, 1981 for further discussion).

The results of this paper show that cosmological events continue to produce energy and entropy in the universe, at least until the cosmological decade $\eta \sim 100$. As a result, cosmological heat death is postponed until after that epoch, i.e., until the Dark Era. After that time, however, it remains possible in principle for the universe to become nearly adiabatic and hence dull and lifeless. The energy- and entropy-generating mechanisms avail-

able to the universe depend on the mode of long-term evolution, as we discuss below.

If the universe is closed (Sec. V.A) or becomes closed at some future time (Sec. V.B), then the universe will end in a big crunch and long-term entropy production will not be an issue. For the case in which the universe remains nearly flat, density perturbations of larger and larger size scales can enter the horizon, grow to nonlinearity, and lead to continued production of energy and entropy through the evaporation of black holes (see Sec. V.F.1). These black holes saturate the Bekenstein bound and maximize entropy production. Cosmological heat death can thus be avoided as long as the universe remains nearly flat.

On the other hand, if the universe is open, then density fluctuations become frozen out at some finite length scale (Sec. V.B). The energy contained within the horizon thus becomes a finite quantity. However, the Bekenstein bound does not directly constrain entropy production in this case because the effective size R grows without limit. For an open universe, the question of cosmological heat death thus remains open. For a universe experiencing a future inflationary phase (Sec. V.E.1), the situation is similar. Here, the horizon is effectively shrinking with time. However, perturbations that have grown to nonlinearity will be decoupled from the Hubble flow. The largest nonlinear perturbation will thus define a largest length scale λ and hence a largest mass scale in the universe; this mass scale once again implies a (finite) maximum possible amount of energy available to a local region of space. However, the system is not bounded spatially and the questions of entropy production and cosmological heat death again remain open.

To close this paper, we put forth the point of view that the universe should obey a type of Copernican Time Principle which applies to considerations of the future. This principle holds that the current cosmological epoch (η=10) has no special place in time. In other words, interesting things can continue to happen at the increasingly low levels of energy and entropy available in the universe of the future.

ACKNOWLEDGMENTS

This paper grew out of a special course taught at the University of Michigan for the theme semester "Death, Extinction, and the Future of Humanity" (Winter 1996). We would like to thank Roy Rappaport for providing the initial stimulation for this course and hence this paper. We also thank R. Akhoury, M. Einhorn, T. Gherghetta, G. Kane, and E. Yao for useful discussions regarding proton decay and other particle physics issues. We thank P. Bodenheimer, G. Evrard, J. Jijina, J. Mohr, M. Rees, D. Spergel, F. X. Timmes, and R. Watkins for many interesting astrophysical discussions and for critical commentary on the manuscript. This work was supported by an NSF Young Investigator Award, NASA Grant No. NAG 5-2869, and funds from the Physics Department at the University of Michigan.

REFERENCES

Abramowitz, M., and I. A. Stegun, 1972, *Handbook of Mathematical Functions* (Dover, New York).
Adams, F. C., 1993, Phys. Rev. D **48**, 2800.
Adams, F. C., and M. Fatuzzo, 1996, Astrophys. J. **464**, 256.
Adams, F. C., and K. Freese, 1991, Phys. Rev. D **43**, 353.
Adams, F. C., and K. Freese, 1995, Phys. Rev. D **51**, 6722.
Adams, F. C., K. Freese, and A. H. Guth, 1991, Phys. Rev. D **43**, 965.
Albrecht, A., and P. J. Steinhardt, 1982, Phys. Rev. Lett. **48**, 1220.
Alcock, C. *et al.* 1993, Nature **365**, 621.
Aubourg, E. *et al.* 1993, Nature **365**, 623.
Bahcall, J. N. 1989, *Neutrino Astrophysics* (Cambridge University Press, Cambridge, England).
Bardeen, J. M., P. J. Steinhardt, and M. S. Turner, 1983, Phys. Rev. D **28**, 679.
Barrow, J. D., and F. J. Tipler, 1978, Nature **276**, 453.
Barrow, J. D., and F. J. Tipler, 1986, *The Anthropic Cosmological Principle* (Oxford University Press, New York/London).
Bekenstein, J. D., 1981, Phys. Rev. D **23**, 287.
Binney, J., and S. Tremaine, 1987, *Galactic Dynamics* (Princeton University, Princeton, NJ).
Blau, S. K., E. I. Guendelman, and A. H. Guth, 1987, Phys. Rev. D **35**, 1747.
Bond, J. R., B. J. Carr, and C. J. Hogan, 1991, Atrophys. J. **367**, 420.
Brune, D., and J. J. Schmidt, 1974, Eds., *Handbook on Nuclear Activation Cross Sections* (IAEA, Vienna).
Burrows, A., W. B. Hubbard, D. Saumon, and J. I. Lunine, 1993, Astrophys. J. **406**, 158.
Burrows, A., and J. Liebert, 1993, Rev. Mod. Phys. **65**, 301.
Carroll, S. M., W. H. Press, and E. L. Turner, 1992, Annu. Rev. Astron. Astrophys. **30**, 499.
Castano, D. J., and S. P. Martin, 1994, Phys. Lett. B **340**, 67.
Chandrasekhar, S. 1939, *Stellar Structure* (Dover, New York).
Chernoff, D. F., and M. D. Weinberg, 1990, Astrophys. J. **351**, 121.
Clausius, R., 1865, Ann. Phys. (Leipzig) **125**, 353.
Clausius, R. 1868, Philos. Mag. **35**, 405.
Clayton, D. D., 1983, *Principles of Stellar Evolution and Nucleosynthesis* (University of Chicago, Chicago).
Coleman, S., 1977, Phys. Rev. D **15**, 2929.
Coleman, S., 1985, *Aspects of Symmetry* (Cambridge University Press, Cambridge, England).
Coleman, S., and F. De Luccia, 1980, Phys. Rev. D **21**, 3305.
Copeland, H., J. O. Jensen, and H. E. Jorgensen, 1970, Astron. Astrophys. **5**, 12.
Crone, M. M., and M. Sher, 1990, Am. J. Phys. **59**, 25.
D'Antona, F., and I. Mazzitelli, 1985, Astrophys. J. **296**, 502.
Davies, P. C. W., 1994, *The Last Three Minutes* (BasicBooks, New York).
Dicus, D. A., J. R. Letaw, D. C. Teplitz, and V. L. Teplitz, 1982, Astrophys. J. **252**, 1.
Diehl, E., G. L. Kane, C. Kolda, and J. D. Wells, 1995, Phys. Rev. D **52**, 4223.
Dorman, B., L. A. Nelson, and W. Y. Chau, 1989, Astrophys. J. **342**, 1003.
Draine, B. T., and H. M. Lee, 1984, Astrophys. J. **285**, 89.
Dyson, F. J., 1979, Rev. Mod. Phys. **51**, 447.

Dyson, F. J., 1988, *Infinite in All Directions* (Harper and Row, New York).
Eddington, A. S., 1931, Nature **127**, 447.
Ellis, G. F. R. and T. Rothman, 1993, Am. J. Phys. **61**, 883.
Ellis, G. F. R., and D. H. Coule, 1994, Gen. Relativ. Gravity **26**, 731.
Elmegreen, B. G., and R. D. Mathieu, 1983, Mon. Not. R. Astron. Soc. **203**, 305.
Faulkner, J., and R. L. Gilliland, 1985, Astrophys. J. **299**, 994.
Feinberg, G., 1981, Phys. Rev. D **23**, 3075.
Feinberg, G., M. Goldhaber, and G. Steigman, 1978, Phys. Rev. D **18**, 1602.
Frautschi, S., 1982, Science **217**, 593.
Freese, K., 1986, Phys. Lett. B **167**, 295.
Gaier, T., J. Schuster, J. Gundersen, T. Koch, M. Seiffert, P. Meinhold, and P. Lubin, 1992, Astrophys. J. Lett. **398**, L1.
Goity, J. L., and M. Sher, 1995, Phys. Lett. B **346**, 69.
Golimowski, D. A., T. Nakajima, S. R. Kulkarni, and B. R. Oppenheimer, 1995, Astrophys. J. Lett. **444**, L101.
Gott, J. R., III, 1993, Nature **363**, 315.
Gould, A., 1987, Astrophys. J. **321**, 571.
Gould, A., 1991, Astrophys. J. **388**, 338.
Grischuk, L. P., and Ya. B. Zel'dovich, 1978, Sov. Astron. **22**, 125.
Grossman, A. S., and H. C. Graboske, 1971, Astrophys. J. **164**, 475.
Guth, A., 1981, Phys. Rev. D **23**, 347.
Guth, A. H., and S.-Y. Pi, 1982, Phys. Rev. Lett. **49**, 1110.
Hamada, T., and E. E. Salpeter, 1961, Astrophys. J. **134**, 683.
Hawking, S. W., 1975, Commun. Math. Phys. **43**, 199.
Hawking, S. W., 1982, Phys. Lett. **B115**, 295.
Hawking, S. W., 1985, Phys. Lett. **B150**, 339.
Hawking, S. W., 1987, Phys. Lett. **B195**, 337.
Hawking, S. W., D. N. Page, and C. N. Pope, 1979, Phys. Lett. **B86**, 175.
Helmholz, H. von, 1854, *On the Interaction of Natural Forces*.
Henry, T. J., J. D. Kirkpatrick, and D. A. Simons, 1994, Astron. J. **108**, 1437.
Hubbell, J. H., H. A. Glimm, and I. Overbo, 1980, J. Phys. Chem. Ref. Data **9**, 1023.
Hut *et al.*, 1992, Pub. Astron. Soc. Pacific, **104**, 981.
Islam, J. N., 1977, Q. J. R. Astron. Soc. **18**, 3.
Islam, J. N., 1979, Sky Telesc. **57**, 13.
Jungman, G., M. Kamionkowski, and K. Griest, 1996, Phys. Rep., in press.
Jura, M., 1986, Astrophys. J. **301**, 624.
Kane, G. L., 1993, *Modern Elementary Particle Physics* (Addison-Wesley, Reading, MA).
Kane, G. L., and J. D. Wells, 1996, Phys. Rev. Lett., in press.
Kennicutt, R. C., P. Tamblyn, and C. W. Congdon, 1995, Astrophys. J. **435**, 22.
Kippenhahn, R., and A. Weigert, 1990, *Stellar Structure and Evolution* (Springer, Berlin).
Kolb, E. W., and M. S. Turner, 1990, *The Early Universe* (Addison-Wesley, Redwood City, CA).
Krauss, L. M., M. Srednicki, and F. Wilczek, 1986, Phys. Rev. D **33**, 2206.
Krauss, L. M., and M. White, 1992, Phys. Rev. Lett. **69**, 869.
Kumar, S. S., 1963, Astrophys. J. **137**, 1121.
La, D., and P. J. Steinhardt, 1989, Phys. Rev. Lett. **62**, 376.
Langacker, P., 1981, Phys. Rep. **72**, 186.
Larson, R. B., 1973, Mon. Not. R. Astron. Soc. **161**, 133.
Larson, R. B., and B. M. Tinsley, 1978, Astrophys. J. **219**, 46.
Laughlin, G., and P. Bodenheimer, 1993, Astrophys. J. **403**, 303.
Laughlin, G., P. Bodenheimer, and F. C. Adams, 1996, Astrophys. J., in press.
Lee, H. M., and J. P. Ostriker, 1986, Astrophys. J. **310**, 176.
Lemaître, G., 1933, Ann. Soc. Sci. Bruxelles **A53**, 51.
Lightman, A. P., and S. L. Shapiro, 1978, Rev. Mod. Phys. **50**, 437.
Linde, A. D., 1982, Phys. Lett. **B108**, 389.
Linde, A. D., 1983, Nucl. Phys. **B216**, 421.
Linde, A. D., 1988, Phys. Lett. **B211**, 29.
Linde, A. D., 1989, Phys. Lett. **B227**, 352.
Linde, A. D., 1990, *Particle Physics and Inflationary Cosmology* (Harwood Academic, New York).
Loh, E., and E. Spillar, 1986, Astrophys. J. Lett. **307**, L1.
Lyth, D. H., 1984, Phys. Lett. **B147**, 403.
Manchester, R. N., and J. H. Taylor, 1977, *Pulsars* (W. H. Freeman, San Francisco).
Marcy, G. W., R. P. Butler, and E. Williams, 1996, preprint.
Mayor, M., and D. Queloz, 1995, Nature **378**, 355.
Meyer, S. S., E. S. Cheng, and L. A. Page, 1991, Astrophys. J. Lett. **410**, L57.
Mihalas, D., and J. Binney, 1981, *Galactic Astronomy: Structure and Kinematics* (W. H. Freeman, New York).
Miller, G. E., and J. M. Scalo, 1979, Astrophys. J. Suppl. **41**, 513.
Misner, C. W., K. S. Thorne, and J. A. Wheeler, 1973, *Gravitation* (W. H. Freeman, San Francisco).
Mohapatra, R. N., and R. E. Marshak, 1980, Phys. Rev. Lett. **44**, 1316.
Ohanian, H. C., and R. Ruffini, 1994, *Gravitation and Spacetime* (Norton, New York).
Oppenheimer, B. R., S. R. Kulkarni, K. Matthews, and T. Nakajima, 1995, Science **270**, 1478.
Page, D. N., 1980, Phys. Lett. **B95**, 244.
Page, D. N., and M. R. McKee, 1981a, Phys. Rev. D **24**, 1458.
Page, D. N., and M. R. McKee, 1981b, Nature **291**, 44.
Particle Data Group, 1994, Phys. Rev. D **50**, 1173.
Peccei, R. D., and H. R. Quinn, 1977a, Phys. Rev. Lett. **38**, 1440.
Peccei, R. D., and H. R. Quinn, 1977b, Phys. Rev. D **16**, 1791.
Peebles, P. J. E., 1993, *Principles of Physical Cosmology* (Princeton University, Princeton, NJ).
Peebles, P. J. E., 1994, Astrophys. J. **429**, 43.
Perkins, D., 1984, Annu. Rev. Nucl. Part. Sci. **34**, 1.
Phillips, A. C., 1994, *The Physics of Stars* (Wiley, Chichester).
Poundstone, W., 1985, *The Recursive Universe* (Morrow, New York).
Press, W. H., and D. N. Spergel, 1985, Astrophys. J. **296**, 679.
Press, W. H., and S. A. Teukolsky, 1977, Astrophys. J. **213**, 183.
Primack, J. R., and M. Sher, 1980, Nature **288**, 680.
Rajaraman, R., 1987, *Solitons and Instantons* (North-Holland, Amsterdam).
Rana, N. C., 1991, Annu. Rev. Astron. Astrophys. **29**, 129.
Rees, M. J., 1969, Observatory **89**, 193.
Rees, M. J., 1981, Q. J. R. Astron. Soc. **22**, 109.
Rees, M. J., 1984, Annu. Rev. Astron. Astrophys. **22**, 471.
Reimers, D., 1975, in *Problems in Stellar Astrophysics*, edited by B. Bascheck, W. H. Kegel, and G. Traving (Springer, New York), p. 299.
Riess, A. G., W. H. Press, and R. P. Kirshner, 1995, Astrophys. J. Lett. **438**, L17.

Roberts, M. S. 1963, Annu. Rev. Astron. Astrophys. **1**, 149.
Sackmann, I.-J., A. I. Boothroyd, and K. E. Kraemer, 1993, Astrophys. J. **418**, 457.
Sakharov, A. D., 1967, JETP Lett. **5**, 24.
Salpeter, E. E., 1955, Astrophys. J. **121**, 161.
Salpeter, E. E., 1982, in *Essays in Nuclear Astrophysics* (Cambridge University Press, Cambridge, England).
Salpeter, E. E., and H. M. Van Horn, 1969, Astrophys. J. **155**, 183.
Sato, K., H. Kodama, M. Sasaki, and K. Maeda, 1982, Phys. Lett. B **108**, 103.
Scalo, J. M., 1986, Fundam. Cosm. Phys. **11**, 1.
Schuster, J., T. Gaier, J. Gundersen, P. Meinhold, T. Koch, M. Seiffert, C. A. Wuensche, and P. Lubin, 1993, Astrophys. J. Lett. **412**, L47.
Shapiro, S. L., and S. A. Teukolsky, 1983, *Black Holes, White Dwarfs, and Neutron Stars: The Physics of Compact Objects* (Wiley, New York).
Sher, M., 1989, Phys. Rep. **179**, 273.
Shu, F. H., 1982, *The Physical Universe* (University Science Books, Mill Valley).
Shu, F. H., F. C. Adams, and S. Lizano, 1987, Annu. Rev. Astron. Astrophys. **25**, 23.
Smoot, G. F., *et al.*, 1992, Astrophys. J. Lett. **396**, L1.
Stahler, S. W., 1988, Astrophys. J. **332**, 804.
Starobinsky, A. A., 1982, Phys. Lett. B **117**, 175.
Steinhardt, P. J., and M. S. Turner, 1984, Phys. Rev. D **29**, 2162.
Stevenson, D. J., 1991, Annu. Rev. Astron. Astrophys. **29**, 163.
Suzuki, M., 1988, Phys. Rev. D **38**, 1544.
't Hooft, G. 1976, Phys. Rev. Let. **37**, 8.
Timmes, F. X., 1996, unpublished.
Tinney, C. G., 1995, Ed., *The Bottom of the Main Sequence and Beyond* (Springer, Berlin).
Tipler, F. J., 1992, Phys. Lett. B **286**, 36.
Tolman, R. C., 1934, *Relativity, Thermodynamics, and Cosmology* (Clarendon Press, Oxford).
Turner, M. S., 1983, Nature **306**, 161.
Visser, M., 1995, *Lorentzian Wormholes: From Einstein to Hawking* (AIP, Woodbury, NY).
Voloshin, M. B., I. Yu. Kobzarev, and L. B. Okun, 1975, Sov. J. Nucl. Phys. **20**, 644.
Weinberg, M. D., 1989, Mon. Not. R. Astron. Soc. **239**, 549.
Weinberg, S. 1972, *Gravitation and Cosmology* (Wiley, New York).
Weinberg, S., 1977, *The First Three Minutes* (Basic, New York).
Weinberg, S., 1978, Phys. Rev. Lett. **40**, 223.
Weinberg, S., 1980, Phys. Rev. D **22**, 1694.
Weinberg, S., 1989, Rev. Mod. Phys. **61**, 1.
Wilczek, F., 1978, Phys. Rev. Lett. **40**, 279.
Wilczek, F., and A. Zee, 1979, Phys. Lett. B **88**, 311.
Wood, M. A., 1992, Astrophys. J. **386**, 529.
Wright, E. L., *et al.*, 1992, Astrophys. J. Lett. **396**, L13.
Zel'dovich, Ya. B., 1976, Phys. Lett. A **59**, 254.
Zinnecker, H., 1984, Mon. Not. R. Astron. Soc. **210**, 43.
Zuckerman, B., and M. A. Malkan, 1996, *The Origin and Evolution of the Universe* (Jones and Bartlett, Sudbury, MA).