ICARUS 46, 293–327 (1981)

Galactic Civilizations: Population Dynamics and Interstellar Diffusion

WILLIAM I. NEWMAN[1]

Institute for Advanced Study, Princeton, New Jersey 08450

AND

CARL SAGAN

Laboratory for Planetary Studies, Cornell University, Ithaca, New York 14853

Received January 18, 1981; revised April 28, 1981

The interstellar diffusion of galactic civilizations is reexamined by potential theory; both numerical and analytical solutions are derived for the nonlinear partial differential and difference equations which specify a range of relevant models, drawn from blast wave physics, soil science, and, especially, population biology. An essential feature of these models is that, for all civilizations, population growth must be limited by the carrying capacity of the planetary environments. Dispersal is fundamentally a diffusion process; a directed density-dependent diffusivity describes interstellar emigration. We concentrate on two models, the first describing zero population growth (ZPG) and the second which also includes local growth and saturation of a planetary population, and for which we find an asymptotic travelling wave solution. For both models the colonization wavefront expands slowly and uniformly, but only the frontier worlds are sources of further expansion. For nonlinear diffusion with growth and saturation, the colonization wavefront from the nearest independently arisen galactic civilization can have reached the Earth only if its lifetime exceeds 2.6×10^6 years. If discretization can be neglected, the critical lifetime is 2.0×10^7 years. For ZPG the corresponding number is 1.3×10^{10} years. These numerical results depend on our choices for the specific emigration rate, the distribution of colonizable worlds, and, in the second model, the population growth rate; but the dependence on these parameters is entrancingly weak. We conclude that the Earth is uncolonized not because interstellar spacefaring societies are rare, but because there are too many worlds to be colonized in the plausible lifetime of the colonization phase of nearby galactic civilizations. This phase is, we contend, eventually outgrown. We also conclude that, except possibly early in the history of the Galaxy, there are no very old galactic civilizations with a consistent policy of conquest of inhabited worlds; there is no Galactic Empire. There may, however, be abundant groups of $\sim 10^5$ to 10^6 worlds linked by a common colonial heritage. The radar and television announcement of an emerging technical society on Earth may induce a rapid response by nearby civilizations, thus newly motivated to reach our system directly rather than by diffusion.

Alexander wept when he heard from Anaxarchus that there was an infinite number of worlds; and his friends asking him if any accident had befallen him, he returned this answer: "Do you not think it a matter worthy of lamentation that when there is such a vast multitude of them, we have not yet conquered one?"—Plutarch, *On the Tranquility of the Mind.*

"Where are they?" is a famous and possibly even apocryphal question posed by Enrico Fermi at Los Alamos in the late 1940s. In a galaxy with $\sim 10^{11}$ stars; with planets apparently abundant and the origin of life seemingly requiring very general cosmic circumstances; with the selective advantage of intelligence and technology obvious and with billions of years available for evolution, should not extraterrestrial intelligence be readily detectable? This question is most often phrased in its astrophysical context (see, e.g., Shklovskii and Sagan, 1966; Dyson, 1966): should not extraterrestrial civilizations millions or, per-

[1] Present address: Department of Earth and Space Sciences, University of California, Los Angeles, Calif. 90024.

0019-1035/81/060293-35$02.00/0

haps, even billions of years more advanced than our own be capable of so altering cosmic objects and energy sources that a range of astrophysical phenomena should exist which cannot be understood apart from the hypothesis of intelligent origin? An answer is sometimes offered in the negative (Shklovskii, 1976); i.e., the apparent absence of astrophysical phenomena of intelligent origin is taken to demonstrate that no very advanced civilizations exist—either because there are as yet undetermined impediments to the evolution of technical civilizations or because such civilizations inevitably self-destruct early in their histories. Alternatively, it can be argued (Sagan, 1974) that there are a wide range of poorly understood astrophysical phenomena, some involving extremely high energies; and that (Sagan, 1973a) the manifestations of very advanced civilizations would be no more apparent to us than the design and function of human engineering artifacts are to ants crawling upon their surfaces. Any astroengineering activity that is so wasteful of energy as to be observable with our limited technology might, by its very nature, be necessarily short-lived. The fact that we have not yet acquired compelling evidence of such activity since our rise as a technological society—although there are certainly major unexplained astrophysical phenomena—is not evidence for the absence of extraterrestrial civilizations. It may rather be that we are not looking in the right place with the right instruments at the right time. In any case, the sole productive approach is the open-minded pursuit of physical explanations of astrophysical phenomena; only a serious failure of such explanations might be evidence for extraterrestrial intelligence. In our present ignorance, the question remains entirely moot.

However, there is another and more modest side to the issue. Pioneers 10 and 11 and Voyagers 1 and 2 show that even our infant technical civilization is capable of interstellar spaceflight, although at velocities $\sim 10^{-4}$ pc year^{-1}. Accordingly, should not civilizations only a little more advanced than ours be effortlessly plying the spaces between the stars (Sagan, 1963; Shklovskii and Sagan, 1966)? And, if so, should we not on Earth today have some evidence of interstellar visits? Occasional serious attempts to deal with this question always conclude that no persuasive evidence of past visits exists in human legends or artifacts (Shklovskii and Sagan, 1966; Sagan, 1979); and the most widely touted claims of such visits uniformly have another and more plausible explanation—generally in the arena of archaeology or hoax (e.g., Story, 1976). Likewise, the extraterrestrial hypothesis of UFOs—the contention that we are today being visited—also exhibits, despite a great deal of study, no persuasive evidence (Sagan and Page, 1972; Klass, 1974). These discussions by no means exclude past or present visits to the Earth; they merely stress the absence of strong evidence for such visits.

If the presence of extraterrestrial civilizations were to imply visits to Earth then the absence of such visits would imply the absence of such civilizations. This argument was put forward by Hart (1975) who advanced it primarily in the context of interstellar colonization, not exploration. He argues that even if there are energetic difficulties or safety hazards (for example, the induced cosmic ray flux) to relativistic interstellar spaceflight, there should be no serious obstacles to interstellar flight at $\sim 0.1c$, in which case the Galaxy would be traversed in $< 10^6$ yr. Other possible objections to large-scale interstellar spacefaring—such as loss of motivation or self-destruction—Hart argues are unlikely because they must apply to all galactic civilizations to explain the absence of extraterrestrials on Earth.

We believe Hart's analysis is flawed on a number of counts. On the one hand, there are universal social impediments to cosmic imperialism of a sort which should apply to every galactic civilization. Von Hoerner (1973) has remarked that even interstellar colonization at the speed of light cannot solve the present human population explosion on the planet Earth. With our present exponential growth rate of $\gamma = 0.02$ year^{-1} and the colonization sphere expanding at the speed of light, in 500 years the expansion volume will have a radius of 50 pc with all habitable planets in that volume reaching the Earth's present population density; thereafter the population growth rate must decline. This is, of course, an extreme example, but it demonstrates that every society which is to avoid overcrowding and exhaustion of resources must practice stringent population control and actively maintain very small values of γ. Since, in the long run, exponentials defeat power laws on every planet in the Galaxy—independent of local biology, evolutionary history, and social customs—any analysis of this problem must consider low values of γ and, in the limit, zero population growth (ZPG). A similar point has been made by Cox (1976) and by Tinsley (1980). One of the objectives of the present paper is to explore in some detail the consequences of low values of γ for interstellar colonization.

Another possible social impediment to rapid interstellar colonization with some conceivable claim to universality is immortality (Kevles, 1975). In extrapolating to advanced technological societies we often seem willing to imagine daring engineering developments such as interstellar spaceflight with velocities $v \simeq c$, but make only the most modest extrapolations in biology or psychology. If aging is due, e.g., to the accumulation of somatic mutations, it is conceivable that an advanced society will have essentially eliminated both disease and aging, and that the limits to individual longevity will then be set by accidents, and by events in geological and stellar evolution. The motivations which we consider reasonable may seem very unnatural to such a society. To what extent, for example, is our motive for interstellar colonization itself a quest for immortality? Would at least a dedicated galactic imperialism be rendered unnecessary if personal immortality existed? A population of immortal organisms, each leaving a offspring per year, exhibits a growth rate identical to that of a population of organisms in which each lives for a single year and then dies leaving $(a + 1)$ offspring (Cole, 1954). A society of immortals must practice more stringent population control than a society of mortals. In addition, whatever its other charms, interstellar spaceflight must pose more serious hazards than residence on the home planet. To the extent that such predispositions are inherited, natural selection would tend in such a world to eliminate those individuals lacking a deep passion for the longest possible lifespans, assuming no initial differential replication. The net result might be a civilization with a profound commitment to stasis even on rather long cosmic time scales and a predisposition antithetical to interstellar colonization.

Colonial blue-green algae are among the most ancient organisms on Earth, and it seems likely that colonial and territorial evolutionary strategies develop early on many planets. In humans, there is some evidence that territoriality is controlled in the R-complex of the brain, first evolved hundreds of millions of years ago (MacLean, 1973). More recent evolution of the neocortex, which comprises some two-thirds of the total brain mass, has tempered, although by no means eliminated, the disposition to territoriality. A strong commitment to territoriality is probably inconsistent with survival after the advent of weapons of mass destruction. Civilizations so old that natural selection has had time to operate after the invention of advanced technology, and civilizations so advanced as to control their own evolution may make great strides in subduing the ancient territorial drives. Civilizations, say, $\gtrsim 10^6$ years more advanced than ours may have little interest in interstellar colonization (while retaining an understandable concern about younger, and more aggressive civilizations).

With even modest annual growth rates in science and technology, it is clear that a civilization 10^6 or 10^8 years in advance of ours would have technological capabilities which would for us be indistinguishable from magic, as pointed out by Arthur C. Clarke. The possibility exists that such a civilization might impose a stringent galactic hegemony. The establishment of an unbreakable *Codex Galactica,* imposing strict injunctions against colonization of or contact with already populated planets (Shklovskii and Sagan, 1966, p. 451), is by no means excluded. This idea, which has been called the Zoo Hypothesis by Ball (1973), is not irrelevant to our concerns merely because it is by definition unprovable; the detection, for example by radio astronomy, of civilizations on planets of other stars would make the hypothesis falsifiable. It is clearly impossible to prognosticate reliably on the social behavior of a hypothetical advanced civilization, but these examples are put forth to suggest that there may be many still less apparent social impediments to extensive interstellar colonization.

On the other hand, Hart feels that the effective velocity v of a colonization wavefront is at least one-third the velocity, v_s, of individual spacecraft (see also Jones, 1976; Cox, 1976). Never in human history has the ratio of colonization wavefront to individual velocities been so large. ("Rome was not built in a day"—although one can cross it on foot in a few hours. For Rome, $v/v_s \sim 10^{-6}$.)

In the history of human exploration and colonization, it is not a single exploratory society which launches all such ventures, but colonies of colonies and higher-order descendents (as Hart himself points out). For example, it might be argued that the exploration of the outer solar system and the first interstellar vehicles built by humans represent a fifth-order Phoenician exploratory venture, the Phoenicians having settled Carthage, the Carthagenians having settled Iberian seaports, the Iberians having discovered America, and the United States exploring the solar system with Pioneer 10 and 11 and Voyager 1 and 2. Lebanon and Tunisia and Spain do not today harbor spacefaring civilizations. There was a waiting time of centuries to millenia before the colony acquired sufficient resources to initiate independent exploration or colonization, during which period the parent civilization declined. If we consider the nth-order colonization of a new planet, the time for acquiring an $(n + 1)$th-order independent colonization capability will be substantially longer.

Jones (1976, 1978) claims to show from a Monte Carlo calculation that, even allowing for this colonization waiting time, it is very difficult to make the colonization wavefront move more slowly than about 10% of the starship velocity—implying that a single expansionist power will colonize the galaxy in $\sim 10^7$ years if the starship velocities are $0.1c$. Colonization for Jones provides an escape valve for an overpopulated civilization. Consequently, his colonization front progresses rapidly. Below we present population dynamics arguments on the effective velocity of such interstellar colonial ventures. We shall show that, for Jones' calculations, the ratio of the colonization wave speed to the speed of an individual starship is roughly the ratio of the interstellar travel time to the sum of the travel time plus the growth time for a new colony to be capable of launching its own interstellar colonial ventures. Moreover, we show that Jones' estimate of the time required by an imperial star system to colonize the Galaxy assumes population growth rates atypical of human and animal populations over long periods of time, but that more accurately describe anomalous population growth transients, induced by short-lived social and technological developments. It is clear that a comprehensive approach to the problem must include considerations of the population growth rates; the related waiting time for the acquisition of an independent colonial capability on a newly colonized planet; the number of colonists that arrive on a virgin world with each expedition (the larger this number, the smaller the waiting time until the next-order colonial venture); the starship velocities and search strategies; the abundance of untenanted planets available for colonization; the lifetime of the colonial civilization and its descendants; and, not unrelated to lifetimes, the possibility of interaction between two independent expanding colonial wavefronts originating on separate worlds.

We wish to assess the assumption, implicit in previous speculations on this subject, that the colonization wavefront advances with an effective velocity comparable to that of the individual starships. In our analysis, we adopt some of the methods conventionally employed in describing growth processes and dispersal mechanisms in terrestrial ecosystems. Their mathematical development provides a rigorous description of the stochastic properties of such processes and permits the use of potential theory. Consequently, we can trace the time-dependent distribution of a community without following the detailed behavior of individual members.

The models we presently describe can be used to simulate the growth, saturation and expansion of an interstellar civilization in the mean. Since the mathematical formulation of the problem is accomplished via potential theory, our models are completely specified by a family of nonlinear partial differential and difference equations. The properties of these equations can be examined analytically, as well as numerically, and the results are directly amenable to scaling. To make the discussion entirely clear it will be necessary first to review some elementary arguments.

In biological applications, these models have proved to be sufficiently general in describing all but the most catastrophic population dynamics phenomena.[2] Moreover, the results they provide in the interstellar context are remarkably insensitive to the detailed description (within limits) of the model and the associated input parameters. Consequently, the solutions obtained below are, we believe, less vulnerable to charges of terrestrial chauvinism than many previous studies.

We now review the population dynamics processes relevant to our discussion, explore various features of the associated mathematical models, and discuss the implications of the model results. The most significant of these is that the expansion velocity of the colonization front is several orders of magnitude smaller than had been previously anticipated. Thus, the answer to the question "Where are they?" may well be that only now are they about to arrive.

1. HOMOGENEOUS PROCESSES

We first consider the population dynamics processes that occur in a homogeneous environment. [In this instance, we ignore factors that contribute to immigration and emigration, for example in humans. In animal species, biologists describe such movements as "dispersal," a term used to describe all factors that lead to the displacement of an organism from its place of birth to the location at which it reproduces or dies. See, e.g., Howard (1960).] The simplest description of growth is that the time rate of change of a population ν is proportional to the population, or

$$d\nu/dt = \gamma\nu, \qquad (1)$$

[2] Implicit to this work and that of Jones, among others, is the view that population growth and diffusion in the *long term* is a relatively smooth, predictable progression of events. The course of human history, particularly since the Industrial Revolution, has been filled with many temporary upheavals. Our view here is that these can be considered perturbations in human affairs, and can be averaged out. For example the terrible deaths in the European Black Plague of the 14th century caused an almost imperceptible perturbation in the European population growth curve. On the other hand, one could argue that such upheavals have been the controlling influence on the evolution of modern civilization. In that case, the course of history would be a progression of catastrophic events whose outcome, almost by definition, is unpredictable.

where the growth rate, γ, is the difference between the birth and death rates. Exponential population growth is often associated with Malthus' name, although the idea that populations increase geometrically seems to have evolved in the Middle Ages, if not in classical antiquity.[3]

The human population growth rate (Coale, 1974) has remained relatively uniform, with a numerical value $\gamma \sim 5.6 \times 10^{-4}$ year^{-1}, from the dawn of civilization until the middle of the 18th century. Since that time, in a period representing only some 0.02% of the time since humans first arose, our numbers have increased 10-fold with γ now $\sim 2 \times 10^{-2}$ year^{-1}. Demographers have speculated that, for the past two centuries, we have been undergoing a *demographic transition*. It began with a decline in the death rate precipitated by advances in public health and nutrition. Later, the birth rate declined, primarily because of changes in the perceived value of having children. In the interim, the growth rate steeply increased and only now shows signs of stabilizing. Malthus in his *Essay on Population*, published in 1798, recognized the fundamental incompatability between a geometrical progression in population and the finite capacity of the environment to support life. He argued that, in the absence of "moral restraint," other forms of birth control then being unknown or considered unconscionable, the combined ravages of war, famine, and pestilence would impose a limit to growth.

Human history has been characterized by a strict balance between birth and death rates as a result of natural selection and social forces. We underscore this point because the application of unrealistic growth rates would result in a galaxy teeming with colonists in some tens of thousands of years. Indeed, if we were to place a hypothetical Adam and Eve on an Earth-like planet and they and their offspring were to reproduce at the current (but transitory) human growth rate of 0.02 year^{-1}, their planet would burgeon with 4.5 billion humans (our current population) in only 1087 years. In examining the present human predicament and generalizing Malthus' Law, Tinsley (1980) has argued that unless rather fundamental technological changes are made (e.g., in energy conservation) the population growth rate will necessarily adjust itself to effect a population decline, perhaps a catastrophic one.

A conceptually straightforward generalization of (1) which is frequently employed in describing this situation is

$$d\nu/dt = \gamma\nu f(\nu), \tag{2}$$

where $f(\nu)$ has the properties

$$f(0) = 1; \qquad df(\nu)/d\nu \leq 0; \qquad f(\nu_s) = 0, \tag{3}$$

and where ν_s describes the carrying capacity of the environment, i.e., the maximum population that it can support. The function $f(\nu)$ decreases in a manner that reflects the diminishing ability of a particular environment to support population growth. A simple but representative example of this function is

$$\begin{aligned} f(\nu) &= 1 - \nu/\nu_s, \qquad && \nu < \nu_s, \\ &= 0, && \nu \geq \nu_s. \end{aligned} \tag{4}$$

The corresponding differential equation for population growth (the Pearl–Verhulst law),

$$d\nu/dt = \gamma\nu(1 - \nu/\nu_s), \tag{5}$$

has a general solution given by

$$\nu(t) = \nu(0)\nu_s e^{\gamma t}/\{[\nu_s - \nu(0)] + \nu(0)e^{\gamma t}\}, \tag{6}$$

[3] Leonardo Pisano (Fibonacci) in the year 1202 attempted to reintroduce into Europe the study of algebra. He posed the problem, complicated by overlapping generations, of how rapidly rabbits would reproduce from an initial pair in the first month and then a second pair in the second month before becoming infertile. The number of pairs capable of reproducing in a given month corresponds to the Fibonacci series 1, 2, 3, 5, 8, 13, ... Asymptotically, the number of fertile rabbits increases by $[(5^{1/2} + 1)/2]$ each month.

where $\nu(0)$ is the initial value of $\nu(t)$. This curve has a characterstic S shape and is referred to as the logistic. This curve shares the qualitative properties common to virtually all models of homogeneous processes: exponential (Malthusian) growth for $\nu \ll \nu_s$, followed asymptotically by saturation at ν_s.

We employ the Pearl–Verhulst law extensively in what follows to describe homogeneous processes. A particularly readable account of the appropriate mathematical theory (as well as equations describing multispecies interaction or "predator–prey" systems) may be found in Davis (1962). Coale (1974) provides an interesting introduction to the history of the human population.

2. HETEROGENEOUS PROCESSES

Many factors contribute to the rates of dispersal in animal species. Certain factors are innate; others are environmental in origin (Howard, 1960). The simplest of these, random dispersal, is a mechanism that evolved, at least in part, to reduce the incidence of homozygotes. An animal such as the muskrat, *Ondatra zibethica L.*, will head off in a randomly selected direction over some characteristic distance before it mates, reproduces, and dies. Since inbreeding is in this way avoided, the frequency of genetic recombination of deleterious recessive mutations is reduced. From a probabilistic point of view, each muskrat executes (in the absence of population growth) a random walk or Brownian motion. On the other hand, the behavior of the muskrat *population distribution* is nonrandom and is described by the diffusion equation (in either its continuum or discrete forms).

The equivalence of collective Brownian motion and diffusion is well known (see, for example, Chandrasekhar, 1943; or Karlin and Taylor, 1975). Since a firm understanding of these phenomena is essential for our discussion, and has been incompletely appreciated in some previous discussions of the problem, we consider a simple and standard example—the drunkard's walk. We imagine a linear array of lampposts, each a distance l from the next, that light the street on which both a bar and a drunkard's home are situated. The drunkard is able only to walk the distance between one post and the next before he collapses. He then arises a time τ later and sets out toward another lamppost. The time interval between consecutive drunkard–lamppost collisions is then τ. Unfortunately, he has no sense of direction and it is equally likely that he will stagger to the left as to the right. If $\rho_n{}^k$ is the probability that he is at lamppost n at time interval k, then the equation governing his behavior is readily observed to be

$$\rho_n^{k+1} = \tfrac{1}{2}\rho_{n+1}^k + \tfrac{1}{2}\rho_{n-1}^k. \tag{7}$$

(When posed in this manner, the drunkard's walk is an example of a Markov process; Eq. (7) is known as the Chapman–Kolmogorov relation.) Rewriting, we obtain

$$(\rho_n^{k+1} - \rho_n{}^k)/\tau = (l^2/2\tau) \times [(\rho_{n+1}^k - 2\rho_n{}^k + \rho_{n+1}^k)/l^2]. \tag{8}$$

Taking the limit as l and τ go to zero (but in a manner whereby $l^2/2\tau$ remains fixed) and noting that, from a first- and second-order Taylor expansion,

$$\partial\rho(x = nl,\ t = k\tau)/\partial t = \lim_{\tau \to 0} (\rho_n^{k+1} - \rho_n{}^k)/\tau,$$

$$\partial^2\rho(x = nl,\ t = k\tau)/\partial x^2 = \lim_{l \to 0} (\rho_{n+1}^k - 2\rho_n{}^k + \rho_{n-1}^k)/l^2 \tag{9}$$

(where we have explicitly made the transition from a discrete to a continuous system), we find that Eq. (8) becomes the diffusion equation

$$\partial\rho/\partial t = D(\partial^2\rho/\partial x^2), \tag{10}$$

where the diffusion coefficient D satisfies the relation

$$D = \tfrac{1}{2}(l^2/\tau). \tag{11}$$

Generally, l and τ are called the collision mean free path and the collision time, respectively. The mathematical limiting process just described can be performed rigorously. The essence of this demonstration is that there exists a precise method for describing random, stochastic processes via partial differential equations, a method known as potential theory. However, the ability to take the Chapman–Kolmogorov relation (7) to its limiting form (10) is not a requirement but rather a convenience shared by many diffusion problems. In some problems, collision mean free paths are macroscopic. To employ partial differential equations in describing diffusion may produce anomalous results, a feature discovered by Einstein (1905) in his classic work on Brownian motion. Potential theory, in its discrete implementation, provides a simple remedy for quantization of spatial or temporal intervals. In practice, potential theory affords a much simpler and more reliable means of determining the behavior of a given population (whether it be composed of muskrats, drunkards, or, we claim, extraterrestrials) than the associated random process (which otherwise is accessible only to Monte Carlo methods).

Before returning to the specific problem of muskrat dispersal, we explore some properties of the diffusion equation. As a generalization of (10), the diffusion equation can be written

$$\partial\rho(\mathbf{x}, t)/\partial t = \nabla \cdot \{D(\mathbf{x}, t, \rho)\, \nabla\rho(\mathbf{x}, t)\}, \quad (12)$$

where $\rho(\mathbf{x}, t)$ now describes the population density. Here, the dimensionality of the process is increased in order to best represent the problem at hand (typically, terrestrial problems are two dimensional). Moreover D, the diffusion coefficient, need not be constant but, by this formulation, the conservation of the total population $\int \rho(\mathbf{x}, t)d\mathbf{x}$ is assured. (In such conservative cases, it is common practice to employ the terms "population density" and "probability density" interchangeably, although these quantities are of course not equal. The ratio of the first to the second is fixed, and the proportionality constant is the total population.) By allowing the diffusion coefficient to vary, we can, for example, represent the influence posed by geographical or environment factors. When muskrats face a mountain range, their advance is slowed. In the present model, it follows that they will travel over a shorter distance l in a time τ than would be the case for a flat plain; the diffusion coefficient declines correspondingly. As a result, the observed migration does not follow straight lines but describes an extremum path of least resistance. (Some sources tend to discount diffusive terms in population dynamics. They offer, e.g., the erroneous counterexample that Minnesota would have been colonized by Europeans long before California, if the diffusion equation were employed. However, they overlook the fact that D should be decreased to represent the difficulties faced in overland travel by settlers coming to Minnesota and increased to describe the relative ease of ocean travel to California.) We see later how a population-dependent diffusion coefficient can be employed to describe social and territorial influences on emigration. For the sake of simplicity in our later discussion of interstellar colonization, we assume that space is homogeneous over a suitably large scale so that D will have no spatial variation. We shall, however, permit the diffusion coefficient to be population-density dependent.

Since we will later be concerned with the problem of assessing the velocity of the colonization front, it is important to note that the velocity associated with the diffusion process described by Eq. (10) is not well defined. For example, if we define $\langle x^2 \rangle$ using the probability density $\rho(x, t)$ as

$$\langle x^2 \rangle = \int_{-\infty}^{\infty} x^2 \rho(x, t)dx, \quad (13)$$

and integrate Eq. (10) by parts twice, we find

$$d\langle x^2 \rangle/dt = 2D. \quad (14)$$

Thus, the quantity $(d/dt)\ \langle x^2 \rangle^{1/2}$, which might be considered representative of the velocity, varies at $t^{-1/2}$. Einstein (1905), in his classic microscopic treatment of Brownian motion, showed that this anomaly arises only when time scales shorter than the collision time are considered. He demonstrated rigorously that the correct velocity at small times is l/τ. Another way of characterizing the velocity $\mathbf{v}$ is to compare the continuity equation (where $\mathbf{j} = \rho\mathbf{v}$ is the flux),

$$(\partial\rho/\partial t) + \nabla \cdot \mathbf{j} = 0, \qquad (15)$$

with (12); i.e., we make the association (known as Fick's law) that

$$\mathbf{j} = -D\nabla\rho \qquad (16)$$

or, alternatively,

$$\mathbf{v} = -D\rho^{-1}\nabla\rho. \qquad (17)$$

Thus, the velocity associated with a diffusion coefficient independent of population density varies inversely as the density e-folding length but has no dependence on the density itself. In both these descriptions, each measure of velocity possesses certain undesirable qualities that can be attributed to the fact that the solution to the diffusion Eq. (10) is spatially unconfined, because the Green's function

$$G(x, t) = (4\pi Dt)^{-1/2} \exp[-x^2/4Dt] \qquad (18)$$

causes any initial configuration to be propagated and redistributed instantaneously over all space. We later see how nonlinear diffusion can overcome this velocity anomaly.

In practical situations, the diffusion coefficient is determined using Fick's law, i.e., as the quotient of the flux and the density gradient. In theoretical contexts, the diffusion coefficient must sometimes be determined using other considerations. For example, discretizing (10) and using length and time scales Δx and Δt, respectively, we find

$$\rho_n^{k+1} = D(\Delta t/\Delta x^2)\,(\rho_{n+1}^k + \rho_{n-1}^k) + \rho_n{}^k[1 - (2D\Delta t/\Delta x^2)]. \qquad (19)$$

Thus, the effective probability P of leaving site n for either site $n - 1$ or $n + 1$ in a time Δt is given by

$$P = 2D\Delta t/\Delta x^2. \qquad (20)$$

Alternatively, the diffusion coefficient in m dimensions may be determined using the probability P, viz.,

$$D = P\Delta x^2/2m\Delta t. \qquad (21)$$

Despite our misgivings about the linear diffusion equation, it has proven remarkably successful in describing the dispersal of, e.g., muskrats. Skellam (1951), citing earlier work, shows how well this model depicts the spread of *Ondatra zibethica L.* since its introduction in central Europe in 1905. But muskrats are not extraterrestrials and the model must be adapted to include social influences on emigration patterns as well as the effects of population growth and saturation. (Skellam's treatment corrects for population growth, but only *a posteriori*.) The possibility of an advanced civilization which practices zero population growth must be explored. Moreover, all this must be done in the context of estimating the velocity of an expanding population front.

3. DENSITY-DEPENDENT DIFFUSION

Consider, now, the case of a diffusion coefficient proportional to some power of the population density,

$$D(\rho) = D_0(\rho/\rho_0)^N. \qquad (22)$$

For $N > 0$, this generalization of the diffusion equation (12) has several conceptual advantages. From (17), we obtain a velocity that now has a population-density dependence. From the probability expression (20) for the drunkard's walk, the likelihood that a drunkard will leave for somewhere else now depends on how many other drunkards have converged on his lamppost. This is precisely the feature we desire: the probability of mounting a colonization venture is a rapidly increasing function of density, even if emigration cannot relieve the population pressure. Moreover, the flux $\mathbf{j} = -D(\rho)\nabla\rho$ is directed *away* from population concentrations. Not only does the population in this model "remember" where it has been (since such movements of population are always in the direction that most readily decreases the density), but also it responds to population pressure in a strongly nonlinear fashion, not at all reminiscent of the conventional drunkard's walk. [A more pronounced density dependence would arise by using an exponential dependence on the density. Crank (1956) has investigated this problem numerically in certain restricted cases. The analytic properties of the solution, however, are not well-understood. For our purposes, the power law (*22*) will suffice in describing a representative cross-section of density dependences.]

For the moment, we continue to neglect the homogeneous processes of growth and saturation. (We will not, at this time, discuss the effects of discretization, leaving that to Section 7 and Appendix I. Discretization effects do not significantly change the time scales for a colonizing civilization which practices zero population growth.) This is not an unreasonable assumption in describing an advanced civilization. It seems likely that such a society would have undergone a long period of zero population growth while it developed the technological capability for interstellar flight. Otherwise, exponential growth would have diverted all available resources to the support of a burgeoning population. ZPG must be an essential ethic of any society which is able to colonize other planetary systems.

In the spherically symmetric three-dimensional case, the population-dependent diffusion equation becomes

$$\partial\rho/\partial t = (1/r^2)(\partial/\partial r)\{r^2 D_0(\rho/\rho_0)^N(\partial\rho/\partial r)\}. \quad (23)$$

In this form, the equation describes the "porous medium problem" and is frequently encountered in hydrology and soil science. [In the latter application, one considers the flow of a fluid through a porous medium. The mass of the fluid is conserved and it is assumed to obey a polytropic equation of state. Instead of the Euler force equation, the fluid is said to obey Darcy's law which provides for a velocity proportional to the pressure gradient. Muskat (1937) provides a definitive treatment of the problem. Philip (1970) offers an historical review of the subject.]

A formally identical equation is encountered in the physics of high-temperature gases, the "thermal wave" problem:

$$\partial T/\partial t = (1/r^2)(\partial/\partial r)\{r^2 a T^N(\partial T/\partial r)\}. \quad (24)$$

Here, T is the temperature and the diffusion coefficient is temperature dependent, as would be the case in, for example, an ionized gas. Such a circumstance arises immediately after a high-yield nuclear explosion, when temperatures are sufficiently great that the velocity associated with radiative energy transport far exceeds the sound velocity. Since diffusion is now much more efficient at higher temperatures, it quickly acts to equilibrate the temperature distribution of the radiatively heated gas. There is a thin boundary layer separating the hot gas from the ambient atmosphere that has not yet been heated by the radiation. In this boundary layer, called the "thermal wavefront," the temperature undergoes a precipitous decline. Consequently, the thermal wavefront is slow to conduct heat and the net temperature distribution changes slowly.

The overall result of these various effects is that the temperature distribution preserves a constant shape. Qualitatively, the temperature is relatively uniform from the origin to the thermal wavefront, where it declines rapidly to zero. Quantitatively, the temperature (expressed in units of the temperature measured at the origin) is a function only of the radius (expressed as a fraction of the thermal wavefront radius). To illustrate, consider the one-dimensional analogue of (24),

$$\partial T/\partial t = (\partial/\partial x)\ \{aT^N(\partial T/\partial x)\}. \quad (25)$$

From this equation, we see that the quantity

$$Q = \int_{-\infty}^{\infty} T(x, t)dx \quad (26)$$

is conserved (Q is proportional to the total thermal energy). There is only one dimensionless combination of the coordinate x and the time t that can be obtained in terms of a and Q using (25) and (26):

$$\xi = x/(aQ^N t)^{1/(N+2)}. \quad (27)$$

The quantity $(Q^2/at)^{1/(N+2)}$ has the dimensions of temperature, and a solution to (25) which preserves its shape is

$$T(x, t) = (Q^2/at)^{1/(N+2)} f(\xi). \quad (28)$$

The solution for $f(\xi)$ (see Zel'dovich and Raizer, 1967, for details) is

$$f(\xi) = \{N\xi_0^2/[2(N+2)]\}^{1/N}[1 - (\xi/\xi_0)^2]^{1/N}, \quad |\xi| < \xi_0,$$
$$= 0 \quad |\xi| > \xi_0, \quad (29)$$

where

$$\xi_0 = [(N+2)^{1+N}2^{1-N}/N\pi^{N/2}]^{1/(N+2)} \times [\Gamma(\tfrac{1}{2} + 1/N)/\Gamma(1/N)]^{N/(N+2)} \quad (30)$$

and Γ is the gamma function.

For the case $N = 0$ [i.e., Eq. (10) taking $D \equiv a$],

$$f(\xi) = (4\pi)^{1/2}e^{-\xi^2/4} \quad (31)$$

The normalization employed provides

$$\int_{-\infty}^{\infty} f(\xi)d\xi = 1 \quad (32)$$

and, for $N > 0$, the position of the thermal wavefront, using (27), is just

$$x_f = \pm\xi_0(aQ^N t)^{1/(N+2)}. \quad (33)$$

In Fig. 1, $f(\xi)$ is shown for $N = 0$, 1, and 2. From Eqs. (29) and (31), we see that:

(a) if $N = 0$, the distribution is unconfined;

(b) if $0 < N < 1$, the distribution has a finite cutoff, where the temperature gradient vanishes;

(c) if $N = 1$, the temperature distribution has a finite cutoff with a finite, nonvanishing temperature gradient;

and

(d) if $N > 1$, the temperature distribution has a finite cutoff with an infinite temperature gradient.

The particular curves given in Fig. 1 correspond to a Gaussian, a parabola, and an ellipse, respectively. Apart from a differing normalization factor, the $f(\xi)$ profiles for

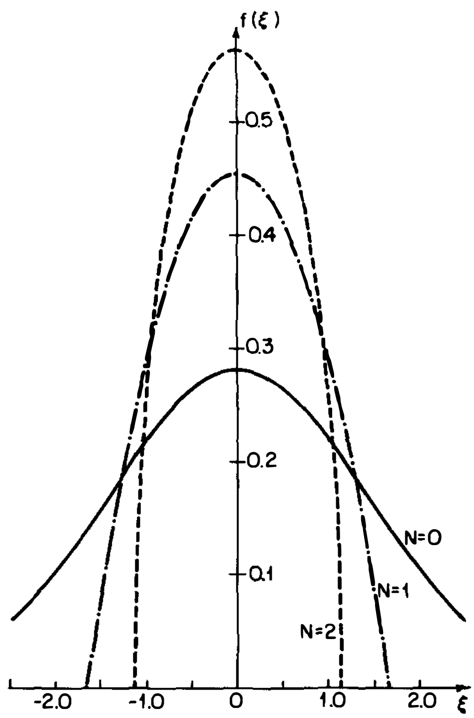

FIG. 1. Thermal wave profiles for diffusion coefficients with a power law density dependence with exponent N.

spherical symmetry [Eq. (24)] are unchanged. This solution to the porous medium or thermal wave problem was obtained independently by Barenblatt (1952) and by Pattle (1959). A thorough, physically motivated discussion of this problem is given in Zel'dovich and Raizer (1967). Knerr (1977) reviews some of the mathematical questions that arise in the one-dimensional version of this problem.

As we have already noted, the shape of the temperature distribution is preserved in this solution. Such a solution is called "self-similar." [In the Russian literature, solutions of this type are called "automodelled" solutions. Courant and Friedrichs (1948) were the first to consider such shape-preserving solutions in application to gas dynamic shocks; they employed the term "progressive waves."] Generally speaking, similarity solutions are characterized by an equation of the type

$$r_f = At^{\alpha} \tag{34}$$

where r_f describes the radius of the phenomenon being studied (e.g., a shock front or a thermal wave), and A and α constants.

There are two types of self-similar solutions. In the first, α is obtained by dimensional analysis and A is selected in order to satisfy a physical conservation law. The thermal wave problem is a case in point and Eqs. (30) and (33) display the relevant features of the scaling.[4]

A second type of self-similar problem is characterized when α is selected to insure the *existence* of a mathematical solution. Since α is not known from dimensional considerations, the self-similar solution is usually obtained by exhaustive numerical tests. The standard example of such a problem is the implosion physics studied by Guderley (1942). In his case, $\alpha < 0$ and trial values of α were varied (while numerically integrating the fluid dynamics equations) until a value of α was found that satisfied his boundary conditions. A more pertinent example of Type II self-similar solutions is that associated with the semilinear diffusion equation in one dimension. (This equation is often used in population dynamics modelling; we examine it in detail shortly.) In that case, the solution for the density $\rho(x, t)$ can be expressed as a travelling wave $\rho(x - vt)$, where v is the velocity of the travelling wave ("uniform propagation regime" in the Russian literature). If we take $x = \bar{x} \ln \chi$ and $t = \bar{t} \ln \tau$, then the solution can be written as $\rho(\xi)$, where

$$\xi = \chi/\tau^{\tilde{v}} \tag{35}$$

and

$$\tilde{v} = v\bar{t}/\bar{x}.$$

Comparing this with (34), we find that

$$\chi_f = \xi_f \tau^{\tilde{v}} \tag{36}$$

describes the position of the travelling wavefront, where ξ_f is a constant and $\tilde{v}$ plays the same role as α.

Although self-similar solutions are special solutions characterized by their dependence on a single dimensionless variable, ξ, they appear to describe the "intermediate-asymptotic" behavior of a much broader class of initial, boundary, and mixed problems. In practice (for example, in the case of a thermal wave), the solution often loses all memory of the initial and boundary

[4] Another example of such so-called Type I self-similar solutions is that obtained for a strong isotropic blast wave, as when the thermal wave described earlier cools sufficiently for gas dynamic energy transport to be more efficient than diffusive dissipation. This solution was first obtained by Sedov (1959) in an application to supernovae. He argued that the shocked gas would lose all memory of its initial conditions due to convection and would approach a solution characterized only by the total energy E_0 of the original blast and the mass density ρ_0 of the unperturbed gas. In that case, $r_f = \xi_f(E_0/\rho_0)^{1/5}t^{2/5}$ describes the radius of the shock, where ξ_f is a dimensionless number of order unity selected to assure that the blast energy equals the total internal and kinetic energy of the shocked gas. It is not without some irony that the methods for treating explosive processes, such as thermal or blast waves, and, in particular, nuclear explosions, are valuable in studying population dynamics.

conditions. Then, the solution evolves into one characterized only by, e.g., conserved quantities (such as Q) and parameters that specify the problem (such as a). This property of self-similar equations has been rigorously demonstrated only for the linear diffusion equation (10) and the semilinear diffusion equation. Numerical and experimental results (where available) display the convergence of solutions to self-similar form under a broad array of initial conditions. Thus, self-similarity is not of interest simply because it facilitates the calculations, but also because it describes a seemingly genuine asymptotic feature of the solution (see Barenblatt and Zel'dovich, 1972).

To recapitulate, we observe that a population-dependent diffusion coefficient overcomes many of the problems inherent in the conventional diffusion equation. From the self-similar solution, we observe that the distribution is *confined*. Moreover, the velocity of advance of the wavefront is no longer anomalous but can be calculated directly from Eq. (33), for one dimension, or (for $N > 0$) from

$$r_f = \xi_1 (aQ^N t)^{1/(3N+2)}, \tag{37}$$

where

$$\xi_1 = [(3N + 2)/(2^{N-1} N \pi^N]^{1/(3N+2)} \times \{\Gamma[\tfrac{5}{2} + (1/N)]/\Gamma[1 + (1/N)]\,\Gamma(\tfrac{3}{2})\}^{N/(3N+2)} \tag{38}$$

in three dimensions (Zel'dovich and Raizer, 1967). The velocity monotonically decreases with time, since the central density decreases causing the diffusion rate to diminish accordingly. Although this model ignores homogeneous processes (and, in this treatment, discretization effects) it is a useful description of situations where the overall population is conserved (ZPG), and provides useful insight into models that combine both homogeneous and heterogeneous processes.

4. DIFFUSION AND GROWTH

As a first step toward simulating the combined effects of homogeneous and heterogeneous processes in the continuous case, consider the equation

$$\partial\rho/\partial t = \alpha\rho + a(1/r^2)\,(\partial/\partial r)\,[r^2\rho^N(\partial\rho/\partial r)]. \tag{39}$$

This is similar to an approach first proposed by Gurney and Nisbet (1975). A population-dependent diffusion coefficient, they suggest, can describe one nonrandom cause of dispersal, the "directed motion" associated with territorial drives. (They confined their attention to the case $N = 1$, although they did allow for an additive third term describing random motion; this they called the biased random motion model.)

Amplifying on this theme, Gurtin and MacCamy (1977) proposed a method of solution for the one-dimensional analogue of (39),

$$\partial\rho/\partial t = \mu\rho + (\partial^2/\partial x^2)k\rho^\alpha. \tag{40}$$

After making the substitutions

$$\tilde{\rho} = e^{-\mu t}\rho,$$
$$\tau = [(\alpha - 1)\mu]^{-1}\{\exp[(\alpha - 1)\mu t] - 1\}$$
$$x' = x(k\alpha)^{-1/2}, \qquad N = \alpha - 1, \tag{41}$$

we obtain

$$(\partial/\partial\tau)\tilde{\rho} = (\partial/\partial x')\,\{\tilde{\rho}^N(\partial\tilde{\rho}/\partial x')\}. \tag{42}$$

We have already discussed the solution to the latter equation [via Eqs. (25) through (30)]. However, because of the scalings employed in (41), we observe that $\rho(x = 0, t)$ must ultimately increase exponentially with time. To eliminate this anomaly, we must prevent the population from growing above the local carrying capacity of the environment.

5. THE SEMILINEAR DIFFUSION EQUATION

We wish to describe the behavior of an organism with advantageous genes that undergoes random dispersal while increasing in population locally according to the Pearl–Verhulst logistic law. In one dimension, this can be written

$$\partial\rho/\partial t = \gamma\rho(1 - \rho/\rho_s) + (\partial/\partial x)\{D(\partial\rho/\partial x)\}, \quad (43)$$

where γ is the local growth rate, ρ_s is the local carrying capacity (or population-density saturation level) of the environment, and D is the local diffusion coefficient (assumed here to be population independent). Further, it is assumed that γ, ρ_s, and D are constant and do not vary spatially or temporally. Making the transformations

$$\rho' = \rho/\rho_s, \quad t' = \gamma t, \quad x' = x(D/\gamma)^{-1/2}, \quad (44)$$

we find Eq. (43) becomes

$$\partial\rho/\partial t = \rho(1 - \rho) + (\partial^2/\partial x^2)\rho \quad (45)$$

(where, for convenience, the primes have been dropped).

Fisher (1937) and Kolmogoroff *et al.* (1937) independently proposed that Eq. (45) has a right-going travelling wave solution. [The relationship of the travelling wave to self-similar solutions was reviewed in the discussion leading to Eqs. (35) and (36).] Assuming, then, that ρ depends only on $(x - vt)$ in (43), we first transform the velocity v into

$$\tilde{v} = v(D\gamma)^{-1/2}, \quad (46)$$

where $\tilde{v}$ is a dimensionless quantity, and, employing (44) and (45), obtain the ordinary differential equation

$$-\tilde{v}(d\rho/dx) = \rho(1 - \rho) + (d^2/dx^2)\rho. \quad (47)$$

We wish now to find the range of $\tilde{v}$ for which Eq. (47) possesses a solution. (Recall that Type II similarity solutions are characterized by a dimensionless constant, in this instance $\tilde{v}$, whose value is chosen in order to insure the *existence* of a mathematical solution.) Since we assumed a right-going travelling wave, we should expect the density ρ to decrease monotonically in the direction of increasing x. Further, we assume that

$$\lim_{x \to +\infty} \rho(x) = 0, \quad \lim_{x \to -\infty} \rho(x) = 1, \quad (48)$$

i.e., that the density vanishes in the region unaffected by the travelling wave, while the source of the organism (infinitely far off to the left) is supported at the carrying capacity. Since ρ is assumed monotonic in x, we can replace the derivatives in (47) according to

$$d/dx = -q(d/d\rho), \quad q \equiv -d\rho/dx \quad (49)$$

(q is defined so that it is necessarily positive valued). Then, (47) becomes

$$\tilde{v}q = \rho(1 - \rho) + q(d/d\rho)q. \quad (50)$$

Near $\rho = 0$, this equation has the asymptotic representation

$$\tilde{v}q \simeq \rho + q(dq/d\rho). \quad (51)$$

Since we expect that q vanishes when $\rho = 0$, consider a solution of the form

$$q = b\rho^n \quad (52)$$

where b and n are real-valued positive constants. Substituting into (51), we find that $n = 1$ and

$$\tilde{v}b = 1 + b^2. \quad (53)$$

Solving for b in terms of $\tilde{v}$, we obtain

$$b = [\tilde{v} + (\tilde{v}^2 - 4)^{1/2}]/2. \quad (54)$$

Therefore, in order to guarantee that b be real valued, the discriminant must be positive and

$$\tilde{v} \geq 2. \quad (55)$$

(The case $\tilde{v} \leq -2$ corresponds to a left-going travelling wave.)

Although Fisher and Kolmogoroff *et al.* obtained this result, Kolmogoroff *et al.* went on to show that the solution to the partial differential Eq. (45), initialized as a step function, converges to a right-going travelling wave with the minimal velocity $\tilde{v} = 2$. Although a continuous spectrum of propagation velocities $\tilde{v} \geq 2$ is in principle possible, only the solution corresponding to the extreme point of the spectrum ($\tilde{v} = 2$) can be an asymptotic solution to the partial differential Eq. (45). Therefore, the velocity of the front of advance of the advantageous genes, in Fisher's formulation of the problem, "converges" to $2(D\gamma)^{1/2}$.

Although this model from population genetics can be made more general by employing a broader description of *in situ* growth processes than that provided by the Pearl–Verhulst law, such models are characterized by *random* dispersal, i.e., a population-independent diffusion coefficient (see, e.g., the review article by Hadeler, 1977). Characteristically, they have travelling wave solutions which, when normalized in a similar fashion, also give $\tilde{v} = 2$. Thus, despite the presence of a *linear* diffusion term, the associated velocity of advance is well defined [as contrasted with the absence of a self-consistent velocity for the linear diffusion model in the absence of population growth, Eq. (10)].

There is, however, one feature of this model (apart from the population-independent diffusion coefficient) that is undesirable. From Eq. (54), we see that $b = 1$ when $\tilde{v}$ is minimal and, from (49) and (52), we find

$$\rho(x) \propto e^{-x} \tag{56}$$

when ρ is small. Therefore, the population's travelling wave front is unconfined. [This again parallels the result obtained for the linear diffusion model (10).]

This model has been employed in describing the interaction of several species. For example, in a homogeneous environment, the equations describing a predator, denoted by ρ_1, and its prey, denoted by ρ_2, might be written

$$\begin{aligned} d\rho_1/dt &= -c_1\rho_1 + c_2\rho_1\rho_2, \\ d\rho_2/dt &= c_3\rho_2 - c_4\rho_1\rho_2, \end{aligned} \tag{57}$$

where c_1, c_2, c_3 and c_4 are positive constants. In the absence of prey, the number of predators will decay exponentially. On the other hand, in the absence of predators, the number of prey will grow exponentially. When both predators and prey are present in small numbers, the number of predators will decay less rapidly, while the number of prey will grow less rapidly. In fact, there exists an exact solution to this problem that is *stationary*, $\rho_2 = c_1/c_2$ and $\rho_1 = c_3/c_4$. (In this case, the number of predators keeps the number of prey exactly in check, while there is a sufficient number of prey to maintain a constant number of predators.) This formulation of the predator–prey relationship is generally associated with the names of Lotka and Volterra (see Davis, 1962, for a simple treatment of the mathematical theory), and is frequently used to describe a number of well-observed ecosystems. The solution to Eqs. (57) describe closed orbits around that stationary solution in the $\rho_1 - \rho_2$ plane and are akin to the solution for a harmonic oscillator in that ρ_1 and ρ_2 will oscillate in a pattern strictly determined by initial conditions. When Eqs. (57) are augmented by linear diffusion terms $D\ \partial^2\rho_1/\partial x^2$ and $D\ \partial^2\rho_2/\partial x^2$, respectively, our equations assume more of the character of the semilinear diffusion Eq. (43). An example of the predator–prey relationship is seen in herbivorous copepods (zooplankton) that feed on phytoplankton. In performing a linearized perturbation analysis for such marine planktonic communities, Steele (1974) noted that all finite-wavelength perturbations will be damped and the ecosystem will tend to spatial homogeneity. Nonlinear effects, however, could conceivably destabilize the situation leading to "patchiness" and other developments beyond the scope of this discussion. Levin (1976) and McMurtrie (1978) have reviewed the multi-species population dynamics problem.

In the case just described, diffusion acts to stabilize the system (at least in a linearized representation). In other circumstances, diffusion can produce a travelling wave similar, in certain respects, to that of the semilinear diffusion Eq. (43). This feature was employed by Noble (1974)[5] in devising a model for the Bubonic Plague of 1347. His equations are

$$\partial I/\partial t = KIS - \mu I + D\nabla^2 I,$$
$$\partial S/\partial t = -KIS + D\nabla^2 S, \qquad (58)$$

where S and I are the densities of the susceptible and the infected populations, μ is the mortality rate of the disease, and D is the diffusion coefficient. The coefficient K describes the rate at which the plague is transmitted locally. [There is, in fact, a third "species" implicitly included here, the density of individuals B who have contracted the disease and either perished or recovered. The "mortality" rate μ describes the rate at which this third population density grows, $\partial B/\partial t = \mu B$. The total population $\int (I + S + B)\, d^2x$ then remains constant.] Equation (58) represents the rate of change of the number of infectives within a small area as the rate of transitions (KIS) from the susceptible population minus the removal rate (due to mortality, $-\mu I$, and dispersal $D\,\nabla^2 I$). Meanwhile, the rate of change of the susceptible population within a small area is a net loss due to the transition to the infected population ($-KIS$) and dispersal ($D\nabla^2 S$). Nobel performed a numerical integration of the one-dimensional analogue of Eqs. (58). His simulation quickly evolved into a pair of travelling waves, one describing the infectives and the other, the susceptibles. The qualitative shape of the two population distributions is depicted in Fig. 2. Thus, we see that semilinear diffusion acting in a multispecies problem can produce a set of travelling waves

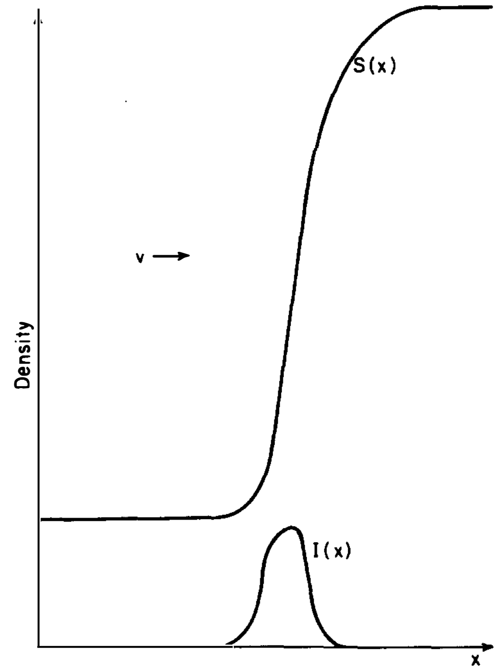

FIG. 2. Wave of advance of the bubonic plague, after Noble (1974).

just as we observed in the single-species case.

6. DENSITY-DEPENDENT DIFFUSION WITH GROWTH AND SATURATION

We have already described the need for a population dynamic model that combines local growth and saturation processes with spatial dispersal mechanisms that are population-density dependent (and yet have an implicit memory of the population's point of origin). This model, in one dimension, may be expressed as

$$\partial\rho/\partial t = \gamma\rho(1 - \rho/\rho_s) + (\partial/\partial x)[D(\rho/\rho_s)^N(\partial\rho/\partial x)], \qquad (59)$$

where the diffusion constant D is scaled in order to provide the diffusion coefficient when $\rho = \rho_s$.[6] From the transformations

[5] Noble compared his model predictions with the information available on the Black Death (Langer, 1964). His results appear to match what is known about the geographic advance of the disease. In commenting on the applicability of his model to other epidemiological problems, he suggested that the spread of certain mass sociopsychological phenomena (such as new religions) might be sufficiently similar to that of plagues to be worth investigating by similar methods. It is possible that the spread of UFO belief systems might be described in this way.

[6] Density-dependent diffusion does not make the net emigration a response to population pressure alone. Since the *flux* of the population is given by $-D(\rho/\rho_s)^N\partial\rho/\partial x$, the density gradient plays an all-important role. Thus, e.g., there is little *net* movement of a population that is approximately uniform in density, even if the density is nearly saturated.

(44), we obtain an equation analogous to (45), namely,

$$\partial\rho/\partial t = \rho(1-\rho) + (\partial/\partial x)[\rho^N(\partial\rho/\partial x)]. \quad (60)$$

We have investigated the properties of the solution to Eq. (60) for various N by numerical means. The equation was integrated using a conventional four-point explicit integration scheme (Richtmeyer and Morton, 1967). The density distribution, following Kolmogoroff *et al.* (1937), was initialized as a step function,

$$\rho(x, t{=}0) = 1, \quad x < 0,$$
$$= 0, \quad x > 0. \quad (61)$$

Since we were also interested in the spherically symmetric case, we numerically integrated

$$\partial\rho/\partial t = \rho(1-\rho) + (1/r^2)(\partial/\partial r)\{r^2\rho^N(\partial\rho/\partial r)\} \quad (62)$$

using a stable differencing scheme developed by Eisen (1967) in applications to the spherically symmetric diffusion equation. The initial conditions employed were

$$\rho(r, t{=}0) = 1, \quad r < r_f,$$
$$= 0, \quad r > r_f, \quad (63)$$

where r_f is some arbitrary radius. As a test of the integration methods, the case $N = 0$ was included. In both geometries, for $N =$ 0, 1, and 2, the numerical solution evolved into a travellilng wave. (The asymptotic correspondence between the one-dimensional and spherically symmetric cases is not surprising. The only difference between their respective equations is $(2/r)\rho^N\,\partial\rho/\partial r$. Assuming that the solution tends to the form $\rho(r - vt)$, we see that this term makes a contribution of order t^{-1}, which at large times is negligible.) Since, for $N = 0$, the numerically evolved travelling wave has a velocity $\tilde{v}$ very near 2, we consider the asymptotic convergence of the solutions (for $N = 1$ and 2) to travelling waves to be significant. Figure 3 illustrates the density

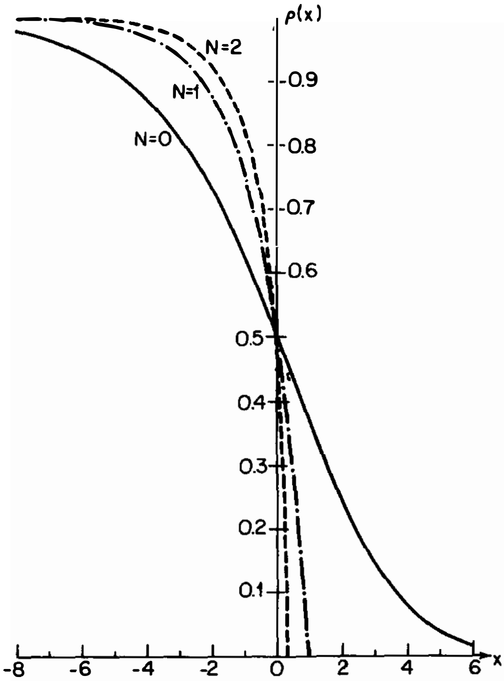

FIG. 3. Population travelling wavefront for three model equations, with N as the exponent of the population density. The Pearl–Verhulst Law applies. Figure 1 is the ZPG version of Fig. 3.

profile for these three cases. In order to facilitate a comparison among the wavefronts, the density profiles are constructed so that density is one-half at the origin.

Several other interesting features of the solution emerge from the computer analysis. First, the velocity $\tilde{v}$ of the travelling wave [Eq. (46)] decreases with increasing N. The values obtained for $N = 1$ and 2 were approximately 0.70 and 0.45, respectively. (Higher values of N were also investigated. Fluctuations in $\tilde{v}$ made these cases difficult to estimate, but the decline of $\tilde{v}$ with increasing N was unmistakable.) Secondly, the density profile near the point of vanishing density resembled the self-similar thermal wave profiles for $N = 1$ and 2 which appear in Fig. 1.

Consider, therefore, the problem of finding a travelling wave solution to (60). Employing the methods used earlier in our discussion of the semilinear diffusion equation, we write the equation for $\rho(x)$ as

$$-\tilde{v}(d\rho/dx) = \rho(1-\rho) + (d/dx)\{\rho^N(d\rho/dx)\}. \quad (64)$$

Then, making the same assumptions here as were employed in obtaining (50), we have

$$\tilde{v}q = \rho(1-\rho) + q(d/d\rho)(\rho^N q). \quad (65)$$

As before, we obtain an approximate solution as a power law [Eq. (52)], namely,

$$q \simeq \begin{cases} \tilde{v}\rho^{1-N}, & N > 0, \\ \frac{1}{2}[\tilde{v} + (\tilde{v}^2 - 4)^{1/2}]\rho & N = 0. \end{cases} \quad (66)$$

Near $\rho = 1$, the solution is given approximately by

$$q \simeq \tfrac{1}{2}[(\tilde{v}^2 + 4)^{1/2} - \tilde{v}](1-\rho) \quad (67)$$

for all N. Integrating (67), we see that the density exponentially decays to unity as we move inside the travelling wave. Near the front of the travelling wave, the density (for the $N = 0$ case) decays exponentially. However, for N nonzero, the solution is of the form

$$\rho(x) \simeq [N\tilde{v}(x_c - x)]^{1/N}, \quad (68)$$

where x_c denotes a cutoff point to the right of which the density is identically zero. This behavior is essentially the same as we observed in the thermal wave problem for $N > 0$ [see Eq. (29) and subsequent discussion] and explains the striking similarity observed in comparing the $N = 1$ and 2 curves in Figs. 1 and 3 (the thermal and travelling wavefronts, respectively). What has occurred is that for $N > 0$, the growth and saturation term in Eq. (64) has no influence on the shape of the density profile near the cutoff point and a solution reminiscent of the self-similar thermal wave problem emerges. The growth and saturation terms, however, have a controlling influence on the determination of $\tilde{v}$.

Upon finding this asymptotic behavior, we noted that (67) was an *exact* solution to (65) for $N = 1$ if we set $\tilde{v} = 2^{-1/2}$. A straightforward integration then gives

$$\rho(x) = \begin{cases} 1 - \exp[2^{-1/2}(x - x_c)], & x < x_c, \\ 0, & x \geq x_c. \end{cases} \quad (69)$$

This profile corresponds, within the numerical limitations posed by the integration scheme, to the computed solution and the predicted travelling wave velocity of $2^{-1/2}$ corresponds well with the numerical result of 0.70. Although we are not at present able to demonstrate the stability of this solution (in the sense of Kolmogoroff *et al.*), the convergence observed in the numerical solution to the profile given by Eq. (69) is a relatively convincing demonstration of some form of asymptotic stability. However, unlike the case of the Fisher–Kolmogoroff problem, we have explicitly obtained a closed-form expression for the population-density profile.[7] [One of us (Newman, 1980) has expanded upon some of the methods we have just discussed and found a class of solutions to some problems that appear in population genetics and combustion (slow flame propagation).]

Before continuing, we must consider the possibility that the nonrandom source of dispersal (directed Brownian motion on a microscopic scale) cannot be described adequately by a partial differential equation. In particular, we are concerned with the case where the wavefront predicted by the partial differential equation [e.g., that given by Eq. (69)] is narrower than the mean

[7] Lax (1969), in presenting the Ninth John von Neumann Lecture of the Society for Industrial and Applied Mathematics remarked on how fascinated von Neumann had become by the possibility that patterns disclosed by numerical calculations might reveal entirely unsuspected properties of solutions of nonlinear differential equations. A case in point is the Kortweg–de Vries Equation, which describes long waves over water and some wave phenomena in plasma physics. In his lecture, Lax went on to describe how Kruskal and Zabusky discovered the existence of certain solitary wave solutions after studying motion pictures of the computations. In our case, the computed solutions to Eq. (59) pointed out the existence of an asymptotic travelling wave solution and, in one instance, led to a closed-form analytic expression for the profile and its associated travelling wave velocity.

separation distance between colonizable planetary systems, but where the wave crossing time between such systems is much longer than the interval between successive colonization ventures. [This point was brought to our attention by Jones (1980), who noted this feature in a discretized numerical simulation of our model in a restricted region of parameter space. He observed an elevation of almost an order of magnitude in the wave speed.] This problem parallels the anomaly studied by Einstein in his classic microscopic treatment of Brownian motion [cf. our discussion of (13)]. Several functionally equivalent methods are available to us that can easily remedy this problem; the most famous is the modelling of the microscopic Brownian motion by an equation of Langevin type (the route Einstein followed). Alternatively, we can discretize the density-dependent diffusion equation including growth and saturation. This latter approach has been very successful in applications to population genetics (Maruyama, 1972, 1977) and is often referred to as the "island model" (Nagylaki, 1977). We pursue this approach since it most closely parallels our own [see, e.g., the derivation of the diffusion equation from the Chapman–Kolmogorov description for the drunkard's walk, Eq. (7) *et seq.*].

We employ discrete spatial differences in describing the diffusion term while maintaining the time derivative in the description of temporal evolution. The effect of spatial discretization on travelling wave speed has not attracted significant attention in mathematical biology. However, travelling waves in discrete systems have been studied exhaustively in solid state physics. In particular, one can consider a monatomic lattice to be an ensemble of masses connected by springs described by some generally anharmonic potential. The Euler–Lagrange Equations which describe this mechanical system have a similar appearance (in certain respects) to the discretized version of our nonlinear diffusion equation. In what is called the Toda chain, the force term characteristic of each mass point is an exponential. Toda (1967) was able to calculate the exact solution to the problem in terms of the Jacobian elliptic functions; however, he also showed how one can obtain, in the linearized limit, a bound on the velocity of the travelling wave. In Appendix I, we employ a variation of this technique to determine the minimum wave speed for the discretized version of the semilinear diffusion equation, comparing this result with that due to Fisher and to Kolmogoroff *et al.* discussed in Section 5.

7. SUMMARY OF MATHEMATICAL APPROACHES

We have reviewed the advantages (and limitations) of potential theory in describing population dynamic processes. We have observed that dispersal is essentially a diffusion process, and that population growth (if it takes place) is restricted by the carrying capacity of the environment. A density-dependent diffusion coefficient was shown to be capable of describing the desirability of emigration (all other factors being equal). Essentially, two basic models emerge. One model describes a condition of zero population growth. The properties of its solution may be obtained from the porous medium or thermal wave problems. The second model also includes local growth and saturation. (In Appendix I, we examine the influence of discretization in this second model.) A numerical study of its properties disclosed the existence of an asymptotic travelling wave solution.

Two features stand out in our population dynamics modelling. First, the velocity of the colonization wavefront can be estimated by dimensional analysis using the physical and other parameters that influence the evolution of the civilization, including the growth rate γ and the diffusion coefficient D. The latter, in turn, is related by Eq. (21) to the effective probability P that an individual member of a civilization will embark on such a venture, the mean distance Δx between habitable solar systems, and the effective time span Δt between successive colonization ventures (discussed below). Second, the dimensionless multiplying coefficients that appear in the expressions for the velocity (ξ or $\tilde{v}$ for the ZPG and travelling wave solutions, respectively) are ~ 1 for *all* continuous models, and the explicit dependence of velocity on the physical parameters is characterized by power laws with fractional exponents. This manifests itself in dimensionless velocities that are remarkably and encouragingly insensitive to the particular model and parameters employed. (Discretization effects complicate this feature somewhat, an issue addressed below and in Appendix I.) In the next section, we see how both models yield a much lower colonization wave velocity than has previously been predicted.

Before continuing, however, we must point out why our models provide results that contrast so strongly with those of Jones (1976): He describes the growth and expansion of a species which exercises population control by

$$\partial \nu/\partial t = (1 - \nu/\nu_0)(\gamma - \eta)\nu + I \quad (70)$$

[cf. Eq. (5)]. Here, ν is the population per planet at a given distance from the home world, ν_0 the "optimum population" (more precisely, the carrying capacity), γ the growth rate, η the emigration rate, and I the immigration rate and different from the I in Eq. (58). (Jones' η is not necessarily comparable to our Ψ below, because η is independent of external factors while Ψ depends on both ν and $d\nu/dx$.) His numerical experiment was performed by scattering unpopulated worlds at random in a test volume and choosing values of η/γ and of ν_L/ν_0, where ν_L is a limiting population above which immigration is not permitted for a given colony, and below which there is no emigration. The volume density of worlds, ρ, is also a parameter. Colonies with ongoing emigration send ships alternately to the two nearest open colonies ($\nu < \nu_L$).

In Jones' model, the distribution of colonial civilization assumes an approximately spherical shape, where the core of the sphere is nearly saturated. Thus, his colonists arise near the surface of the sphere. In our model, the flow of colonists goes as $\rho^N(\partial\rho/\partial r)$, and is also significant only in a thin layer at the population wavefront. In both our model and that of Jones the rate of total increase, $d\nu/dt$, is proportional to the number of colonists present. The number of colonists, on the other hand, is proportional to the surface area of the saturated sphere containing the expanding interstellar civilization. Since $\nu \propto r_f^3$, where r_f is the radius of the sphere, both models predict a population growth rate $d\nu/dt \propto \nu^{2/3}$. This integrates to give $\nu \propto t^3$ or $r_f \propto t$, and both models yield a spherical colonization wavefront undergoing uniform expansion.

Since we expect the core of the civilization to be saturated, almost all available resources would by definition be committed to the continued maintenance of the local ecosystem and few, if any, resources would be available for the construction of starships capable of transporting and supporting colonists across interstellar space. Interstellar spaceflight for a substantial community of colonists at $v \sim c$ must, for any civilization, be enormously expensive in propulsion and shielding. At $v \ll c$, the cost remains enormous because the long transit times require elaborate life-support systems. Contemporary colonies of $\sim 10^4$ persons in cislunar space (O'Neill, 1975, 1978) are estimated optimistically to cost $\gg \$10^{11}$. Even with generous allowances for future technological progress, a larger colony designed to traverse parsecs safely might be argued to cost substantially more than the available resources of the planet Earth ($\sim \$10^{13}$ gross planetary product). It is clear that, for either velocity regime, colonization ventures to more distant targets would be correspondingly more expensive. Except for extremely advanced and long-lived technical societies, colonization is possible only to the nearest star systems. Thus colonization is initiated from the periphery of an empire for two reasons. First, the transit distances and times are shorter. Second, materials are not in such short supply and the launching of colonization ventures would not place as unbearable a strain on the outposts of the empire as on its central worlds. In both discussions, a random walk is executed (recall that Jones' calculations were performed via a Monte Carlo program) but, unlike the conventional drunkard's walk, there is a directional bias—away from the population centers. We derive below approximate analytic formulae that closely reproduce the numerical results obtained by Jones (1978, 1980) in his Monte Carlo program. In Appendix I, we show how our nonlinear diffusion continuum model must be modified in order to accommodate discretization effects, since collision lengths are ~ parsecs. There, we observe that our model for multiplicative nonlinear diffusion becomes indistinguishable from Jones' when the population growth rate (among other factors) becomes sufficiently large. Thus, what distinguishes our results from those of Jones is the choice of parameters which describe the demographic behavior of our respective colonial civilizations. We consider Jones' model to be a limiting case of our multiplicative nonlinear diffusion model in the presence of discretization effects. We now derive Jones' results analytically to first order (cf. Appendix I for a more exact derivation).

Clearly, for Jones' model, the ratio of the effective velocity to the ship velocity is given by

$$v/v_s = t_{\text{travel}}/(t_{\text{travel}} + t_{\text{growth}}), \quad (71)$$

where t_{travel} is the ship's travel time from the world supplying the emigrants to the world absorbing them, a distance Δx away ($t_{\text{travel}} = \Delta x/v_s$). The time t_{growth} is a measure of the interval between the landing of colonists on an untenanted world and the time when it grows to a population ν_L and begins to launch its own colonial ventures.

There are two steps required to estimate t_{growth}. First, we must determine how many colonists a virgin world receives during the emigration phase. From Eq. (70), the new world receives $(\eta/2\gamma)(\nu_O - \nu_L)$ colonists for $\eta \ll \gamma$. This follows since, as the colonizing world grows from population ν_L to ν_O, a fraction η/γ of its population increase, namely, $\nu_O - \nu_L$, emigrates. The factor $\frac{1}{2}$ arises since only one-half of the colonists go to a given world. [We have ignored here the possibility that a given colony is receiving immigrants from more than one planet. Jones (1980) argues that the average distance travelled by colonists is close to the average interstellar separation. Thus the assumption that colonists are received at a given time from only one source will not significantly affect our results.] Accordingly, each new colony has a "seed population" $\sim(\eta/2\gamma)(\nu_O - \nu_L)$ from which it develops. Given this seed population, we must estimate how long it would take for the colony to grow to population ν_L. In detail, this depends on the Riccati Equation (70) and the source term I. We can, however, obtain a crude estimate (within a factor of 2 or 3) of this time scale by the following argument: The e-folding time for growth is γ^{-1}. [This is a modest underestimate because of the factor $(1 - \nu/\nu_O)$ in Eq. (70).] The number of e-folding times involved is just the logarithm of ν_L divided by the seed population. (This is also a modest underestimate since it presupposes that all colonists arrive at one time, rather than during some interval.) Therefore

$$t_{\text{growth}} \simeq \gamma^{-1} \ln \{[2\gamma\nu_L/\nu_O]/[\eta(1 - \nu_L/\nu_O)]\}. \quad (72)$$

Note that for Jones' preferred value, $(\nu_L/\nu_O) = \frac{1}{2}$, this is just $\gamma^{-1} \ln(2\gamma/\eta)$. Since

$t_{\text{travel}} = \Delta x/v_s$, combining Eqs. (71) and (72) yields

$$v/v_s \simeq (\gamma\Delta x/\{\gamma\Delta x + v_s \ln[(2\gamma\nu_L/\nu_O)/\eta(1 - \nu_L/\nu_O)]\})$$
$$\simeq \{\gamma\Delta x/[\gamma\Delta x + v_s \ln(2\gamma/\eta)]\} \quad (73)$$

for $\nu_L/\nu_0 = \frac{1}{2}$. Since $\gamma\Delta x \ll v_s \ln(2\gamma/\eta)$, we expect that power law fits for v/v_s would go crudely as $\gamma^1\eta^0$. The logarithmic term in (73) would decrease and increase the powers of γ and η, respectively, by the same amount. The $\gamma\Delta x$ term in the denominator would reduce the power of γ somewhat further. Jones (1980) has found that $v/v_s \propto \gamma^{0.83}\eta^{0.13}$ from his Monte Carlo calculations. Equation (73) predicts very similar effective velocities, approximately two times greater for the reasons given.

Now that we have explained Jones' Monte Carlo results in a mathematical sense we must understand the difference between the models on population dynamics grounds. In nonlinear multiplicative diffusion, dimensional analysis provides a natural distance scale, l, from the associated rate constants D and γ:

$$l = (D/\gamma)^{1/2}. \quad (74)$$

For example, the width of the travelling wavefront in the continuum problem is of this order. By dimensional considerations, we can associate Jones' immigration rate η with $D/\Delta x^2$, where Δx is the distance between habitable worlds, so

$$D \sim \eta\,\Delta x^2. \quad (75)$$

Combining (74) and (75), we find

$$l \sim (\eta/\gamma)^{1/2}\,\Delta x. \quad (76)$$

In animal populations $\eta \gg \gamma$ since almost all animals undergo dispersal during their lifetime. (Here, we have used η as generally representative of immigration or emigration rates. In detail, however, there are significant differences between Jones' definition of immigration rates and our own.) Thus, $l \gg \Delta x$ and discretization effects are negligible in describing animal populations. For human populations on Earth, on the other hand, it is often the case, especially in historical times, that $\Delta x \gg l$. Hence, discretization effects must be considered in modelling, for example, European colonization of America. In Fig. 4 (described in detail in Appendix I), we plot the dimensionless velocity $\bar{v}$ against $\Delta x/l$. (Consider here the curve labelled "standard case.") It is our contention, in the concluding section, that the most probable choice of rate coefficients will induce discretization effects that increase colonization front speeds at most by an order of magnitude. The extraordinarily high values of γ and η that Jones prefers are characteristic of expansionist populations and yield unrealistically high wave speeds. We believe that the economics of colonization and other factors will rule out a determined galactic imperialism, as discussed further below.

8. RESULTS

We now collect some relevant results of the preceding sections, apply numerical values, and address ourselves to the central question of whether the Earth is likely to have been visited during geological time by extraterrestrial civilizations.

The steady state number N' of extant civilizations in the Milky Way Galaxy more advanced than our own can be written, after Drake, as

$$N' = fL, \quad (77)$$

where L is the mean lifetime of galactic civilizations in years and f is a factor that

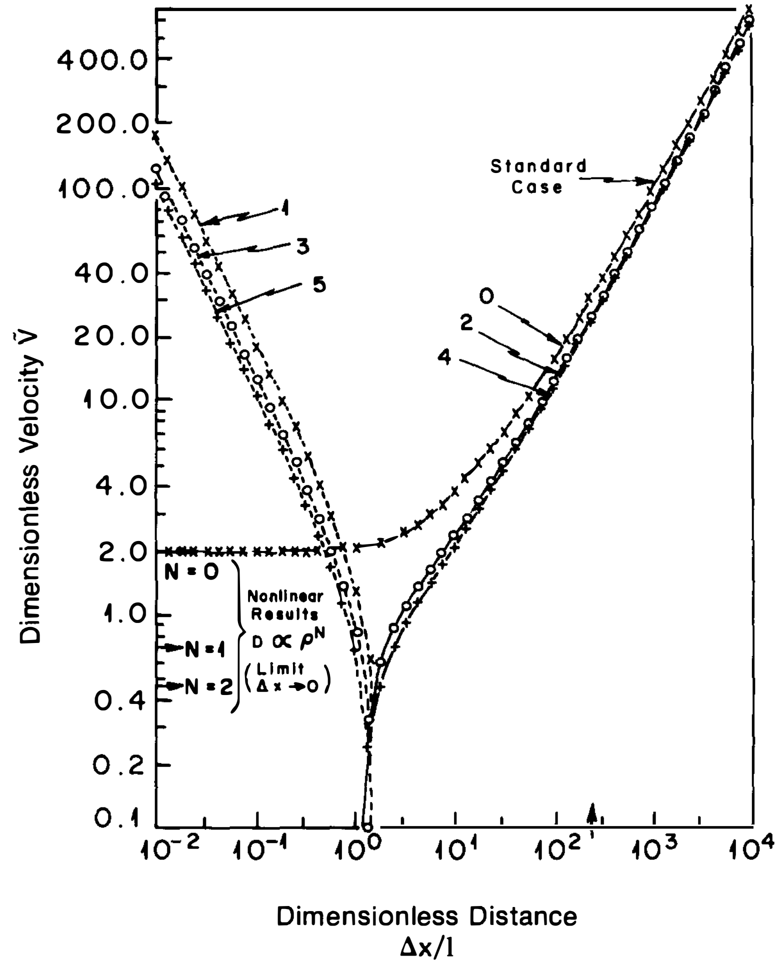

FIG. 4. Influence of discretization in semilinear diffusion on wave speed (analytic). The numbers on the curves are the modes corresponding to the eigenvalues of Eq. (I-18).

combines the rate of star formation, the fraction of stars with planetary systems, the number of ecologically suitable planets per such system, the probability of the origin of life on a given otherwise suitable planet, and the likelihood of the evolution of intelligence and technical civilization (for extensive discussions, see, for example, Shklovskii and Sagan, 1966; Sagan, 1973b). A conventional estimate, which is, however, no more than a semi-informed guess, is $f \sim 10^{-1}$ year^{-1} (*op. cit.*). However, in the case of colonization, the probability of the emergence of intelligent life and technical civilization on a given world approaches unity, and the factor increases perhaps to $f \sim 10$ year^{-1} (Shklovskii and Sagan, 1966, p. 451).

Assuming stars in the Milky Way Galaxy to have a mean separation of 1 parsec (= 3.26 light years), we find the mean distance between advanced civilizations to be

$$\Lambda \simeq (2.5 \times 10^{11}/N')^{1/3} = (2.5 \times 10^{11}/fL)^{1/3} \text{ pc.} \quad (78)$$

We here assume 250 billion stars in the Galaxy. Henceforth, distances will be measured in parsecs and times in years. With a conventional value (*op. cit.*) for independently arisen civilizations of $N' = 10^6$, $\Lambda = 63$ pc $= 205$ *l.y.* Equation (78) is based on a spherical distribution of stars in the Galaxy;

but the inaccuracies for distances larger than the thickness of the Galaxy (~100 pc) are still small enough that we may neglect them in this problem.

For our density-dependent diffusion problem with growth and saturation, we have found that the velocity v of the colonization wavefront can be represented by

$$v = \tilde{v}(D\gamma)^{1/2} \quad (79)$$

where $\tilde{v}$ is a dimensionless constant (of order unity unless discretization is important); D is a diffusion coefficient that, when multiplied by the population density gradient, yields the correct outward flux of population as it nears saturation; and γ is the population growth rate. For the human population on the Earth today, $\gamma \sim 10^{-2}$ year^{-1}. In the centuries prior to 1750, γ was $\sim 6 \times 10^{-4}$ year^{-1} (Coale, 1974). In the late Pleistocene, γ was much smaller still, perhaps 10^{-7} or 10^{-8} year^{-1}. We reemphasize that human history has been characterized by a strict balance between birth and death rates, apart from short-lived periods of rapid technological change. The Earth is currently recovering from such a perturbation and the prevailing growth rate of $\gamma \sim 10^{-2}$ year^{-1} must be viewed as a transient anomaly.

We have found [Eq. (21)] from finite differences in the standard m-dimensional diffusion equation that the probability of diffusing to another planet Δx away in time Δt is $P = 2mD\,\Delta t/\Delta x^2$. This is in fact the probability of the entire population emigrating to another solar system at a distance Δx over an effective time Δt. This is in strict agreement with the behavior of a particle undergoing Brownian motion. However, since the diffusion equation describes the behavior of a *distribution* of particles, it is conceptually useful to think of P as the expectation value of the fraction of the population that will emigrate to another world at a distance Δx in a time Δt. In the sense of an expectation value, then, we define the specific emigration rate, Ψ, by

$$\Psi = P/\Delta t = 2mD/\Delta x^2 \quad (80)$$

[cf. Eq. (75)]. This quantity describes the fraction of the population that emigrates per unit time to the next habitable star system. Here, Δt is the time interval between successive colonial ventures. If the transit time exceeds this time interval then we substitute the transit time for Δt. Conversely, the problem is paced by the time interval between successive colonizations for very fast transit times, so that even the unlikely contingency of faster-than-light travel (through multiply connected space in the vicinity of black holes, say) does not alter this analysis significantly.

It is impossible to make a very reliable estimate of Ψ for extraterrestrial civilizations. Very high values should be stupifyingly expensive. Very low values cannot provide even short-term relief of population pressure at the periphery of the civilization. In the third century B.C. the Chin emperor approved a proposal of Hsu Fu to launch a colonization mission "with several thousand young men and maidens to go and look for the abodes of the immortals hidden in the Eastern Ocean" (Needham, 1971). The fleet of sailing ships was never heard from again and the emperor complained bitterly about the cost. Over the following centuries a number of further expeditions to the Pacific Ocean to find an elixir of immortality were mustered, but none on the scale of the venture of Hsu Fu (*op. cit.*). All these expeditions may have perished; alternatively some may have colonized Japan and have been one source for the non-Ainu population there today. At this time the total population of China was $\sim 3 \times 10^7$. The next major distant emigrations from East and Northeast Asia were the Mongol invasions and the exploration and trade by the Ming navy in the Indian Ocean in the 15th century. These numbers give $\Psi \sim 10^{-7}$ year^{-1}. During the 18th century European colonization of North America, Ψ was $\sim 3 \times 10^{-4}$ year^{-1} (Potter, 1965). Proposals have been made recently for the launching into Earth orbit of self-contained closed ecological systems with $\sim 10^4$ inhabitants (O'Neill, 1975, 1978). This represents $\sim 10^{-6}$ of the population of the Earth. A very difficult undertaking in the next few centuries would be to launch one such Space City into interstellar space every century. This gives $\Psi \sim 10^{-8}$ year^{-1}. We adopt 3×10^{-4} year$^{-1} \geq \Psi \geq 10^{-8}$ year^{-1}, with a bias to the smaller value for a young interstellar civilization.

The existence of abundant extrasolar planetary systems seems very plausible but is still undemonstrated. Presumably a civilization wishing to colonize the Earth comes from an Earth-like world; although, since schemes have been proposed to terraform Venus and Mars—that is, to convert them from rather different environments to ones which resemble the Earth (Sagan, 1960; Sagan, 1973c)—xenoforming the Earth into some rather different environment might not be beyond the capabilities of an interstellar spacefaring civilization. We note that such xenoforming will add to the waiting time for a colony to develop an independent colonizing capability. Simple theoretical models of the formation of planetary systems from solar nebulae suggest that a wide variety of planetary systems can be formed, depending on, for example, the initial mass density distribution function of the nebula; in some schemes terrestrial planets appear to be formed abundantly (Dole, 1964; Isaacman and Sagan, 1975). In this case the distance between systems with Earth-like planets is probably ≤ 3 pc. Hart (1978, 1979) has argued that only an extremely unlikely set of circumstances has preserved the Earth's environment—particularly abundant liquid water—from the opposing threats of a runaway greenhouse effect and a global ice age. These arguments ignore plausible variations in some climatic parameters and neglect negative feedback loops in terrestrial climatology (Schneider and Thompson, 1980); we consider them unlikely. But if they are valid they increase Δx possibly to ~ 300 pc. Accordingly, we will allow Δx^2 to range from 3 to 10^3 pc^2, with a preferred value of 10 pc^2.

If we adopt the dimensionality index, $m = 3$, $\Delta x^2 = 10$ pc^2, and $\Psi = 10^{-8}$ year^{-1} we find a typical value of the diffusion coefficient to be $D = 2 \times 10^{-8}$ pc^2 year$^{-1} = 5 \times 10^{21}$ cm^2 sec^{-1}, an absolutely enormous diffusion coefficient by planetary atmospheres standards.

We are now ready to examine the apparent absence of extraterrestrial colonies on or near the Earth. Let $t_{\min}$ be the minimum time for the advanced technical civilization nearest us in space to reach us at a diffusion wavefront velocity v. Then $vt_{\min} = \Lambda$. If $t_{\min} > L$, there should be no such colonies, as observed. Consequently, there is a critical lifetime for a spacefaring interstellar civilization

$$L_c = t_{\min} = \Lambda / v. \tag{81}$$

For the Earth to have been colonized, the colonial civilization must have a lifetime in excess of L_c.

Combining Eqs. (78)–(81) we find

$$L_c = 1.95 \times 10^7 (f/0.1)^{-1/4} (\tilde{v}/2)^{-3/4} \times [(\gamma/10^{-4})(\Psi/10^{-8})(\Delta x^2/10)]^{-3/8} \text{ years.} \tag{82}$$

We have accommodated the possible effect of discretization through the inclusion of the $(\tilde{v}/2)^{-3/4}$ term. In order to estimate $\tilde{v}$, we can employ the "standard case" curve in Fig. 4, where $\tilde{v}$ is plotted with respect to $\Delta x/l$, which is just $(2m\gamma/\Psi)^{1/2}$ [from Eqs. (74) and (80)]. For comparison, we combine Eqs. (23), (24), (26), (78), and (80) to obtain the critical lifetime for a thermal wavefront with $N = 2$, and find

$$L_c = 1.3 \times 10^{10} (f/0.1)^{-8/11} \times (\Delta x^2/10)^{-12/11} (\Psi/10^{-8})^{-3/11} \text{ years.} \tag{83}$$

This is the case of strictly observed ZPG. [Recalling from the discussion following Eq. (80) that discretization effects are a function of $\Delta x/l \sim (\gamma/\Psi)^{1/2}$, we conclude that they are insignificant in ZPG problems at late times because then, crudely speaking $\gamma \simeq$

0.] Our free parameters, f, γ, Δx^2 and Ψ have normalized standard values already discussed, with distance measured in parsecs and time in years.

We have now obtained the critical lifetime L_c for the cases $L_c \ll \gamma^{-1}$ (zero population growth) and $L_c \gg \gamma^{-1}$ (nonlinear population diffusion with growth and saturation). From Eq. (82), we see that L_c increases as γ decreases. When γ is set equal to L_c^{-1} in (82), we obtain a value for L_c of 4.3×10^9 years when f, Ψ, and Δx assume their normalized standard values, neglecting the effect of discretization. Although (82) is not strictly valid in this case, this value for L_c suggests, as we intuitively expect, a smooth transition between our two models; i.e., as γ decreases, L_c increases until it assumes the ZPG value.

We see that for the colonization wavefront of the nearest technical civilization to have reached the Earth, the lifetime of that civilization must exceed 20 million years for nonlinear diffusion with growth and saturation. Such a civilization will have been intensively occupied in the colonization of more than 200,000 planetary systems before reaching the Earth, some 63 pc away. Thus, many colonial empires, of vast extent by terrestrial standards, may still occupy only an insignificant volume of the Milky Way Galaxy, and not embrace the nearest independently arisen technical civilization, whether it has colonial ambitions or not.

If strict ZPG is observed, L_c is of the order of the age of the universe and it would be unlikely in the extreme for us to observe such a population. Variations in f would almost certainly increase the critical lifetime, while variations in Ψ will have an almost negligible effect. Therefore, strict ZPG readily explains our failure to observe nearby extraterrestrial civilizations and we will confine our remaining considerations to cases including population growth and saturation.

Our conclusions for the case of density-dependent diffusion with population growth and saturation seem to be interestingly insensitive to the choice of input parameters. Reasonable variations in f, to allow for extensive colonization, change L_c by a factor ~ 3. We believe that γ cannot be much larger than 10^{-4} year^{-1} to be consistent with the very powerful population pressures that any colonial empire must have experienced earlier in its history than the time of extensive interstellar colonization. (In addition, one possible means of interstellar transport—the "generation ship" in which a much later generation of colonists arrives than that which originally set out—imposes an extremely strict regime of ZPG on the settlers.) Δx^2 might be as large as 10^3 pc^2, meaning that the colonists must be in the practice of searching comparatively large volumes of space before establishing a colony. Ψ conceivably could be $\gg 10^{-8}$ year^{-1}, although values as large as those for the European colonization of North America in the 18th century seem prohibitive for interstellar spaceflight. But, even adopting $\Delta x^2 = 10^3$ pc^2 and $\Psi = 3 \times 10^{-4}$ year^{-1}, we find that $L_c \sim 10^5$ years. Even with massive interstellar colonization efforts by a nearby civilization, involving 10% of the base population every century, the Earth would not have been visited unless the colonizing civilization were very long-lived.

The effective velocity of the colonization wavefront is given by

$$v = 2.6 \times 10^{-6}\,(\tilde{v}/2)[(\gamma/10^{-4}) \times (\Psi/10^{-8})(\Delta x^2/10)]^{1/2}\ \text{pc year}^{-1}, \quad (84)$$

where we have accommodated the possible effect of discretization through the inclusion of the $(\tilde{v}/2)$ term. Note from dimensional analysis that this is the only combination of γ, Δx^2, and Ψ which will make a velocity. With nominal values of γ, Δx^2 and Ψ, $\tilde{v} \sim 30$, and the colonization wavefront has an effective velocity $\sim 4 \times 10^{-5}$ pc year^{-1}—slower even than the interstellar velocities of the Pioneer 10 and 11 and Voyager 1 and 2 spacecraft.

A velocity of 4×10^{-5} pc year^{-1} corresponds to only 40 km sec^{-1}. Thus, differential galactic rotation and peculiar stellar velocities can significantly influence a civilization's long-range colonial strategy. Differential rotation would shear a civilization sphere, conceivably producing a tube oriented along a spiral arm. Peculiar velocities would further diffuse the interstellar boundaries of the civilization, potentially increasing the velocity of the expanding colonization front. It is unlikely that the velocity of expansion would be much greater than typical stellar peculiar velocities. Significantly, even the peculiar velocities of stars are two orders of magnitude less than the colonial wavefront velocities proposed by Jones (1976).

Civilizations with lifetimes $\sim 10^9$ to 10^{10} years, on the other hand, could have colonized the Earth by now even if they originated some considerably greater distance from us within the Milky Way Galaxy. However, for reasons we have already mentioned—concerning the evolution of very advanced societies—we believe that their motivations for colonization may have altered utterly, and that their science and technology may be extremely different from anything we can recognize or even imagine.

We have found that only technical civilizations with lifetimes $\sim 10^5$ to 2×10^7 years or longer could have initiated a colonial wavefront which has reached the Earth from the nearest technological star system. The lifetimes of extraterrestrial technical civilizations are, of course, highly uncertain, but many have hypothesized that only a tiny fraction of civilizations survive beyond 10^5 to 10^7 years (see, for example, Shklovskii and Sagan, 1966). In addition, there is the question of phase relations in the evolution of technical civilizations. In the solar neighborhood there are unlikely to be many stars with habitable planets that are significantly older than the Sun, and we certainly expect a substantial waiting time—perhaps several times 10^9 years—for the indigenous evolution of technical societies. There are unlikely to be any colonizing civilizations in the solar neighborhood $\gg 10^6$ years old.

It may be useful to divide technical civilizations into two categories: young ($\ll 10^6$ years old) and old ($\gg 10^6$ years old). Old civilizations are improbable colonial powers in the sense we are describing. Young civilizations may be embarked on extensive colonial ventures but, compared to the volume of the Milky Way, these are of very limited scope. There may be empires of tens or even hundreds of thousands of worlds. But it is implausible that a true Galactic Empire exists: there are too many worlds to conquer.[8]

The number of habitable planets in the Galaxy may exceed the number of individuals on the Earth. Might the political evolution of the Galaxy to some extent parallel that of the Earth, in which individuals start out in family groups, and eventually form tribes, city-states, nations, superpowers, and, perhaps eventually, a single global state? In that case the Earth is at the present time still in the situation of a small family group—one which even wonders if there are any other groups at all. The stage we have been describing in this paper, when there are tens or hundreds of thousands of colonial worlds for each interstellar civilization, corresponds to a much later stage in human history, roughly that of the first city-states. It is only at this point that groups interact. Thus Star Wars, if there are any, occur at the level of city-states—Athens versus Sparta. Actually, the likely disparity in technological capability of any two newly interacting colonial powers is so great that "wars" in our sense cannot hap-

[8] The possibly extensive use of automated interstellar probes to scout candidate worlds does not change these considerations substantially. After finding, say, $\gg 10^3$ habitable worlds lacking indigenous intelligence, detailed investigation and subsequent colonization of the discovered worlds is a more likely response than an accelerated program of automated exploration of new worlds.

pen. The stronger civilization will simply overwhelm the weaker if a conflict arises.

What is the likelihood of such encounters? For a random, uniform distribution of civilizations the probability p of an interaction would behave as

$$p \lesssim (d/\Lambda)^3, \qquad (85)$$

where Λ is the mean distance between civilizations [Eq. (78)] and d is the mean value of the maximum radius of an interstellar empire; i.e.,

$$d = vL = \bar{v}(D\gamma)^{1/2}L. \qquad (86)$$

Using (78), (81), (83), and (86) we find

$$p \lesssim (L/L_c)^4 \qquad (87)$$

for $L < L_c$ where, for simplicity, we have ignored the effect of discretization. Thus, the likelihood of two civilizations interacting is remote unless $L \sim L_c$. (This argument depends of course on a given civilization growing in a manner that is independent of its neighbors.)

If the probability of Star Wars goes as L^4; if the canonical lifetime for contact to be just occurring is 10^7 years; and if we were to imagine all civilizations at $L = 10^6$ years, then only 1 in 10^4 civilizations would be interacting. Since $N' = 10^6$, ~100 will be interacting at a given time. If $L = 10^5$ years, then 1 in 10^8 civilizations will be interacting and, for $N' < 10^6$ civilizations, there are no Star Wars. Thus, if warfare is an illness which is outgrown in the first 10^5 years or so of the existence of a galactic civilization, then there are never interstellar conflicts; in any case, they would be very brief because of the technological disparities.

At the end of Section 7 we distinguished between interstellar civilizations expanding because of high population densities and high population gradients together, and those expanding because of high population densities alone. The latter are societies bent on conquest. For such an imperialist galactic civilization [Eqs. (72) and (73) for $\nu_L/\nu_O = \frac{1}{2}$] and for our above choices of γ, Δx, and η (taken here as $\simeq\Psi$), t_{growth} ~10^5 years and $v/v_s \sim 3 \times 10^{-3}$ for $v_s = 0.1c$, while $v/v_s \sim 3 \times 10^{-4}$ for $v_s = c$. In these cases the Galaxy would be colonized in 3×10^8 to 3×10^7 years, respectively.

An advanced civilization *intent* on colonization and military conquest of the entire Galaxy might therefore succeed. But we have considered exploration and colonization as an extension of a civilization's thirst for knowledge, or as insurance against self-destruction, not as a relief valve for overpopulation (which it logistically cannot be) or as an instrument of territorial conquest and expansion. The apparent absence on Earth of representatives of advanced galactic civilizations thus argues against the presence of any such expansionist civilizations devoted to domination of the Galaxy. It is curious that the solution to the problem "Where are they?" depends powerfully on the politics and ethics of advanced societies. Why are there no dedicated galactic imperialists, since we anticipate an enormous range of biologically distinct civilizations to arise on their separate home worlds? It may be that such behavior is universally outgrown in the first 10^5–10^7 years of the lifetime of a technical civilization, e.g., because of the common invention of the technology of personal immortality, or because of the selective disadvantage of the drive for territoriality after the development of weapons of mass destruction, as discussed earlier. Or it may be that the first civilizaiton to colonize the Galaxy, perhaps long before the formation of the Earth, imposed certain boundary conditions on the expansion of those interstellar civilizations which subsequently arose. These two possibilities have different, but in both cases encouraging, implications for the long-term future of the human species.

The slow propagation speed of the colonization wavefront does not exclude specific exploratory missions to targets of particular astrophysical or biological interest. The wavefront of the nearest colonial civilization may be tens or even only a few parsecs from the Earth, and able to visit our solar system in relatively short times in the likely case that starship velocities $v_s \gtrsim 0.1c$ exist. They have not visited the Earth because there is no motivation, no apparent way for a neighboring civilization to know of the existence of our just-emerging interstellar civilization. This situation may change in the relatively near future. Military and astronomical radar systems and, especially, commercial television have been for the last several decades generating an unmistakable radio signature, now ~ 10 pc away, of the existence of a new technical society in the vicinity of the Sun. This may trigger a radio response in the not too distant future. Depending on the disposition of nearby civilizations, their colonies and their exploratory vessels, and how carefully they are examining the radio spectrum in an eavesdropping mode, the leakage radiation from our technology might also stimulate an exploratory mission from the nearest galactic civilization within the next few centuries or less.

APPENDIX I: THE EFFECTS OF DISCRETIZATION IN NONLINEAR DIFFUSION

In this paper, we have described nonlinear diffusion as the outcome of a random-walk process. Our discussion focused on the limiting case that collision mean free paths and collision times were small compared with "macroscopic" spatial and temporal scales. By the analytic and numerical study of the resulting partial differential equations, we were able to deduce the principal feature of nonlinear diffusion in the continuous limit: the existence of a travelling wave which propagates out from the population center at a velocity determined by the rate constants associated with *in situ* growth and diffusion. Moreover, the velocity of advance at late times is relatively unaffected by the initial conditions and by the details of the particular nonlinear diffusion model employed.

The diffusion process, as a representation of Brownian motion, is not limited to the case that mean free paths $\ll$ than the spatial scales that otherwise describe the problem. A natural scale is the mean distance between habitable star systems. There is, fortuitously, no natural time scale that must also be included. Thus, we must rederive, after Chandrasekhar (1943), the equations that describe random walk where spatial discretization may be important. To illustrate, consider the semilinear diffusion equation (43) in one spatial dimension described earlier:

$$\partial\rho(x, t)/\partial t = \gamma\rho(x, t)[1 - \rho(x, t)/\rho_s] + (\partial/\partial x)\{D\partial\rho(x, t)/\partial x\}. \quad \text{(I-1)}$$

This equation can be regarded as the continuum limit of a Markov process such as that developed in Eqs. (7)–(11). The diffusion term in (I-1) was obtained from (9),

$$\partial^2\rho(x = nl, t)/\partial x^2 = \lim_{l\to 0}[\rho_{n+1}(t) - 2\rho_n(t) + \rho_{n-1}(t)]/l^2, \quad \text{(I-2)}$$

where l is the collision path length in

$$\rho_n(t) \equiv \rho(nl, t). \quad \text{(I-3)}$$

It is natural, in this problem, to associate with the collision length l the distance Δx between habitable systems. Thus, in order not to suppress the effects of discretization, we write (I-1) in the Kolmogorov–Chapman form

$$d\rho_m(t)/dt = \gamma\rho_m(t)[1 - \rho_m(t)/\rho_s] + (D/\Delta x^2)\{\rho_{m+1}(t) - 2\rho_m(t) + \rho_{m+1}(t)\}, \quad \text{(I-4)}$$

where $\rho_m(t)$ denotes the population density at time t at the mth site (i.e., spatial position $m\Delta x$). The site index m is different from the dimensionality index m in Eqs. (21) *et seq.*

Can this discretized semilinear diffusion model support a right-going travelling wave? (See Section 5 for a parallel discussion of the continuous case.) In particular, for which range of values of the travelling wave velocity v will a solution to Eq. (I-4) *exist*? Assuming a travelling wave solution,

$$\rho_{m+1}(t) = \rho_m(t - \Delta x/v),$$

$$\rho_{m-1}(t) = \rho_m(t + \Delta x/v). \quad \text{(I-5)}$$

Then, Eq. (I-4) assumes the form

$$d\rho_m(t)/dt = \gamma\rho_m(t)[1 - \rho_m(t)/\rho_s] + (D/\Delta x^2)\{\rho_m(t - \Delta x/v) - 2\rho_m(t) + \rho_m(t + \Delta x/v)\}. \quad \text{(I-6)}$$

Any solution must exist in the asymptotic regime where $\rho_m(t) \to 0$, i.e., where $[1 - \rho_m(t)/\rho_s]$ can be approximated by unity. Therefore, we consider the linearized approximation

$$d\rho_m(t)/dt = \gamma\rho_m(t) + (D/\Delta x^2)\{\rho_m(t - \Delta x/v) - 2\rho_m(t) + \rho_m(t + \Delta x/v)\}, \quad \text{(I-7)}$$

and Fourier analyze; we consider solutions to (I-7) of the form $\exp[i\omega t]$ so that

$$\rho_m(t + \tau) = \exp[i\omega\tau]\rho_m(t). \quad \text{(I-8)}$$

Then, (I-7) becomes

$$i\omega = \gamma + (D/\Delta x^2) \times [\exp(i\omega\Delta x/v) - 2 + \exp(-i\omega\Delta x/v)]. \quad \text{(I-9)}$$

For the sake of simplicity, we render all quantities dimensionless by the transformations

$$\omega' = \omega/\gamma,$$

$$\Delta x' = \Delta x(D/\gamma)^{-1/2},$$

$$\tilde{v} = v(D\gamma)^{-1/2} \quad \text{(I-10)}$$

[cf. Eqs. (44) and (46)]. Equation (I-9) becomes

$$i\omega = 1 + (\Delta x)^{-2}\{\exp(i\omega\Delta x/\tilde{v}) - 2 + \exp(-i\omega\Delta x/\tilde{v})\}, \quad \text{(I-11)}$$

(where, for convenience, the primes have been dropped).

We now wish to find the value(s) of ω which satisfy this equation. Clearly, this is equivalent to finding the roots of the function $F(\omega)$ which we define to be

$$F(\omega) \equiv 1 + (\Delta x)^2\{\exp(i\omega\Delta x/\tilde{v}) - 2 + \exp(-i\omega\Delta x/\tilde{v})\} - i\omega. \quad \text{(I-12)}$$

Anticipating the possibility that ω is complex, we write it as $\Omega - i\nu$, where both Ω and ν are real. Similarly, since $F(\omega)$ can be complex, we write F as $f + ig$, where both f and g are real-valued functions. Separating the real and imaginary terms in (I-12), we obtain the dispersion relations

$$f = 1 + (\Delta x)^{-2}[2\cosh(\nu\Delta x/\tilde{v}) \times \cos(\Omega\Delta x/\tilde{v}) - 2] - \nu;$$

$$g = (\Delta x)^{-2}\, 2\sinh(\nu\Delta x/\tilde{v}) \times \sin(\Omega\Delta x/\tilde{v}) - \Omega. \quad \text{(I-13)}$$

Now, consider the behavior of f and g as functions of ν (i.e., holding Ω fixed). If $\omega = \Omega - i\nu$ is a solution to the latter dispersion relations, so is $-\Omega - i\nu$. (Physically, this corresponds to the fact that a left-going wave travelling with the same velocity is also possible.) So, we consider solutions with $\Omega \geq 0$ without loss of generality.

Since g is an odd function of ν and the sign of $\partial g/\partial\nu$ does not change, g has a single root in ν. (If $\Omega = 0$, $g = 0$ is automatically satisfied.) The function f, on the other hand, is concave with respect to ν since $\mathrm{sgn}(\partial^2 f/\partial\nu^2) = \mathrm{sgn}[\cos(\Omega\Delta x/\tilde{v})]$. By inspection, $f(v)$ has either two or no roots. Consider the value of ν, say ν_0, for which $\partial f/\partial\nu = 0$. If there are two roots, ν_0 will be situated between them and $\mathrm{sgn}[f(\nu_0)] = \mathrm{sgn}[\cos(\Omega\Delta x/\tilde{v})]$. (There is a degenerate case, which establishes a limiting velocity for $\tilde{v}$, where the two roots coalesce into ν_0. In this degenerate case, there is only one root and it satisfies $f = 0$ and $\partial f/\partial\nu = 0$, simultaneously.) Finally, there are no roots if $\mathrm{sgn}[f(\nu_0)] = \mathrm{sgn}[\cos(\Omega\Delta x/\tilde{v})]$.

We must simultaneously adjust Ω and ν so that the single root of g and one of the two roots of f coincide. This solution for Ω and ν also depends on the value of $\tilde{v}$ employed. As $\tilde{v}$ is decreased, the two roots of f will move together and will coalesce at some critical velocity, say $\tilde{v}_c$. Below that velocity $\tilde{v}_c$, no simultaneous solution for both Ω and ν exists. Thus, $\tilde{v}_c$ is the minimum velocity at which a travelling wave can propogate in this low density (linearized) regime. [If we were to take the limit $\Delta x \to 0$ at this point, we would recover the result (55) due to Fisher and Kolmogoroff *et al.* that $\tilde{v}_c = 2$.] In order to identify the critical velocity, we must introduce the condition under which the two roots of f coalesce, namely,

$$\partial f/\partial \nu = [2 \sinh(\nu\Delta x/\tilde{v}) \cos(\Omega\Delta x/\tilde{v})/\Delta x\tilde{v}] - 1 = 0. \quad \text{(I-14)}$$

Therefore, we can identify Ω and ν as well as the minimal velocity $\tilde{v}_c$ by requiring that (I-13) and (I-14) simultaneously vanish.

Comparing the latter with the second of Eqs. (I-13), we obtain

$$\tan(\Omega\Delta x/v) = \Omega\Delta x/v. \quad \text{(I-15)}$$

For convenience, let

$$y \equiv \Omega\Delta x/v, \quad \text{(I-16)}$$

so that we need the eigenvalues y which satisfy

$$\tan y = y. \quad \text{(I-17)}$$

The first few are 0, 4.49341, 7.72525, 10.90412, . . . and the corresponding values of cos y, which we need later, are 1, −0.21723, 0.12837, −0.09133, ...

We now solve, for each eigenvalue y, the first of Eqs. (I-13) and (I-14)—provided that such a solution exists. (There is no guarantee that it will, apart from $y = 0$, if Δx is sufficiently different from zero.) For convenience, let us define a variable q,

$$q \equiv \exp(\nu\Delta x/\tilde{v}). \quad \text{(I-18)}$$

From (I-14), we obtain

$$q - (1/q) = [\Delta x\tilde{v}/\cos y] \quad \text{(I-19)}$$

or

$$q = \tfrac{1}{2}\{(\Delta x\tilde{v}/\cos y) + [(\Delta x\tilde{v}/\cos y)^2 + 4]^{1/2}\}. \quad \text{(I-20)}$$

Writing ν as $(\tilde{v}/\Delta x)\ln q$, we rewrite the first of Eqs. (I-13) as

$$0 = 1 + \{[(q + 1/q)\cos y - 2]/\Delta x^2\} - (\tilde{v}/\Delta x)\ln q, \quad \text{(I-21)}$$

where q is given in terms of v and Δx as well as the eigenmode y by (I-20). This is now an equation in a single variable, namely, $\tilde{v}$ (real valued and positive) which can be obtained numerically. The only issue remaining is the existence and the uniqueness of solutions to (I-21).

Multiplying it by Δx^2 and writing

$$u \equiv \tilde{v}\Delta x/\cos y \quad \text{(I-22)}$$

we obtain

$$0 = \Delta x^2 - 2 + \cos y\ \{(u^2 + 4)^{1/2} - u \ln\{[u + (u^2 + 4)^{1/2}]/2\}. \quad \text{(I-23)}$$

Further, by writing

$$s = \sinh^{-1}(u/2), \quad = \sinh^{-1}(\tilde{v}\Delta x/2 \cos y), \quad \text{(I-24)}$$

the dispersion equation (I-23) reduces, after some algebra, to

$$(2 - \Delta x^2)/\cos y = 2[\cosh s - s \sinh s]. \quad \text{(I-25)}$$

We note that for $s > 0$ ($s < 0$ corresponds to a left-going wave with the same speed), the right-hand side of (I-25) is strictly decreasing and has a maximum value of 2. Thus, the solution, if it exists, is unique and depends on the left-hand side of (I-25) via Δx and the sign of cos y, where y is an eigenvalue of (I-17). For example, if cos y is negative, no solution exists if

$$\Delta x > [2(1 - \cos y)]^{1/2}. \quad \text{(I-26)}$$

It is convenient to refer to the y eigenvalues, 0, 4.49341, 7.72525, etc., as modes 0, 1, 2, etc. We can prove that the eigenvalues of (I-17) are such that successive values of cos y alternate in sign and converge uniformly to zero. Odd-numbered modes decay ($\nu < 0$) while even-numbered modes grow ($\nu > 0$). Thus, odd-numbered modes are uninteresting since they will enjoy only a brief existence. Of the even-numbered modes, only mode 0 ($y = 0$) does not oscillate and it corresponds to the Fisher and Kolmogoroff *et al.* travelling wave in the $\Delta x \rightarrow 0$ limit. Since the even-numbered modes grow, we must understand the possible role of the oscillatory modes, 2, 4, . . . , etc. This is especially important since modes 2, 4, etc. could propagate much more slowly than the non-oscillating mode 0. The important feature of the even-numbered modes, therefore, is the relationship between their exponential growth rates. In particular, we can show that modes 2, 4, etc. grow less rapidly than mode 0. Thus, mode 0 will dwarf all other even-numbered modes, as well as the odd-numbered ones, in the longterm.

In Fig. 4, we plot the dimensionless critical velocity $\tilde{v}_c$ against the dimensionless discretization distance $\Delta x' = \Delta x/l$ for the different modes, where l describes the "natural" distance scale $(D/\gamma)^{1/2}$ [cf. Eq. (I-10)]. (In the absence of discretization, the wavefront would have an effective thickness $\sim(D/\gamma)^{1/2}$, a feature observed in Section 5.) Of particular interest is the case $\Delta x' \sim 245$, which is highlighted by an arrow on the horizontal axis. (This corresponds to the situation we consider likely, described in Section 8, where, in three dimensions, $\psi = 10^{-8}$ year^{-1} and $\gamma = 10^{-4}$ year^{-1}.) In Fig. 4, we see clearly how $\tilde{v}_c$ tends to 2 for the mode zero (the "standard case," since it corresponds directly to the result of Fisher and of Kolmogoroff *et al.*). Although unimportant in the long term, the critical velocity is plotted for modes 1 through 5 in order to elucidate the existence and early role of oscillation.

These dispersion relation results describe the velocity below which no wave can persist at low density in the discretized Fisher problem (where D = const.). The dispersion relation results are a *necessary* condition for the existence of a wave at low density, but say nothing about the existence of a wave at higher densities where nonlinear saturation effects occur.

In order to resolve this issue, we have performed nonlinear simulations to determine whether the nonlinearities could cause the wave to travel any faster than the low-density limit would predict. We also employed the simulations as a test of the sensitivity of the wave velocity to the density dependence that we wish to introduce into the diffusion coefficient [i.e., $D(\rho) \propto \rho^N$ or $D(\rho) = D(\rho/\rho_s)^N$]. (Figure 4 also displays the nonlinear results obtained in Section 5 for N = 0, 1, and 2 in the limit $\Delta x \rightarrow 0$.) For this, we performed a number of nonlinear simulations for $\Delta x' = \Delta x/l$ in the range 100 to 1000 for N = 0, 1, and 2.

Figure 5 summarizes the nonlinear results. Points on the figure are numerical results (subject to a 5% relative error) for diffusion with N = 0, 1, 2. The solid line is mode 0 ("standard case") from the previous figure. Several trends are manifest: (a) The wave speed is insensitive to our choice of N. (b) The wave speed decreases slightly with increasing N. (c) The dispersion relation for the discretized, linearized semilinear diffusion equation provides an

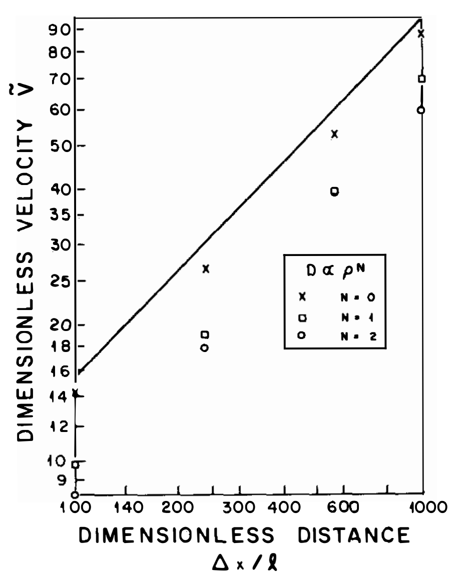

FIG. 5. Influence of discretization in nonlinear diffusion on wave speed (numerical). The solid line is the mode 0 standard case from the analytic solution (Fig. 4).

order-of-magnitude velocity estimate for the *nonlinear* problem.

Our dispersion relation approach has yielded remarkably accurate estimates of wavefront velocities ranging from the discrete to the continuous limit for a variety of (nonlinear) models. Indeed, the mode 0 results are also a good approximation to the Monte Carlo computations performed by Jones. Earlier, in Section 7, we developed a good approximation to Jones' results based on a crude representation of his model. We conclude this Appendix by demonstrating that this approximation is equivalent to solving the dispersion relation in the large-Δx limit.

By inspection of Eq. (I-23) for Δx large and $y = 0$ (mode 0), we see that only the $u \cdot \ln \{[u + (u^2 + 4)^{1/2}]/2\}$ term is important. Thus, we write

$$\Delta x^2 \simeq u \ln\{[u + (u^2 + 4)^{1/2}]/2\} \simeq u \ln u. \quad \text{(I-27)}$$

For $y = 0$, we recall from (I-22) that $u = \tilde{v}\,\Delta x$. Employing this transformation in (I-27), we obtain the implicit relation

$$\tilde{v} \simeq [\Delta x/(\ln \Delta x + \ln \tilde{v})] \simeq \Delta x/2 \ln \Delta x. \quad \text{(I-28)}$$

In order to compare this result with the crude approximation in Eq. (73), we must make them dimensionally compatible. First, we note that $t_{\text{growth}} \gg t_{\text{travel}}$. (As was pointed out earlier, a single pair of organisms reproducing at the current and precipitous human rate of $\gamma = 0.02$ year^{-1} would require $>10^3$ years to reach the present population of the Earth. Meanwhile, the travel time between stars would be measured in decades.) Thus, (73) can be approximated (in dimensionless units) as

$$v \simeq \gamma \Delta x/\ln(2\gamma/\Psi), \quad \text{(I-29)}$$

where Δx now describes the distance between interstellar habitats. (Here, we have approximated Jones' immigration rate η by our Ψ, although this is only approximately correct.) Expressing (I-28) in dimensional units, we can show that

$$v \simeq \gamma \Delta x/\ln(2m\gamma/\Psi). \quad \text{(I-30)}$$

The discrepancy between (I-29) and (I-30), which is very small, is essentially due to our development of the dispersion relations in a one-dimensional geometry (instead of $m = 3$ dimensions).

In this Appendix, we have mathematically developed a link between the conventional continuum limit of a broad array of nonlinear diffusion processes and their behavior in a regime where the random walk step size is significant. In the limit that the step size is very large, we showed that the velocity of expansion could be simply related to the ratio of travel to growth times. Moreover, the colonization front velocity dependence was observed to be remarkably insensitive to the details of the model.

ACKNOWLEDGMENTS

We are grateful to S. A. Levin of the Division of Ecology and Systematics, H. Kesten of the Department of Mathematics, R. V. E. Lovelace of the Department of Applied Physics, and G. V. Chester of the Department of Physics, all at Cornell University; to C. Alcock, S. Tremaine, F. J. Dyson, and H. Woolf of the Institute for Advanced Study; to R. M. May of the Department of Biology, Princeton University; and, especially, to Steven Soter of the Laboratory for Planetary Studies at Cornell University and Eric M. Jones of Los Alamos Scientific Laboratory for stimulating discussions; and to F. D. Drake, M. H. Hart, B. M. Oliver, and M. Stull for comments. This research was supported by the Planetary Biology Division of the National Aeronautics and Space Administration under Grant NGR 33-010-101 and by NSF Grants AST 75-21153 and PHY 77-20612.

REFERENCES

BALL, J. A. (1973). The zoo hypothesis, *Icarus* **19**, 347–349.

BARENBLATT, G. I. (1952). On some unsteady motions of a liquid and a gas in a porous medium. *Prikl. Mat. Mekh.* **16**, 67–78. (In Russian)

BARENBLATT, G. I., AND ZEL'DOVICH, YA. B. (1972). Self-similar solutions as intermediate asymptotics. *Ann. Rev. Fluid Mech.* **4**, 285–312.

CHANDRASEKHAR, S. (1943). Stochastic problems in physics and astronomy. *Rev. Mod. Phys.* **15**, 1–89.

COALE, A. J. (1974). The history of the human population. *Sci. Amer.* **231**, 12–25.

COLE, L. C. (1954). Population consequences of life history phenomena. *Quart. Rev. Biol.* **29,** 103–137.

COURANT, R., AND FRIEDRICHS, K. O. (1948). *Supersonic Flow and Shock Waves,* Interscience, New York.

COX, L. J. (1976). Correspondence: An explanation for the absence of extraterrestrials on earth. *Quart. J. Roy. Astron. Soc.* **17,** 201–208.

CRANK, J. (1956). *The Mathematics of Diffusion.* Oxford Univ. Press (Clarendon), London.

DAVIS, H. T. (1962). *Introduction to Nonlinear Differential and Integral Equations.* Dover, New York.

DOLE, S. H. (1970). Formation of planetary systems by aggregation: A computer simulation. *Icarus* **13,** 494–508.

DYSON, F. J. (1966). The search for extra-terrestrial technology. In *Perspectives in Modern Physics; Essays in Honor of Hans Bethe on the Occasion of his Sixtieth Birthday.* (R. E. Marshak, Ed.), Interscience, New York.

EISEN, D. (1967). On the numerical solution of $u_t = u_{rr} + (2/r)\, u_r$. *Num. Math.* **10,** 397–409.

FISHER, R. A. (1937). The wave of advance of advantageous genes. *Ann. Eugen.* **7,** 355–369.

GUDERLEY, G. (1942). Starke kugelige und zylindrische Verdichtungsstösse in der Nähe des Kugelmittelpunktes bzw der Zylinderachse. *Luftfahrtforschung* **19,** 302–312.

GURNEY, W. S. C., AND NISBET, R. M. (1975). The regulation of inhomogeneous populations. *J. Theor. Biol.* **52,** 441–457.

GURTIN, M. E., AND MACCAMY, R. C. (1977). On the diffusion of biological populations. *Math. Biosci.* **33,** 35–49.

HADELER, K. P. (1977). Nonlinear Diffusion Equations in Biology. In *Proceedings, Conference on Partial Differential Equations, Dundee, Scotland, 1976.* Lecture Notes in Mathematics No. 564. Springer-Verlag, Heidelberg/Berlin/New York.

HART, M. H. (1975). An explanation for the absence of extraterrestrials on Earth. *Quart. J. Roy. Astron. Soc.* **16,** 128–135.

HART, M. (1978). The evolution of the atmosphere of the Earth. *Icarus* **33,** 23–39.

HART, M. (1979). Habitable zones about main-sequence stars. *Icarus* **37,** 351–357.

HOWARD, W. E. (1960). Innate and environmental dispersal of individual vertebrates. *Amer. Midl. Nat.* **63,** 152–161.

ISAACMAN, R., AND SAGAN, C. (1977). Computer simulations of planetary accretion dynamics: Sensitivity to initial conditions. *Icarus* **31,** 510–533.

JONES, E. M. (1976). Colonization of the galaxy. *Icarus* **28,** 421–422.

JONES, E. M. (1978). Interstellar colonization. *J. Brit. Interplan. Soc.* **31,** 103–107.

JONES, E. M. (1980). Personal communication.

KARLIN, S., AND TAYLOR, H. M. (1975). *A First Course in Stochastic Processes.* Academic Press, New York.

KEVLES, D. (1975). Private communication.

KLASS, P. J. (1974). *UFO's—Explained.* Random House, New York.

KNERR, B. F. (1977). The porous medium equation in one dimension. *Trans. Amer. Math. Soc.* **234,** 381–415.

KOLMOGOROFF, A., PETROVSKY, I., AND PISCOUNOFF, N. (1937). Etude de l'équation de la diffusion avec croissance de la quantité de matrière et son application a un problème biologique. *Bull. Univ. Moskou. Ser. Internat. Sect. A,* **1**(6), 1–25.

LANGER, W. L. (1964). The Black Death. *Sci. Amer.* **210,** 114–121.

LAX, P. D. (1969). Nonlinear partial differential equations and computing. *SIAM Rev.* **11,** 7–19.

LEVIN, S. A. (1976). Population dynamic models in heterogeneous environments. *Annu. Rev. Ecol. Syst.* **7,** 287–310.

MACLEAN, P. D. (1973). *A Triune Concept of the Brain and Behaviour.* Univ. of Toronto Press, Toronto.

MCMURTRIE, R. (1978). Persistence and stability of single-species and prey–predator systems in spatially heterogeneous environments. *Math. Biosci.* **39,** 11–51.

MARUYAMA, T. (1972). The rate of decay of genetic variability in a geographically structured finite population. *Math. Biosci.* **14,** 325–335.

MARUYAMA, T. (1977). *Stochastic Problems in Population Genetics.* Lecture Notes in Biomathematics No. 17. Springer-Verlag, Berlin.

MUSKAT, M. (1937). *The Flow of Homogeneous Fluids Through Porous Media.* McGraw-Hill, New York.

NAGYLAKI, T. (1977). *Selection in One- and Two-Locus Systems,* Lecture Notes in Biomathematics No. 15. Springer-Verlag, New York.

NEEDHAM, J. (1971). *Science and Civilization in China,* Vol. V, Part 2. Cambridge Univ. Press, Cambridge.

NEWMAN, W. I. (1980). Some exact solutions to a nonlinear diffusion problem in population genetics and combustion. *J. Theor. Biol.* **85,** 325–334.

NOBLE, J. V. (1974). Geographical and temporal development of plagues. *Nature* **250,** 726–728.

O'NEILL, G. (1975). Space colonies and energy supply to the Earth. *Science* **190,** 943–947.

O'NEILL, G. (1978). *The High Frontier* Bantam, New York.

PATTLE, R. E. (1959). Diffusion from an instantaneous point source with a concentration-dependent coefficient. *Quart. J. Mech. Appl. Math.* **12,** 407–409.

PHILIP, J. R. (1970). Flow in porous media. *Annu. Rev. Fluid Mech.* **2,** 177–204.

POTTER, J. (1965). The growth of population in America, 1700–1860. In *Population in History* (D. V. Glass and D. E. C. Eversley, Eds.). Arnold, London.

RICHTMEYER, R. D., AND MORTON, K. W. (1967). *Difference Methods for Initial-Value Problems*. Academic Press, New York.

SAGAN, C. (1961). The planet Venus. *Science* **133,** 849–858.

SAGAN, C. (1963). Direct contact among galactic civilizations by relativistic interstellar spaceflight. *Planet. Space Sci.* **11,** 485–498.

SAGAN, C. (1973a). On the detectivity of advanced galactic civilizations. *Icarus* **19,** 350–352.

SAGAN, C. (1973b), Ed. *Communication with Extraterrestrial Intelligence (CETI)*. MIT Press, Cambridge, Mass.

SAGAN, C. (1973c). Planetary engineering on Mars. *Icarus* **20,** 513–514.

SAGAN, C. (1974). An introduction to the problem of cosmic communication. In *Cosmic Communication* (C. Ponnamperuma and A. G. W. Cameron, Eds.), Houghton Mifflin, Boston.

SAGAN, C. (1979). *Broca's Brain* Random House, New York.

SAGAN, C. AND PAGE, T., Eds. (1972). *UFO's: A Scientific Debate* Cornell Univ. Press, Ithaca, N. Y./London.

SCHNEIDER, STEPHEN H. AND THOMPSON, S. L. (1980). Cosmic conclusions from climatic models: Can they be justified? *Icarus* **41,** 456–459.

SEDOV, L. I. (1959). *Similarity and Dimensional Methods in Mechanics* Academic Press, New York.

SHKLOVSKII, I. S. (1976). Could intelligent life in the universe be unique? *Voprosy Filosofii* **95,** 80–93. (In Russian)

SHKLOVSKII, I. S. AND SAGAN, C. (1966). *Intelligent Life in the Universe* Holden–Day, San Francisco.

SKELLAM, J. G. (1951). Random dispersal in theoretical populations. *Biometrika* **38,** 196–218.

STEELE, J. H. (1974). Spatial heterogeneity and population stability. *Nature* **248,** 83.

STORY, R. (1976). *The Space-Gods Revealed: A Close Look at the Theories of Erich von Däniken.* Harper & Row, New York.

TINSLEY, B. A. (1980). Technical development and colonization as factors in the long-term variation in limits to growth. *Cosmic Search* **2,** 10–12.

TODA, M. (1967). Vibration of a chain with nonlinear interaction. *J. Phys. Soc. Japan* **22,** 431–436.

VON HOERNER, SEBASTIAN (1973). Population explosion and interstellar expansion. National Radio Astronomical Observatory, Green Bank, W. Va. Reprint Series B, No. 380. [See also von Hoerner, in Sagan (1973b).]

ZEL'DOVICH, YA. B., AND RAIZER, YU. P. (1967). *Physics of Shock Waves and High-Temperature Hydrodynamic Phenomena,* Vol. II, Chapter X. Academic Press, New York.