

IS DARWINIAN EVOLUTION A MATHEMATICALLY STABLE PROCESS?

Douglas S. Robertson
15 Pavilion Drive
Gaithersburg, MD 20878

Received 13 May 1993, 24 November 1993

ABSTRACT: The answer is no. Darwinian evolution is unstable against the effects of runaway feedback loops. But there is a two-fold reason for the somewhat disingenuous question of the title: First, the tacit assumption of stability is common in the literature, and this assumption often leads to results that are at variance with the real world; second, an understanding of the nature and causes of the instabilities that are inherent in the Darwinian natural selection process may provide important insight into the patterns that are seen in the evolution of species as observed both in the fossil record and in extant ecological systems.

* * * * *

STATIC AND DYNAMIC MODELS

To understand the instabilities inherent in Darwinian processes it is necessary to develop a dynamic theory of the behavior of the natural selection process, a theory that deals explicitly with the causes and effects of time variations in fitness values. A static theory is commonly and often unconsciously used in the analysis of problems in evolutionary biology; it can lead to unnecessary difficulties whose cause may be unclear, even unsuspected. The extent of the differences between a static and a dynamic theory of the natural selection process is not widely understood or appreciated.

The development of this dynamic theory will center on "pure" Darwinism. By this I mean an evolutionary theory that encompasses only random variation plus natural selection. For the moment I will ignore most of the rest of biology, including explicitly such considerations as genetics and developmental constraints. None of this discussion will preclude these considerations, but they will not be specifically addressed here. They should be added later by those who are knowledgeable in the appropriate areas. This "stripped-down" theory contains all that is needed to understand an important component of the chaotic instabilities inherent in evolutionary processes. We will find that pure Darwinian evolution does not generally behave in the manner that is commonly supposed. Chaotic instability is an unavoidable mathematical consequence in Darwinian evolution, and the evidence for such an instability is abundant in the fossil record; many of the puzzles and conundrums contained in that record have simple explanations that involve the unavoidable instabilities of the evolutionary process.

To develop a dynamic theory of Darwinian evolutionary processes, we begin by examining some of the difficulties of a static theory, that is, a theory in which fitness functions are considered to be time-invariant. Of course fitness values cannot be perfectly static. Fitness changes between night and day, between summer and winter, and on many other characteristic timescales. This time-variability can lead to serious theoretical difficulties in the mathematical development of Darwinian theory, difficulties so formidable that it is often necessary to ignore them and assume, as a first approximation, that fitness functions are time-invariant; often this is the only way to make progress with problems that would otherwise be intractable. However, the assumption of time invariance can lead to serious difficulties, if not outright errors, unless it is kept carefully in mind.

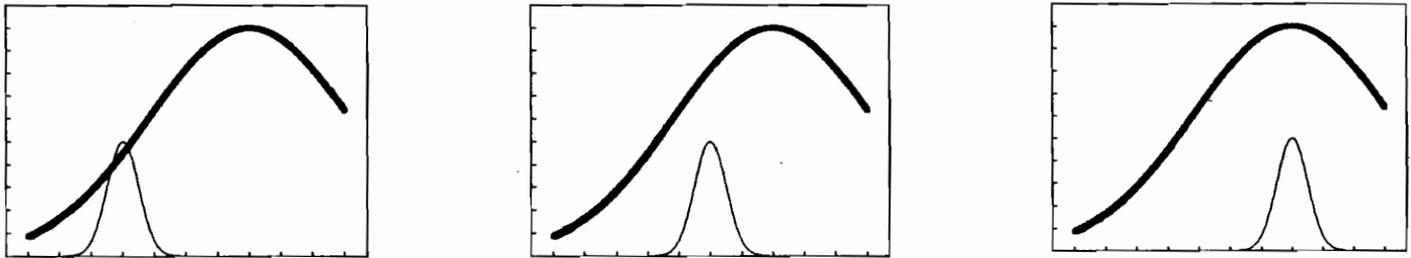
An example of the difficulties created by the assumption of time invariance in fitness functions is found in Sewall Wright's "shifting balance theory." Wright devoted a great deal of effort to devising mechanisms that would allow evolving organisms to cross the "saddle points" that separate the peaks of a complex fitness function [See, for example, Wright, 1965]. The saddle point problem requires the context of a time-invariant fitness function, for if the function were time-varying, then the saddles would not necessarily be permanent features of the fitness "landscape," and crossing them might not entail any permanent difficulty.¹ Wright, in other words, has developed a perfectly correct analysis of a problem which may

¹ Of course Wright often explicitly noted that fitness functions were not time-invariant. Yet he apparently ignored, or failed to appreciate, the significance of this variability for the saddle-crossing problem.

* * * * *

not often occur in nature. Other examples of difficulties introduced by implicit or explicit assumptions of time-invariance are not difficult to find.

The conventional picture of evolution under static or time-invariant fitness functions is illustrated schematically in the following figures:



Figures 1a, 1b, and 1c. These frames illustrate the evolution of a characteristic phenotypic parameter (for example, body size), plotted at three successive points of time. The abscissa represents the phenotypic parameter value, the thick line at the top represents the (time invariant) fitness function, and the thin curve at the bottom represents the distribution of values of that parameter in some population. For all of the figures in this paper the ordinate for the thick curve is fitness, that for the thin curve is population density.

In figures 1a to 1c the fitness function does not change with time; it is the same in all three frames. Meanwhile the average value of the observed parameter distribution, driven by the process of natural selection, is seen to move gradually toward the maximum value of the fitness curve, attaining that maximum at the time shown in the third frame.

This picture is perfectly consistent with Darwinian theory; indeed, there may be some who think that this is the *only* way that Darwinian evolution can operate. As McMenemy put it: "Many paleontologists seem to assume that the adaptive landscape is a fixed and static surface, underneath which plant and animal taxa jockey for access to the highest peaks." [1990, p. 97].

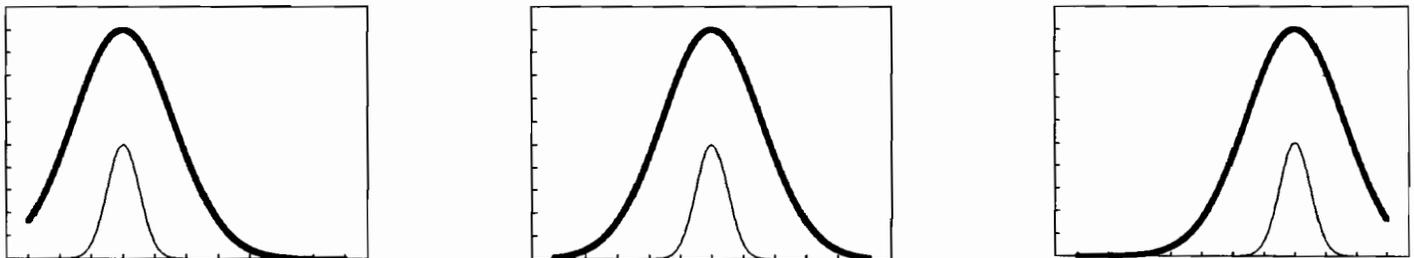
Part of the reason that the process diagrammed in figure 1 is seductive is that it is not entirely wrong, and the basic problem with it is not at all obvious. The problem lies in the timescale required for the process. Although there is no single and universal timescale for the operation of natural selection under all possible conditions, there are several lines of argument that strongly suggest that its characteristic timescale will generally be extremely short, in geological terms. The main argument involves the similarity of natural and artificial selection processes, a similarity that Darwin examined at some length. Artificial selection must operate on timescales that are infinitesimal in geological terms. It is perhaps surprising that many people, who are perfectly aware that *artificial* selection operates on timescales of tens of years, remain comfortable with the idea that *natural* selection requires millions of years.

It might be argued that artificial selection operates on shorter timescales than natural selection because it is more efficient; however, the difference in efficiency, if any, is likely to be small. There are many examples of natural selection processes that can be found to operate on timescales very similar to artificial selection. One example is the classic case of the moth *Biston betularia* that developed a dark-winged variety from light-colored relatives in habitats that had been darkened by industrial pollution [see Kettlewell, 1961]. In this case, natural selection is seen to have caused significant changes in phenotype on a time scale of about a century or two. And Grant describes an example of selection effects operating on Galápagos finches on a time scale of only a few seasons [Grant, 1991]. Thus artificial and natural selection are both seen to produce significant phenotypic changes on timescales that are roughly comparable to human lifetimes.

Is Darwinian Evolution a Mathematically Stable Process?

Of course, none of these artificial or natural processes have produced sufficient change to create a new species. What do we know of the timescales required for speciation? Gould suggests that the characteristic timescale might be of the order of 5,000-10,000 years [Gould, 1991b, p 14]. Mayr cites species of cichlids in Lake Victoria that appear to have evolved in about 5,000 years [Mayr, 1976, p 56]. These values are comfortably larger than the timescales available for artificial selection processes, but are still infinitesimal in geological terms. And even if these values are found to be too small by an order of magnitude there would still be essentially no chance of observing the effects of natural selection on such time scales in the fossil record.

Thus in general the changes in the bottom curve illustrated in figure 1 should happen too fast to have any significant probability of leaving a fossil trace. As Gould noted: "millions of years is too generous, too much for sustained unidirectional change . . . All known and empirically studied cases of gradualism at ecological scales would be completed in a geological twinkling of an eye." [1990, p 7]. Ironically, Darwinian selection may prove to be too powerful an engine to show up directly in the fossil record; looking for its effects might be analogous to trying to see a spinning airplane propeller in a photograph taken with a one-second exposure. On geological timescales, therefore, selection operates so rapidly that its main effect will be to lock the characteristic parameters of phenotypes very closely to the *current* extrema of the fitness functions. This suggests a second model for the operation of natural selection processes depicted in figure 2.



Figures 2a, 2b, and 2c, showing a schematic illustration of the development of a parameter under a time-varying fitness function.

In this model the peak of the fitness function changes with time, and the peak of the distribution of the observed parameter values, also driven by natural selection in this model, closely follows the changing fitness peak. As McKinney put it: "If ceratopsian dinosaurs (or any lineage) became larger it is because the optimum ('adaptive peak') moved along. . . ." [1990, p 104]. Mayr wrote that: "Selectionists have found that the individuals near the mean value of the population in their morphological characters often show higher fitness than individuals with greater deviations from the mean." [Mayr, 1976, pp. 50-51]. This is exactly what would be expected if natural selection generally operates by maintaining populations very close to the *present* local extrema of the fitness function.

This may seem like a small difference from the model shown in figure 1, but it is not; the dynamics of the process are totally different. There are at least two major differences between these two models that need to be thoroughly understood; one is conceptual and the other is practical. The conceptual difference concerns the effects of natural selection. In the case illustrated in figure 1, the effect of natural selection is to improve the fitness of the organisms. This is the common interpretation of selection in Darwinian theory. As stated by Gould: "The assumption of adaptive advantage for traits defining a trend has been at the same time, both the most pervasive assumption of our literature and the most frustrating and refractory to adequate demonstration." [1990, p 20]. But in the case illustrated in figure 2 the population is always at a local maximum of the fitness function; there is no way that selection can improve fitness.

This does not mean that fitness cannot change, but it can only change through a change in the fitness function, and then it can change in any direction, toward increased or decreased fitness, or no change in fitness at all, (as illustrated in figure 2) as the adaptive environment changes. There is an apparent paradox here, for if the second model is the one commonly found in evolutionary systems, then selection, which is commonly supposed to improve fitness values, instead does not generally act even to change them, let alone to increase them. Thus under this model Gould's "pervasive assumption" is generally false; this would account for its being "frustrating and refractory to adequate demonstration."

The practical difference between these two models is that the second model, which contains a time-varying fitness function, has a far richer spectrum of mathematical possibilities than the model whose fitness function is time-invariant. Indeed, the principal problem with the time-invariant model is that it possesses only a very limited range of behavior. The only factors that can be varied easily within the framework of this model are the width and the rate of change of the peak of the lower curve. The range of possibilities available under this model is probably inadequate to explain many of the changes observed in the fossil record. The resulting paucity of behavior patterns may have led some theorists away from ideas of competition at the level of individual organisms toward speculation about a variety of hypotheses, including orthogenesis and selection at the group or species level.

Deficiencies of the time-invariant model cannot be rectified by simple changes, such as assuming a more complicated shape for the constant fitness function. The only change that might be of any great consequence would involve the presence of multiple extrema instead of the single extremum diagrammed in figure 1. But if a constant fitness function has more than one extremum then a serious problem develops immediately. Natural selection cannot easily bridge the gaps between the extrema of a constant fitness function, whence the literature on peak shifts.

In contrast, the properties of dynamic or time-variable fitness functions can lead directly to a highly complex range of behavior. Although the model diagrammed in figure 2 was constructed so as to produce a set of variations very similar to those seen in figure 1, these changes do not begin to exhaust the range of possibilities available with that type of model. The range includes such things as systematic changes in the location of the extremum of the fitness function, changes in its height (sometimes to the point of its vanishing altogether), and even bifurcation of the extremum. All of these possibilities and more are likely to be needed to explain the variations in phenotypes observed in the fossil record.

If most of the changes that are observed in the fossil record are therefore produced by selection acting to lock parameters closely to changing extrema of fitness functions rather than by the direct action of selection under constant fitness functions, then selection may be only indirectly responsible for most of the variations seen in the fossil record. This suggestion is in sharp contrast to the conventional picture of Darwinian evolution, in which natural selection is regarded as the primary non-random agent in the development of species. Simpson commented: "selection . . . is *the only nonrandom or antichance evolutionary factor that is objectively demonstrated to exist.*" [Simpson, 1952, p 267 (Simpson's italics)]. Similarly, Gould says: ". . . the cardinal tenet [of strict Darwinism] is that natural selection is *the creative force in evolution.*" [Gould, 1983, p 138 (Gould's italics)]. Instead, systematic time variations in the extrema of the fitness function must be the primary non-random or "creative" force in evolution on geological timescales. A full understanding of these systematic variations is the goal (perhaps overly ambitious) of the development of a dynamic theory of Darwinian processes.

DYNAMICS OF THE FITNESS FUNCTION

In the development of such a dynamic theory the primary stumbling block lies in the extreme difficulty of determining the shape and characteristics of the fitness function. Fitness is not a parameter that is easy to measure, like voltage or barometric pressure. It is a subtle concept involving the various intertwined probabilities of individual survival and the production and survival of progeny. Measuring or even defining the fitness function is a difficult-to-impossible problem even in currently extant ecosystems that can be studied in great detail, and it is even more difficult in fossil ecosystems for which far less information is available. The difficulties of defining and measuring the fitness function are almost certainly among the

Is Darwinian Evolution a Mathematically Stable Process?

principal reasons that its time variability has not been thoroughly studied to date. However, the fact that the function is difficult to measure does not lessen its importance.

Progress in understanding the dynamics of fitness functions may depend on developing a good working definition for them. Following Van Valen [1976] and Sober [1984, p 208], among others, we might *define* fitness to be the function that is maximized by the process of natural selection. Selection itself is not too difficult to understand heuristically. Some organisms survive and produce abundant progeny similar to themselves, and others do not; over time, the types that survive and produce progeny will tend to predominate. The definition of fitness in terms of the action of the selection process may appear tautological, and indeed tautology has often been used as a criticism of Darwinian theory (“survival of the survivors”). Yet tautology is a criticism that could be leveled at most of mathematics; it does not bother mathematicians much, and neither should it be of great concern in the biological community. *All* definitions involve tautology--this is inherent in the definition of “definition.” The question, therefore, is not whether the definition is a tautology, but whether it leads to important new insights. I believe that this definition can lead to an improved understanding of the time variations of fitness functions through an analysis of the forces involved in the natural selection process.

These forces fall into two basic classes: physical processes of non-biological origin and biotic processes that are produced by interactions among organisms. The physical processes include such things as climate change, sea level fluctuations, and diurnal and seasonal variations. Many of these physical variations, especially the diurnal and seasonal fluctuations, are periodic or quasi-periodic on a variety of time scales, and will not generally introduce chaotic effects into evolutionary processes (although it is possible to produce chaos with periodic driving forces). Other physical processes, such as climate fluctuations, volcanic explosions, and asteroid impacts, are inherently irregular and may introduce a strong chaotic component into Darwinian processes. Despite its overall importance as a source of instability in evolution, this is not the class of chaotic instabilities that I am interested in analyzing here, because the chaos is not inherent in the evolutionary process itself.

The chaotic instability that *is* inherent in the internal mathematical structure of Darwinian processes results from the second or biotic class of forces. To understand the general properties of such biotic forces we should begin by addressing a major error of omission in the discussion up to this point. It is a simple error, all too common in the literature: The necessary interactions between the curves shown in figures 1 and 2 have been neglected. In particular I have assumed that arbitrary changes in distributions of populations shown in the lower curves in figures 1 and 2 do not affect fitness values shown in the upper curve. But this is obvious nonsense. Fitness is a direct and strong function of the distribution of extant populations because organisms are a critical component of their own selective environment [see Robertson, 1991, p 470]. For example, a population of cheetahs cannot be added to an ecosystem without affecting the fitness values of gazelles. More precisely for the case shown in figures 1 and 2, adding a population of cheetahs will affect the fitness of an individual cheetah. The dynamics of the interaction between populations and fitness can be expressed in two deliberately symmetric statements:

1. Changes in fitness functions can cause changes in the distributions of phenotypes.
2. Changes in the distributions of phenotypes can cause changes in fitness functions.

The first statement is merely an assertion of natural selection. The second is a direct consequence of the effect of organisms on their own selective environment. Darwin understood the importance of this biological component of the selective environment. He was unaware of the chaotic mathematical consequences of this fact, at least partly because the theories of both feedback and chaos were not developed until many decades after his lifetime.

It is important to realize that neither of these two statements alone causes any problem; it is only when they are taken together that they imply a significant instability in Darwinian processes. These two statements allow (and probably require) the existence of feedback loops, which are then capable of producing both hyper-stability and chaotic instability, depending only on whether the feedback is positive

ROBERTSON

or negative. The existence of feedback in evolving systems implies chaotic instability [see the discussion in Robertson, 1991, p 482-483]. And the theory of feedback processes provides the conceptual basis upon which a dynamic theory of the natural selection process can be founded.

A few simple principles are needed to help understand these feedback effects. Although these ideas do not begin to exhaust the impact on fitness functions of changes in distributions of phenotypes, they are a useful beginning:

1. There is no advantage to being exactly average, to being in the middle of a crowd.
2. There is commonly an advantage to being slightly bigger, faster, stronger, taller, better armed and armored, etc., than the average phenotype of a given population.

Using these basic ideas, we can develop a third model for the operation of natural selection, following the discussion in Robertson [1991]. The model is illustrated in figure 3. The first frame shows a distribution of organisms that is locked to the local extremum of a fitness function, as in figure 2. In the middle frame we add the effects of the advantage that results from small parameter increases, increases in size, for example. The extra fitness advantage is illustrated by the thick curve near the bottom of the frame. Its peak, controlled by the population distribution, must be to the right of the peak of population distribution. This fitness component is added to the fitness values from the previous frame to calculate the total fitness shown in the thick curve above. Selection will still shift the population distribution toward the peak of the total fitness function, but that peak will move with the changing population distribution, as shown in the third frame.

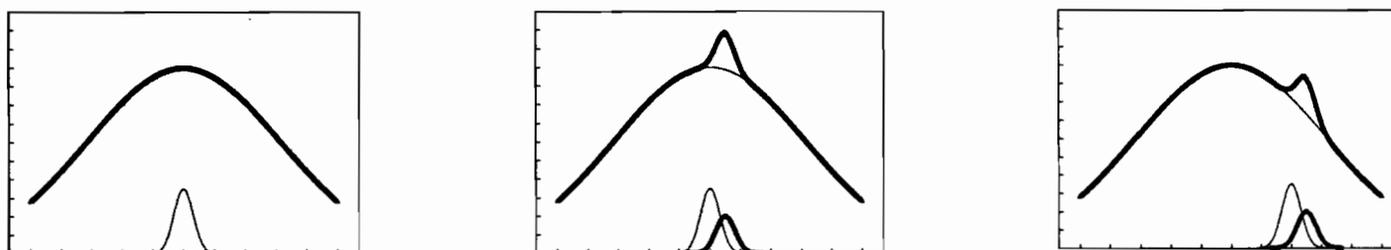


Figure 3. Similar to figure 2, illustrating the effect of the dependence of fitness on the distribution of phenotypes in a population. Where the fitness curve at the top has been significantly modified, the original fitness values are indicated for clarity by a thin line.

With these simple considerations, we have developed a model which is inherently unstable, in contrast to the stable model of figure 2. In this model natural selection will try to lock the population distribution closely to the local extremum of the fitness function, but it can never do so because the location of that extremum is controlled by the population distribution. The very action of selection, moving the population distribution, will shift the peak of the fitness function, which will cause a further shift in the population distribution; a classic runaway feedback loop. Analytically there is no reason the feedback loop could not continue indefinitely. In practice it may end when the limits of physical possibility are reached, or it may end when further motion is impeded by the presence (in the corresponding region in morphospace) of another species that successfully occupies a “blocking” ecological niche. (An ecological niche is a synonym for a local extremum of the fitness function.)

If this mechanism is pervasive in systems that are evolving by Darwinian natural selection (as is suggested by the generality of the assumptions that generate it) then the overall effect of selection will be to decrease the fitness of populations, as in figure 3. This simple fact opens up a new line of explanation

Is Darwinian Evolution a Mathematically Stable Process?

for a wide variety of puzzling phenomena observed both in the fossil record and in extant ecosystems (see Robertson [1991]), including Cope's rule of phyletic size increase, neoteny, periodic extinctions, even speciation and punctuated equilibrium. Note also that although the fitness of a population is routinely decreased with this model, that population must remain in the close vicinity of a fitness extremum, or ecological niche. So organisms that are driven to low fitness values will remain very well adapted to their local niche, even as they are driven toward extinction.

Perhaps the best evidence for the pervasiveness of this model in evolutionary processes is found in the fact that, for many generalized body plans, organisms can be found that are close to the physical limits that are possible for that body plan. Examples are found in organisms as diverse as condors, flying foxes, pterosaurs, sauropods, ground sloths, giraffes, sequoias, whales, perhaps salps and giant squids. The existence of these organisms could be explained conventionally by assuming that fitness functions just happen to have peaks close to the appropriate physical limits for a wide variety of phenotypes; however there are at least two problems with this explanation. First, it requires a certain amount of special pleading, explicit assumptions about the form of fitness functions for an extraordinarily wide range of organisms. More seriously, it runs afoul of the observed fact that during periods of mass extinction (such as at the end of the Cretaceous, or the end of the Pleistocene), these (large) organisms that push the limits of physical possibility are preferentially destroyed. Again, with additional special pleading, it could be argued that all of the special fitness peaks responsible for these organisms just happen to have low fitness values or that for some reason all of these organisms are unable to adapt to changed conditions. But under the action of runaway feedback loops as diagrammed in figure 3, natural selection should be expected to drive phenotypes to the edges of physical possibility and at the same time reduce the population fitness values, possibly to the level that extinction becomes probable.

One of the few areas of evolutionary theory in which runaway feedback is often considered as a significant mechanism in evolution concerns runaway sexual selection (see, *e.g.*, O'Donald, 1980; Lande, 1981; 1982). However, the decreases in fitness found under these runaway feedback processes are commonly explained in terms of a trade-off of improved reproductive success compensating for decreased survival probability. In contrast the full theory of feedback allows (and almost requires) general reductions in fitness under pure natural selection without any need for compensating trade-offs.

CHAOS THEORY AND DARWINIAN EVOLUTION

As noted in Robertson [1991], the presence of feedback loops in a complex system will generally produce chaos. Feedback is not the only mathematical process that will produce chaos: Almost any non-linear process will do the trick. Chaos is almost impossible to avoid in any system that is even slightly more complicated than the two-body problem in Newtonian mechanics. But the intimate connection between feedback, evolution, and chaos theory can be illustrated by examining the following equations [Stewart, 1989, pp 20, 155, 321, and 200]:

$$x_{n+1} = kx_n^2 - 1 \quad (1)$$

$$x_{n+1} = kx_n(1-x_n) \quad (2)$$

$$x_{n+1} = kx_n(1-x_n^2) \quad (3)$$

$$x_{n+1} = k \sin(x_n) \quad (4)$$

These recursive equations exhibit the general form that often occurs in analysis of feedback (generically, the next state of the system is some non-linear function of the previous state) and also the form that would occur in the analysis of evolutionary systems (similarly, the next generation (or ecosystem) is some non-linear function of the previous generation (ecosystem)). But these equations were not borrowed from works on feedback or evolution--they were taken from a discussion of chaos theory. All of them exhibit chaotic behavior for certain values of k and x_n .

I do not mean to argue that these are exactly the equations that would describe any particular evolutionary process. Such equations would be extremely complex and difficult (essentially impossible) to

derive; even for the simplified model of figure 3 they would require detailed and realistic information about the curves shown in figure 3 and the variation process. But the curious thing about chaotic behavior is that the exact analytic form of the equation has surprisingly little effect on the behavior of the iterative calculation. The analytic forms of equations 1-4 are very different, yet they all show very similar behavior. The only important property is the presence or absence of a local extremum in the functional form. If such an extremum is not present, then the iteration process will be unbounded, and will tend toward infinity (not a very useful or interesting result). However, in the presence of a local extremum, where the results are bounded, essentially any iterative equation of this form will show a curious interleaving of regular and chaotic behavior. Thus on very general grounds we can expect feedback to produce chaotic behavior in evolutionary systems, no matter what form of equation the analysis of the underlying biological system requires [see Stewart, 1989, p 197-208]

In many common situations, such as weather and climate forecasting, chaos theory denies or sharply restricts predictability. Nevertheless, chaos theory is an indispensable component of the theorist's arsenal, because so many processes in the real world are genuinely chaotic, or at least a good approximation of chaos. If the predictive properties of theories are thereby compromised or eliminated, then we will probably have no choice but to change the objectives of the theory process. As Stewart put it, in discussing the limitations of chaos theory: "If you can't win, move the goalposts." [1989, p 286]. What may seem like a backward step is often the only way to make progress. In the context of a somewhat different problem in chaos theory, Stewart noted:

We started out not understanding [the chaotic motion of Saturn's satellite] Hyperion. Now we do not even understand $2x^2 - 1$. In mathematical terms that constitutes stunning progress. [1989, p. 21]

The difficulties of dealing with chaotic processes have been known for a long time. Euler encountered them over two centuries ago, commenting that "If it is not permitted to us to penetrate to a complete knowledge concerning the motions of fluids, it is not to mechanics, or to the insufficiency of the known principles of motion that we must attribute the cause. It is analysis itself which abandons us here." [quoted in Stewart, 1989, p 38.]

If evolutionary processes are genuinely chaotic, then we must recognize that radical changes may be needed in our approach to many problems. Pious hopes for understanding the details of evolutionary processes may have to give way to more general and less detailed principles. Rather than trying to predict the exact course of evolutionary processes, which may be even more futile than long-range weather forecasting, we should try instead to understand the nature of the underlying chaotic processes, and try to understand the patterns that should be expected to result from these processes. There are at least four important principles of chaotic behavior that should be kept in mind:

1. The butterfly effect--chaotic processes often show extreme sensitivity to small perturbations, as suggested by Robert May's classic metaphor: "The flapping of a butterfly wing in Asia can cause a hurricane in the Bahamas."
2. Flip-flop or bifurcation effects--chaotic systems commonly switch rapidly between two or more distinct modes of behavior, with no discernable cause. The cause is real, but it may be buried in some tenth or hundredth decimal place where it would be impossible to ferret out.
3. Rates and averages may not be locally defined--in chaotic, fractal, systems, the behavior of curves may not be "smooth" enough for derivatives ("rates") to have any meaning (although long-term rates may be bounded). Theorists often speak of "rates" of evolution and even "average rates" without considering whether the concepts have any meaning in the context of evolutionary biology. They may be asking the wrong questions. They should be asking instead if the process is stable enough that the word rate has any meaning. I think the answer to that question will generally be "no," although that

may be difficult to prove in specific instances. (Bookstein [1987] made similar comments in the context of random walk processes rather than fractal processes.)

4. Even punctuated equilibrium may turn out to be just another natural mode of chaotic behavior. In discussing a study of chaos in the orbits of planets in the solar system, one commentator noted: "A curious feature of the chaos emerged. Orbit changes tend to occur in small spurts of a few percent, rather than in a smoothly exponential manner." [Anonymous, 1993]

DISCUSSION

There are a number of specific examples that may illustrate some of the insights gained from the idea that evolutionary processes are dominated by the unstable, chaotic effects of runaway feedback loops. Robertson [1991] described examples of such effects including Cope's rule of phyletic size increase, neoteny, punctuated equilibrium, orthogenesis, and periodic extinctions. We can add here a number of other cases that illustrate the utility of these concepts. All of these cases are susceptible to explanation by conventional evolutionary theory without recourse to feedback. However, most of the conventional explanations resort to special pleading or ad-hoc arguments. Feedback provides a simple, natural and unavoidable framework that effortlessly provides simple and obvious explanations for all of them, without recourse to any ad-hoc special conditions.

Consider, for example, the "butterfly effect" of extreme sensitivity to small perturbations of the system. In ordinary mathematical chaos theory this sensitivity can be tested by computer simulation. If insignificant changes in initial conditions produce exponentially growing differences in the result, then the process is said to be chaotic. In biological systems we do not have the luxury of running easy computer simulations, but the patterns of diversity seen in present ecosystems strongly suggest that evolution possesses such chaotic sensitivity. The principal evidence for this is seen in the divergence of species following a geographic separation, such as is produced by continental drift. Mutually isolated ecosystems do not converge to the same "solution." There are no kangaroos native to South America, for example. Gould describes a number of other examples of divergent evolution and captures the essence of the matter with the apt metaphor of the history of life as a tape recording, which, if replayed, does not produce the same result twice [Gould, 1989, p 293-323].

The "flip-flop" or bifurcation type of instability is also commonly seen in the fossil record, most prominently at the base of the Cambrian, but also after a number of later mass extinction events. The fossil evidence is perfectly consistent with a tendency of ecosystems to shift quickly between different stable states, as is typical for chaotic systems.

Next consider the concept of "overspecialization," *i.e.*, the fact that organisms too highly specialized for a particular ecological niche are often more vulnerable to extinction than their less specialized cousins. The problem with this phenomenon is that in the context of stable, time-invariant fitness functions the most specialized organisms should be the *least* likely candidates for extinction rather than the most likely, because they should be closer to the maximum of the fitness function. In other words, the problem is that the best-adapted organisms often seem to be the ones that are particularly prone to failure. Of course, there would be occasional cases in which one might expect such overspecialized types to fail from sheer bad luck. However, when bad luck becomes the rule, as in this case, one should immediately suspect that something besides random luck is involved.

There is even more of a paradox here, for not all overspecialized organisms fail. There are many cases in which such groups are successful; sometimes they even represent the dominant groups at various time periods. Reptiles and mammals, for example, could be considered to be overspecialized varieties of fish.

In the context of unstable evolutionary processes driven by runaway feedback loops the explanation for these apparent paradoxes is clear. Under the action of runaway feedback loops, fitness will generally decrease, rather than increase, even though it is driven by natural selection. Overspecialization leads directly to extinction in the cases in which a runaway feedback loop has driven a species so far from its underlying fitness maximum that extinction has become fairly likely [see Robertson, 1991, figures 1-4]. The remaining cases in which overspecialization has led to the development of an organism that has flourished

and become dominant, often in a novel environment, are examples of situations in which the underlying fitness function has multiple extrema. Here the runaway feedback loop has pushed the species across a "saddle" or local minimum in the fitness function, toward a possibly higher maximum that may have been unattainable without the runaway feedback effects. Thus not only do *random* shifts in the adaptive landscape essentially eliminate the problem associated with such saddle crossings (as noted earlier), but, as we see from figure 3, runaway feedback effects are capable of producing systematic changes in that landscape that would drive species directly across such saddles by pure Darwinian selection.

Mayr describes four different types of evolution which he labels progressive evolution, maintenance evolution, switch evolution, and speciation [Mayr, 1976, p 58]. Unstable evolutionary processes may be needed to understand at least the first three of these categories. By progressive evolution, Mayr means evolution "toward a condition of ever-greater perfection," the behavior illustrated in figure 1 of this article. He might be surprised to find that many of these cases that are conventionally thought to illustrate this effect actually represent the situation seen in figure 2, in which organisms remain close to time-variable peaks in the fitness function, and the overall fitness of the organisms may either increase, decrease, or remain constant, as illustrated. Mayr's second type, maintenance evolution, is simply the case in which the extrema of the fitness function are static in time, and natural selection locks the organisms close to the static peaks. What Mayr terms "switch evolution" occurs when "an evolutionary line invades a new niche . . . [which] may turn out to be the entrance to a whole new adaptive universe." This is simply the case in which the fitness curve has multiple extrema, and moving local extrema of the fitness curve, driven by feedback effects, have forced species across the "saddle point" between these underlying extrema.

Mayr's fourth group involves speciation, *i.e.*, the tendency of organisms to "clump" into distinct and discrete species rather than to approach the (single) global maximum of a time-invariant fitness function. The interpretation of this phenomenon may lie in the tendency of feedback loops to produce bifurcations of extrema of fitness functions, and to produce moving local extrema, as in the "dripping faucet" model (see Robertson, 1991, pp. 473-475). If competition between feedback loops is able to freeze those extrema into a relatively long-term stasis then the species could remain at the locations of those extrema for extended periods. The instabilities inherent in positive feedback loops, combined with the stasis that can be produced by competing feedback loops, lead naturally to the idea that organisms will clump into separate and discrete species in the manner that is clearly observed in both extant and fossil organisms.

The next example of an insight provided by the concept of unstable evolutionary processes involves the fact that evolution generally requires species to have short lifespans. Evolution would not work if either individuals or species were immortal [see, *e.g.*, Raup, 1991, pp. 16-20]. Although there is a general mechanism, albeit poorly understood, for eliminating individual organisms after a finite lifespan, there is no corresponding mechanism other than random bad luck for removing species under the assumption of constant fitness functions. But the instabilities resulting from fitness variations caused by runaway feedback loops provide a natural mechanism for regularly driving species toward extinction. In the overall scheme of evolution, this mechanism for eliminating species may prove to be as important as the finite lifespan of individuals.

The literature of the mathematical theory of natural selection is replete with examples of assumptions of constant fitness values; the analysis is fine, but the assumptions render the conclusions questionable. For example, in Haldane's discussion of the fitness of gene pairs, constant fitness values are assigned to the four genotypes under consideration [Haldane, 1932, p 108]. In a more realistic model that entails feedback effects--the idea that organisms are themselves significant components of the environment to which they are adapting--the fitness values should be functions of the distribution of alleles, as:

$$\begin{aligned} AB &= f_1(x,y) \\ aaB &= f_2(x,y) \end{aligned}$$

etc. But this immediately leads to two difficult analytic problems. First, it is not obvious what the form of these unspecified functions should be. Second, even under the simplest assumptions, such as bivariate

Is Darwinian Evolution a Mathematically Stable Process?

Gaussian functions, the analysis becomes intractable. Thus the reason for the simplifying assumption is apparent--it allows further progress with the analysis. Unfortunately, by denying feedback effects and any time-variability in the fitness values, Haldane renders the application of his analysis to situations in the real world somewhat dubious.

The final case in which the concept of unstable evolutionary processes provides important new insights is different from the others in that it does not involve evolutionary theory *per se*. Instead, it involves the perversion and misuse of that theory for malicious or malevolent social purposes. Gould notes that “. . . Darwinism had been widely portrayed as a defense of war, domination, and domestic exploitation . . .” [Gould, 1991a, p 428], particularly, though not exclusively, by members of the German high command during the first World War, who argued that Darwinism justified all of the horrors of the war. Gould presents several good reasons why these arguments are incorrect, particularly because moral lessons cannot be drawn from the functioning of the natural world. In this I concur completely. But perhaps a better response would be to show that the argument represents a serious misapprehension of the chaotic nature of the behavior of Darwinian selection, particularly of its tendency to drive species or societies directly toward extinction. Sadly, competition among human societies is subject to the same type of runaway feedback loops and chaotic instabilities as the similar competition among the “lower” animals. Thus in the first World War the Imperial German high command was engaged in a runaway competitive feedback loop that was capable of driving all of the participants *away* from the peak of their fitness curve and directly toward extinction, instead of improving their fitness as they thought. This might have been obvious to them if they had had a better understanding of the instabilities inherent in Darwinian theory.

These examples and the others cited above demonstrate the some of the insights that can be developed from the simple idea that Darwinian evolution is inherently unstable against the effects of runaway feedback loops. Evolution will not be completely understood until we fully understand the dynamics of fitness functions, which will require an understanding of the feedback that results from the effects of organisms on their own adaptive environment (in addition to other effects). Without a detailed knowledge of the shape and time variations of the function that is optimized by the process of natural selection, we will not understand the development of species. Natural selection is only half of the picture, and possibly the less important half. To use a nautical or avionic metaphor, selection is the propeller that drives evolution, but it is not the rudder. That is, selection is the critical driving force, but it is not the factor that controls the direction of evolution. The biologically and mathematically interesting portion of evolution problems often lies in the dynamics of the fitness function.

This conclusion is in no sense anti-Darwinian, but rather an essential extension of Darwin's work. Darwin assumed the existence of fitness functions, but he made little attempt to deal with their dynamics and time variations. The omission does not reflect any fault on Darwin's part; many of the theoretical concepts and mathematical structures needed to approach the problem, including feedback theory and chaos theory, were not developed until many decades after Darwin's death. We have done no more than begin to work on this difficult problem. The full resolution of these questions and the development of reliable techniques to observe and measure changes in fitness functions are matters that are likely to require a great deal of both theoretical and experimental work. Today, when many of the theoretical tools needed to analyze the problem are reasonably mature and ripe for application to biological problems, the time appears to be right to begin a thorough investigation of these problems.

REFERENCES

- Anonymous, (1993) in *Sky and Telescope*, **85**, #2, 14.
 Bookstein, F.L., (1987), Random Walk and Evolutionary Rates, *Paleobiology*: **13**: 446-447.
 Gould, S.J., (1983) *Hen's Teeth and Horses Toes*, New York: W.W. Norton.
 Gould, S.J., (1989) *Wonderful Life*, New York: W.W. Norton.
 Gould, S.J., (1990) Speciation and Sorting as the Source of Evolutionary Trends, or 'Things Are Seldom What They Seem,' in *Evolutionary Trends*, K.J McNamara, ed., U. of Arizona Press, Tucson.
 Gould, S.J., (1991a) *Bully for Brontosaurus*, New York: W.W. Norton.

ROBERTSON

- Gould, S.J., (1991b, August) Opus 200, *Natural History*.
- Grant, P.R., (1991) Natural Selection and Darwin's Finches, *Scientific American*, **265**, #4, 82-87.
- Haldane, J.B.S., *The Causes of Evolution*, Longmans, Green & Co, 1932; reprinted by Princeton U. Press, 1990.
- Kettlewell, H.B.D., (1961) The Phenomenon of Industrial Melanism in Lepidoptera, *Annual Review of Entomology*, **6**: 245-262.
- Lande, R., (1981) Models of Speciation by Sexual Selection on Polygenic Traits, *Proc. Natl. Acad. Sci.*, **78**: 3721-5.
- Lande, R., (1982) Rapid Origin of Sexual Isolation, and Character Divergence in a Cline, *Evolution*, **36**: 213-223.
- Mayr, E., (1976) *Evolution and the Diversity of Life*, Belknap Press, Cambridge, MA.
- Mayr, E., (1988) *Toward a New Philosophy of Biology*, Cambridge, MA: Harvard University Press.
- McKinney, M.L., (1990) Trends in Body Size Evolution, in *Evolutionary Trends*, K.J McNamara, ed., U. of Arizona Press, Tucson.
- McMenamin, M., (1990) The Cambrian Explosion, *Palaios*, **5**, # 2, 97.
- O'Donald, P., (1980) *Genetic Models of Sexual Selection*, Cambridge U. Press, New York.
- Raup, D.M., (1991) *Extinction--Bad Genes or Bad Luck*, W.W. Norton, New York.
- Robertson, D.S., (1991) Feedback Theory and Darwinian Evolution, *Journal of Theoretical Biology*, **152/4**: 469-484.
- Simpson, G.G., (1953) *The Major Features of Evolution*, Columbia, New York.
- Sober E., (1984) *The Nature of Selection: Evolutionary Theory in Philosophical Focus*, MIT Press, Cambridge, Mass.
- Stewart, I., (1989) *Does God Play Dice? The Mathematics of Chaos*, Basil Blackwell, Cambridge, MA.
- Van Valen, L, (1976) Energy and Evolution, *Evolutionary Theory*, **1**: 179-229.
- Wright, S., (1965) Factor Interaction and Linkage in Evolution. *Proc. Roy. Soc. Lond. B*, **162**: 80-104.