Biology and Philosophy **14:** 83–101, 1999.

Causality in Complex Systems

ANDREAS WAGNER
University of New Mexico
Department of Biology
167A Castetter Hall
University of New Mexico
Albuquerque, NM 87131-1091
U.S.A.

Abstract. Systems involving many interacting variables are at the heart of the natural and social sciences. Causal language is pervasive in the analysis of such systems, especially when insight into their behavior is translated into policy decisions. This is exemplified by economics, but to an increasing extent also by biology, due to the advent of sophisticated tools to identify the genetic basis of many diseases. It is argued here that a regularity notion of causality can only be meaningfully defined for systems with linear interactions among their variables. For the vastly more important class of nonlinear systems, no such notion is likely to exist. This thesis is developed with examples of dynamical systems taken mostly from mathematical biology. It is discussed with particular reference to the problem of causal inference in complex genetic systems, systems for which often only statistical characterizations exist.

Key words: ceteris paribus, epistasis, nonlinear gene interactions, polygenic diseases, regular causes

When Russell (1913) argued that the notion of causation should be abandoned, he reasoned that it should be replaced by that of functional relations among the state variables of a system. He also argued that the notion was not really necessary, since the physical sciences were doing well without it. Aside from the latter assertion, which is disputed by Suppes (1971), it is clear that the solution can not be that simple. As Hume already noted, causation is a notion fundamental to human cognition, so fundamental that it is unlikely to ever be eradicated. This is obvious from its pervasiveness in everyday language, but also from the prominence of debates centered around the notion of causality in the philosophy of science in this century. Curiously, however, the attention paid to this issue varies considerably across fields. In economics, and in the social sciences in general, intense debates on the nature of causality have been going on for decades. In these areas, the main issues regard the possible inference of causal relations from statistical data. Biology, on the other hand, seems virtually untouched by such debates. For example in the ten year period from 1985 to 1995, merely one article pertaining

to the issues considered here appeared in the journal “Biology and Philosophy” (Francis 1990). Clearly, questions regarding the notion of causality lurk behind most contemporary topics in biology, but rarely are they made explicit. There is ample room for speculation regarding the possible reasons for such differences. A prominent candidate is the relevance of the social sciences and especially of economics to policy decisions. For policy issues the use of causal language, such as “investment causes increased production of capital goods”, is central. The debates then result from attempts to clarify the terminology, and from the manifold ambiguities in identifying causes and their effects. Examples for such ambiguities are abundant, and for the purpose of illustration one shall be mentioned. It is described by Rossi et al. (1980), and discussed further by Glymour et al. (1987). It concerns a study of newly released felons in Georgia and Texas who received unemployment benefits for six months after their release. They were prevented from working during that period. It was found that the rearrest rates were nearly the same between the study groups and control groups that did not receive any payments. Two qualitatively different conclusions were made based on the statistical data. One stated that payments reduced recidivism, but that unemployment increased it, and that the effects canceled in the study group. The other conclusion stated that payments had no effect on recidivism. Clearly, any policy decisions based on these two conclusions would be very different. Identifying “correct” causal relations among economic variables, and eliminating “spurious” correlations from statistical data is therefore an activity with potential impact on the lives of many individuals. The dire need for some kind of factual support of economic policy is also obvious from the nature of many economic theories. They combine reliance on plausibility assumptions that are difficult to validate (e.g., “rational expectations”) with a sparse data support to an extent that would place them in the realm of mere speculation, say, in the physical sciences. Still, policy decisions have to be made, and theories with poor explanatory or predictive power are better than no theories. Debates centering around causation are far from academic in these areas. They may affect the lives of people in obvious ways, which might contribute to their continuing importance.

The situation in the life sciences has long been different, in that most research may have been of less immediate practical applicability. However, this is about to change, and the enormous success of biomedical research is the reason for this change. Diseases with a hereditary component have been known for a long time, and the influence of genetic factors on complex traits such as “aggression” or “intelligence” was postulated and debated long before the biochemical nature of genetic information was known. Since the identification of DNA as the carrier of genetic information, powerful tools

have been developed to analyze and manipulate this information. This has not only lead to a new dimension of insight into the mechanistic basis of many biological processes, it has also lead to a gold rush for applications of this knowledge. Disease "causing" genes or disease susceptibility genes are found at an ever increasing pace. Examples include genes for susceptibility to breast cancer, muscular dystrophy, obesity, and many more (e.g., Lander and Schork 1994). Investigations in this area have not stopped short of trying to elucidate the genetic basis of fairly complex conditions, such as schizophrenia. In these cases, the condition itself can often not be identified unambiguously. The question to what extent complex genetic factors "cause" a disease becomes a very practical one, if one considers its impact on the health care system. Should carriers of susceptibility genes pay higher health insurance premiums or not receive insurance at all? How about access to employment of individuals who carry susceptibility markers for certain noxious environmental influences? Such questions are becoming prominent social policy issues (e.g., Wadman 1996). Whatever the outcome of ongoing debates in this area, societal changes affecting the lives of large parts of the population are likely. At the root of the issue is the notion of the disease causing gene, and for certain "monogenic" diseases this notion may be unproblematic. However, as technology advances it becomes clear that many diseases are influenced by multiple genetic and non-genetic factors. Polygenic diseases are the rule rather than the exception. This is not surprising if one considers that most phenotypic traits emerge from the action of large numbers of genes, embedded in metabolic pathways and regulatory networks of enormous complexity. Embryonic development and "housekeeping" of multicellular organisms involves the interaction of an immense number of individual gene products. Many of these gene products and qualitative features of their interactions are currently being identified. However, the state of the art in this area is a far cry from a quantitative understanding of biochemical and regulatory pathways. Clearly, any of the genes in such pathways, or many of the genes acting together might bring about a given change. What does it mean then that one gene becomes associated with a disease? How often are such associations spurious or context-dependent, i.e., dependent on the genetic background? And the central question, is it possible to speak of causality meaningfully in systems involving that many factors?

The practical importance of these problems alone would be sufficient reason to pay more attention to issues of causality. However, what has been said so far is only one possible motivation for what is to follow. Various complementary motivations for taking a fresh look at causality could be given. Examples include ecology, where a major issue is the extent to which human influence may cause ecosystem instability, or neurobiology, where

the relation of neural activity to consciousness is now seemingly becoming amenable to experimental analysis (Crick 1996). This much broader arena justifies a decoupling of a discussion of causality from the above issues. Here, the discussion will be cast in a general mathematical framework which transcends even the broader biological arena.

One of the major developments in the history of biology has been a shift from descriptions of natural history to an attempted "physicalistic" explanation of biological processes, i.e., from properties of the constituents of biological systems. An important aspect of this development has been the increased usage of mathematical formalism, and the concurrent development of mathematical methods suitable to address biological problems. This has brought some areas of biology closer to the methodology of the physical sciences. Examples include evolutionary biology, where the "synthetic theory" unified Darwinian concepts of evolution with Mendelian genetics. The result is the sophisticated apparatus of mathematical population genetics whose applications range from breeding programs to molecular phyrogenetics. In developmental biology, the mathematical analysis of pattern formation has been enormously successful in explaining phenomena as dissimilar as the morphogenesis of primitive algae, and the banding patterns of snail shells. In neurobiology, mathematical models of neural networks have shown that even networks of a few hundred neurons can display features that were previously thought to be the exclusive domain of the human mind, such as generalization and abstraction. The adaptation of economic game theory to evolutionary problems has proven invaluable in explaining many features of competition and cooperation in animals. In some instances, mathematical biology has contributed to discoveries whose impact was felt far beyond the life sciences. A prominent example is theoretical ecology, where a numerical analysis of an equation describing logistic population growth has stimulated the development of the theory of chaotic dynamical systems. These systems are now known to be important on all levels of organization of both animate and inanimate matter.

All these areas of biological investigation have one aspect in common: they are concerned with biological processes, which implies that they have to represent temporal and spatial changes in biological systems. Almost universally, they use a particular mathematical representation for such change. They represent the state of a system at some time t by a set of state variables $\vec{x}(t) = (x_1(t), \ldots, x_n(t))$ whose interpretation strongly depends on the nature of the problem. State variables may represent gene frequencies in population genetics, concentrations of chemical reagents in a pattern formation process, or electrical activities of neurons. The range of possible values of $\vec{x}$ defines

the state space S of the system. Changes in the state variables are most often represented by a differential equation, such as

$$\frac{d\vec{x}(t)}{dt} = f_1[\vec{x}(t), t, \Lambda]. \tag{1}$$

This kind of equation describes the change of the state variables in terms of some function f_1 of the system state, $\vec{x}(t)$, of time t, and of some set of system parameters Λ. Such parameters may represent, for example, the fitness of individuals with a given genotype in models of gene frequency change in a population, or the strength of trophic interactions among species in an ecological model whose state variables x_i stand for species abundance. A completely analogous, alternative representation is that of a difference equation

$$\vec{x}(t+1) = f_2[\vec{x}(t), t, \vec{\lambda}]. \tag{2}$$

The choice of (1) or (2) depends on whether a discrete or continuous notion of time is more appropriate for the problem at hand. The dynamics of either system is completely determined by its parameters, and by a set of initial conditions $\vec{x}(t_0)$, i.e., by the state of the system at some time t_0. For these dynamical systems, an important basic distinction is that of linear and non-linear equations. Illustrated for the example of an ordinary autonomous differential equation, linear equations are of the form

$$\frac{d\vec{x}}{dt} = \Lambda\vec{x}, \tag{3a}$$

with components

$$\frac{dx_i(t)}{dt} = (\Lambda\vec{x})_i = \sum_{j=1}^{n} \lambda_{ij} x_j(t) \tag{3b}$$

where Λ is a matrix of parameters $\Lambda = (\lambda_{ij})$. Any equation that can not be represented in this form is called non-linear. A cursory glance at any contemporary mathematical biology text, such as Murray's (1989) "Mathematical Biology", shows that the vast majority of mathematical models in this area are non-linear. Whether it is developmental biology, neurobiology, ecology, or population genetics, nonlinear systems are pervasive. Unfortunately, the large number of studies of non-linear systems available in many areas of science have not resulted in an all-encompassing mathematical theory, but in a sizable zoo of disparate techniques to analyze them. The result is a large collection of case studies, more or less well understood equations, hopefully representing important aspects of the modeled system. Therefore, most often

the only way to illustrate important principles governing such equations is the presentation of example cases.

Given the universal use of mathematical representations like (1), and given also its enormous success, what would be more natural than to cast the problem of causality in terms of such models? Well, one might argue, it is necessary to understand the notion of causality in reality, and not in a mathematical formalism that is at best a caricature of this reality. To this one might respond that everything one can say about reality relies on mental constructs which are models of the world. Even the most basic concepts produced by our sensory systems are only models of the world. They rely on many implicit assumptions, as evidenced by cognitive psychology, a field that exposes these assumptions by finding conditions under which they fail. Because they rely on implicit assumptions, such sensory models of the world are easily *identified* with reality. They are our windows into reality, and they can thus be distinguished from other, more abstract (e.g., mathematical) models because they function as a link between the more abstract models and the reality they represent. Assigning them priority over mathematical models in discussing causality, however, would be fallacious. If anything, the mathematical constructs used in science have an advantage over implicit perceptual constructs. Many of them have much more predictive power than any other mental construct. Their solutions often approximate a system's behavior well beyond the original intention and range of applicability of a model. This is especially remarkable because the solutions sometimes have counterintuitive properties that are not obvious from the structure of the equation, and that may not have been foreseen by whoever first discovered or described the equation. It is hard to escape the conclusion that such models capture something profound about a system's nature. Moreover, because they are phrased in a language with many fewer ambiguities than everyday language, one can expect that definitions of cause and effect can be defined more clearly and analyzed more easily for mathematical models. These observations suggests that, if anything, mathematical models are a much more appropriate domain of an investigation into the nature of causality than everyday language.

A scientist's notion of causation is based on regularities in natural phenomena. If the activity of an enzyme in a metabolic pathway is increased, the rate at which substances are metabolized in this pathway is likely to increase. Changes in enzyme activity cause changes in metabolic rates. This notion of causation is not restricted to changes in a system's states, but also applies to the states itself, e.g., a mutant gene in the genome of an individual is said to cause a disease. As the similar phrasing of these two examples shows, the distinction between these two cases is somewhat blurry, and it is conceivable that they could be lumped into one by a suitable definition.

However, a profound problem exists with this notion of causality, which was first recognized by Hume. Simply stated, it is that causation as a necessary connection between two events can not be inferred from correlation or association, an issue that is also central to debates about causality in the social sciences and in epidemiology. Partly driven by this problem, a number of attempts have been made to render the notion of causality exact. They include notions of probabilistic causality (Suppes 1971), analysis of causality using counterfactuals (Mackie, 1985), and a representation of causes as "INUS-conditions" (*i*nsufficient but *n*on-redundant parts of *u*nnecessary but *s*ufficient conditions; Mackie 1974). The latter example shows the considerable terminological sophistication that characterizes some of these attempts. Unfortunately, such sophistication often seems to be negatively correlated with operational usefulness. The notion of cause based on regularities, on the other hand, is operationally useful, and it has historically been quite successful. In addition, for the purpose of my argument, the "spurious correlation" condition frequently cited as a problem of the regularity view is irrelevant. This is because here mathematical representations of a system are used to study the issue. In these representations, functional relations among state variables of a system are provided, and the question becomes whether the notion of cause and effect can be meaningfully used *given* these relations.

Before this question can be addressed, some conceptual issues have to be clarified. They regard the kinds of entities that qualify as causes and effects in a dynamical system's representation of a process. The values of individual state variables $x_i(t)$ at some time t are the elementary states of the system. Elementary events are changes in these state variables, which may be continuous or discrete, depending on how time is represented. A special case are changes in a component of the initial conditions, $x_i(t_0)$. Of the numerous conceivable notions of causation, a notion based on events will be used here. Any change in a state variable $x_i(t)$ shall qualify as a cause. In systems like (2), the corresponding effect is a change in one or more state variables at time $t + 1$. In (1), effects are represented by changes in derivatives dx_i/dt. These notions of effect are more similar than they appear at first glance, if one notes that (2) can be written in terms of the change of state variables $\Delta\vec{x} = \vec{x}(t + 1) - \vec{x}(t)$ as well. Also, the continuous time system (1) can be viewed as a limiting case of the discrete system. If one takes this view, the fact that cycles of causation (effects acting back on their causes) may occur is unproblematic, since there is always a time lag (finite or infinitesimal) between cause and effect (cf. also Malinvaud 1966, p. 59). The important common feature is that changes in state variables cause changes in the subsequent behavior, i.e., the *trajectory* of the system. This highly formalized notion of cause and effect is not new. Similar concepts are used in

econometrics (Malinvaud 1966). It is useful to distinguish two different kinds of effects in (1) and (2). A local (short-term) effect is a change in derivatives or state variables at time t or $t + 1$, for (1) and (2), respectively. A global (long-term) effect concerns the qualitative dynamics of these systems, i.e., the dynamics $\vec{x}(t)$, as $t \to \infty$. The examples given below will serve to illustrate the latter notion.

Thus far, there has been an emphasis on elementary events as causes. However, one might argue that changes in the entire state $\vec{x}(t)$ of a system should qualify as causes as well. This issue is related to a classical philosophical distinction, namely the distinction between regular causes and singular causes (e.g., Koch 1990). Regular causes are best understood as regularities in a systems behavior. In the statement, "if at least x organisms in a population are carriers for a disease, the disease will spread through the population" a high number of disease carriers is a regular cause for an epidemic. Singular causes are particular, unique, often historic events that change the behavior of a system (e.g., the invention of the steam engine profoundly changed the world economy). Although singular causes may be involved in a system's behavior, only regular causes will lead to prediction or explanation, which is what scientists try to achieve. In the framework used here, changes in the entire state $\vec{x}(t)$ of a system are a special kind of singular cause, and only elementary events represent potential regular causes. The reason has to do with how mathematical models are used in scientific practice, and with the *ceteris paribus* condition of identifying regular causes. When a natural process is modeled, say the spreading of an epidemic in a population of organisms, a choice is made by the investigator as to the factors important in understanding the process. These factors, e.g., the number of individuals carrying a disease, enter the model as state variables x_i or as parameters λ. All other aspects of reality are excluded from the model. Loosely speaking, for the purpose of modeling a particular process, the dynamical system is taken as a model of the world. The *ceteris paribus* clause states that "everything else being equal", an event (cause) shall always be followed by the same effect. It is at least implicitly used whenever regular causes are to be identified in a scientific context. If the dynamical system is taken as a model of the world, and if the cause already encompasses a change of the entire system state $\vec{x}$, then there is no notion of *ceteris paribus* that can be meaningfully applied. However, when considering elementary events, say the change in one state variable x_i, *ceteris paribus* applies to all other state variables x_j unaffected by the event. This restriction to elementary events can be relaxed somewhat to allow changes in a number of state variables small compared to the size n of the system.

While the *ceteris paribus* condition is necessary for an event to be considered a regular cause, a much stronger condition is often implicitly used by scientists. It is a condition on the *separability* of the action of variables relevant to a system. I call $x_i(t)$ separable from $x_j(t)$, if the effect of a change in $x_i(t)$ is independent of the value of $x_j(t)$. It is necessary to incorporate this notion when identifying regular causes, because scientists are frequently not in the position to control all the potentially relevant variables of a system. What is more, they may not even know some of these variables. It is therefore important that some aspects of a system can be ignored while other aspects are investigated. Take the example of a biochemical investigation into some metabolic pathway by analyzing mutants of genes acting in that pathway. Often, geneticists have very limited means of controlling the "genetic background" of the organism they are working with, i.e., all the genes not of immediate interest to the studied pathway. If, however, the effects of this background on the pathway are sufficiently weak, they can be ignored. Thus, whether one can make causal statements implying regularity, such as "a change in the activity of enzyme E changes the metabolic flux by a factor x" may depend critically on a notion of context independence or separability of the action of state variables. If the actions of state variables are separable, then it may not matter whether one knows all of them, or whether their values can be held constant, i.e., whether the *ceteris paribus* condition can be applied.

Consider now the case of linear systems as given by (3). In all that follows, phenomena will be neglected that occur only (i) in parts of the state space S with measure (i.e., volume) zero, or (ii) in parts of the set of all possible parameters with measure zero (e.g., in the set where $\mathrm{tr}\Lambda = \lambda_{11} + \lambda_{22} = 0$). Albeit important for mathematical analysis, such phenomena are largely irrelevant for applications of the equations to the modeled process, since parameters and state variables can in practice not be restricted to such sets. What is the effect of an elementary event, i.e., the effect of changing the value of one state variable, say $x_k(t)$, to $x_k'(t)$? The linear relation $\dot{x}_i = \sum_j \lambda_{ij} x_j$ implies that, whatever the value of the other state variables is, the effect on the trajectory (as given by the derivative $\dot{x}_i$) will always be proportional to the amount of change, i.e., it is $[x_k'(t) - x_k(t)]\lambda_{ik}$. Local effects are completely independent of the context of the other variables, and they depend only on the constant parameters Λ of the system. A given cause will always have the same magnitude of effect. Similarly, global (long-term) effects on the behavior of the system are largely independent of the context or "background" provided by the other state variables. This is due to basic results from the theory of linear dynamical systems which imply that a representation of the system exists that makes the global effects of any elementary event independent of the context. This shall be illustrated by only one example, that of a two-variable linear

system with a (2×2) matrix Λ whose eigenvalues are real-valued and have different signs. A phase diagram of this system, indicating sample trajectories for a few different initial conditions, is shown in Figure 1a. Except along the straight lines intersecting at the origin, each state variable approaches either $+\infty$ or to $-\infty$. To which of these values it converges depends on the value of the other state variable. However, there exists a simple representation of the system that eliminates this dependency. Technically speaking, one has to change the coordinate system via a linear transformation of the state variables. In graphical terms, one has to rotate the $x - y$ axes in the figure by different angles, such that each axis coincides with one of the invariant diagonal lines (arrows) shown in the figure. Then, the long term behavior of the system for any one variable is independent of the value of the other variable. A completely analogous line of reasoning holds for all possible cases of (2×2) matrices Λ, and extends to higher dimensions as well.

In sum, it is always possible in a linear system to find a representation such that local or global effects of elementary events become independent of the context of the other state variables. Importantly, the appropriate coordinate transformation (if necessary) is itself linear, such that the linear character of the system is not changed. In this representation, effects will only depend on the fixed set of parameters Λ. Comparable results hold for linear difference equations. Thus, for linear systems, and based on the terminology used here, the notions of cause and effect do not pose major problems.

The situation is quite different in nonlinear systems. Two examples will be used to illustrate the main issue. Consider first the nonlinear differential equations

$$\dot{x}_1 = x_1(1 - r)(r - 2) - x_2 \tag{4a}$$

$$\dot{x}_2 = x_2(1 - r)(r - 2) + x_1 \tag{4b}$$

where $r^2 = x_1^2 + x_2^2$. This is a simple nonlinear oscillator with two state variables x_1 and x_2, and an angular velocity equal to one. Its dynamics are illustrated in a phase diagram (Figure 1b), which depicts the qualitative patterns of change in the state variables. Each arrow in the figure represents a sample trajectory of the system. For initial conditions $x_1(t_0), x_2(t_0)$ such that $r(t_0) > 2$, the system ultimately (as $t \to \infty$) approaches a limit cycle with radius $r = 2$, i.e., both state variables oscillate such that the resulting motion is circular in a Cartesian representation of (x_1, x_2) coordinates. In this asymptotic state, both variables are characterized by sinusoid periodic behavior with angular velocity one and amplitude two. Starting with $r(t_0) < 1$, the system approaches the origin, which is a stable fixed point, i.e., a point where no more change in the system occurs. Starting close to $r = 1$, the system will either

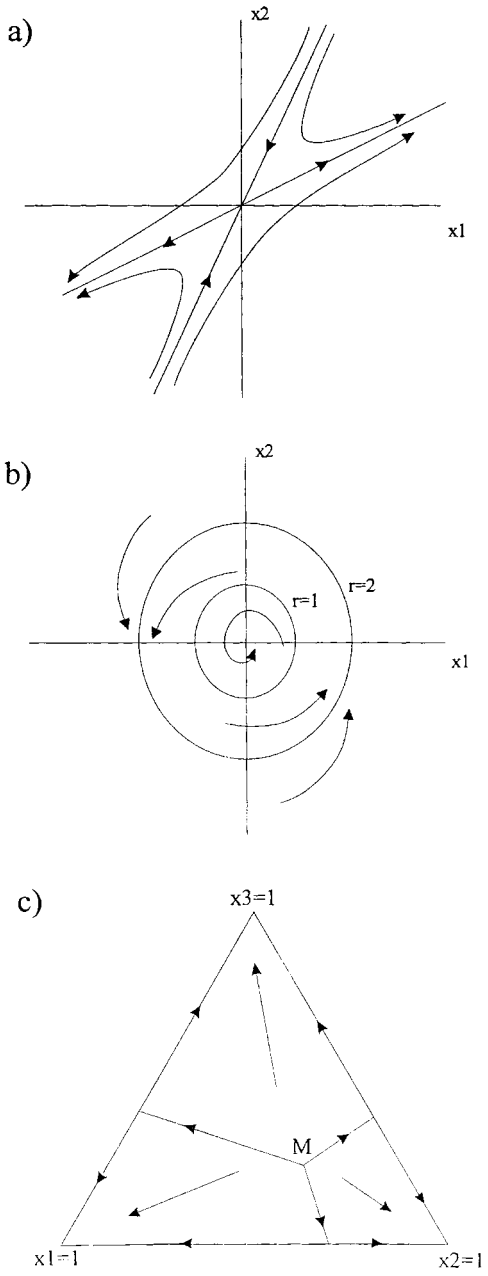

Figure 1. a) Phase diagram of the linear dynamical system (3) with $n = 2$ and real eigenvalues of opposite sign. b) Phase diagram of the nonlinear oscillator (4). c) Phase diagram of equation (5) depicting concentration dynamics of three autocatalytic molecules in a flow reactor. See text for details.

approach the circle with radius $r = 2$ or the origin, depending on whether $r(t_0)$ is greater or less than one. The circle with $r = 1$ is an unstable limit cycle. Small deviations from $r = 1$ will move the system away from $r = 1$. What are the local effects of changes in a state variable of this system? If x_1 is changed to x_1', the resulting change in r is $\Delta r = \sqrt{x_1'^2 + x_2'^2} - \sqrt{x_1^2 + x_2^2}$. Since r enters into (4), and because of the general form of (4), it is clear that local effects depend on both x_1 and x_2, and not only on x_1. Thus, the magnitude of change in $\dot{x}_1$ is clearly dependent on the context, which in this case is simply the state variable x_2. Given the magnitude of the change in x_1 (the cause), it is the context that determines the effect. This holds also for the global effect of a change in x_1. If r is changed from a value $r > 1$ to a value $r < 1$, the system will approach a fixed point where it previously approached a limit cycle. Both the magnitude of the change in x_1 and the context (x_2) are important in this regard. In fact, the effect of an elementary event depends on the context to the extent that any notion of causality based on regularity becomes meaningless. *Ceteris paribus* means that x_2 has to be known. But this knowledge already completely specifies the state of the system. Thus, one can only speak meaningfully of singular causes which are of little use. However, if one restricts one's view to a smaller range of values, say x_1 and x_2 being such that $r < 1$, then the behavior of the system becomes *topologically* equivalent to a linear system. *Topologically* here means that the trajectories of the system (4) around the origin can be "straightened out" (i.e., smoothly transformed), such that the dynamics is the same as that of some linear system around the origin (e.g., Jordan and Smith 1987). This property is heavily exploited in the analysis of nonlinear systems. Here it is important because it means that problems with regular causes arise in those areas where the nonlinear system behaves qualitatively different from the linear system. In nonlinear systems, regular causes may be (but need not be) meaningful in small regions of the state space. Wherever nonlinear systems behave "qualitatively" nonlinear, the notion of regular causes may not be of much use.

Consider as a second example the system of equations

$$\dot{x}_i = x_i(\lambda_i x_i - \sum_{j=1}^{n} \lambda_j x_j^2) \quad 1 \leq i \leq n. \tag{5}$$

Each state variable x_i represents the relative concentration of an autocatalytically self-reproducing (RNA) molecule in a flow reactor containing n such molecules (Hofbauer and Sigmund 1988, p. 90). The λ_i's are a measure of the replication rate of molecule species i. (Structurally similar equations describe the interactions of species in an ecosystem, in which case x_i represents the abundance of organismic species i in the system.) The dynamics of

(5) is illustrated for 3 molecule species by the phase diagram of Figure 1c. In this figure, each point in the interior of the equilateral triangle (a simplex) corresponds to a combination of relative concentrations, such that $\sum x_i = 1$ holds. On the edges of the simplex, only two of the three molecule species have concentrations different from zero. The vertices of the simplex correspond to a situation where only one species exists. Three unstable equilibria are located on the edges, and one (denoted by M) in the interior. The vertices of the simplex are the only stable equilibria, and their basins of attraction are separated by the straight lines connecting them to M. The location of M in the simplex is determined by the parameters λ_i. Consider as an example for the local effect of a change in a state variable the effect of a change in x_1 on x_j, $j \neq 1$. It follows from (5) that $\Delta\dot{x}_j = x_j(x_1^2 - x_1'^2)\lambda_1$. Therefore, the local effect of a change in one state variable on the trajectory $\dot{\vec{x}}$ does not exclusively depend on that state variable, but on all other state variables as well. For the global effects, it is clear that a small change in, say, x_1 will not affect the convergence of the system to $x_1 = 1$ (where only molecule species 1 is present) for trajectories well in the interior of its basin of attraction. However, the same change at a point close to the boundary of this basin, and specially close to M, may result in convergence of the system to $x_2 = 1$ or $x_3 = 1$. Whether this happens, depends not only on the direction of change in x_1, but critically on the values of all the other state variables. Thus, the same reasoning as for the above example applies to (5), including (topological) equivalence to a linear system if one restricts the dynamics to any one basin of attraction.

The examples given here are very simple, and one might argue that in most practical cases it would be possible to identify the mathematical relations among only two or three state variables. However, it would be easily possible to illustrate the same point with much more complicated dynamical systems, whose behavior is not nearly as intuitively clear as that of the examples given here. One only has to bear in mind that dynamical systems of realistic dimensions may be vastly more complex. For example, Lewontin (1974) gives an example from population genetics in which a dynamical system is used to model the change in frequencies of alleles at five gene loci in a population of diploid organisms. The system has of the order of 10^9 alternative equilibria. In such systems, knowledge of the local or global dynamics of the system may depend on detailed knowledge of the values of all state variables, precluding any meaningful definition of regular causes.

In systems with many variables and complex behavior, an additional factor may become important, namely the "density" of interactions among state variables. It may be crucial to know whether all state variables in the system interact with each other or whether interactions among state variables involve

only small groups of variables. This point is illustrated by the Ising spin model, which originated in solid state physics, but has recently become of great importance in the theory of neural networks (Amit 1989). Consider a system with discrete state variables x_i assuming values of $+1$ and -1 only. These state variables may represent the spins (magnetic moments) of individual atoms in a magnetic material, or the activity states of neurons in a neural network. The dynamics of the system is given by

$$x_i(t+1) = \sigma\left[\sum_{j=1}^{n} \lambda_{ij} x_j(t) + h\right] \quad 1 \leq i \leq n \tag{6}$$

Here σ is the sign function ($\sigma(x) = 1$ for $x > 0, \sigma(x) = -1$ for $x < 0$, and $\sigma(x) = 0$ for $x = 0$), and h is some real constant corresponding to an external magnetic field in the magnetic model, and a "firing" threshold in the neural model. $\Lambda = (\lambda_{ij})$ is a matrix of "connectivity" parameters specifying the interactions among state variables. This is a nonlinear difference equation, whose nonlinearity stems from the function σ. The number n of spins may be quite large. In fact many of the mathematical results for this class of models are derived for the thermodynamical limit $n \to \infty$.

Consider first a simple special case, that of a circular chain of spins, in which each spin interacts only with its nearest neighbor. Then, one has

$$x_i(t+1) = \sigma[x_{i-1}(t) + x_{i+1}(t) + h]$$

Here $i-1$ is identified with n if $i = 1$, and $i+1$ is identified with 1 if $i = n$. Consider the elementary event of a change in a state variable, say x_1. Its effect on another state variable x_2 is completely determined if one knows the state of x_3 at time t. A cause will have the same effect *ceteris paribus* means that only x_3 has to be kept constant. One need not know the state of the entire system. Because interactions among spins are only local, regular local causes can be meaningfully defined.

A far more important case, for example for neural network modeling, is that of a fully connected Ising spin model, in which all parameters λ_{ij} are different from zero. In this case, (6) implies that the effect of a change in a state variable may depend on the values of all other state variables. The context in which such a change happens is therefore essential, and the *ceteris paribus* condition is not useful for the identification of regularities. There is clearly a continuum of systems between the local interactions of the linear Ising chain, and the global interactions of the fully connected model. As one moves from local to global interactions, *ceteris paribus* implies that an increasing number of state variables has to be kept constant. Conversely,

changes in individual state variables become less and less separable from the state of the entire system in the sense introduced above. This illustrates that the cut-off point between regular and singular causes is quite arbitrary for these systems.

In sum, a notion of causation based on regularities can only be meaningful for areas of the state space, in which the behavior of a nonlinear system is topologically equivalent to that of a linear system. Local interactions among state variables may sometimes render the notion of regular causes meaningful, but as these interactions become global, its usefulness evaporates.

When discussing nonlinear dynamical systems it is tempting to include systems with chaotic dynamics. In this area examples of almost arbitrarily pathological behavior can be constructed (e.g., Sommerer and Ott 1993). The sensitivity of such systems to changes in state variables may allow one to illustrate how profoundly problematic a regularity theory of causality is in nonlinear systems. However, the above examples show that one does not have to resort to the rather peculiar features of chaotic systems, although systems similar to (6) may show chaotic behavior.

In light of what has been said about causality in nonlinear systems, it is instructive to return to the example of the social sciences for a moment. Because the mathematical terminology in these areas is somewhat different from that used here, I will briefly clarify how the two are related. Often, in the social sciences the mathematical framework is not one of differential equations, but of simultaneous (linear) equations in some state variables which have to be solved. However, such equations can often be viewed as representing some assumed equilibrium of a dynamical system. A dynamical system representation would then prescribe the trajectory of the system towards that equilibrium (e.g., Malinvaud 1966, p. 52). Further terms that are not usually found in models of the life sciences and social sciences are those of endogenous and exogenous variables (Malinvaud 1966). Exogenous variables are variables that can influence the state of the system, as represented by all endogenous variables, but not vice versa. In the dynamical systems framework used here, endogenous variables correspond to state variables, and exogenous variables correspond to system parameters. The kinds of causes studied in these systems are of a somewhat different nature as well. Typically, it is asked how changes in system parameters (exogenous variables) affect the system's equilibria. An important task of causal models is to determine the relation among state variables such that one can predict how a change in a system parameter will affect the state of the system in equilibrium. While the mathematical examples given above all concerned changes in state variables, it would be straightforward to conceive of similar

examples involving parameter changes (e.g., the general class of dynamical systems involving Hopf bifurcations; Murray 1989). Thus, while superficially distinct, it is straightforward to cast issues regarding causal relations in the social sciences in the terminology used here. Characteristically, discussions regarding the identification of causal relations in the social sciences mostly restrict themselves to linear interactions among variables (e.g., Glymour et al. 1980; Simon 1971). It is tempting to speculate that this basic commitment to linear models is another important reason why discussions of causality have continued to be important in this area. In linear causal models, the assumption is made that parameter changes translate linearly into changes of the state variables in equilibrium. If this is the case, path analysis (e.g., Sokal and Rohlf 1981) provides a powerful and widely used tool to analyze and quantify these interactions. However, in nonlinear systems no comparable tool is likely to exist. Taking again the dynamical systems viewpoint, it is well known that changes in system parameters (exogenous variables) need not smoothly change individual state variables as they do in linear systems. They may change the stability of equilibria, they may cause discontinuous shifts in equilibria from one position to another in the state space, they may turn stable equilibria into limit cycles, or even induce chaotic dynamics (Murray 1989; Jordan and Smith 1987). As in the examples shown here, it may be necessary to know the value of all state variables and parameters to determine the systems behavior. Again, a regularity notion of causation is not likely to be meaningful, except in areas of the (state or parameter) space where a linear approximation is appropriate. Such an approximation can again be understood in the sense of a topological equivalence between linear and nonlinear dynamics. It is interesting to note that nonlinear or curvilinear regression in statistical theory exploits this very ability of moderately nonlinear systems. Variable transformations or presumed functional relations among variables which are then fit to data are usually smooth, e.g., logarithmic or polynomial.

Why is this digression into causal models in the social sciences relevant to biological models? The reason is that there is an instructive commonality between the two areas. In the social sciences, it seems that the linearity of causal models is often due to a lack of insight into functional relations among the variables of a model. This lack of insight probably exists for good reasons, such as the inability to do experiments, and methodological difficulties in collecting even observational data. Given this lack of insight, the most parsimonious assumption about a system is that it is in equilibrium, and that the interactions among its state variables and parameters in equilibrium are linear. An additional incentive to using these assumption is that they render a regularity notion of causation meaningful, either in terms of everyday language, or in the more sophisticated language of path analysis. In some areas of biology,

notably those that may become of great political importance in the near future, a similar situation holds, insofar as assumptions of linearity substitute for insight. This is often the case for population genetics, on which medical genetics relies heavily. Little is known about the precise functional interactions among gene products that lead to most phenotypic traits. In light of this fact, and considering that many phenotypic characters show a considerable amount of additive genetic variation (a statement about *statistical* relations among variables; Falconer 1981), the parsimonious assumption is regularly made that the *functional* relations among genetic variables are also linear. Of course, this assumption neglects many qualitative results from molecular biology which suggest that interactions among gene products usually have highly nonlinear features (e.g., cooperativity in transcriptional regulation of gene expression). It also neglects the importance of epistatic variance, which is often found experimentally, and caused by nonlinear interactions among genes. However, given the little available quantitative information about gene interactions, population genetic models often do not have viable alternatives to assumptions of linearity. These assumptions only become deeply problematic if causal roles are to be attributed to individual genetic factors, e.g., in genetic diseases. If, as one has every reason to assume, nonlinear phenomena are important, the association of a particular genetic marker with a disease may crucially depend on the context, i.e., on the genetic background of the population in which the association was established. It need not hold then, that background effects "average out" over the individuals in a population, as they may in the linear case. Associations found between genetic markers and a disease in one population may not exist in a different population, and loci explaining (in a statistical sense) phenotypic variation in one population, may not do so for other populations. This is because the value of all relevant state variables, corresponding to alleles at all relevant loci, can be so important in nonlinear systems. As in the nonlinear dynamics examples given here, one may need to know the state of the entire system with all its state variables to make predictions. The importance of such nonlinear interaction effects may be an important reason why the analysis of complex genetic diseases is hampered by numerous problems. Sometimes even apparently "simple" genetic diseases reveal complex patterns of inheritance. Lander and Schork (1994) quote the example of sickle cell anemia, a textbook case of a "simple" genetic disease caused by a particular mutation at the β-globin locus. As it turns out, individuals with the same genotype at that locus may be affected to quite different extents. Part of the reason are influences exerted by the genetic background, in this case at least two other loci. A similar example involves schizophrenia, a disease with an undoubtedly complex genetic component (Baron et al. 1990). Here linkage between the disease and genetic markers on

the human chromosome 5 were found in one study, but could not be replicated in a different pedigree. Numerous other examples of the importance of background effects could be given (Baron et al. 1990; Lander and Schork 1994), and they are not surprising for nonlinear systems. At the very least, independent studies in unrelated populations are necessary to assess the importance of the genetic background, but statistical reasoning can not replace insight into functional relations among relevant variables in these cases.

In sum, there are good reasons not to abandon the notion of causality, as suggested by Russell (1913), because it is useful in systems that behave qualitatively linear in the sense used here. For such systems, powerful statistical tools exist to delineate causal interactions. Because these tools in general rely on measures of linear associations, they are likely to fail for qualitatively nonlinear systems. In these systems statistical reasoning can not replace insight into functional relations among variables, as given by a mathematical formalism describing their interactions. To argue that the notion of regular causes has severe limitations in nonlinear systems may seem like an academic problem. It is not in those areas where important decisions are likely to be made on the basis of statistical criteria alone.

Acknowledgments

I am grateful to Mary Morgan and Margaret Morrison for numerous helpful discussions and critical comments. The financial support of the Institute for Advanced Study in Berlin, and the Santa Fe Institute is gratefully acknowledged.

References

Amit, D.J.: 1989, *Modeling Brain Function*, Cambridge University Press, Cambridge.

Baron, M., Endicott, J. and Ott, J.: 1990, 'Genetic Linkage in Mental Illness', *British Journal of Psychiatry* **157**, 645–655.

Cartwright, N.: 1989, *Nature's Capacities and their Measurement*, Oxford University Press, Oxford.

Crick, F.: 1996, 'Visual Perception: Rivalry and Consciousness', *Nature* **379**, 485–486.

Falconer, D.: 1981, *Introduction to Quantitative Genetics*, Longman, London.

Francis, R.C.: 1990, 'Causes, Proximate and Ultimate', *Biology and Philosophy* **5**, 401–415.

Glymour, C., Scheines, R., Spirtes, P. and Kelly, K.: 1987, *Discovering Causal Structure*, Academic Press, Orlando, FL.

Hofbauer, J. and Sigmund, K.: 1988, *The Theory of Evolution and Dynamical Systems*, Cambridge University Press, Cambridge.

Jordan, D.W. and Smith, P.: 1987, *Nonlinear Ordinary Differential Equations*, 2nd ed. Oxford University Press, New York, NY.

Koch, G.: 1990, *Kausalität, Determinismus und Zufall in der wissenschaftlichen Naturbeschreibung*, Duncker & Humbolt, Berlin.

Lander, E.S. and Schork, N.J.: 1994, 'Genetic Dissection of Complex Traits', *Science* **265**, 2037–2048.

Lewontin, R.C.: 1974, *The Genetic Basis of Evolutionary Change*, Columbia University Press, New York, NY.

Mackie, J.L.: 1974, *The Cement of the Universe*, Oxford University Press, Oxford.

Mackie, J.L.: 1985, *Logic and Knowledge*, Selected Papers Vol. 1. Oxford University Press, Oxford.

Malinvaud, E.: 1966, *Statistical Methods in Econometrics*, North-Holland, Amsterdam.

Murray, J.D.: 1989, *Mathematical Biology*, Springer, New York, NY.

Rossi, P.H., Berk, R.A. and Lenihan, K.J.: 1980, *Money, Work, and Crime*, Academic Press, New York.

Russell, B.: 1913, 'On the Notion of Cause', *Proceedings of the Aristotelian Society*, New Series **13**, 1–26.

Simon, H.A.: 1971, 'Spurious Correlation: A Causal Interpretation', in Blalock, H.M. (ed.), *Causal Models in the Social Sciences*, Atherton, Chicago, IL, pp. 5–18.

Sokal, R.R. and Rohlf, F.J.: 1981, *Biometry*, Freeman, New York.

Sommerer, J.C. and Ott, E.: 1993, 'A Physical System with Qualitatively Uncertain Dynamics', *Nature* **365**, 138–140.

Suppes, P.: 1970, *A Probabilistic Theory of Causality*, North-Holland, Amsterdam.

Wadman: 1996, 'US Health Politics May Give Insurers a Respite on Genetic Information', *Nature* **380**, 91.

Wagner, A.: 1996, 'Can Nonlinear Epigenetic Interactions Obscure Causal Relations between Genotype and Phenotype?', *Nonlinearity* **9**, 1–23.