1

In the beginning...

1.1 ... there was Poincaré

If modern nonlinear dynamics has a father, it is Henri Poincaré (1854–1912). Dynamic studies, prior to his studies in the 1880s, concentrated on obtaining analytic solutions of dynamic equations, as characterized by many astronomical investigations of planetary motions and by Lord Rayleigh's ubiquitous studies of nearly every moving mechanical system. Many great names in analysis are associated with these studies – Newton, Leibniz, Euler, Gauss, Lagrange, Laplace, Jacobi, Lie and, of course Poincaré, among others. One of the 'grand problems' of classical dynamics, whose solution had withstood the efforts of many people, was the gravitational three-body problem. While ten integrals of the motion for this system had long been known, all attempts to find any of the remaining eight integrals had ended in failure. In 1887 Bruns proved that the ten classic integrals are, in fact, the only algebraic integrals of this system. This was followed by Poincaré's more famous, and frequently misunderstood, theorem (1890) concerning the nonexistence of integrals which are analytic in a perturbation parameter (see Historical outline). Finally, Painlevé (1898) extended Bruns' theorem to N-bodies, and generalized the spatial possibilities. These theorems, which represent a turning point in the analysis of dynamic systems, are discussed in some detail by Whittaker (1944). One can say that the theorems represented a certain loss of innocence. No longer was there any hope of solving all dynamic systems in terms of indefinite integrals involving elementary functions and uniformly valid power series. Other approaches had to be developed to understand the secrets hidden within the dynamic equations.

It was the genius of Poincaré which supplied us with many of our present methods for exploring the unexpected wonders of even 'simple' (low-order) dynamic systems. In particular, he emphasized the importance of obtaining a global, qualitative understanding of the character of a system's dynamics. Many of his suggestions were subsequently refined and extended by others, but our debt to his imagination and insight can hardly be overstated. Unfortunately, his contributions in this area were largely ignored by scientists for about 50 years, but fortunately not by many mathematicians, who extended his concepts to such areas as topology, dimension theory, asymptotic series, various maps and their fixed points, bifurcations, and also proved a number of his conjectures. This wealth of ideas was subsequently enriched by the studies of Birkhoff, who further characterized possible dynamic complexities, and by such concepts as the structural stability of equations, introduced by Andronov and Pontriagan. While the strict application of these concepts proved to have a limited practical impact, they produced an appreciation for the concept of 'robustness' and stimulated other avenues of investigation whose import are still being determined. Some of Birkhoff's abstract concepts were found to occur in surprisingly simple physical systems, as first experimentally detected by van der Pol and van der Mark (1927), and analyzed by Cartwright and Littlewood, and by Levinson in the 1940s. This in turn stimulated talented mathematicians, such as Smale, and thus produced a variety of new concepts.

In the 1940s an entirely new tool of analysis also came on the scene – the digital computer. In addition to its obvious brute-force capabilities of grinding out numerical solutions of differential equations, both von Neumann and Ulam foresaw some of its more subtle applications, as a flexible, interactive tool for the purpose of discoveries. This interplay between computations and analysis, which Ulam called 'synergetics', has indeed proved to be of great importance, and represents one of the major methods of uncovering dynamic properties. In more recent years, the term 'synergetics' has also been applied to a field of nonlinear dynamics, in which the system consists of many subsystems, of possibly very different characters (see 'coexistence' in Section 1.5). Indeed, the area of computer science has rapidly progressed to a state in which it can now make fundamental contributions to our knowledge, rather than acting only as the servant of other methods of analysis.

The period following 1950 has been particularly rich in new ideas and in the growing application of these ideas to a wide variety of disciplines, such as physics, chemistry, biology, neurology, astronomy, geophysics, meteorology, aeronomy, economics. In addition, many new perspectives have been introduced from mathematics, along with the innovative and potentially basic contributions of computer science. The purpose of this book is to stimulate your imagination, by surveying a wide variety of these concepts, and to illustrate them as explicity as possible by simple dynamic models (some of which, hopefully, are relevant to the real word!). Examples will be drawn from any of the above areas which will help to illustrate the relevance of some concept to real dynamic problems.

Unfortunately, the present limited survey omits concepts which may prove to be of great importance but it is hoped that the flavor and thrust of modern developments will be made clear in what follows. The objective here is not to present an established and completed format for analyzing nonlinear dynamics, for that does not exist but, rather, to illustrate the rich variety of concepts which are presently known, in the hope that it will stimulate new ideas and experiments which will further delineate those dynamic aspects that are of basic importance. This book is therefore to be viewed as a 'living' text in the process of growth; it is full of questions whose answers are unknown to the author and which, he trusts, will be both stimulating and fun for the reader.

1.2 What are 'nonlinear phenomena'?

Before considering any details of equations which describe 'nonlinear dynamics', the meaning of a 'nonlinear process', or 'nonlinear phenomena', should be examined. The latter expressions are introduced to make it clear that what is of greatest interest in the present study are 'phenomena'; that is to say observable physical variables and the dynamics of these variables. This is to draw the important distinction between the physical question: 'What are nonlinear phenomena?'; and the relatively trivial mathematical question: 'What are nonlinear equations'? In particular, a response to the former question which might first come to mind:

> nonlinear phenomena are those phenomena which are described by nonlinear equations (1.2.1)

is not necessarily correct and, even when true, fails to clarify a basic aspect of the orginal question.

In order that there is no misunderstanding concerning the mathematical side, recall that a linear operator $L(\phi)$ is defined to be one such that linear superposition holds:

$$L(a\phi + b\psi) = aL(\phi) + bL(\psi) \tag{1.2.2}$$

for any two constants a, b and all (vèctor) functions ϕ, ψ with suitable regularity properties (e.g., if $L(\phi)$ is a differential operator, then functions need to be in C^m, the space of m-continuously differentiable functions).† We shall then define a nonlinear operator $N(\phi)$ to be any operator which does not satisfy (1.2.2) for some a, b or some ϕ, ψ in C^m.

While this might appear to be an obvious definition for a nonlinear operator, the literature does not conform uniformly to this definition. For example, linear parametric equations are frequently included in studies of nonlinear oscillations, whereas nonlinear equations which are known to be linearizable by transformations, such as 'quasilinear' partial differential equations (PDEs), are sometimes considered to be linear (see below). We shall, however, use the above definition for nonlinear operators.

Now, returning to the original physical question, we can see that (1.2.1) is not generally a satisfactory answer because of the simple observation that the fundamental many-body Schrödinger equation is a linear equation. Moreover, in the classical

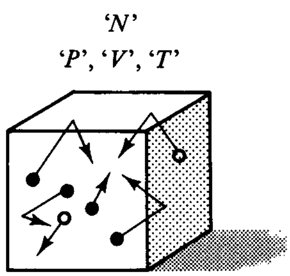
Mathematical notation (e.g. C^m will be kept to a minimum but some is necessary for a degree of precision. A glossary defining the notation is provided is Appendix A context, the Liouville equation, which encompasses all of Hamiltonian dynamics, is also a linear equation. Since any(!) mechanical phenomena can presumably be described by such basic linear equations, it follows that nonlinear equations are in principle not necessary to describe nonlinear phenomena! What then are 'nonlinear phenomena'? Moreover, this raises the interesting question:

> Why use nonlinear equations if the nonlinear phenomena can be described by linear equations?

This confronts a common misconception which we will discuss in Section 1.3.

As noted before, the key word is 'phenomena', or even 'phenomena of interest', with the implication that the variables are observable, and 'of interest'. It has of course been appreciated for a long time that the Schrödinger equation or the Liouville equation for many interacting particles contains dynamical information which probably will never (or, possibly, can never) be of any observational interest, and certainly is not of present interest. The phenomena of interest usually involve the behavior of certain macroscopic properties, which can be viewed as 'projections' or 'reductions' from the detailed microscopic states of the system. For example, in equilibrium systems we typically describe the system in terms of its density, pressure, volume, and temperature (Fig. 1.1).

Fig. 1.1



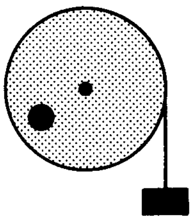
These concepts are all 'projections' of the microscopic state of those systems involving various space–time average properties. However, we will be interested in time dependent systems, for which such projections are more difficult. In any case, it is the spatial–temporal behavior of projections which forms the nonlinear dynamics of our present interest.

A fundamental theoretical problem arising from such a reduced description (projection) of the underlying microscopic dynamics is the determination of the collection of projected variables whose dynamics can be 'approximately' described by a (deterministic) system of nonlinear equations. Simple examples of such systems of equations occur in the fields of mechanics, or electrical circuits.

Exercise 1.1. The following systems may contain more than 10^{25} particles, which we typically project down to two or three variables, plus assorted constants.

(a) What is a set of dynamic variables, and a set of related physical parameters, which are presumably deterministic in simple electric circuits?
(b) A mechanical system consists of a solid wheel, which can rotate about a central axle (Fig. 1.2). An off-center mass and another mass hanging from a rope wrapped around the wheel are attached. Name a set of presumably deterministic dynamic variables, and associated parameters. State two major assumptions about these 10^{25} particles?

Fig. 1.2



A system can have both nonlinear and linear properties:

(i) A nonlinear property in the present system is to determine its equilibrium configuration. Does it exist? If so, what is it?
(ii) A linear property is to determine the frequency of small oscillations about its equilibrium configuration. Determine this angular frequency.

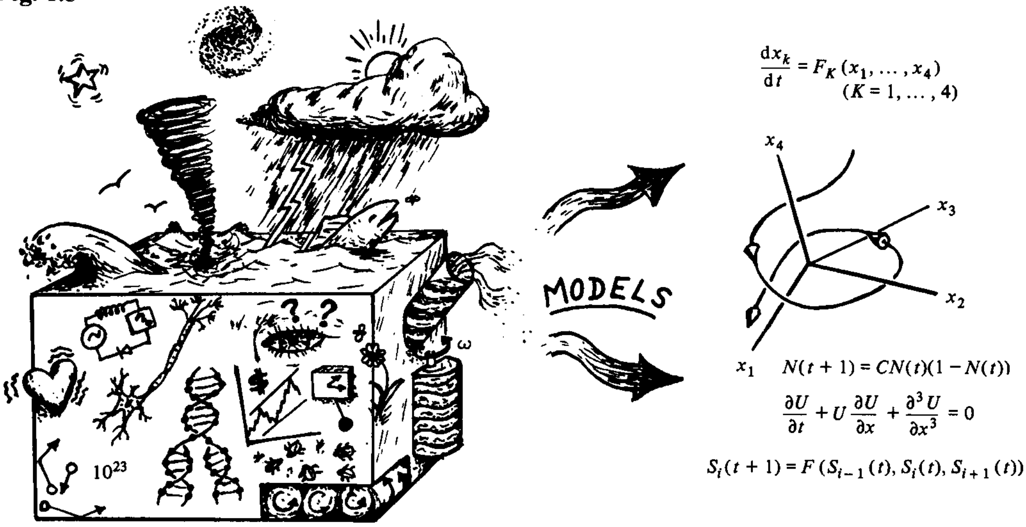
More sophisticated projections occur in various fluid equations of motion, which result from different methods of truncating (closing) an infinite system of moment equations (see Exercise 1.4). The limitations implicit in such truncations are not usually known with any precision, but are more or less loosely formulated. However, the physical nonlinear equations of interest are obtained by some 'projection and truncation' procedure, which yields a dynamic model of the system. The most common of types of models are:

ordinary differential equations (ODE); e.g. $\mathrm{d}x/\mathrm{d}t = ax + bx^2$;
partial differential equations (PDE); e.g. $(\partial u/\partial t) + u(\partial u/\partial x) + (\partial^3 u/\partial x^3) = 0$;
difference equations (DE), *or Maps*; e.g. $x(t+1) = ax(t) + bx^2(t)(t = 0, 1 \ldots)$;
cellular Automata (CA): discrete time, space, and functional values.

Attempts to obtain more fundamental projections, particularly in the area of nonequilibrium solids, liquids and gases can be found in the literature (e.g., see Grabert, 1982). Such a fundamental systematic approach is impossible in more complicated systems (chemically reactive, biological, sociological, economic, etc.), and we must rely on fairly simple models (Fig. 1.3).

In any case these ideas yield one (somewhat abstract) definition of what we will mean by 'nonlinear phenomena', namely

Fig. 1.3



Dynamics in the real world

This process contains the physical insight ('artistry') of the theorist in attempting to describe real phenomena

Nonlinear dynamics in the phase space of the physical variables

Nonlinear phenomena concern processes involving 'physical' variables, which are governed by nonlinear equations. These models have been obtained, by some approximate 'projection' rationale from presumably more fundamental microscopic dynamics of the system. (1.2.3)

It is possible, of course, that a reasonable projection may yield simple linear equations in some approximation, and that indeed is usually the first approximation that is attempted. Classic examples of this are sound waves in gases, water waves, the 'phonons' in solids, or 'plasmons' in ionized systems, which define 'collective' variables. Note that these approximations yield linear equations in the physical variables. This is to be distinguished from linear equations in nonphysical variables which may be obtained from a mathematical transformation of the original physical nonlinear equations. The phenomenon is linear only if the equations are linear in the physical variables.

It is useful to augment the rather abstract definition (1.2.3) with a more physical, operational definition. One possibility is:

Physical phenomena concern the interrelationship of a set of physical variables which are deterministic (within some accuracy). Nonlinear phenomena involve those sets of variables such that an initial change of one variable does not produce a proportional change in the behavior of that variable, or some other variable. In other words, the ratio (action/reaction) is not constant. (1.2.4)

To further emphasize (and clarify) some of the mathematical interrelationships between linear and nonlinear equations, consider the following facts:

(A) The general solution of the linear PDE

$$\frac{\partial \phi}{\partial t} + \sum_{k=1}^{n} F_k(x,t)\frac{\partial \phi}{\partial x_k} = 0 \quad (x \in R^n) \tag{1.2.5}$$

for the function $\phi(x,t)$ of $(n+1)$ independent variables is (locally in space/time) equivalent to solving the nonlinear system of ODE

$$\dot{x}_k = F_k(x,t) \quad (k = 1, \ldots, n) \tag{1.2.6}$$

(e.g., see R. Courant and D. Hilbert, *Methods of Mathematical Physics*, Vol. 2, Chapter II, §2). That is, any function $\phi(K(x,t))$ of the constants of the motion K_i, of (1.2.6) (so that $\mathrm{d}\phi/\mathrm{d}t \equiv 0$; see (2.4.4)) is a solution of (1.2.5), and any solution of (1.2.5) yields a constant of the motion of (1.2.6).

(B) A system of nonlinear PDE of the special form

$$\sum_{i=1}^{n} A_i(x,\phi) - \partial\phi_j/\partial x_i = B_j(x,\phi) \quad (j = 1, \ldots, m) \tag{1.2.7}$$

is called a quasi-linear system ($x \in R^n$, $\phi \in R^m$). If one defines $x_{n+s} \equiv \phi_s$ and $A_{n+s} \equiv -B_s (s = 1, \ldots, m)$, then (1.2.7) is equivalent to the single homogeneous linear PDE (see Courant and Hilbert)

$$\sum_{i=1}^{n+m} A_i(x)\partial\phi/\partial x_i = 0 \quad (x \in R^{n+m}), \tag{1.2.8}$$

which, of course, is the same as (1.2.5). Hence (1.2.7) is equivalent to (1.2.5).

(C) This yields the theory of characteristics for general first order PDE, $F(x, u, u_{x_1}, \ldots, u_{x_n}) = 0$ (where $u_x \equiv \partial u/\partial x$). This can be replaced by a system of $(n+1)$ quasi-linear PDE, of the form (1.2.7); n equations are obtained by considering $\mathrm{d}F/\mathrm{d}x_i = 0$ $(i = 1, \ldots, n)$ and setting $\phi_i \equiv u_{x_i}$. Then $\mathrm{d}F/\mathrm{d}x_i = \partial F/\partial x_i + \partial F/\partial u \phi_i + \sum \partial F/\partial\phi_k \partial u_{x_k}/\partial x_i = 0$. But $\partial u_{x_k}/\partial x_i = \partial\phi_i/\partial x_k$. So there are n equations of the form (1.2.7), namely

$$\sum_k \left(\frac{\partial F}{\partial \phi_k}\right)\frac{\partial \phi_i}{\partial x_k} + \left(\frac{\partial F}{\partial u}\right)\phi_i + \frac{\partial F}{\partial x_i} = 0 \quad (i = 1, \ldots, n).$$

Since $F = F(x, u, \phi)$ we need one more equation for u. This is obtained trivially, because

$$\sum \left\{\left(\frac{\partial F}{\partial \phi_i}\right)\frac{\partial u}{\partial x_i} - \left(\frac{\partial F}{\partial \phi_i}\right)\phi_i\right\} = 0.$$

Thus the general (and hence possibly very nonlinear) first order PDE

$$F\left(x, u, \frac{\partial u}{\partial x_1}, \ldots \frac{\partial u}{\partial x_n}\right) = 0 \tag{1.2.9}$$

can be related to the linear PDE (1.2.5).

More interconnections between linear and nonlinear equations are known (e.g., see the Riccati equation), and are continually being uncovered (e.g., the Hopf–Cole transformation, the inverse scattering transform, etc.).

Exercise 1.2. (This exercise requires some background in statistical mechanics). Write the linear PDE (1.2.5) for a system of N noninteracting particles with charges $q_i (i = 1, \ldots, N)$ in an electric field E (constant). What does this equation refer to in the field of statistical mechanics (i.e., what does $\phi(\mathbf{r}_1, \ldots, \mathbf{r}_N, \mathbf{v}_1, \ldots, \mathbf{v}_N, t)$ represent)? By solving equations (1.2.6) obtain the general solution of (1.2.5), using the $2N$ constants of the motion (initial conditions $(\mathbf{r}_i^0, \mathbf{v}_i^0)$ – see Chapter 2 if you have difficulty). Note that in this simple case even (1.2.6) is a linear system. Does ϕ (in its statistical mechanical context) represent an observable ('physical') variable? What types of quantities, related to ϕ, are common physical variables? What are the 'projections' in these cases? Obtain some PDE equations of motion for these physical variables from (1.2.5). Are they deterministic? If not, what must one do to get a deterministic system? Is that physically reasonable for the above system?

1.3 Two myths

It was noted in the last section that the clarification of the meaning of nonlinear phenomena raised the point that these phenomena presumably can be described by more basic linear equations (however, not in the 'physical' variables). If nonlinear phenomena can be described by linear equations, why would we want to use a nonlinear equation? After all, it is widely 'known' that

MYTH 1: linear equations are easier to solve than nonlinear equations.

This widely held misconception may have its origin in the fact that 'linear equations', to many scientists, connote equations for harmonic oscillators, free particles, the (very special!) Schrödinger equations for the hydrogen atom or a particle in a box, or possibly even a system of ordinary linear equations with constant coefficients. These linear equations are, of course, elementary to solve, but they are obviously not characteristic of most linear equations.

The property of linear superposition (1.2.2) is generally useful only when one can obtain a complete set of solutions $\{\phi_n\}$ of $L(\phi_n) = 0$, from which one can then construct general solutions. At least one can construct a series representation of the solution, but its general properties may remain unknown. In any case, the determination of this basis set $\{\phi_n\}$ is generally not aided in the slightest by the linearity of the operator. The point to be emphasized is that the fundamental linear equations of physics must in fact generally be more difficult to solve than their 'projected' nonlinear equations which describe the nonlinear phenomena of interest. (Those who may still doubt this should peruse the book, *Perturbation Theory of Linear Operators*, by T. Kato.)

Finally, the last paragraph raises the important fact that it is not generally true that

> *MYTH 2: an analytic solution of an equation gives, if it exists, the most useful information.*

One may have an analytic solution and yet have a very incomplete understanding of the behavior of the system. To make this point clear, it is only necessary to consider the simple equation

$$\frac{d^2x}{dt^2} + x = 0. \tag{1.3.1}$$

This has the solution

$$x = A\cos(t) + B\sin(t), \tag{1.3.2}$$

where A and B are arbitrary constants, provided that the functions in (1.3.2) are defined by

$$\begin{aligned} \cos(t) &= \sum_{n=0}^{\infty} (-1)^n t^{2n}/(2n)! \\ \sin(t) &= \sum_{n=0}^{\infty} (-1)^n t^{2n+1}/(2n+1)! \end{aligned} \tag{1.3.3}$$

It is, of course, 'known' that these functions are periodic, satisfying

$$\cos(t + 2\pi) = \cos(t); \quad \sin(t + 2\pi) = \sin(t). \tag{1.3.4}$$

This periodicity is obviously a very basic and important feature of the solutions of (1.3.1), and establishing it is therefore of crucial importance.

Exercise 1.3. It is left as a challenge for you to prove that (1.3.4) follows from the solutions (1.3.3) (i.e., to show that a 'π' exists such that (1.3.4) holds).

This 'simple' feature (1.3.4) does not come most readily from the analytic expressions (1.3.3) but, rather, from other methods. These methods may, moreover, be applied as easily to the nonlinear oscillator (e.g., if $F(-x) = -F(x)$ and $F(x) > 0$ if $x > 0$)

$$\frac{d^2x}{dt^2} + F(x) = 0 \tag{1.3.5}$$

as they can be applied to (1.3.1), and hence a basic property of $x(t)$ satisfying (1.3.5) is not made any more transparent if $F(x) = x$ (i.e., if the equation is linear).

It might be tempting to argue that, given (1.3.3) and a computer, it should be relatively easy to establish (1.3.4) (and determine π?!). It is obvious, however, that a computer (alone) can never prove such an equality. It can only show that

$$\cos(t + 2\text{'}\pi\text{'}) \simeq \cos(t) \tag{1.3.6}$$

for some 'π', and to within some accuracy. The difference between (1.3.6) and (1.3.4) becomes of more than theoretical interest if we want to know if

$$x(t + n2\text{'}\pi\text{'}) \simeq x(t)$$

for a large number, say $n = 10^5$. Moreover, since the answer can be established without using a 'sledgehammer on a peanut' approach, it is clear that the computer (alone!; see below) is quite inappropriate for such questions, and indeed for many 'qualitative' features of solutions.

This point illustrates the importance of using a variety of approaches to analyzing systems, as outlined in the last section. In particular, the periodicity property of (1.3.5) and (1.3.1) is established by a simple phase plane ('topological') analysis. However, particularly in more complicated cases, the use of the computer together with continuity and topological considerations can be a very useful blending of the approaches in analyzing systems. The computer can be used to determine effectively whether some condition has changed between two regions of space, from which conclusions can be drawn about the situation at some intermediate point, based on continuity arguments. Numerous examples of this useful method will be discussed later.

Exercise 1.4. Higher order equations require more ingenuity. For example, three independent solutions of $d^3x/dt^3 + x = 0$ are, in analogy with (1.3.3), $\text{ain}(t) \equiv t - t^4/4! + t^7/7! - \cdots$, $\text{bin}(t) \equiv t^2/2! - t^5/5! + t^8/8! - \cdots$, and $\text{cin}(t) \equiv 1 - t^3/3! + t^6/6! - t^9/9! - \cdots$. Are any or all solutions of this equation periodic?

1.4 Remarks on modeling

The art of modeling physical systems by suitable systems of equations is, of course, an old and yet surprisingly new practice. It is not only that new systems on which to practice this art are always being considered, but also that the possible objectives of such modeling have taken on a new dimension in recent years. As any good book on modeling will tell you (see the references for a start), we begin by trying to identify the physical variables which we believe are responsible for the phenomena in question, and their interrelations, in order to construct a deterministic system of equations. This step is clearly of major importance. We need only to consider the complexity of gases and fluids, economic predictions, weather forecasting, complex chemical reactions, neurological networks (the brain!), to appreciate this step. The idea is then to check