Foundations of Mechanics

Now that we have become familiar with the content and applications of mechanics, we are prepared to examine its conceptual foundations systematically. This calls for an explicit formulation and analysis of all presuppositions of the theory. It goes beyond a mere statement of Newton's laws to an analysis of the status of laws in a theory and nature of scientific theories in general. This kind of study belongs to the philosophy of science, but it is no mere academic exercise. The profound revolutions in physics due to Newton and Emstein were changes in the conceptual foundations resulting from careful analysis. So it takes a study of foundations to fully understand the evolution of physics, or, if the facts demand, to instigate a new revolution. Improvements in the foundations are truly revolutionary, because they are so rare and their repercussions are so extensive, bearing on every application of the theory.

Newton's original formulation of mechanics nearly 300 years ago is followed with little change in most mechanics books even today. Nevertheless, it is not entirely satisfactory for several reasons. First, it is incomplete in the sense that not all major assumptions of the theory are explicitly spelled out. Second, in the last century Newtonian theory has undergone profound modifications and extensions which should be taken into account. To begin with, Einstein's Theory of Relativity has revolutionized the scientific concepts of space and time. We now know that any adequate formulation of space and time has empirical content with testable consequences. So a clear and explicit formulation of these concepts is scientifically as essential to Newtonian mechanics as it is to relativity theory. Pedagogically, it is needed to help students distinguish between their own vague intuitions of space and time and an objective scientific formulation of these concepts. Fortunately, the formulation can be designed so a small change in the concept of simultaneity generates a smooth transition from Newtonian mechanics to relativistic mechanics.

Another big change in mechanics since Newton has been brought about by the development of the field concept. Even introductory physics courses move rapidly from interactions between particles to interactions of particles with electric and magnetic fields. We need a formulation of Newton's laws which readily accomodates this profound theoretical change. We need to provide for a smooth transition from pure particle mechanics to the classical theory of fields and particles.

A modern formulation of mechanics should also incorporate profound changes in the concept of a theory which have evolved since Newton. Today it is widely recognized that physics is concerned with constructing and testing mathematical models of physical systems. Thus, the concept of a mathematical model is central to the modern conception of a scientific theory. Yet physics textbooks scarcely mention models, let alone explain that mathematical modeling is the essential core of the scientific method.

1. Models and Theories

> Philosophy is written in that great book which ever lies before our eyes – I mean the Universe – but we cannot understand it if we do not first learn the language and grasp the symbols in which it is written. This book is written in the mathematical language, and the symbols are triangles, circles, and other geometrical figures, without whose help it is impossible to comprehend a single word of it; without which one wanders in vain through a dark labyrinth.*
>
> — Galileo Galilei

This magnificent passage is the capstone of Galileo's great intellectual achievements. It is the first incisive formulation of a philosophical viewpoint which played a crucial role in the development of modern science. This viewpoint has been so thoroughly assimilated into modern science that most scientists take it for granted without recognizing that a profound issue is involved. On the other hand, it is still debated endlessly in philosophical circles, where it is called *scientific realism*. The importance that Galileo himself attached to the above passage is clear from his order that it be placed at the head of his collected works.

Scientific realism must be distinguished from the *naive realism* of common sense. The presumption common to all forms of realism is that a "real world" of things exists independently of any person to observe them. According to common sense, things in the real world are just as we see them; they are known to us directly through experience, provided the senses are operating properly so the view is not distorted. But, as Galileo puts it, scientific realism holds that the real world is known only indirectly; it is merely posed to us through the senses as a cipher, so to know real things we must decode the messages of experience. Moreover, the code can be broken only by recognizing that geometrical properties of things are primary, and we can know them only conceptually by representing them mathematically.

Galileo's profound scientific realism evolved from long contemplation and a variety of astute observations. Throughout his writings Galileo was occupied with an analysis of experience to distinguish the "primary properties" essential to real objects from "secondary properties" which depend on the mode of human sensation. The analysis was continued by Descartes and Boyle among others, and it was a crucial preliminary to Newton's definitive formulation of mechanics in the *Principia*, from which all reference to secondary properties was banished. This decisive step severed psychology cleanly from physics, enabling physics to progress without being distracted by the complexities of subjective experience. It is the basis today for such distinctions as between the perceived color of light (a secondary property) and the frequency of light (a primary property), or the pitch of a tone and its frequency. The properties ascribed to objects by physics, such as mass, velocity, force and frequency, are very different from the directly perceived properties of things. Physical properties are primary properties which can be represented as quantities. Thus, the distinction between primary and secondary properties was a crucial preliminary to developing a mathematical theory of the real world.

In this chapter we adopt a modern version of scientific realism, which holds that objective knowledge about the real world is obtained by developing validated mathematical models to represent real objects. Scientific realism maintains a sharp distinction between a physical thing and its model, between the real world of physical things and the mental world of concepts. One should realize, however, that *this dualism is only methodological*. It by no means requires that the physical and mental worlds exist independently of one another. It is entirely compatible with an explanation of mental phenomena in terms of physical brain states. Indeed, the distinction between primary and

* Translation from p. 67 of E. A. Burtt, *The Metaphysical Foundations of Modern Science*, Routledge and Kegen Paul LTD, London (1932). Burtt gives a historical account of the origins of scientific realism.

secondary properties opens the possibility of explaining secondary properties in terms of primary properties. But this is an issue for neuropsychology to investigate. What matters here is that scientific realism holds that a clear distinction between physical things and their models can be made and must be maintained against the contrary tendencies of natural language which is infected with naive realism.

Scientific realism has been vigorously challenged recently by physicists and philosophers who hold that it is incompatible with quantum mechanics. They claim that quantum mechanics does not allow a sharp separation between the state of a real object and an observer's knowledge of that state. We cannot get involved in that debate here. Suffice it to say that the issue has not been resolved to the satisfaction of all concerned physicists. Without further apology, in this chapter we strive for a sharply formulated theory of scientific knowledge from the viewpoint of scientific realism.

Models

The term "model" is often used in the scientific literature with only a vague meaning. To sharpen the concept of model, we need terminology which expresses clear distinctions and specifications. We assume that a *model* is a conceptual representation of a real object. The represented object is said to be a *referent* of the model. A model may have more than one referent. For example, a model of the hydrogen atom has all hydrogen atoms for referents, while a model of the solar system has a single referent. The set of all referents of a model is called its reference class. If its reference class is empty. a model is said to be *fictitious*. An assignment of a particular referent or *reference class* to a given model is called a *factual interpretation* of the model, or a *physical interpretation* if the model belongs to physics. A single model may be given many different factual interpretations, especially in a mature science like physics. For example, the one-dimensional harmonic oscillator may be interpreted as a model for such diverse objects as an elastic solid, a pendulum, a diatomic molecule or an atom.

We are concerned here with mathematical models, though much of our discussion applies more generally. A **mathematical model** has four components:

(1) A set of names for the object and agents that interact with it, as well as for any parts of the object represented in the model.
(2) A set ot descriptive variables (or descriptors) representing properties of the object.
(3) Equations *of the model*, describing its structure and time evolution.
(4) An interpretation relating the descriptive variables to properties of objects in the reference class of the model.

Each of these components needs some explication.

Numerals are often used as object names; thus, we may speak of "particle 1" and "particle 2." Descriptive variables are functions of the object names, since each descriptor represents a property of a particular object. For example, the velocity descriptor $\mathbf{v}_k$ for the kth object in a system is an explicit function of the object name k. Often, however, the dependence of descriptors on object name is tacitly understood, as when we write $\mathbf{v}$ for the velocity of some object.

There are **three types of descriptors:** object variables, state variables and interaction variables.

Object variables represent intrinsic properties of the object. For example, mass and charge are object variables for a material particle, while moment of inertia and specifications of size and shape are object variables for a rigid body. The object variables have fixed values for a particular object, but they have different values for different objects, so they are indeed variables from a general modeling perspective.

State variables represent intrinsic properties with values which may vary with time. For example, position and velocity are state variables for a particle. A descriptor regarded as a state variable in one model may be regarded as an object variable in another model. Mass, for example, is a state variable in a model that allows it to change, though it is usually constant in particle models. Thus, object variables can be regarded as state variables with constant values.

An **interaction variable** represents the interaction of some external object (called an *agent*) with the object being modeled. The basic interaction variable in mechanics is the force vector; work, potential energy and torque are alternative interaction variables.

Different kinds of property can be distinguished by characteristics of their representations as descriptive variables. A property is said to be *quantitative* if it can be represented by mathematical *quantities*, such as elements of the Geometric Algebra. Otherwise it is said to be *qualitative*. Physics is concerned with a particular set of quantitative properties called *physical properties*. The corresponding descriptors are called *physical variables* or *physical quantities*. The equations of a mathematical model describe relations among quantitative properties. Equations determining the time evolution of the state variables are called *dynamical equations*, or *equations of motion* in mechanics. In a mature scientific theory, the equations are derived from laws of the theory. Otherwise, they must be assumed as hypotheses subject to verification.

It is common practice in the literature to say that a particular dynamical equation constitutes a mathematical model. This should be recognized as a loose use of language, for an equation represents nothing unless its variables are given factual interpretations.

The **interpretation** of a model is specified by a set of attribute functions for its properties. The set of objects with a given property is called the *scope* or *reference class* of that property. The *attribute function* for a property assigns particular values of the *descriptive variable* to objects in its reference class. When specific numerical values are assigned to certain variables, these variables are said to be *instantiated*. As examples of *instantiation* in particle mechanics, we have the assignment of a particular mass to a particle or particular initial conditions for its trajectory.

When, for specific instantiations, the equations of a model are sufficient to determine specific values for all its descriptors, the model is said to be a *specific model*. A specific model can thus describe a particular object under particular circumstances.

Theory

Evidently we have tacitly employed a theory of some sort in specifying the general characteristics of a model. A vaguely defined theory of this sort is frequently called *Systems Theory* in the scientific literature; although it is seldom formulated in the generality we need here. We may regard Systems Theory as a theory of theories, or more specifically, a general theory of mathematical models. Thus, Systems Theory specifies the characteristics of models common to all scientific theories. Consider, for example, the distinction between state variables and interaction variables in a model. That distinction was first sharply drawn in mechanics. But, as other theories developed, many people noticed that the distinction has a wider significance if the concepts of state and interaction are suitably generalized. Based on this distinction, Systems Theory goes on to describe how complex objects can be modeled as systems of interacting parts. Thus, it provides a general theory of structure and composition of objects of any kind. This too is a generalization of concepts developed in physics. A complete development of Systems Theory will not be attempted here. However, the general characterization of a scientific theory, to which we now turn, may be regarded as part of Systems Theory.

A **scientific theory can be regarded as a system of design principles for modeling real objects**. The theory consists of:

I. A **framework** of generic and specific laws charcterizing the descriptive variables of the theory.
II. A **semantic base** of correspondence rules relating the descriptive variables to properties of real objects.
III. A **superstructure** of definitions, conventions and theorems to facilitate modeling in a variety of situations.

The mathematical language used to formulate a theory is usually taken for granted. However, it should be recognized that most of the mathematics used in physics was developed to meet the theoretical needs of physics. In Chapter 1 of NFCM, we saw that this is true of the real number system and its generalization to Geometric Algebra. Moreover, differential equations were first invented to formulate dynamical laws of physics. The moral is that the symbolic calculus (mathematics) employed by a scientific theory should be tailored to the theory, not the other way around.

The key concept in a scientific theory is the concept of a scientific law, so it should be explicated carefully. A *scientfic law* is a relation or system of relations among descriptive variables presumed to represent an objective relation or pattern among the corresponding properties. If the relation is among physical variables, it is called a *physical law*. Most physical laws are formulated as mathematical equations. Scientific realism maintains that it is important to distinguish between a law and the objective pattern it represents, because the latter is an unchanging property of the real world while the former may be changed when we understand the world better. Moreover, a law may be true or false or approximately true, but the property pattern it is presumed to represent just "is". To qualify as a law, a relation among descriptive variables must represent a property which is *universal* in the sense that its scope is not limited to a finite number of objects, and it must be *corroborated* in some empirical domain by scientific methods. A proposed law which has not been experimentally tested and confirmed is called a *hypothesis*. Thus, a law is a corroborated hypothesis.

There are several types of law. The *generic laws* of a theory define the basic descriptive variables of the theory. The generic laws of Classical Mechanics fall into two groups: (a) The *Zeroth Law*, which defines the concepts of position, motion and composition of bodies, and (b) *Dynamical Laws* (Newton's Laws), which implicitly define the concepts of mass and force. The Zeroth Law is so general that it belongs to every physical theory; indeed, it is presumed (tacitly at least) in *every* scientific theory. The Dynamical Laws apply only to material objects. In Sections 2 and 3, these laws will be formulated and discussed in detail.

The specific laws of a theory specify relations among the descriptive variables defined by the generic laws. As a rule they apply only to special circumstances, whereas the generic laws are presumed to hold in every application of the theory. The specific laws of *Classical Mechanics are interaction laws* such as Coulomb's Law, Newton's Law of Gravitation and Stokes' Law of fluid friction.

Taking the Zeroth Law for granted, the other basic laws of any scientific theory can be classified into *dynamical laws*, which determine the time evolution of state variables, and *interaction laws*, which interrelate the state variables of different objects.

The *basic laws* of a theory are included in the theory by assumption. The superstructure of the theory also contains *derived laws*, such Galileo's law of falling bodies. As a rule, the scope of a basic law is much wider than the scope of a derived law.

We must be clear about what it means to say that concepts like motion and mass are *defined* by generic laws. All sorts of unnecessary difficulties are caused by a sloppy or inadequate concept of definition, so it will be worth our while to explicate the concept. The purpose of a definition is to establish the meaning of a concept (or the term (symbol) which designates it) by specifying its relation to other concepts (terms). When this has been done, we say that the concept (term) is well-defined. There are two ways to do it, yielding *two kinds of definition: explicit and implicit*.

A concept is *defined explicitly* by expressing it in terms of other concepts. This is the conventional notion of definition, used, for example, in defining the kinetic energy K by the equation $K = mv^2/2$.

A concept (term) is *defined implicitly* by a set of *axioms* which relate to it other concepts (terms). Thus, the concept of "point" is defined by the axioms of geometry which specify its relations to other points, lines and planes. Similarly the concept of "vector" is defined implicitly by specifying how to add and multiply vectors. In each case axioms define concepts by specifying relations. Axioms are set apart from other statements or equations by accepting them as definitions, so they need not be proved. Nevertheless, terms like "point" and "vector" introduced by axioms are commonly said to be "undefined terms." This is a misleading expression that ought to be discarded. Novices often interpret it in the sense of "ill-defined" or "obscure." At least they find it unnecessarily mysterious. Evidently it conflicts with established usage of the term "well-defined." It would be better to say that "some terms in a theory must be defined implicitly" rather than "some terms must be undefined."

Generic laws are axioms defining basic descriptive variables. Our definition of model might have given the impression that descriptive variables can be defined independently of any laws. But why are descriptive variables scalar- or vector-valued, that is what makes them quantitative? It will be seen that this is a consequence of the Zeroth Law, which introduces geometrical attributes into every physical theory. The generic laws of space and time are usually taken for granted, so they are seldom mentioned in the formulation of a model. They are essential, nevertheless. A variable which is undefined by laws is completely nondescript; it is no more than a name. To be definite concepts, descriptive variables must be well-defined by laws.

Newton's Laws are sometimes called axioms. That invites confusion between the purely mathematical concept of an axiom and the factual concept of a law. A law is an axiom, but the converse is not true. *A physical law is an axiom with a physical interpretation.*

The correspondence rules of a theory determine factual interpretations for its descriptive variables and laws, and so for models designed with it. They include operational procedures for *measurement*, that is, the assignment of particular values for the descriptors of particular objects. Thus, they determine attribute functions relating descriptors to the properties they represent. The correspondence rules are not independent of physical laws; rather they are specified in accordance with the laws. For example, any operational procedure for measuring length must be consistent with the Euclidean properties of physical space, as specified by the Zeroth Law. Moreover, the laws of physics often enable us to measure the same physical quantity in many different ways, so the results of measurement must be independent of the particular procedure employed.

A correspondence rule for measuring a physical quantity is often called an "operational definition." But this is an abuse of language, confusing the concepts of definition and measurement A definition, whether explicit or implicit, relates concepts to concepts, not concepts to things. Mario Bunge has suggested that the term "operational definition" be replaced by "operational referition," since it is concerned with the semantic concept of reference; it relates a descriptor (a concept) to its referents (things).

The set of real objects which can be modeled with a theory is called the *reference class* of the theory. The reference class of Classical Mechanics is enormous. The set of all material bodies. Yet the generic laws of Mechanics model a very small number of properties. The theory asserts that these are properties that all bodies have in common, so we call them *basic properties*. The fact that the generic laws describe only basic properties does not mean that other properties cannot be described by the theory. A composite body has new properties not possessed by its parts which emerge when it is assembled. They are called *emergent properties*. This challenges theory to explain emergent properties in terms of basic properties. Indeed, it challenges physicists to explain all physical properties of matter—geometrical, mechanical, electrical, thermodynamic, optical—in terms of a small number of basic properties. This grand challenge has long been a major motivation for research.

The emergent geometrical properties of size and shape can be explained in terms of basic properties by the Zeroth Law, which incorporates the physical content of Greek geometry. Geometry can be regarded as the theory of size and shape. This may be obvious, but it is far from trivial, as witnessed by the whole field of architectural design. The Kinetic Theory of Gases is a subtheory of Mechanics which explains temperature as an emergent property. The problem of explaining all thermodynamic properties as emergent from physical properties of molecules is so complex that a separate theory, Statistical Mechanics, has been developed to handle it. More specialized theories like Plasma Physics, Solid State Physics and Theoretical Chemistry are also concerned with explaining emergent properties. All these theories are founded on Classical Mechanics as well as Quantum Mechanics.

Having discussed the general features of models and theories, let us turn now to a formulation of generic laws for Classical Mechanics.

2. The Zeroth Law of Physics

Everyone has well-developed notions of space and time abstracted from personal experience. Perceptual categories of space and time are essential for sorting out sensory data. However, perceptual space and time must sharply be distinguished from the concepts of physical space and time. The former is a *modus operandi* of the human brain — the proper study of psychology, psychophysics and neuroscience. It provides an intuitive base for the physical concepts. But the concepts of physical space and time are objective rather than intuitive. Intuitive concepts are subjective, which is to say that they vary from person to person; whereas objective concepts are the same for everyone. Objectivity is achieved in science by providing concepts with explicit mathematical definitions and factual interpretations in terms of rules which might be applied by anyone, or by a computer for that matter. Of course, everyone's conception of space and time combines intuitive and objective components. But only the objective component will concern us here.

Objective concepts evolve with changes in their definitions and interpretations. Since Newton's day two major improvements in the concepts of space and time have evolved which should be incorporated into the foundations of mechanics. First, we have learned to distinguish between mathematical and physical geometries. Scientific realism regards physical geometry as a feature of the real world which we model with a mathematical geometry. Thus, our model geometries should be subjected to empirical tests. In Newton's day no one had conceived of an alternative to Euclidean geometry or the idea of testing it, though, of course, it had been subjected to many crude informal tests when employed in architectural design and construction. Alternatives to Euclidean geometry were first conceived by mathematicians in the nineteenth century, but none was incorporated into a viable physical theory until Einstein's General Theory of Relativity in the twentieth century. We shall formulate a Euclidean model of physical geometry, since that is appropriate for classical mechanics. But we aim to do it in a way which makes its "physical content" explicit, and allows for easy generalization to "relativistic theories."

The second major improvement in concepts of space and time is due mainly to Einstein. He recognized that the concept of *distant simultaneity* is an essential part of the time concept which had not previously been explicitly defined in classical physics. Rather, physicists had unwittingly adopted an implicit concept of simultaneity which was inconsistent with ideas of causality and experimental fact. By supplying an appropriate definition of distant simultaneity and analyzing its consequences, Einstein created his Special Theory of Relativity. Thus, the Special Theory is best regarded as a completion of classical physics with a full elucidation of the time concept.

The change instituted by Einstein in the classical time concept appears to be comparatively small, but its consequences are immense. It implies that space and time are relative concepts which cannot be defined independently of one another and do not correspond to unique features of the real world. It implies that the real physical geometry is a non-Euclidean geometry of a 4-dimensional entity space-time, with respect to which the separate concepts of space and time only describe the viewpoint of a particular observer. Thus, a small change in the time concept has profoundly altered the physicists' conception of reality.

The Special Theory of Relativity is discussed in the second edition of NFCM. Here we will be content with preparing the way for a smooth transition to the modern spacetime concept by elucidating the classical concepts of space and time. We begin with the concept of space.

The problem of providing the concept of space with a precise mathematical formulation has been solved to nearly everyone's satisfaction. But physicists are still far from agreement on the physical status of space. Is space a thing or a property of things? Or is it a property of the human mind, a "category of the understanding," as the philosopher Immanual Kant proposed? Every kind of answer can be found in the literature. This attests to widespread confusion about the conceptual foundations of physics. Confusion is perpetuated by an outmoded concept of space which infects our natural language. Thus, we speak of physical objects *in* space as if space were a container with an existence independent of its contents. The literature shows that physicists are not immune to this infection, but a cure can be achieved by a careful conceptual analysis. The source of the infection is easy to identify. The natural language was developed to describe features of perceptual experience, which it can do with remarkable fidelity. The brain does, indeed, contain a *sensorium*, a carrier of perceptions which exists independently of its contents. This is reflected in perceptual experience and so in the natural language. Thus, a cure for confusion about the nature of space begins with a clear distinction between the perceptual space of subjective experience and the objective concept of physical space. The complete cure requires a rigorous formulation of the physical concept in perfect accord with experimental practice.

To ascertain a suitable physical interpretation for the concept of space, we must examine the role of geometry in experimental practice. We note that every measurement of distance determines a relation between two objects. Every measurement of position determines a relation between one object and some other object or system of objects. In accordance with the standpoint of scientific realism, we regard such measured relations as representations of real properties of real objects. These are mutual (or shared) properties relating one object to another. We call them *geometrical properties*. We are now prepared for an explicit formulation of physical space as a system of relations among physical objects.

To begin with, we recognize two kinds of objects, *particles* and *bodies* which are composed of particles. Given a body $\mathcal{R}$ called a *reference frame*, each particle has a geometrical property called *its position with respect to* $\mathcal{R}$. We characterize this property indirectly by introducing the concept of *Position Space*, or *Relative Space*, if you prefer. **For each reference frame $\mathcal{R}$, a position space $\mathcal{P}$ is defined by the following postulates:**

A. **$\mathcal{P}$ is a 3-dimensional Euclidean space**.
B. **The position (with respect to $\mathcal{R}$) of any particle can be represented as a point in $\mathcal{P}$**.

The first postulate specifies the mathematical structure of a position space while the second postulate supplies it with a physical interpretation. Thus, the postulates define a physical law, for the mathematical structure implies geometrical relations among the positions of distinct particles. Let us call it the **Law of Spatial Order**.

Notice that this law asserts that every particle has a property called position and it specifies properties of this property. But it does not tell us how to measure position. Measurement is a separate matter, since it entails correspondence rules as well as laws. In actual practice the reference frame is often fictitious, though it is related indirectly to a physical body. Our discussion is simplified by feigning that the reference frame is always a real body.

We turn now to the problem of formulating the scientific concept of time. We begin with the idea that time is a measure of motion, and *motion is change of position with respect to a given reference frame*. The concept of time embraces two distinct relations: temporal order and distant simultaneity. To keep this clear we introduce each relation with a separate postulate.

First we formulate the **Law of Temporal Order:**

> **The motion of any particle with respect to a given reference frame can be represented as an orbit in position space.**

This postulate has a semantic component as well as a mathematical one. It presumes that each particle has a property called motion and attributes a mathematical structure to that property by associating it with an *orbit* in position space. Recall that an orbit is a continuous, oriented curve. Thus, a particle's orbit in position space represents an ordered sequence of positions. We call this order a *temporal order*, so we have attributed a distinct temporal order to the motion of each particle.

To define a physical time scale as a measure of motion, we select a *moving* particle which we call a *particle clock*. We refer to each successive position of this particle as an *instant*. We define the *time interval* Δt *between two instants* by

$$\Delta s = c\Delta t\,,$$

where c is a positive numerical constant and Δs is the arclength of the clock's orbit between the two instants. Our measure of time is thus related to the measure of distance in position space.

To use this time scale as a measure for the motions of other particles, we need to relate the motions of particles at different places. The necessary relation can be introduced by postulating the **Law of Simultaneity:**

At every instant, each particle has a unique position.

This postulate determines a correspondence between the points on the orbit of any particle and points on the orbit of a clock. Therefore, every particle orbit can be parametrized by a time parameter defined on the orbit of a particle clock.

Note that this postulate does not tell us how to determine the position of a given particle at any instant. That is a problem for the theory of measurement.

So far our laws permit orbits which are nondifferentiable at isolated points or even at every point. These possibilities will be eliminated by Newton's laws which require differentiable orbits. We include in the class of allowable orbits, orbits which consist of a single point during some interval. A particle with such an orbit is said to be *at rest* with respect to the given reference frame during that interval. Of course, we require that the particles composing the reference frame itself be at rest with respect to each other, so the reference frame can be regarded as a rigid body.

Note that the speed of a particle is just a comparison of the particle's displacement to the displacement of a particle clock. The speed of the particle clock has the constant value $c = \Delta s/\Delta t$, so the clock moves uniformly by definition. In principle, we can use any moving particle as a clock, but the dynamical laws we introduce later suggest a preferred choice. It is sometimes asserted that a periodic process is needed to define a clock. But any moving particle automatically defines a periodic process, because it moves successively over spatial intervals of equal length. It should be evident that any real clock can be accurately modeled as a particle clock. By regarding the particle clock as the fundamental kind of clock, we make clear in the foundations of physics that the scientific concept of time is based on an objective comparison of motions.

We now have definite formulations of space and time, so we can define a *reference system* as a representation $\mathbf{x}$ for the possible position of any particle at each time t in some time interval. Each reference system presumes the selection of a particular reference frame and particle clock, so $\mathbf{x}$ is to be interpreted as a point in the position space of that frame. Also, a reference system presumes the selection of a particular origin for time and space and particular choices for the units of distance and time, so each position and time is assigned a definite numerical value. The term "reference system" is sometimes construed as a system of procedures for constructing a numerical representation of space and time.

After we have formulated our dynamical laws, it will be clear that certain reference systems called *inertial systems* have a special status. Then it will be necessary to supplement our *Law of Simultaneity* with a postulate that relates simultaneous events in different inertial systems. That is the critical postulate that distinguishes Newtonian theory from Special Relativity, but we defer discussion of it until we are prepared to handle it completely. It is mentioned now, because our formulation of space and time will not be complete until such a postulate is made.

It is convenient to summarize and generalize our postulates with a single law statement, the **Zeroth** (or *Spatiotemporal*) **Law of Physics:**

Every real object has a continuous history in space and time.

To explicate this law, we assert that it has four major components:

1. *The Law of Spatial Order.*
2. *The Law of Temporal Order.*
3. *The Law of Simultaneity.*
4. *The Generic Law of Composition.*

The **Generic Law of Composition** asserts:

> **The properties of any real object can be represented mathematically by the values of a state function defined on the position and time variables of a given reference system.**

A *Specific Law of Composition* imposes some condition on the nature of the state function for a particular object or class of objects. The successive values of the state function as a function of time describes the history of the object. The Zeroth Law does not specify the *history* of any object; dynamical laws determining the state function are needed for that. But it does assert that every object has a history.

In classical physics, *every model of a real object is one of three kinds: particle, body or field.* Each model is distinguished by a particular state function. We have already specified the state function for a particle, namely, the function $\mathbf{x} = \mathbf{x}(t)$ for its orbit in position space. A *material particle* also has a property called *mass*, so a complete state function must specify any time variation of the mass.

A *body* is an *extended object*, which is to say that more than one point in Position Space is required to specify its location. We have modeled bodies as systems of particles. In this case, the *location* of a body is the set of positions of its particles, and its history is the set of particle histories. Alternatively, a *material body* (or *material medium*) can be modeled as a spatially continuous object which does not have a unique decomposition into particles.

A *field* is also an extended object, but its state function as well as its physical interpretation is quite different from that of a body. Let us note here that the *Classical Theory of Fields* asserts the existence of real objects called electric fields, each of which can be represented by a vector-valued state function $\mathbf{E}(\mathbf{x}, t)$. The Zeroth Law applies to Quantum Mechanics as well as classical physics, but the state functions for particles are different.

The Zeroth Law is the most universal of all scientific laws. It asserts that every real thing that ever existed or will exist has definite spatiotemporal properties, that is, definite spatiotemporal relations to every other real thing. Some aspect of the Zeroth Law is presumed in every scientific theory and investigation.

Other scientists can take the Zeroth Law for granted, but physicists are responsible for refining its formulation and testing its consequences. The present formulation has been designed for compatibility with the Special Theory of Relativity, and it can be directly generalized to the "curved spacetime" of the General Theory of Relativity. The mathematical structure attributed to space and time in our formulation is widely accepted by physicists, but the physical interpretation is controversial. We have adopted a *relational view*, interpreting space and time as a system of relations among real objects. But some physicists prefer a *material view*, interpreting spacetime as a primal material out of which all things are composed, so that objects can be regarded as local variations in the properties of spacetime. Thus, the material view interchanges the objects and properties of the relational view. Is there a definite empirical distinction between these two interpretations so we can decide on one over the other? That is a profound question which will not be easily answered. At least both interpretations are consistent with scientific realism.

3. Generic Laws and Principles of Particle Mechanics

The spatiotemporal properties of real objects are described by the Zeroth Law. To produce a complete physical theory, the Zeroth Law must be supplemented by a set of dynamical laws which describe the nature and effect of interactions between objects. In Particle Mechanics the *interaction property* is represented by force functions. A set of *generic laws* implicitly define the concepts of mass and force and assign them a physical interpretation. To produce a specific model of interacting particles, the generic laws must be supplemented by *specific force laws* which specify definite force functions.

Our formulation of the general theory consists of four generic laws, one hypothesis and three generic principles. Let us present them all at once, and then comment on each one separately. Of course, our formulation presumes the Zeroth Law, so the notions of particle, time, position, velocity and acceleration are all well-defined. In addition, the formulation is presumed to hold only for a certain kind of reference system called an *inertial system*, which is implicitly defined by the First Law. Now we are ready.

First Law (*Law of Inertia*):

> **In an inertial system, every free particle has a constant velocity**. A particle is said to be *free* if the total force on it vanishes.

Second Law (*Law of Causality*)

> **The total force exerted on a particle by other objects at any specified time can be represented by a vector f such that**
>
> $$\mathbf{f} = m\mathbf{a}\,,$$
>
> where $\mathbf{a}$ is the particle's acceleration and *m is a positive scalar constant* called the *mass* of the particle.

Third Law (*Law of Reciprocity*):

> **To the force exerted by any object on a particle there corresponds an equal and opposite force exerted by the particle on that object.**

Fourth Law (*Superposition Law*):

> **The total force f due to several objects acting simultaneously on a particle is equal to the vector sum of the forces $\mathbf{f}_k$ due to each object acting independently, that is**
>
> $$\mathbf{f} = \sum_k \mathbf{f}_k\,.$$

To relate formulations of the laws in different inertial systems, we adopt the

Hypothesis of Absolute Simultaneity:

> *Local events which are simultaneous in one inertial system are simultaneous in every inertial system.*

A *local event* is defined as a change in the position or velocity of a particle.

Specific force laws need not be regarded as part of the general theory. However, they are restricted in form by generic principles. The principles function as laws when force functions are unknown. In particular, they sharpen the general concept of force defined by the generic laws. However, when we have specific force laws that satisfy the principles, the principles are superfluous. For this reason we do not call them laws. In Section 3-1 we introduced *The Principle of Analyticity.* Two other principles are important:

The Principle of Local Interaction:

The force on a particle at any time is a unique function of particle position and position time derivatives; it is independent of the particle's past or future history.

The Principle of Relativity:

The laws of mechanics have the same functional form in all inertial systems.

Comments on the First Law

The First Law implicitly defines a time scale for an inertial system. For it requires that displacements of the system's particle clock are proportional to displacements of a free particle. This amounts to requiring that equal intervals of time be *defined* by equal displacements of a free particle. Thus, the motion of a free particle determines the time scale for an inertial system up to a multiplicative constant. This fundamental kind of scale is called an *inertial time scale*.

Besides determining a time scale, the First Law associates straight lines in an inertial system with free particle motion. So within an inertial system, the deviation of any particle from uniform motion in a straight line can be attributed to the action of other objects in accordance with the Second Law. The First Law defines an inertial system implicitly by specifying a physically-grounded criterion which distinguishes it from noninertial reference systems. It tells us that an inertial frame can be identified in principle by examining the motion of free particles. In practice, such a procedure is usually impossible. For inertial frames do not occur naturally, so most measurements are done with respect to an accelerated reference frame. And free particles are not usually available for experiments either. Such practical difficulties should not be construed as casting doubt on the utility of the First Law. Rather, they pose the experimentalist with the problem of distinguishing real forces, due to the interaction of objects, from pseudoforces, due to the acceleration of his reference frame. He needs the First Law to make such a distinction, but he needs the other Laws as well.

The First Law is evidently not independent of the other Laws, because they are needed to define what is meant by "free particles."

Comments on the Second Law

The formula $\mathbf{f} = m\mathbf{a}$ by itself is frequently presented as a complete statement of the Second Law. Although such a formula is an acceptable mathematical axiom, a law statement should include a physical interpretation of the mathematical terms it employs. Of course, an interpretation for $\mathbf{f}$ could be supplied by a separate postulate, but it is best included in the Second Law since that is where $\mathbf{f}$ first appears in the theory.

Thus, our formulation asserts that $\mathbf{f}$ represents a physical property called "force" which a particle shares with other objects. The vector $\mathbf{f}$ itself is commonly referred to as "the force on a particle," so $\mathbf{f}$ serves also as a name for the property it represents. The term "objects" in our law statement is presumed to be defined previously by the Laws of Composition as part of the Zeroth Law. As said before, we recognize three kinds of objects: particles, bodies and fields, and this reduces to two basic kinds if bodies are modeled as systems of particles. To produce a pure particle theory, one need only replace the word object with the word particle in all our law statements. But our formulation generalizes particle theory by allowing interactions with fields. Indeed, in a pure field theory particles never interact directly, but only through the intermediary of a field. In that case the term "object" in our law statements should always be interpreted as "field," and we need additional laws to fully describe the properties of fields.

It is often claimed that $\mathbf{f} = m\mathbf{a}$ is a definition of force. On the contrary, an explicit definition of force is impossible. Rather, the complete set of generic laws is required to define $\mathbf{f}$ implicitly by specifying the common characteristics of all forces. The equation $\mathbf{f} = m\mathbf{a}$ represents only one characteristic of force. It relates the general property of interaction to the general spatiotemporal property of motion. The $\mathbf{f}$ represents the action of the universe on a particle while the $m\mathbf{a}$ represents the particle's response with a change in its state of motion. This provides us with a physical interpretation of mass as a measure of the strength of a particle's response to a given force. No other definition or interpretation of mass is needed in the theory.

Comments on the Third Law

For two interacting particles, the Third Law can be written

$$\mathbf{f}_{12} = -\mathbf{f}_{21} \,,$$

where $\mathbf{f}_{12}$ is the force of particle 2 on particle 1. This is called the *weak form* on the Third law in Section 6-1 of NFCM. One readily verifies that this relation is satisfied by Newton's gravitational force law and the similar Law of Coulomb. However, it fails for direct magnetic interactions between charged particles (see Exercise 5).

This failure of the Third Law for a force law of such great physical importance raises a serious problem of determining precisely under what conditions the Third Law can be expected to hold and what is responsible for its failure. The problem is best addressed by considering the Third Law in a different form. For a 2-particle system, the Second Law gives us

$$\frac{d\mathbf{p}_1}{dt} = \mathbf{f}_{12} \qquad \text{and} \qquad \frac{d\mathbf{p}_2}{dt} = -\mathbf{f}_{21} \,,$$

where $\mathbf{p}_1$ and $\mathbf{p}_1$ are momenta of the particles. So the Third Law can be written

$$\frac{d\mathbf{p}_1}{dt} = -\mathbf{f}_{21} \,.$$

Thus, the Third Law can be interpreted as a *Law of Momentum Exchange*. Hence a failure of the Third Law would be a failure of momentum conservation. Today, physicists regard the *Law of Momentum Conservation* as more fundamental than Newton's Laws because it holds in Quantum Mechanics as well as Classical Mechanics with no known exception. Any apparent violation of momentum conservation prompts the question: "What happened to the missing momentum?" On several occasions attempts to answer this question have led to the discovery of new physical objects, of which the elementary particle called the *neutrino* is a spectacular example.

Classical Field Theory accounts for the apparent failure of magnetic interactions to satisfy momentum conservation by attributing momentum to the electromagnetic field. We are not prepared for a quantitative discussion of this matter using Field Theory, so we must be content with qualitative remarks. Electromagnetic Field Theory allows a particle to interact only with fields at the position of the particle. This extends our stated *Principle of Local Interaction* to include field variables. It precludes the possibility of instantaneous interparticle interactions except as an approximation. Rather, the interaction between particles is indirect with the field as intermediary. It proceeds by a transfer of momentum from one particle to the field; then the field transports some of the momentum at the speed of light to the position of the second particle where it can be transferred from field to particle, while the rest of the momentum may travel freely as electromagnetic radiation. The point to be made here is that the Third Law is completely consistent with Field Theory if we extend the Principle of Local Interaction and interpret the "object" in the law statement as a field. Physicists do not ordinarily speak of "a force exerted by a particle on a field" as in the law statement. But this just means "rate of momentum transfer from particle to field," which is a conventional expression.

For a unified view of physics, particle mechanics should be regarded as an approximation to Classical Field Theory. In this approximation, then, the Third Law can be applied to particles acting instantaneously at a distance, as in Newton's theory of gravitation.

Comments on the Fourth Law

This Law is sometimes regarded as part of the Second Law, but it deserves an independent formulation to emphasize its importance. It helps us "divide and conquer" in mechanics by allowing us to decompose complex forces into simpler parts for separate analysis, just as the Law of Composition allows us to decompose extended bodies into particles. Conversely. it allows us to lump a great many forces into a single force to be analyzed as a unit. In a word, the Third and Fourth Laws are the main mathematical tools for assembling and disassembling interactions.

Comments on Absolute Simultaneity

The Hypothesis of Absolute Simultaneity is best regarded as a supplement to the First Law. It implies that an inertial time scale set up in one inertial system can be employed in any other inertial system, so one time scale suffices for all inertial systems. This is equivalent to Newton's assumption that there exists a unique absolute time variable that can be employed in any reference system. Absolute simultaneity is called a hypothesis rather than a law here, because it is now known to be empirically false, though it is approximately true in a large empirical domain. Explicit formulation of this hypothesis, which is implicit in Newtonian theory, shows us exactly where Relativity differs from Classical Mechanics. Einstein replaced absolute simultaneity with the

Law of Light Propagation:

The speed of light is constant with all inertial systems.

With this law we can use an idealized light pulse or photon to construct a model particle clock, a *photon clock*. The photon clock establishes an inertial time scale which is the same for all inertial systems and uniquely relates the time scale to the distance scale. Moreover, the Law of Simultaneity which we introduced as part of the Zeroth Law can now be reduced to a mere *definition of simultaneity*. All this leads to a conceptual fusion of space and time into a unified concept of *spacetime*. The mathematical formulation and analysis of these ideas using geometric algebra will be developed in the second edition of NFCM. It should be mentioned here that the Light Propagation Law requires a small but significant alteration of the Second Law because it modifies the concept of time. But no other changes in the laws are needed to give us relativistic mechanics.

Comments on the Local Interaction Principle

The Principle of Local Interaction is implicit in every treatment of mechanics, yet it has not been singled out peviously as a postulate of the general theory. It is essential if we are to conclude from the generic laws that specific forces determine definite differential equations for particle orbits. And our aim is to formulate the general presumptions of mechanics as explicitly and completely as possible.

Our formulation of Local Interaction allows the force to be a function of time derivatives of the position vector to any order. As a rule, the velocity is the only time derivative to appear in a specific force law. But there is an exception of great theoretical importance, namely, the *radiative reaction force* due to the reaction of electromagnetic radiation on a particle emitting it. This force law depends on the third time derivative of position. However, this is not the place to study it.

We have already noted the close relation of Local Interaction to the Third Law. In a pure particle theory we can combine these postulates to draw conclusions about the functional form of the two particle force. Thus, for a force that depends only on position and velocity we find

$$\mathbf{f}_{12} = \mathbf{f}[\mathbf{x}_1(f), \mathbf{v}_1, (t), \mathbf{x}_2(f), \mathbf{v}_2, (t)] = -\mathbf{f}[\mathbf{x}_2(f), \mathbf{v}_2, (t), \mathbf{x}_1(f), \mathbf{v}_1, (t)]\,. \tag{4.1}$$

The Relativity Principle restricts the function form still further. This significantly restricts the force laws to be considered in a pure particle theory.

Comments on the Relativity Principle

The effect of a change in reference system on the equations of motion for a particle is discussed in Section 5-5 of NFCM. Here we will merely comment on its general theoretical implications in accordance with the Relativity Principle. We saw that the most general transformation of position vectors relating one inertial system to another has the form

$$\mathbf{x} \to \mathbf{x}' = R^\dagger(\mathbf{x} + \mathbf{a} + \mathbf{u}(t + t_0))R\,, \tag{4.2}$$

where R is a unitary spinor and $\mathbf{a}$, $\mathbf{u}$ and t are constants. This transformation is a composite of a space translation, a Galilean transformation, a time translation and a rigid rotation. It maps a particle orbit $\mathbf{x} = \mathbf{x}(t)$ onto an orbit $\mathbf{x}' = \mathbf{x}'(t') = \mathbf{x}'(t + t_0)$. Differentiation therefore gives the general *velocity addition theorem*

$$\dot{\mathbf{x}} \to \dot{\mathbf{x}}' = R^\dagger(\dot{\mathbf{x}} + \mathbf{u})R\,. \tag{4.3}$$

Another differentiation gives us the transformation of the Causality Law,

$$\ddot{\mathbf{x}} = \mathbf{f} \to m\ddot{\mathbf{x}}' = \mathbf{f}'\,, \tag{4.4}$$

showing that its form is unchanged, as required by the Relativity Principle, provided the force undergoes the induced transformation

$$\mathbf{f} \to \mathbf{f}' = R^\dagger \mathbf{f} R\,. \tag{4.5}$$

The Relativity Principle requires more, however. It requires that the functional form of the force law must be preserved by the transformation (4.2). In particular, the two particle force law (4.1) must be of the more restricted form

$$\mathbf{f}_{12} = \mathbf{f}[\mathbf{x}_1 - \mathbf{x}_2, \dot{\mathbf{x}}_1 - \dot{\mathbf{x}}_2]\,; \tag{4.6}$$

it must depend only on the relative position $\mathbf{x}_1 - \mathbf{x}_2$ to be form invariant under translations, and on the relative velocity $\dot{\mathbf{x}}_1 - \dot{\mathbf{x}}_2$ to be invariant under Galilean transformations. Moreover, invariance under time translations implies that $\mathbf{f}$ cannot be an explicit function of time. To be even more specific, if $\mathbf{f}_{12}$ is an algebraic function of $\mathbf{x}_1 - \mathbf{x}_2$ and $\dot{\mathbf{x}}_1 - \dot{\mathbf{x}}_2$, then (4.5) is automatically a consequence of (4.2) and (4.3). This is, indeed, characteristic of the most fundamental force laws we have considered.

Clearly the Relativity Principle is an important modeling principle. It tells us that our models should be independent of our chosen (inertial) reference system, so interactions should be functions only of *relative* positions and velocities. We have interpreted the transformation (4.2) as a *passive* change in descriptive variables without altering the state of motion of any object. Alternatively, for $t_0 = 0$, we can regard (4.2) as a change in description due to an *active* rigid displacement and boost in velocity of a single reference body (or frame). If our models are to be unaffected by such a shift of the reference body, as required by the Relativity Principle, we conclude that the reference body must not be interacting with real objects. In other words, the reference body must be regarded theoretically as fictitious. Of course, real objects are needed as reference bodies in experiments. So the Relativity Principle serves as a guide to the idealizations required for a theoretical description of experiments.

Another profound implication of the Relativity Principle is found by interpreting (4.2) as a transformation with respect to a single reference system. The transformation (4.2) maps any orbit $\mathbf{x} = \mathbf{x}(t)$ onto an orbit $\mathbf{x}' = \mathbf{x}'(t + t_0)$ at a different time and place. According to the Relativity Principle, these orbits describe physically equivalent (or congruent) processes. Thus, the *Relativity Principle can be regarded as a general congruence law*, providing a precise criterion for the equivalence of different physical processes at different places and times. This makes it possible to compare results of different experiments performed at different places and times. Thus, the Relativity Principle provides a theoretical basis for the *reproducibility* and *predictability* of physical results.

It should be noted that the Relativity Principle is a semantic principle, because it is concerned with the interpretation of descriptive variables, that is, with the relation of models to their referents. It is appropriate to regard the Relativity Principle as a "congruence law," because it describes an equivalence relation under rigid transformations in space and time, so it generalizes the notion of congruence from elementary geometry. This geometrical character of the Relativity Principle shows that it should be grouped together with the Zeroth and First Laws. These three laws together determine the model of space and time used in classical mechanics, and they must all be modified to characterize the model of space-time proposed in Einstein's General Theory of Relativity.

Comments on the Theoretical Structure of Mechanics

We have completed our formulation of the generic laws and principles of Particle Mechanics. These laws and principles compose an axiom system from which all results of the theory can, in principle, be derived as theorems. We say "in principle" because no one has bothered to develop the theory as an orderly system of theorems and proofs based on well-defined axioms. A major reason for this has been the lack of a complete and appropriate set of axioms. Under the influence of recent mathematical fashion, some authors have developed axiomatic formulations of mechanics using set theory. But set theory is not the right mathematical tool, because it is too general. Consequently, theorems and proofs in this approach are inordinately unwieldy. Geometric algebra is a better tool, because it was designed for the geometrical job. And our formulation of the axioms conforms well to physical practice.

Of course, we have already derived the results of major interest in mechanics in an informal way, so there is no point to embarking on a formal development here. However, it is worth pointing out that formalization of mechanics should have some advantages. It can be expected to clarify the structure of the theory, eliminate unnecessary redundancy and make results more accessible for applications. On the other hand, it must be recognized that the organization of mechanics should be dictated by physical rather than mathematical considerations. For the purpose of theory is to make specific models.

4. Modeling Processes

Scientific knowledge is of two kinds, factual and procedural. The *factual knowledge* consists of theories, models, and empirical data interpreted (to some degree) by models in accordance with theory. A theory is to be regarded as factual, rather than hypothetical, because the laws of the theory have been corroborated, though theories differ in range of application and corroboration. The *procedural knowledge* of science consists of strategies, tactics and techniques for developing, validating, and utilizing factual knowledge. It is commonly referred to under the rubric of *scientific method* . The structure of factual knowledge has been explicated in our general discussion of models and theories and our detailed analysis of classical mechanics. Our aim in this section is to explicate the structure of procedural scientific knowledge. The subject is complex, so we cannot hope to produce much more than an outline. We will do well to identify organizing principles which give the subject some coherence.

The key to an explication of scientific method is recognizing that the central activity of scientists is the development and validation of mathematical models. Thus we need to analyze the processes of mathematical modeling. We can distinguish two types of modeling process: model development and model deployment. The first is concerned primarily with theoretical aspects of modeling, while the second is concerned with empirical aspects. Theoretical and empirical aspects are often interrelated, so the distinction between development and deployment is a matter of emphasis rather than sharp separation. Let us proceed to a discussion of each process in turn.

Model Development

A model is a surrogate object; it depicts or portrays a real object by representing its properties. The properties of a real object are known only through their representation in a model; they are never experienced directly. Moreover, our knowledge of any real object is always incomplete. Every model is an *idealization* or *partial representation* of its referent, which is to say that some but not all properties are represented in the model. Nevertheless, physicists strive to construct complete models of the most elementary constituents of matter, such as electrons. (These are the only objects that might be simple enough to model in all detail — but that is pure speculation).

Deliberate idealization is a method of simplification. A model which fails to represent known properties of its referent is often useful when those properties are regarded as irrelevant or uninteresting. Thus, we model the Earth as a particle when concerned with its motion in the solar system.

The method of deliberate idealization generalizes to the *method of successive refinements*, which is one of the major modeling strategies in science. Beginning with a simple model, a sequence of increasingly complex models is constructed by successively incorporating additional attributes to represent the object with increasing detail. Thus, the simple particle model of the Earth is refined by modeling it as a rigid body to describe its rotation, further refined by modeling it as an elastic solid to account for the effects of tidal forces; then it may be assigned a model atmosphere and molten core to account for its thermal properties. The modeling is never finished, as any geophysicist or climatologist can attest.

The process of developing a mathematical model can be analyzed into four essential stages: (I) *Description*, (II) *Formulation*, (III) *Ramification*, and (IV) *Validation*. The stages are implemented consecutively, though backtracking to revise the results of an earlier stage is not uncommon. The entire model development process is outlined schematically in Figure 4.1 to indicate the kind of information processing in each stage. The figure can be regarded as the outline of a modeling strategy as well as a description of the modeling process. Moreover, it can be regarded as a problem solving strategy, since, by and large, physics problems are solved by developing models.

The modeling strategy outlined in Figure 4.1 is sufficiently general to apply to any branch of physics, indeed, to any branch of science. Therefore it can be regarded as a *general scientific method*. However, the implementation of each stage in a particular model is theory-specific, that is, the tactical details in modeling vary from theory to theory. To understand how the strategy applies to mechanics, we need to elaborate on the details of each modeling stage.

(I) The Description Stage begins with a choice of objects and properties to be modeled. The theory to be used in modeling depends on the kinds of property to be modeled — physical, chemical or biological, for example. When an appropriate theory has been chosen, the theory provides a system of principles which constrain and direct the modeling process.

Object description is the first step in modeling. The object description begins with a decision on the type of model to be developed. For example, a given solid object could be modeled either as a material particle, or as a rigid body. Mechanics provides subtheories to facilitate the modeling of objects of each type. Complex objects are modeled as composite systems of interacting parts, for example, a system of particles or rigid bodies. In that case, the object description must specify

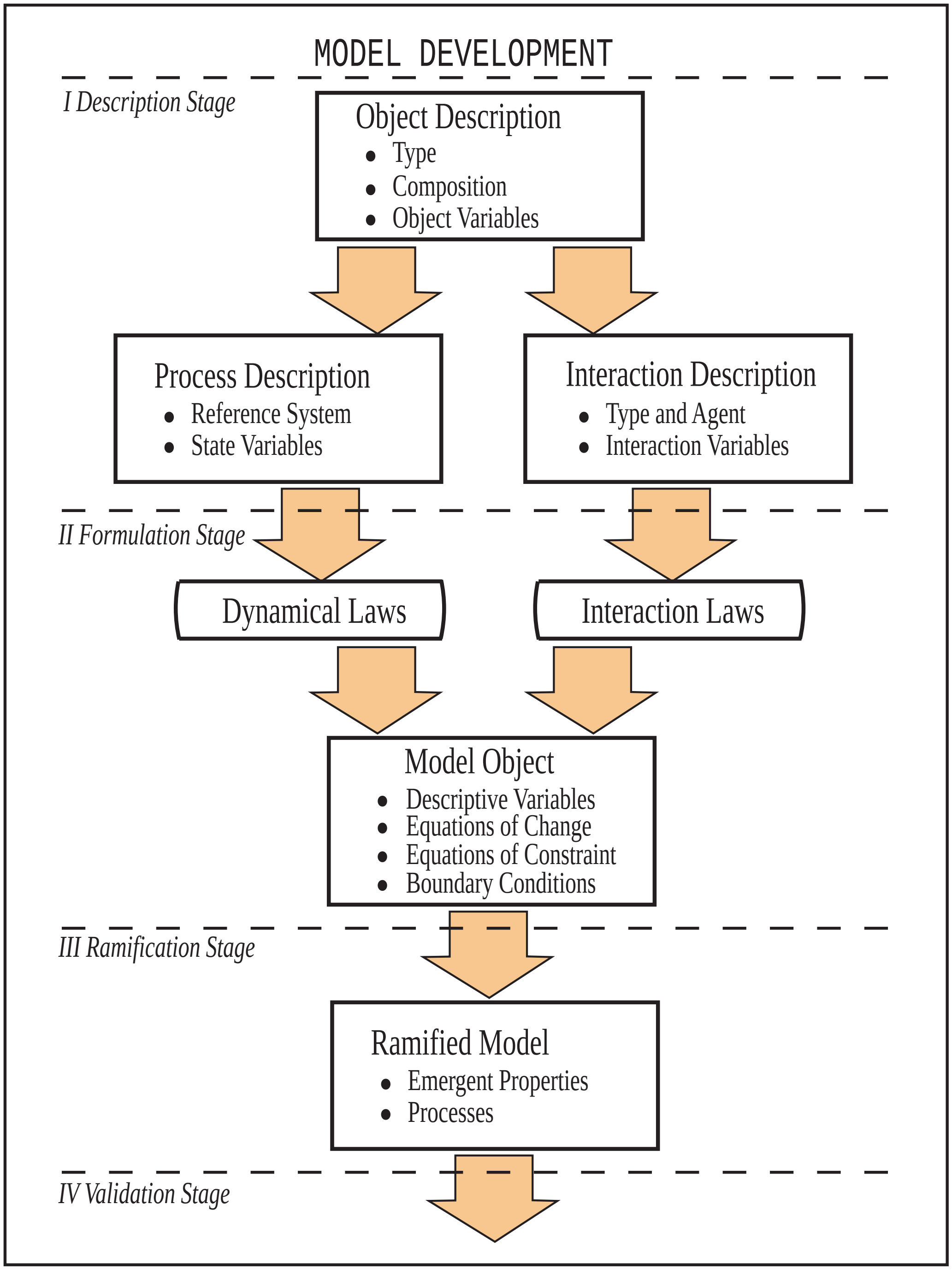

Figure 4.1.

the composition of the system and the model type of each part. Each part can then be modeled separately, and the model for the whole system is determined by the way the interacting parts are assembled.

In a process description the state variables of the model are specified. The state variables may be either basic or derived. *Basic variables* are defined implicitly by the generic laws (including the

Zeroth Law). *Derived variables* must be defined explicitly in terms of basic variables. In mechanics, position and velocity are basic variables, while momentum, kinetic energy and angular momentum are derived variables. A process description necessarily employs the Zeroth Law, so some reference system must be adopted, even if it is not mentioned explicitly.

A *process* is defined as the time evolution of some set of state variables. Motion is the basic process in mechanics. The energy conservation law makes it convenient, sometimes, to consider the process of energy flow independently of the objects processing the energy. In such a case, one is modeling a process rather than an object. A *process model* omits reference to objects underlying the process.

Graphical or diagrammatic methods are often useful in a process description. As a rule, only a qualitative graph of the process can be made in the description stage of modeling; although a few points, such as initial and final states may be specified completely. A quantitative graph is usually possible in the ramification stage of modeling.

An interaction description specifies the interaction type and agent for all interactions in the model, along with appropriate interaction variables, basic or derived. This includes internal interactions among the parts of a composite system, as well as interactions with external agents. The interaction description must be coordinated with the process description; a consistent set of variables must be chosen, and any changes in interactions between different stages of the process must be indicated. In mechanics, for example, use of kinetic energy as a state variable calls for use of potential energy and work as interaction variables. And the description of interactions differs in the processes of projectile motion and collision.

To sum up, the descriptive stage produces complete lists of object names and descriptive variables for the model and supplies the model with a *physical interpretation* by providing referential meanings for the variables.

(II) In The Formulation Stage, the laws of dynamics and interaction are applied to get definite equations of change for the state variables. Within a given theory, the appropriate choice of laws depends on the type of model and descriptive variables, as is clear in examples from mechanics. In a particle model, Newton's Second Law is the dynamical law relating basic descriptive variables, but conservation laws for energy, momentum and angular momentum may be more appropriate when derived variables are used. For a system of particles with interactions described by equations of constraint, we have seen that Lagrange's equation is the most convenient dynamical law. In a rigid body model, we employ separate dynamical laws for translational and rotational motion of the body. These laws belong to the superstructure of mechanics, being derived from the basic laws and the definition of a rigid body. The derivation is a special exercise in model formulation which can be carried out once and for all. The results can then be applied directly to the formulation of any rigid body model.

Besides equations of change, a model may include equations of constraint (as indicated in Figure 4.1). The *equations of constraint* in a model are functional relations among descriptive variables (rather than differential equations). There are many different kinds, including the so-called *Constitutive Relations or Equations of State*, such as the ideal gas law ($PV = nRT$) in thermodynamics and fluid mechanics.

Implementation of the formulation stage produces an *abstract model object* consisting of the set of descriptive variables and equations of change and constraint sufficient to determine values of the state variables. The adjective "abstract" signifies that in an abstract model the descriptive variables are detached from the referential meanings determined in the descriptive stage. Thus, the descriptive variables in an abstract model describe nothing in particular. The adjective "descriptive" remains appropriate, however, because in principle a descriptive variable can always be interpreted by associating it with a referent.

An abstract model does not represent a particular object. A model of a particular object consists of an abstract model together with an interpretation of its descriptive variables; in brief, *a concrete model is an interpreted abstract model*. The detachment of an abstract model from any physical interpretation is a step of major scientific and psychological importance. For the abstract model takes on a theoretical life of its own which can be studied apart from the complexities of a real physical situation. This process of *model abstraction* is crucial to scientific understanding. Paradoxically, physical insight into a given physical situation is achieved by sharply separating the perceived situation from its conceptual representation, that is, by constructing an abstract model. The physicist uses the same abstract model of a particle subject to a constant force to represent many different physical situations, such as a falling body or a body sliding on a rough surface. Thus, the model abstraction process enables the physicist to recognize common elements in different physical situations. Undoubtedly, it plays a role in the discovery of general physical laws from *ad hoc* models constructed wthout the help of general laws.

(III) In The Ramification Stage the special properties and implications of the abstract model are worked out. The equations of change are solved to determine trajectories of the state variables with various initial conditions; the time dependence of significant derived descriptors, such as energy, is determined; results may be represented graphically as well as analytically to facilitate analysis. Let us refer to a model object together with one or more of its main ramifications as a *ramified model*.

The ramification process is largely mathematical, but the analysis of results is just as important. Especially important is the identification of *emergent properties* in composite systems, such as resonances, stabilities and instabilities.

A large part of this book has been devoted to ramifications. We were able to work out ramifications of the gravitational two body problem at length, because the equations of motion can be solved exactly. On the other hand, we found that the ramifications of the gravitational three body problem are only partially known.

(IV) The Validation Stage is concerned with evaluating the ramified model by comparing it with some real object-in-situation which it is supposed to describe. This may range from a simple check on the reasonableness of numerical results to a full-blown experiment test. Validation is a model deployment process, so we will see it in perspective as we analyze model deployment.

Model Deployment

Model deployment is the process of matching a ramified model to a specific empirical situation. The result is a *concrete model* that *represents* objects and/or processes in that situation. We say, then, that the situation has been *modeled* by the scientific theory from which the model was developed. The *match* of the model to the situation is a correspondence between the values of descriptive variables in the model and properties of objects in the situation. The correspondence is established by measurement procedures, including the so-called operational referitions for the variables measured. Measurement involves error and uncertainty, so the match of model to situation must be characterized by some measure for "goodness of fit" and criteria for an adequate match must be set up. These issues are handled by a theory of measurement, which can be regarded as one part of a general theory of model dcployment. We mention measurement here only to indicate how it fits into modeling theory.

Different kinds of model deployment can be classified according to different purposes they subserve. A model may be deployed for the purposes of scientific explanation, prediction or design. Indeed, we say that an empirical phenomenon can be *explained* scientifically if and only if it can be adequately modeled by a scientific theory. Scientific predictions are generated by process models which relate the values of property variables at different times. Scientific design involves the development of models to be deployed as plans for the construction of physical systems with specified properties. The assertion that scientific explanation is a kind of model deployment deserves some further comment, since scientific explanation is not ordinarily characterized that way. There are two common kinds of scientific explanation: causal and inferential. A *causal explanation* of an event A is supplied by identifying its cause, consisting of agents and conditions sufficient to produce A.

An *inferential explanation* of A is supplied by identifying a mechanism (or law) which accounts for A. Each kind of explanation employs one of the essential ingredients of a model. Thus they employ partial models and should be regarded as partial explanations only. A complete explanation requires a complete model.

Empirical tests of a scientific theory are variants of the three major kinds of model deployment we have just discussed. A theory can be tested only indirectly by testing for the empirical adequacy of models developed from the theory. A particular hypothesis can be tested only as part of a theory which is sufficient for the design of testable models and only against an alternative hypothesis which is a candidate to replace it. A test is made by comparing the adequacies of models generated with the alternative hypotheses.

This discussion of model deployment was necessarily brief, because a systematic theory of deployment processes is yet to be developed. The subject is complex, but the concept of model deployment appears to be the thread needed to tie up a lot of loose ends in the methodology of science.

Exercises

1. Does the Zeroth Law imply the existence of a unique physical entity which we might identify as *physical space*?
2. Are space and time objectively real in the sense that they exist independently of any human mind?
3. Develop an explicit formulation of a *Law of Molecular Composition*, providing suitable definitions for the key terms, and carefully distinguishing between mathematical structure and physical interpretation. Discuss the scope and validity of the law.
4. Design a thought experiment for determining if a given reference system is an inertial system.
5. According to the *Biot-Savart Law*, a moving particle with charge q, produces a magnetic field
$$\mathbf{B}(\mathbf{x},t) = \frac{q}{c}\,\frac{\mathbf{v}_1 \times (\mathbf{x}-\mathbf{x}_1)}{|\mathbf{x}-\mathbf{x}_1|^3}\,,$$
where c is a constant, $\mathbf{x}_1 = \mathbf{x}_1(t)$, and $\mathbf{v}_1 = \mathbf{v}_1(t)$ is the velocity of the particle. Examine the magnetic interaction between two charged particles and show that the Law of Reciprocity is not satisfied. How is this result affected by including electric interactions? Evaluate the rate at which this two particle system transfers momentum to the electromagnetic field. What if one particle is initially at rest?
6. Suppose that during all of recorded history the earth was surrounded by a dense cloud cover so that the sun, moon and stars could not be seen. Suppose also that Newtonian mechanics had developed in spite of this handicap. Explain how earthbound physicists could nevertheless detect the rotation of the earth and the orbital motion about the sun and thus separate the associated pseudoforces from real forces.
7. Examine the change in form of the Second Law induced by changing to a time variable which is an arbitrary monotonic function of inertial time.
8. Discuss the change in form of equations of motion when transformed from an inertial system to an accelerated reference system.
9. Discuss the following assertion by J. L. Synge:
   > "It is futile to ask whether nature is ultimately discrete or continuous, for 'discrete' and 'continuous' are categories of the understanding, not properties of nature."
10. Make a thorough critique of Eisenbud's influential article on mechanics (below), comparing it in detail with the formulation of mechanics in this chapter. Note how the concept of definition is used. Carefully distinguish between explicit and implicit definitions, interpretations, correspondence rules and measurements.

    L. Eisenbud, 'On the Classical Laws of Motion,' *Am. J. Phys.* 26, 144–159 (1958).