A. Introduction

Chapter 1. Terminology and Notation

Chapter Introduction

A. Why?

Humpty Dumpty was very emphatic: "When *I* use a word...it means just what I choose it to mean — neither more nor less."

"The question is," said Alice, "whether you *can* make words mean so many different things."

"The question is," said Humpty Dumpty, "which is to be master — that's all."[1]

There is no question that we *can* make words mean what we wish them to mean, but communication requires that the speaker and the listener attach the *same* meaning to words. Thus we must be able to define our terms precisely, and we must be sure the listener has in mind the same definition as the speaker.

B. What to Look For

We introduce here a few basic terms that carry "normal" meanings, but more sharply defined than is customary. We also look at other ways of conveying information, such as choices for symbols and typography.

C. Pre-test for Prior Comprehension[2]

Do you feel comfortable with the following terms? (Can you define them?)

1. physics 2. system 3. surroundings 4. constant of the motion
5. mole 6. process 7. state 8. path
9. field 10. physical particle 11. preservation 12. project-based learning

How would you choose to represent a constant? a variable? a counting integer? in choice of symbol and by typography?

D. Inquiry Questions

a. A few theorems and discoveries stand out in the recent history of physics. Explain

---

[1] For additional insight, see *The Annotated Alice*, by Lewis Carroll, with an introduction and notes by Martin Gardner, Bramhall House, New York, 1960; pp. 268-9.

[2] Answers to pre-tests and example problems are given at the top of the following verso page (even numbered page). The pre-tests tell you whether you are sufficiently familiar with the material so that you may skim the material in the chapter. Examples provide a test of whether you understand the material well enough to apply it in a straight-forward problem. Try to solve the problems *before* you look at the answers, to take advantage of the learning/testing opportunity.

Short answers for pre-test on definitions and customs.
1. *Physics* describes the behaviors of objects and their interactions (the forces between them).
2. The *system* is whatever part of the universe is of special interest at the moment.
3. The *surroundings* is all the rest of the universe (or everything *except the system* that may be affected in the *process* under consideration. (The surroundings need not surround the system.)
4. A *constant of the motion* is any variable quantity that is *constant* (or "preserved") for the system during the process under study.
5. A *mole* is a fixed number of particles (Avogadro's number, 6.022×10^{23}). The symbol is mol. It is necessary in each instance to explain what is being counted.
6. The *state* of a system is described by all the physical properties of that system, although typically most properties are linked so only two properties may be changed independently.
7. A *process* is a change in the state of a system, including the "path" of the change.
8. A *path* is the succession of steps, or the successive *states* of the system, in a process, or change of state.
9. A *field* is a convenient description of a region of space in which forces may act on a body that is not in physical contact with the source, or cause, of the force; *e.g.*, a *gravitational field,* an *electrical field,* or a *magnetic field*.
10. A dictionary definition of *particle* is typically a minute, or "smallest conceivable", bit of matter, treated as a point. In physics, we often deal with *physical particles* defined as an object of *any* size that can change kinetic energy *only*.
11. The *conservation* laws are extremely important, but more often we deal with quantities that are *preserved*, or "constants of the motion", constant for the system under specified conditions.
12. Learning depends more on what the student does than on what the instructor does. Shifting emphasis to "student-based learning" typically involves engaging students in projects and thus has also been called "project-based learning".
A numerical quantity, whether fixed (a *constant*) or changeable (a *variable*) is customarily represented in *italics*. Constants are customarily represented by symbols (letters) from the beginning of the alphabet; variables are customarily represented by symbols from the end of the alphabet. Counting integers are customarily represented by letters (also often italic) from the middle of the alphabet (*i* to *n* or even *q*).

*****

how the following names, effects, or principles have a common link.

i. Condon (Franck-Condon) *ii*. Hall *iii*. Mössbauer
iv. deBroglie

b. What links the following individuals?

i. Benjamin Franklin *ii*. Michael Faraday

#####

To avoid constant interruption of discussions with definitions and explanations, many customs have been adopted that enable the informed reader to pick up a discussion and immediately recognize what is being discussed and what most of the symbols represent. Some of the more important examples are given here as an introduction.

1.1 Terminology. Popular writers often strive for variety of expression, choosing alternative words for a single idea to avoid repetition. Technical writing seeks to limit the variety of words (which is typically very large, in any case) to attempt to achieve a one-to-one correspondence between ideas and words. Thus, if you are talking about how fast something moves, you may want to talk about the *rate* of motion to define *speed*, but then call it *speed*, consistently. Don't substitute other words, such as "velocity" (which means something else!) just for variety or false erudition.

New terms can simplify a discussion. We may be describing how an object moves, or the force exerted by one object on another. The object under consideration may be a box, or a ball, or a sample of gas, or a piece of machinery. Often it makes no difference what the actual object is. We can always call a ball a ball, or we could mention the specific object, then refer to the object as "it". We could also call it "the object", but often we include several "objects" together — two masses attached by a spring, or a box interacting with a gravitational field, or a collection of gas molecules, for example. A more satisfactory solution is to call whatever object(s) we have focused our attention on the "system". Then it is easier to distinguish between a force acting *on* the system and a force exerted *by* the system, or between energy transferred *to* the system or energy transferred *from* the system.

The *system* must be carefully specified in any particular instance (is it the gas, or the gas plus its container, for example), but for many purposes it is sufficient to know that whatever bit of the universe we are dealing with will be called "the system". It is false economy to avoid defining the system under consideration. Far more confusion arises from lack of specification of what is or is not included in the system than is generally recognized. Are we looking at the properties of a ball that is falling in a gravitation field (part of the surroundings) or do we want the properties of the ball plus the field? Is "potential energy" energy stored in the system or energy stored in the surroundings? Confusion easily leads to a sign error (telling the reader that adding energy to a system reduces the energy of the system). Even properties of a gas can be confused by including, with the gas, the cylinder and/or the piston and/or the stops at the end of the motion. Remember: anything you say that *can* be misunderstood, *will* be misunderstood by some listener(s).

A system may change in many ways, such as shape, or temperature, or speed, or rotational speed, or volume. We say that, at any given instant, *all* the information available, or potentially available, about the system describes the *state* of the system. Although only a very few values may be needed to adequately describe the state, in principle the state includes everything that can be known about the system. Thus we can specify the state of a mole of O_2 gas by giving two quantities, such as temperature and pressure, but the resultant state includes a density, volume, color, heat capacity, and other measures, apart from any information about how the oxygen may be moving or rotating or where it is.

The same system may be defined in different ways for different purposes. We choose our definition of a state, and its description, to fit the problem we want to discuss. Occasionally we may wish to examine molecular states. More often we include only large-scale, or *macroscopic*, measurements; that is, quantities such as density, volume, pressure, and energy of the entire system. Then, for systems at rest with respect to us and at equilibrium, we need to specify the *system* (including *what* and *how much*) and then, typically, only *two* properties.

Sometimes additional information is required. If the system has not come to equilibrium, there may be internal strains, or a liquid may be swirling with different compositions and different motions in different parts. We will always assume the system is at equilibrium, unless we specifically say otherwise. Even so, other information may be required, as when there are imposed electric or magnetic fields.

When a system changes, we call that change a “process”. A process includes the end points — the initial state of the system and the final state of the system — but it also includes the “path” followed, that is, precisely *how* the system changes from its initial state to its final state.

Sometimes we specify two systems that interact with each other. More often, we are concerned only with one system and “whatever” it interacts with. Then we speak of the system and its “surroundings”. The *surroundings* then represent everything in the universe *except* the system or, better, everything in the universe that may be affected by the process undergone by the system.[3] We will find that the surroundings do not even have to surround the system. The system and surroundings labels are arbitrary, and may be interchanged! We could have two gases intermixed, and call either one the system and the other the surroundings (if there are no other changes, outside the gases).

Often one or more quantities (temperature, pressure, speed, or energy, for example) may remain constant during a process we are examining. Then, borrowing a label from classical mechanics, we may call that quantity a “constant of the motion” (whether or not motion is involved in the process). It will be convenient to say that a constant of the motion is *preserved*. Thus, if speed of a ball is constant, we could say the speed is preserved during the motion. If the temperature of a container of gas does not change, we could say the temperature is preserved. Very often current discussions apply the terms “conservation” or “conserved” when the intended meaning is “preservation” or “preserved” (as in statements, on their face clearly absurd, such as “Energy is conserved in this process”).

We are all familiar with gravity, and the pull of the Earth on a body near the Earth. It is convenient, for reasons we will discuss later, to describe that pull as exerted by a “field”, in this case a *gravitational field*. The concept of a field was described by Einstein as one of the most important ideas in physics. We will meet it often.

1.2 Variables, Constants, and Indices. A nearly universal system of notation saves much time in representing quantities with letters. Variable quantities are represented by letters at the end of the alphabet: x, y, z, w, v, u, and so forth are easily recognized as variables. Constants are represented by letters from the beginning of the alphabet: a, b, c, d, and so forth. Integers for counting or identifying successive values are chosen from the middle of the alphabet, starting with i, and continuing with j, k, l, m, n, *etc*. (Perhaps designation of axes may be considered an exception; coordinate axes are typically x, y, z or, within a smaller frame, a, b, c. To represent the directions of axes, we may choose *unit vectors*, of unit magnitude and direction along the x, y, or z axes. These unit vectors are generally represented as **i**, **j**, and **k**.)

For example, the speed of light in vacuum is a universal constant, so we represent that

[3] Not only would it be inappropriate to include a distant star as part of the surroundings of a gas in the laboratory, it would generally be technically incorrect, because a change in state of the star could not make an otherwise prohibited change in state of the gas possible.

fixed value by *c*. The speed of other objects, or the speed of light in other substances, can take on many different values so we typically represent such speeds by *v*.

There are exceptions to this notation, but usually only when the exception is obvious to the reader or is explained in the immediate context. The most common exceptions are for the well-known universal constants, such as Planck's constant, *h*; the gas constant, *R*; the gravitational constant, *G*; and local gravitational field strength, *g*; and variables chosen for easy identification, such as *A* for area, *L* for length, and *E* for energy. (There are, naturally, also some writers who choose to ignore the convention. For example, because *c* represents the speed of light, they will let *c* represent the speed of sound in a gas, even though that speed depends on temperature.)

There are not enough letters in the English alphabet to represent every quantity, so occasionally we must duplicate letters (*E* may be energy *or* an electric field; *V* may be volume *or* a potential energy, for example). Also, some terms and symbols have come down to us from early writing in Latin or Greek. For example, a distance covered is often represented by the symbol *s*, from the Latin word for distance between two points, *spatium* (or "space").

Sometimes we avoid duplication by introducing foreign letters, of which Greek letters are most common. In particular, we typically choose Greek letters to represent quantities involved in circular measures of position or motion.

We also follow the standard rules of typography: variables (lengths, times, masses, *etc.*) and constants are presented in *italics*: *x*, *t*, *m; c*, *g*, *G*. Symbols for units, such as gram or pound or second or meter (g, lb, s, or m), as well as prefixes (for milli, centi, and mega, for example) are printed in Roman typeface. This should help in recognizing symbols such as ms, km, and g (millisecond, kilometer, and gram) as units, whereas *m*, *s*, and *g* represent quantities (mass, distance, and gravitational field strength) that may require such units or combinations of units. Thus *mg* would be a product of a mass (*m*) and the gravitational field strength (*g*), a variable and a constant, but mg represents the unit of one milligram. It will take some students a while to know how to recognize these clues. We will also introduce special typeface notations (usually boldface) for vectors to distinguish them from scalars (ordinary numbers).

Many units are named for individuals (Newton, Kelvin, Gauss, Curie, *etc.*) Such units are written without capitalization (newton, kelvin, gauss, curie), although the symbols (N, K, G, Ci) are typically capitalized. These labels are considered to be symbols (like the symbols for chemical elements), rather than abbreviations. Therefore they do not have periods, with the exception of "in." to distinguish the inch from the word "in".

International committees on units and nomenclature have pointed out that the unit should be independent of the magnitude, so 0.1 m and 10 m should each be read as 0.1 meter and 10 meter. Especially in English, this rule is not always followed.

Chapter Summary

Terminology. A difficult task is to keep students informed of what is the subject of the discussion at the moment. (Are you talking about the other side, or the other side?) A major step toward clarification is to carefully define the *system* under discussion, as contrasted from the *surroundings*. Avoid variety of terminology for the sake of variety in technical descriptions.

Notation and Typography Rules. The following typographical rules are intended to aid recognition of symbols.

Quantities with numerical values are represented by italic symbols.

Constants are preferably represented by early letters of the alphabet (or a mnemonic choice).

Variables are preferably represented by late letters of the alphabet (or a mnemonic choice).

Indices (integers) are preferably represented by middle letters of the alphabet.

Quantities describing angles or circular motions are preferably represented by Greek letters.

Units are written in Roman type.

Units named for individuals are written in lower case. The symbols are upper case.

Vectors are written in bold face, or with super arrows.

Chapter 2. The Basic Concepts of Physics

Chapter Introduction

A. Why?

Week after week we hear of new discoveries in science, including many claims of "shaking science to its roots". Other times it may appear that physics has become fragmented, with one set of rules for one type of problem, another set for another problem, and no program guide to explain how the parts fit together. Before discussing any of the parts, it may be helpful to see the broader picture to be able to place these parts in proper perspective. Think of this chapter as a road map, to be consulted *before* you venture into new territory. It is not necessary to be familiar with all the local features to read a road map.

B. What to Look For

We are looking here primarily for connections, including generalizations that will carry over from one type of problem to another.

C. Pre-test for Prior Comprehension

1. How does physics differ fundamentally from mathematics?
2. What general condition determines why things happen?
3. What is the technical (physics) meaning of *conservation*?
4. According to the *equivalence principle*, what is equivalent to what?
5. What is the broad statement of the *principle of relativity*?
6. What is special about an *operational definition*?
7. Why is "uncertainty" considered a "principle"?
8. According to Bohr, what "corresponds" with what?

D. Inquiry Question

If we make the apparently reasonable assumption that physics is the natural, and inevitable, consequence of applying logical deduction based on experimental findings, why is physics today so different in many respects from physics of the 17^{th} century? What differences were there, in the experimental measurements and/or logic, that led to such differences?

Often we find there are important limitations to the measurements of certain quantities under specified conditions (such as the position or the energy of a small particle). Is there an example of a *measurable* quantity that would give a different value (or prediction) under different "branches" of physics (*e.g.*, classical, or quantum, or relativistic physics)?

#####

Short Answers to Pre-Test Questions

1. Mathematics is based on arbitrary sets of *postulates*, no one set of which need correspond to reality. Physics is a description of the universe, and because we have access to only one universe, there can be only one physics.

2. Because all material is made of molecules, that act independently, the direction of change is determined by probability, which is conveniently measured by the function called *entropy*, or by functions that include entropy and other measures of the state of the system and its surroundings.

3. A quantity (necessarily extensive) is said to be *conserved* if the amount remains constant, for the universe (*i.e.*, for the system and its surroundings), for *every possible* process. Conservation is very often confused with preservation (constancy of a property, under specified conditions, for a system).

4. Acceleration of a reference frame is equivalent to, or indistinguishable from, by local measurements, a gravitational field (in the opposite direction).

5. Absolute (linear) motion cannot be detected.

6. An operational definition includes a prescription of how to measure the quantity defined.

7. Any measurement of a system *must* affect the system being measured, so following the measurement we know less (by at least a predictable minimum amount) about the state of the system.

8. When more than one model of physics applies to a system (*e.g.*, in the region of overlap of classical Newtonian and/or special relativistic and/or quantum physics), each model gives the same answers for measurable quantities.

*****

Physics is an area of knowledge that was initially defined[1] in the 17th century. It is a branch of *natural philosophy*, or knowledge of nature, concerned with material objects and their interactions, as contrasted with topics such as religion, morals, literature, and fine arts. Today we may describe physics as dealing with matter (or material) and its structure, its motions and changes of motion, and its interactions, including forces between objects. Applications of physics to practical problems often falls into the realm called engineering. In part because it started much earlier, but more because the basic investigational tools and methods are different, astronomy is usually set apart from physics, as is mathematics, even though they are closely intertwined. Similarly, medicine and biology, including botany and zoology, began quite distinctly and have remained separate from physics (despite overlapping name roots such as physics, physician, and physiology), even though the overlap has increased dramatically in recent decades. In the 18th century, other specialized fields, such as chemistry and geology, were separated out and remain largely distinct, incorporating principles of physics in applications to changes in materials.

---

[1] The term appears in the writings of Aristotle, referring to what would later be called “natural philosophy”, the broad science of the nature of matter and how it interacts. Aristotle’s 13 books were called *Metaphysics*, which was quickly misinterpreted as ‘transcending physics’.

Physics has customarily been taught following a roughly historical development. Classical mechanics comes first, serving to introduce the language, quantities, and units of physics. This is followed by thermodynamics, electricity and magnetism, and optics.[2] Only then are the "newer" topics of the late 19th and early 20th centuries introduced. The advantage of this method, and the reason it continues in most classrooms, it that the foundation stones of understanding are more easily presented, and understood, in this sequence. Ideas subject to direct observation and measurement, such as length, area, volume, mass, speed, acceleration, force, and energy, must be understood before attempting current physics. We begin by talking about comparatively concrete objects that can be seen and felt rather than by talking about fields or invisible particles. Furthermore, the early concepts (other than optics) are required to understand later concepts.

On the other hand, there are two unfortunate consequences of this approach. First, by putting off "the good stuff", some students may lose interest before their attention is caught by challenging thoughts. Second, and perhaps even more important, is that it may give the impression that the physics underlying the newer topics is separate and distinguishable from the "old" physics. Even many experienced physicists have fallen into this trap of dividing physics into irreconcilable parts, accepting the misconception that Newton was wrong.

As more and more detailed information is obtained on how materials behave at the microscopic and submicroscopic level, as well as on the laboratory and astronomical scales, it seems time to provide a better perspective of physics as it is understood at the turn of the millennium. The underlying structure of physics is the same, regardless of the topic to be investigated. Understanding that underlying structure is the key to grasping the unity of physics and being able to carry concepts over from one part of the subject to another.

2.1 The Contrast of Physics and Mathematics.

Physics is closely associated with mathematics, but the fundamental difference is very important. Mathematics begins with a set of postulates, then explores the logical conclusions that may follow. By varying the postulates (*e.g.*, parallel lines never meet, or parallel lines cross, or parallel lines diverge), different schemes of mathematics may be developed. One is "as good" as another. The user may decide which is more appropriate for a particular problem. (Nonsense is often written by journalists who fail to distinguish between different arbitrary models of mathematics.)

By contrast, physics is a description of the physical universe. There is only one observable universe, so we can have only one valid physics. There may be various approximations to that physics, such as classical physics, special relativity, and quantum physics, but each is an approximation to the one whole.

The challenge of physics, then, is to find a description of nature, and how it behaves, that is understandable and quantitative, based on a unified set of rules. We should be able to envision the rules and responses ("if you push on this, that will happen") of any part of nature. We should be able to make quantitative predictions, which will require some mathematics. There will necessarily be differences in techniques (certainly experimental techniques, and often mathematical

---

[2] The first broad discussion of optics was Newton's *Optiks*, so a proper historical sequence would introduce optics, or at least geometric optics, early in a physics course.

methods) in dealing with different classes — very large or very small objects, or very slow and very fast objects — but the rules we apply should essentially be seamless, independent of the details or the scale of the system under examination.

Fortunately, there are only a few really basic principles, upon which all of physics is founded. These are the principles that extend through all branches, including classical physics, special and general relativity, and quantum physics, as well as newer explorations. Seven of these "constants" of physics, selected as most important, are presented here very briefly. Examples of these principles are embedded in all of the following chapters.

2.2 Operational Definitions

How would *you* measure time (other than to look at your wristwatch or a wall clock)? How could you explain the meaning of time, or of a time interval, to a student? What is required to measure the length of an object?

The test of any physical theory is its ability to make quantitative predictions that agree with experiments on the real universe. It follows that we must be able to make measurements, which in turn requires that we know how to measure the quantities that appear in our theories. That is, we must have an *operational definition* of each quantity. It has no real meaning to define a length, or a mass, or a time, unless we can state how we measure length, mass, or time. It is not enough to be able to measure time for slow-moving objects within a small laboratory. We must have an operational definition that will also apply to fast-moving objects and larger distances.

Similarly, no conclusion of modern science can be considered significant unless it is falsifiable. That is, there must be at least one experiment, or measurement, that will give a different answer if the conclusion is false than if the conclusion is valid.[3] This, again, requires that we have operational definitions of each concept. It was the introduction of operational definitions into physics that led directly to the development in the 20th century of what has been called *modern physics*. Textbooks seldom, if ever, give operational definitions. However, they give definitions from which it should be possible to deduce operational definitions.

2.3 The Principle of Relativity

---

[3] For centuries, philosophers have argued that nothing can be proved, because in the first place almost any definition relies on other definitions, and in the second place, any arbitrary number of confirming experiments does not rule out other models that might lead to the same results. Introduction of operational definitions, and the concept of falsifiable tests, around 1900, cannot totally negate these arguments but effectively makes them unimportant. There are still multiple models for quantities we do not yet fully understand, but we have learned how to choose between, or accommodate, conflicting models, and have better defined our starting points or standards for measurements.

Perhaps the most important insight of Newton,[4] which enabled him to understand the causes of changes of motion, was his recognition of the principle of relativity. He did not emphasize it, as such, but he incorporated it as the critical first step in simplifying a model of motion that allowed him to construct a theory of mechanics.

Newton was very careful to point out that absolute (linear) motion is not detectable. We cannot tell whether an object is moving or is at rest. When it is at rest with respect to one observer, it is moving with respect to another. Motion is only detectable relative to a chosen reference, and that chosen reference may, itself, be moving.

Being in motion and being at rest are the same thing.

This discovery greatly simplified the description of how bodies move and how changes in motion happen.

It was this principle of relativity that enabled Newton to explain how observed motions, such as motions of the planets in the solar system, depend on forces exerted by other objects. Thus Newton went beyond the *kinematics* of Kepler and Galileo (an explanation of *how* things move) to an explanation of *dynamics*, or the *causes of changes* in motion.

When Newton's concept of relativity was combined, near the beginning of the 20th century, with operational definitions of time, length, mass, and speed, it became clear that special care was required in any theory or measurement that involved questions of simultaneity of events in different locations, or when the motions of the observer and the system under measurement were significantly different. These considerations proved to be the key to understanding motions at very high speed and, later, motions affected by large masses.

2.4 The Equivalence Principle

In Newtonian physics, we *can* distinguish accelerated motion from uniform motion in a straight line at a constant speed, whether the acceleration is a change of speed or a change of direction (or both). We can therefore also detect the presence of a force, such as a gravitational field.

One of the mysteries encountered by Newton was the similarity of equations for accelerating a body and for the interaction of the same body with a gravitational field. We all have experienced the effect, *e.g.*, in an elevator. As the elevator starts upward, we feel heavier and it is harder to hold a box or book bag. As the elevator slows at the top of its motion, or as it starts downward, we feel lighter and whatever we are carrying becomes momentarily lighter. It is not possible, by measurements solely within the elevator, to distinguish between the gravitational field and an acceleration (in the opposite direction).

Newton was able to accommodate this apparent coincidence with his definition of mass. That is, he defined *mass* in two ways — as a measure of the gravitational force exerted on a body, at a given point, and as a resistance to change of motion, or acceleration. He gave these two quantities the same symbol, and set them equal to each other (with appropriate proportionality

---

[4] From time to time it is necessary to speculate on what authors meant. The primary issue is not whether the speculation (here or elsewhere) is correct, but whether the meaning attributed to their words provides insight to readers today.

constants built into the definitions of gravity and acceleration). Later, the question was reconsidered and the identity of behavior of gravitational field and accelerated reference frame was labeled as the *equivalence principle*.

Einstein recognized (as had Newton) that Newtonian mechanics provided no clue as to *why* an acceleration and a gravitational field should have similar effects. He concluded there must be some other cause, outside of Newtonian mechanics, which must therefore itself be a fundamental principle. From this start, he was able to extend his own and Newton's principles of mechanics into what is now known as the subject of general relativity, or geodynamics.

Construction of mechanics on the equivalence principle (*i.e.*, the theory of general relativity) provides new insights into the meanings of space, time, forces, and energy. We will get some brief looks at these new interpretations, although we will be severely limited by the mathematical complexities that arise in quantitative treatments.

2.5 The Uncertainty Principle

Careful measurement requires apparatus designed for the purpose. The apparatus to measure position is not the same as that for measurement of speed. The apparatus for measurement of time is different than that for measurement of other quantities, including energy. If we wish to carefully measure two quantities (*e.g.*, position *and* speed), it will generally be necessary to measure them sequentially, first with one set of apparatus, then with the other. That is no problem if the first measurement leaves the system unchanged, but can we always make that assumption?

For example, what equipment, and methods, would you need in order to determine where a book is located? (Assume you can see and can reach the book.) What equipment, and methods, would you need in order to determine whether the book is moving (*e.g.*, with respect to the table)?

How, or when, can you make a measurement of some property of a system *without* changing the state of the system? That is, what is required in order that our detector (call it D) may detect some property of a system (call it S)? Unless S has an effect on D, D cannot know that S is there, much less know the properties associated with S. Newton recognized, and explicitly discussed, a fundamental relationship. If S has an effect on D, then D must have an (equal and opposite, as carefully defined later) effect on S. That is, *any* measurement of a system must, *in principle*, disturb the system.

Therefore, it is not a question of whether the first measurement changes the system, but rather whether the changes are important enough to limit the validity of the subsequent measurement(s) on the same system. When this question was examined, Heisenberg recognized (in the 1920's) that there is a natural limitation, in nature, to the amount of such disturbances. The disturbance cannot be smaller than a fixed value, common for all types of measurements. Consequently, although we can make any *one* measurement as accurate as we choose (subject to our cleverness in designing the apparatus and our patience), any subsequent measurement may be influenced by the first measurement, so we necessarily get a limited description of the system we examine. We may know "exactly" where it was, but then we cannot tell how it was moving; or we may know very accurately how it was moving, but cannot then determine where it was at the time. This limitation is known as the *uncertainty principle.*

It’s Relative

Bob Poe

Sometimes I see a bird go past;
  Then wonder (you know me) —
It seems it’s moving very fast;
  But is it moving? Or are we?

If I were flying through the sky
  And saw the bird below,
Would it still blithely pass me by?
  Or would it backward go?

A tree is standing on the hill,
  Its roots fixed in the ground;
And yet, we know, while “standing still,”
  It goes the whole Earth round.

Twenty-five thousand miles per day.
  How much is that per hour?
A thousand miles per hour, you say?
  With no motor and no power.

It’s out, and back, and out again.
  Motion in circles is easy to track.
Straight-line motion is harder to ken.
  Do I go forward, or it go back?

Can any object, anywhere
  Truly be judged to be “at rest”?
For any object, do we dare
  Its speed or motion to attest?

We echo, then, Sir Isaac’s creed:
  Absolute motion we cannot detect.
Measuring “rest” and measuring “speed”
  Depends on the reference we select.

Newton’s First Law

A body at rest or in uniform motion*, not acted on by any net external force, will remain at rest or in uniform motion.

*as measured relative to any other non-accelerated body.

It follows that we cannot assign exact positions *and* speeds to any object. Perhaps you are thinking, "Sure, the measurement changes things, but the object did have a position *and* a speed before the measurement." Then we must remember the importance of operational definitions. If we cannot measure that position *and* speed, then we cannot construct a meaningful theory based on those concepts. We cannot obtain experimental numbers to put into our theoretical equations to predict an exact outcome of the pair of measurements.

The limitations of measurement of more than one variable imposed by the uncertainty principle are of no consequence for measurements of objects of ordinary size. Because the *same* uncertainty (roughly 10^{-34} kg·m^2/s) applies to every system, planets, elephants, baseballs, and blood cells may be observed without *significant* disturbance. Smaller objects, such as electrons and protons, are affected equally by the measurement process and therefore the measurement *does* significantly influence the state of the particle.

Because the purpose of a theory is to make predictions of behavior, it is necessary for any theory that includes the description of very small objects to include the uncertainty introduced by the measurement process. We will find there are straightforward methods of building this uncertainty into the theory, but the theory then often requires somewhat more advanced mathematical methods than for objects of ordinary size.

2.6 Bohr's Correspondence Principle

It may, someday, be possible to describe all applications of physics, from very small to very large and from very slow to very fast, with a single set of equations. If so, such equations will not be very practical. Approximate equations, that apply very accurately to the particular systems we are interested in at the moment, are generally much easier to work with than more general equations would be.

Nevertheless, the various sets of equations agree when they have to. When we go from very small to moderate sizes, or from moderate sizes to very large, or from slow speeds to very fast speeds, there is some region of overlap in which either of two (or more) methods can be applied. In such regions of overlap, the different methods must give the same answers. This is known as Bohr's *correspondence principle*. If, in the regions of overlapping validity (or applicability), two methods disagree, then one (at least) is inadequate and must be corrected.

In this introductory treatment, we will emphasize the simplest method of analysis that gives satisfactory results for the objects under discussion. However, we will attempt to make clear what the limitations are of the methods discussed and at least the nature of the changes required when we move to larger, or smaller, or faster-moving objects.

2.7 The Probability Principle

Why will a ball roll downhill, but a ball at the bottom of the hill will not, by itself, roll uphill? Why will air escape from a tire, through even a very small opening, but air from outside will not collect inside the tire at the "working" pressure? Why will a drop of water on the floor accumulate sufficient energy from its surroundings to move to a higher-energy state (water vapor)? Why can a ship not similarly gather energy from its surroundings (the ocean) to become a fast-moving ship?

Feynman has claimed that the single most important piece of information about the universe is that everything is made of atoms (or even smaller particles). As justification for that statement, we may observe that we could not explain why things happen without that understanding!

Because matter is made up of very small parts (atoms and molecules), and there are very large numbers of parts in almost all the systems that we study, it is neither possible, nor desirable, to describe such systems in intimate detail. Within the systems, smaller particles are moving at high rates of speed, frenetically exchanging positions and other properties with each other, while the larger system that we are observing remains in what seems to be a placid, stable state. If it makes no difference to the system where molecules A and B are located at any given instant, why should we care?

What is important to us is *how many* particles are present, and *how many* equivalent states are accessible to those particles. Because there are large numbers of particles, it is quite sufficient to look at the probabilities of different macroscopic states. Based on very simple arguments, we can show that there will almost always be one state, or more accurately, a collection of very similar states that are practically indistinguishable and therefore may be treated as a single state, that is so much more probable than any other state we might observe, that we can quite safely assume the system will be in that state, and will remain in that state, unless or until it is disturbed.[5] This statistical behavior of atoms and molecules was first explored by Boltzmann in the second half of the 19th century. The tendency of any system to exist in the most probable state may be called the *probability principle*. It is also called the *second law of thermodynamics*.

The probability principle, or second law of thermodynamics, can be quite adequately treated by consideration of large, or macroscopic, systems, such as you would observe with a meter stick, analytical balance, thermometer, and other common laboratory tools. Looking at the behavior of the "microscopic" (actually, atomic level) particles, however, gives very important insights, not only for the behavior of the small particles but insights into why macroscopic samples behave as they do.

2.8 The Conservation Laws

A *conservative* quantity, or one that is *conserved*, is a property of a system that may or may not be constant for the system under study, but which is *always* constant for the *system plus its surroundings*. Thus as the quantity is passed back and forth between the system and its

---

[5] The nomenclature can easily get confusing. The state of a system involving properties such as pressure, volume, temperature, and total energy, may be called a *macrostate*. The instantaneous description of the system in terms of the positions, speeds, and other properties of individual molecules may be called a *microstate*. Then there would be an extremely large number of microstates that would give indistinguishable values for the macrostate properties. The macrostate is constant with time, for a system in equilibrium, while the system is rapidly shifting among the many microstates. The macrostate changes only undetectably because of these changes of microstates. The chance of a variation in probability (as measured by entropy) of 1 g of helium gas by one part in a million is about 10^{-10} to the 18th power; that is, there would be 10^{180} digits, before the decimal point, in the number! (If each digit occupies 0.07 in. or 1.76 mm, the number would be roughly 10^{165} times longer than the width of the solar system.) That is why we consider such variations to be negligible.

surroundings, the *total* quantity, for system plus surroundings, does not change. We say that such a property is subject to a *conservation law*.

Unfortunately, the labels "conserved" and "conservative" have been applied not only to other topics, such as politics and the environment, but also in physics to quantities that are constant, for the system alone, but only for selected processes; that is, for quantities that we choose to say are "preserved" for the particular conditions considered. For example, the velocity (*i.e.*, the speed and the direction) of a free particle does not change, so velocity has been (mis)labeled as a conserved quantity. It is more accurate, and much more helpful, to describe quantities such as velocity as a "constant of the motion", or a "preserved" quantity, for those motions for which it does not change. There is no conservation law for velocity.

The simplest description of a conserved quantity would be to say that it is constant for the entire universe. This may be satisfactory, but it immediately introduces questions about cosmology for which we may not have adequate answers. The difficulties are easily avoided by rewording the definition.

> A *conserved* quantity is constant for the system plus its surroundings, where by the surroundings we mean all of the universe, except the system, that may be affected by the process under consideration.

A property, such as density or temperature, that is the same for half the (uniform) system as for the entire system, is called an *intensive* property. A property such as weight or energy that is proportional to the amount of the (uniform) system is called an *extensive* property. An extensive property may or may not be conserved. An intensive property cannot be conserved.

An example of an extensive property of a system is the volume. At a constant temperature and pressure, the volume of a liquid is a constant as the liquid is poured from one container to another (without evaporation). At constant temperature, the volume of a copper block is constant. Even the volume of a gas-filled balloon is constant if temperature and pressure are constant. We can describe volume, under these conditions, as *preserved*, or as a *constant of the motion*.[6]

For more general conditions, in which temperature and/or pressure may change, the volume of the liquid, of the block, and, especially, of the balloon may change. What happens, then, to the volume of the surroundings?

The volume of the surroundings is also changed, equal and opposite to the change in volume of the system. If we choose the surroundings to be the room, minus the volume of the system, then the volume of the system plus the surroundings is the volume of the room and does not change when the volume of the system changes. *Volume is a conserved quantity*. The volume of system plus surroundings is constant for all processes, including those for which the volume of the system is changed.[7] Conservation of volume is the easiest conservation law to see and understand.

---

[6] The terminology is borrowed from mechanics. There is no motion involved explicitly.

[7] In more complex processes, such as those involving mixing of liquids or gases, more careful definitions of volume are required to divide total available volume into portions assignable to the system and to the surroundings. (See discussions of partial molal volumes.)

There are very few quantities that are subject to conservation laws.[8] Thus when we find, or recognize, a conservation law, it represents a significant "fact" in the structure of physics. Conservation of momentum was built into Newton's laws. Conservation of energy was first recognized by Mayer, Joule, and others about 1850. Most other conservation laws are much more recent.[9]

Conservation laws are extremely important in all branches of physics. They provide a substantial part of the unifying framework that ties together the different branches of physics. New conservation laws are added from time to time and some have been discarded, but the more fundamental conservation laws have withstood extensive testing and emerged unscathed.

2.9 The Importance and Limitations of the "Constants" of Physics

The seven basic concepts discussed above are independent. None can be derived from any of the others. But they are not sufficient to derive the necessary laws of physics, so we must add rules for specific situations. For example, we will add Newton's second law, which specifies how the motion of a body changes when an external force acts on it. We will add Coulomb's law for the forces exerted on and by electrically charged particles. We will add rules for rotational motion. Each of these applies in specific areas of physics, broad or narrow.

The seven basic concepts may be considered as load bearing walls of the physics structure. Other rules are important, especially in localized sections of physics, but they may be compared with interior walls. We can move partitions around, add doors and windows, and change paint colors from time to time, but the load-bearing principles keep the structure intact and stable. Whatever new discoveries are made in physics, it will require exceptionally strong evidence if new theories violate any of these fundamental principles.

---

[8] Because the labels "conservation" and "conservative" are so often misapplied, the reader must be very skeptical of descriptions involving these terms. Established terminology of physics sometimes dictates misuse of "conserved" and "conservative" when the quantity is not truly constant for the system and surroundings for *all* possible processes. Most textbooks, for example, introduce a far more complex terminology in which quantities that are *preserved* in special circumstances are labeled as (provisionally) "conserved". Then it becomes necessary to classify *processes* with a similar mixture of terminologies. A "conservative force" describes a process for which mechanical energy is preserved (but labeled as "conserved") and "non-conservative forces" act when mechanical energy is not preserved (labeled "not conserved"). Conservation laws are too important to the structure of physics to discard them for conformation to historical misuse.

[9] Emmy Noether specialized in modern abstract algebra, but in 1918, while working in Göttingen with Hilbert and Klein, she wrote an elegant exploration of the basis of conservation laws. She showed that there is a one-to-one correspondence between conservation laws and symmetry properties of space and time, which resolved difficulties encountered by Einstein in his general theory of relativity.

Chapter Summary

Definitions. The following terms are defined in the chapter.
Operational definition. A definition of any quantity that includes information on how the quantity can be measured.
Relativity. The assumption that absolute (linear) motion cannot be determined.
Equivalence. The assumption that a gravitational field is indistinguishable, by local measurements, from effects of an accelerated reference frame.
Uncertainty. The generalization that the products of the uncertainties in conjugate variables can never be less than $h/4\pi$.
Correspondence. The agreement of alternative descriptions of nature in the regions where the alternative approximations overlap.
Probability. The principle that systems are most likely to be found in the most probable (collections of equivalent, accessible) states.
Conservation. A generalization that certain quantities may be transferred between a system and its surroundings but will be constant for the (directly measurable) universe.

Addenda

References:

a. Isaac Newton, *Principia*, Vol. I, U. California Press, 1934. See especially pp. 2, 9.
b. Vladimir A. Fock (В. А. ФОК), *Fundamentals of Quantum Mechanics*, Second Russian Edition, Mir Publishers, 1976, translated by Eugene Yankovsky, 1978, Moscow. See especially chapter 1 on the initial experiment, the final experiment, and the implications for uncertainty.
c. Werner Heisenberg, *The Physical Principles of the Quantum Theory*, translated by Carl Eckart and F.C. Hoyt, 1930, Dover.
d. Robert P. Bauman, *Why Things Happen*, unpublished manuscript, 2007.

Uncertainty and the Schwarz Inequality:

Quantum-mechanical states are represented by mathematical *operators* (such as multiplication, derivatives, and matrix multiplication), which we represent here by Greek symbols, *e.g.*, $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$, which in general are not commutative; $\boldsymbol{\alpha\beta} \neq \boldsymbol{\beta\alpha}$. (This is a common characteristic of derivatives, matrices, and vectors in general.) Non-commuting operators have an important property relevant to their representation of physical operators, given by a mathematical theorem called the Schwarz inequality. Assume the commutator of $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ is some (initially unspecified) operator $\boldsymbol{\gamma}$. That is,

$$[\boldsymbol{\alpha}, \boldsymbol{\beta}] = \boldsymbol{\alpha\beta} - \boldsymbol{\beta\alpha} = i\boldsymbol{\gamma}$$

where $i = \sqrt{-1}$. The *dispersion* of $\boldsymbol{\alpha}$ is defined as the mean-square deviation of $\boldsymbol{\alpha}$ from its average value.

$$\Delta x = \overline{x^2} - \overline{x}^2 = \overline{\left(x - \bar{x}\right)^2} = <\left(x - <x>\right)^2>$$

where the super bar or brackets are alternative ways of representing the average value.

We wish to find the commutator of the dispersions, *i.e.*,

$$\left(<|\alpha^2|> - <|a|>^2\right)\left(<|\beta^2|> - <|\beta|>^2\right)$$

Without loss of generality, we may find the spread in $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ about average values taken as zero. "Turning the crank" on the mathematics shows that the product of the dispersions must be greater than, or equal to, the average value of $\boldsymbol{\gamma}$ over 2.

Interpreting, we find that the commutator of the operators gives us the minimum value of the dispersions, or more simply, if we let two physical quantities, such as position and momentum, be represented by operators for which the commutator is Planck's constant over 2π, as Heisenberg showed necessary to fit experimental measurements, then Heisenberg's uncertainty principle follows.

$$(\Delta x)^2 (\Delta p_x)^2 \geq \frac{\hbar^2}{4}$$

We choose a vector, **V**, that we represent in Dirac's notation. A vector, **i**, is represented by |**i**>, or simply | > (called a "ket"). When multiplied (as a column matrix vector may be multiplied by a square matrix) it is still a vector, ( )| >. When premultiplied by the transpose (conjugate) vector, < | (called a "bra"), we find the expectation value, or average value, for the operator,

$$\overline{\alpha} = \langle\ |\alpha|\ \rangle$$

Note that together they form a bra(c)ket: $\langle\ |\ \rangle$

Let

$$V = \left(\frac{\langle |\gamma| \rangle}{2i\langle |\alpha^2| \rangle} \alpha + \beta \right) |\ \rangle$$

Then the Schwarz inequality tells us that the square of this vector, $V^* V$, must be equal to or greater than zero.

$$|V|^2 = \langle\ \left| \left(\frac{-\langle\ |\gamma|\ \rangle}{2i\langle\ |\alpha^2|\ \rangle} \alpha + \beta \right) \left(\frac{\langle\ |\gamma|\ \rangle}{2i\langle\ |\alpha^2|\ \rangle} \alpha + \beta \right) \right|\ \rangle \geq 0$$

$$= \frac{\langle\ |\gamma|\ \rangle^2}{4\langle\ |\alpha^2|\ \rangle^2} \langle\ |\alpha^2|\ \rangle + \langle\ |\beta^2|\ \rangle - \frac{\langle\ |\gamma|\ \rangle}{2i\langle\ |\alpha^2|\ \rangle} \langle\ |\alpha\beta - \beta\alpha|\ \rangle \quad (f6)$$

But $\alpha\beta - \beta\alpha = [\alpha,\beta] = i\gamma$, so

$$|V|^2 = \frac{\langle\ |\gamma|\ \rangle^2}{4\langle\ |\alpha^2|\ \rangle} + \langle\ |\beta^2|\ \rangle - \frac{\langle\ |\gamma|\ \rangle^2}{2\langle\ |\alpha^2|\ \rangle} = \langle\ |\beta^2|\ \rangle - \frac{\langle\ |\gamma|\ \rangle^2}{4\langle\ |\alpha^2|\ \rangle} \geq 0$$

$$\langle\ |\alpha^2|\ \rangle \langle\ |\beta^2|\ \rangle - \frac{\langle\ |\gamma|\ \rangle^2}{4} \geq 0 \quad (f7)$$

or

$$(\Delta\alpha)^2 (\Delta\beta)^2 \geq \frac{\overline{\gamma}^2}{4} \quad (f8)$$

and

$$\Delta\alpha\ \Delta\beta \geq \frac{\overline{\gamma}}{2} \quad (f9)$$

Chapter 3. The Logic of Physics

Chapter Introduction

A. Why?

Understanding physics requires a type of thinking, or cognitive development, not achieved by a significant fraction of high school students. It may be necessary to stimulate this mental development to be able to communicate effectively with the students.

B. What to Look For

Science relies heavily on “If ... then ...” thinking, which sometimes appears as technical definitions and sometimes as “necessary and sufficient” conditions. Abilities differ widely, also, in spatial visualization, which can be quite important. Without looking for these skills, teaching cannot be matched to the needs of the student.

C. Pre-test for Prior Comprehension

1. Are you familiar with “associationism” and its relationship to implication?
2. Is the following statement True, False, or Uncertain? *Only reptiles are snakes.*
3. How can a Venn diagram represent simple implication?
4. Artists often draw a thin crescent Moon lying on its left side (open to the upper right). When (if ever) might you see such a Moon in the sky?
5. If y is proportional to x, what possible appearances might a graph of x *vs.* y have?
6. What is meant by “necessary and sufficient” conditions?

Inquiry

Excellent practice at development of logical thinking in the context of games can be found in the commercial educational games WFF’N PROOF, EQUATIONS, and ON WORDS, dealing with logic, algebra, and words (somewhat like Canasta), in a format suitable for very small groups or for classes. They are particularly effective because they encourage/require analysis of what strategy the opponent plans.

#####

Many students have recognized that learning physics requires a different type of thinking than most other courses. Teachers know that some students enter the course well prepared to grasp new material and others have difficulty. The difference does not correlate well with grade point average coming into class. By contrast, in at least one study,[1] there was a significant correlation of this cognitive preparation with final grade in introductory physics courses. What is the background difference? (Can it be bottled and sold?)

In tests on very large numbers of young students (up to about age 14, primarily) Piaget

[1] Robert P. Bauman, *J. Col. Sci. Teach.* ***6***, 94-96 (1976).

Short Answers to Pre-test Questions.

1. An “association” is often substituted for, or interpreted as, a cause and effect relationship.
2. There are no mammals, or birds, or fish that are snakes. Only reptiles are snakes. (T)
3. A Venn diagram represents ranges of possibilities as areas, thus clearly showing such properties as inclusion, exclusion, and partial overlap typical of “If...then...” statements.
4. Such a crescent Moon would be seen (in the northern hemisphere) rising in the east shortly before sunrise (soon before a new Moon).
5. If y is proportional to x, the graph must be a straight line passing through the origin.
6. “If and only if” statements express necessary and sufficient conditions.

*****

found a wide variety of thinking skills were correlated.[2] Renner[3] showed that most topics encountered in a high school physics course require the upper level of thinking skills (formal operational), but half or more of the students graduating from high school have not reached that level (“expected” by age 12-14). Other studies[4] showed that the Piaget correlations remain valid for older students, are subject to progression in older students, and are a major contributor to success in learning physics. Yet most high schools assume such progression is automatically achieved and do little, if anything, to move the students along in cognitive development.

Even a few hours of work on problem solving can make a statistically significant change in the ability of students to handle problems at the level of the upper high school/beginning college courses and examinations. Furthermore, experience has shown that students enjoy problem solving in a non-stressful environment. Thus the topic is doubly valuable for a first-class session. It can get a course started in an appropriate environment of “learning is fun”, and point students toward methods they can apply to learning better problem-solving skills.

It is important to remember that physics is not an accumulation of information. Computers are much better at “learning” and “remembering” (storing and retrieving) information than any of us, but no computer has ever passed a physics course, nor has a computer applied physics principles to create and run a business or perform even simple surgery. Physics is the application of logical concepts to some basic information to generate new information as required. Those who can think logically are most successful at mastering basic physics.

3.1 Logic and Problem Solving

The “great leap of understanding” may be a fond wish, but it seldom happens. Success is more

---

[2] Bärbel Inhelder and Jean Piaget, *The Growth of Logical Thinking from childhood to adolescence*, Basic Books, 1958.

[3] J.W. Renner and A.E. Lawson, *J. Col. Sci. Teach.* **5**, 89-92 (1975); *Phys. Teach*. **11**, 273-276 (1973).

[4] R.P. Bauman, T. Wdowiak, and I. Loomis, "Teaching for Cognitive Development", in *Cognitive Process Instruction at the College Level*, edited by John Lochhead and John Clement, Franklin Institute Press, Philadelphia, Pa., 1979, pp. 267-273.

likely to be achieved by breaking the problem down into small steps, each step requiring logical thought at its core.

What do we mean by logic, as it applies to physics? It is, of course, the same as the (valid) logic that may be applied in any other field, but perhaps with more diversity. Applications of logic in physics generally appear in a form that may be considered as problem solving, in the broadest sense.

Attitude is critical, perhaps even more than in sports (because physics is basically non-competitive — everybody *can* win). If you believe you can solve a problem, you have a much better chance of doing so than if you look at it and decide you won't be able to work through it. Even non-mathematicians, for example, found they could solve some mathematics problems better when they were asked, "How would you solve this *if you were a mathematician*?"

Accuracy is important, not in the sense of working to more decimal places (the best answer is sometimes only an order of magnitude) but accuracy in ensuring that each step proceeds logically from what went before. It is of no help to reach a solution if there is a break in the logical sequence.

Activity is often the key. When you seem to be stuck, just looking at a problem seldom helps, but making a drawing or a model, or working backward (ask "What would I have to know to get this answer easily?"), or trying, in parallel, a similar but simpler problem to see how it works, may get you underway.

One important form of activity is to break the problem down into smaller steps. A piece of roast beef you could choke on goes down well in small bites; the wall you cannot leap over is easily surmounted if you find the steps.

Should you guess? Some experts say yes,[5] some say no. The question must be better defined. Guessing an answer is often very helpful because it gives you a clue as to where you are going or what terms will be important. But no one else is much interested in your guess, and certainly you wouldn't want to pass off your guess as your best effort toward a solution. Making a guess before you begin may not only help point you in the right direction, but tell you, when you later compare that guess with the true solution, whether you understood what was happening before you went through the solution process. But *keep the guess to yourself*, while you then proceed by logical steps toward the answer.[6]

3.2 Awareness of the Solution Process

When it was observed that many students at the University of Chicago were unable to pass the comprehensive examinations given for admission to upper-class status (even though the students had been highly selected for ability at admission), Benjamin Bloom undertook to discover a cause. He found a high correlation of exam failure with poor problem-solving habits. Students seemed to have difficulty solving a problem, and then could not say how they had attempted to solve it.

---

[5] See especially Polya, G. *How To Solve It*, Princeton University Press, 1971.

[6] A peculiar exception may appear in some external examinations with multiple-choice answers. When only the total number of correct answers is tabulated, it pays to guess rather than leave the answer to a question blank. Such test-taking skills have little to do with learning.

He also found that a relatively short program directed specifically to this difficulty had a dramatic effect on student performance.[7]

Students were asked to work in pairs. One, acting as solver, would work a problem out loud. The other, acting as listener, would check to see whether

a. the problem was read accurately (it does help to be sure you are solving the correct problem);

b. each step was stated clearly and performed correctly; and

c. the conclusion reached was consistent with the step(s) leading to that conclusion.

On the next problem, the listener and solver exchanged roles.

3.3 Sample Problems

Examples given here[8] are some problems that may be solved by students working in pairs. One student should act as problem solver and the other as listener. The role of listener is as important as that of solver. The goal of such exercises is to become aware of your own thought processes, so they must be expressed aloud in order that the listener can confirm the process. (It has been found helpful to let students work each of these as a pair, then combine to get larger and larger groups (and eventually the full class) to compare answers and, as necessary, compare methods. The methods as well as the solutions in the initial solution process should be recorded by the listener.)

a. Bill is taller than Jack. Bill is shorter than Joe. Who is the shortest, Bill, Jack, or Joe?

b. Jean is fairer than Angela. Jean is darker than Edith. Who is the fairest, Jean, Angela, or Edith?

c. Ann is facing Rex. Ann turns to her right (90^o). Then Ann turns “about face” (180^o). How must Rex now turn to be facing in the same direction as Ann?

d. It takes 1 hour and 20 minutes for the bus to travel from Lafayette to Indianapolis. If it is due to arrive in Indianapolis at 6:05, what time should it leave Lafayette?

e. Three salesmen stopped for the night at the Sleepy Time Motel, where the clerk charged them $60 for their room, which they paid in advance ($20 each). When the manager looked over the records a few minutes later, he pointed out that they should have received the commercial rate of $55, so he told the clerk to take $5 up to their room. The clerk always had trouble with fractions and did not want to try splitting $5 three ways, so he gave each of the salesmen $1 and pocketed the other $2. That way each salesman paid $19 for his room (a total of $57). When the clerk got home he told his wife what he had done, but she pointed out that $57 + $2 is only $59. What had he done with that other dollar? How should the clerk explain the transaction to his wife?

---

[7] See Bloom, B.S., and Broder, L. J. *Problem-solving Processes of College Students*, Univ. of Chicago press, Chicago, 1950. For a later application of these ideas, see Whimbey, A., and Lochhead, J., *Problem Solving and Comprehension*, 4th edition, Erlbaum, Hillsdale, N.J., 1986, and reference 4.

[8] From R.P. Bauman, “A First Course in Physical Science”, Wiley, New York, 1987; p. 17.

f. A very old puzzle is embedded in the following rhyme.

As I was going to St. Ives
I met a man with seven wives.
Every wife had seven sacks;
Every sack had seven cats;
Every cat had seven kits.
Kits, cats, sacks, and wives.
How many were going to St. Ives?

3.4 Association and Implication

Early attempts at understanding logical implication are typically diverted by *associationism*. If the question is asked, "What makes the clouds move?", the most probable answer is "The wind". If asked "What makes the wind blow?" a young child (up to and including some college physics students, at least) may well answer, "The clouds". To these beginners in logical thinking, the word "because" indicates simply that one bit of information is associated with another bit. The wind blows, the leaves move, the clouds move. All are associated. There is no discrimination as to cause and effect.

Consider Figure 3.1, which shows some triangles, some circles, and some squares, some of which are black and some white. Is the statement, "If it's a triangle, then it's white" a true statement?

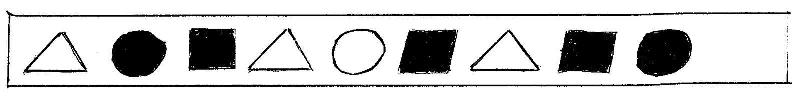


Figure 3.1

To answer the question safely, we note that it says, "If it's a triangle", so we should mark all those small figures that are triangles. It says, "then it's white", so we look just at the figures we have marked (the triangles) to see if they are *all* white. They are. So the statement is true.

We might also ask, for the same drawing, about the statement, "If it's a triangle, then it's not white." When we look at all the triangles, we find none of them are not white, so the statement is false.

Finally, we could consider the statement, "If it's white, then it's a triangle". So we mark (in some different way) all the small figures that are white. Then we look to see if all the newly marked figures are triangles. Some are; some are not. So the statement is true for some of the small figures but false for some of the others. For the drawing as a whole, the statement is *uncertain*.

There are two parts to such a sentence, in this case the part about being white and the part about being a triangle, so there are two choices as to which part goes first (*i.e.*, in the "if" condition phrase). For the "white" part, there are two choices (white or not white) and for the "triangle" part there are two choices (triangular or not triangular). That gives 2 x 2 x 2 = 8 possible sentences. How many of these sentences are true? How many are false? How many are uncertain?

Write out all eight variations on the sentence, "If it's a man, it has a head." (We can agree before hand to neglect decapitated corpses.) How many of the statements are true? How many are false? How many are uncertain? Can you see that there is a strict parallel between the white-triangle sentences and the man-head sentences?

John Venn, a mathematician who lived from 1834 to 1923, pointed out that such English statements can also be nicely represented geometrically. Draw an area that represents all objects with heads (men, women, children, tables, coins, dogs, *etc.*). Within that area, draw another area that corresponds to men (including women and children or not as you prefer). The "men" area is included in the "heads" area, but only part of the "heads" area is included in the "men" area.

English is a very flexible language that can express implication (If ... then ...) statements quite precisely (or, when you wish, quite fuzzily). Look back at the first drawing of white triangles, black and white circles, and black squares (Fig. 3.1). Is the sentence, "Only if it's white, it's a triangle" a true statement?

Again, to proceed with caution, we should mark all the small figures that are white, then look to see whether *only* these are triangles, or are there some other, unmarked figures that are triangles. Only the marked (white) figures are triangles, so the statement is true (even though there are marked (white) figures that are *not* triangles, as well). Notice that the "only" has the effect of inverting the statement: "If it's a triangle, then it's white" can be written "Only if it's white, it's a triangle".

A special type of "If ... then ..." statement is represented in Figure 3.2, a collection of black circles, white triangles, and white squares. Consider the statement, "If it's round, it's black." Consider also the statement, "If it's black, it's round." Both forms are true.

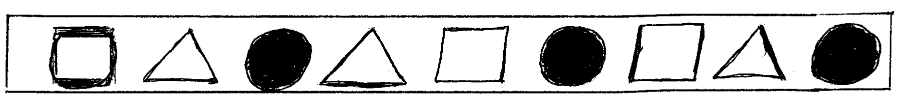


Fig. 3.2

We could therefore describe this last drawing with the two statements, "If it's round, it's black", and "If it's black, it's round", which is equivalent to "Only if it's round, it's black". It is an easy step to shortening the English: "If and only if it's round, it's black" says the same as the pair of statements, "If it's round, it's black" and "Only if it's round, it's black" or, equivalently, the pair of statements, "If it's round, it's black" and "If it's black, it's round".

Such pairs of statements, condensed into one, are extremely important. They describe what we call *necessary and sufficient* conditions. A definition (of a noun) is basically a necessary and sufficient definition, or an "If ... and only if ... " statement. That is, if *and only if* an object fits the description given in the definition, then the object fits the word described. Teachers and authors like to assume that students understand necessary and sufficient conditions.

Try the following examples, preferably working in pairs as solver/listener as before.

a. Consider Figure 3.3:

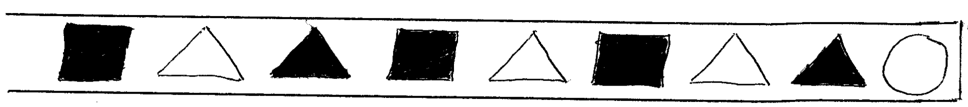


Fig. 3.3

Start with the statement, “If it’s black, it’s square.” Is this True, False, or Uncertain?

What are the seven other variations on blackness and squareness? Identify which are True, which are False, and which are Uncertain.

b. Assume the following statement to be true: If no force acts on an object, the object will not change its speed.

Does it follow, *from this statement*, that “If an object does not change its speed, no force acts on the object”?

c. Evaluate the following statement: “If *A* is true, then *B* is true. We observe that *B* is true. Therefore we may conclude that *A* is true.” (This has been labeled “political logic”.)

d. If an object is not not blue, what color is the object?

3.5 Beyond Simple Implication

Simple implication imposes strict requirements on the possible statements. In order, for example, to have $p \rightarrow q$ (*p* implies *q*), it must also be true that *not q* implies *not p*. But satisfying these requirements, by themselves, would indicate an “If and only if ...” statement. Therefore a simple implication statement must have at least one other possibility, which clearly *cannot* be *p* implies *not q*, nor can it be *not q* implies *p*.

Although simple “If ... then ... “ statements are the basis of much of the logic of physics and of life, there are other, more complex, variations that arise frequently, which most of us have already learned to handle in day to day situations. Without any claim of covering the field of logic, we will look at a different model that offers greater flexibility and may therefore be helpful in some situations.

Simple implication statements, of “If ... then ... ” form, are conveniently represented by symbols: The first phrase is represented by *p*; the second phrase by *q*. Then a simple implication statement could be written $p \rightarrow q$, *p* implies *q*. The negatives are represented by a bar over *p* or *q*.

$\overline{p}$ = not *p*
$\overline{q}$ = not *q*

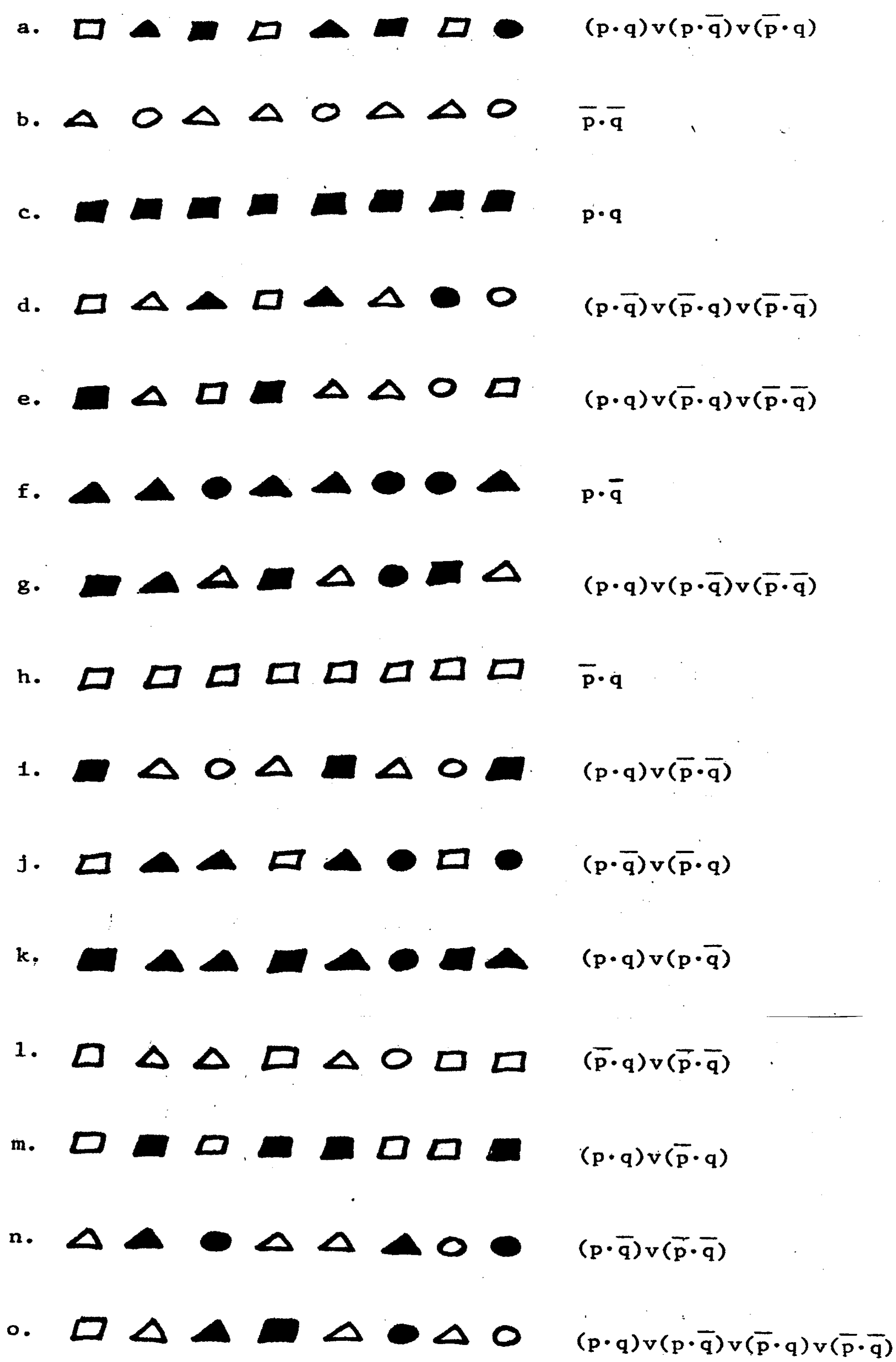
These individual segments can be combined in different ways: *p* may be true and *q* may also be true; then we write “*p* and *q*”, or simply $p{\cdot}q$, letting the dot represent “and” (a *conjunction*), with ($p{\cdot}q$) equivalent to ($q{\cdot}p$).

We may have a situation where *p* may be true *or q* may be true, but they cannot both be true (the *exclusive or*). Other times we mean “and/or”, as in “Do you want something to eat or drink?”. The and/or is called a *disjunction*, and may be represented by v, as in *p* v *q*. (Think of the v as representing *vs.*) The drawings of black and white squares, circles, and triangles in Figure 3.4 can then be represented by the symbolic statements given next to each. We let *p* = it is black, and *q* = it is square.

Note first of all that not all of these sets in Figure 3.4 can be represented by simple implication. Identify from the symbolic representation which of the 15 figures[9] are simple implication statements. Then confirm your analysis by checking the drawings. Identify also

[9] 15 is an odd number. The 16th possibility would be no black or square figures.

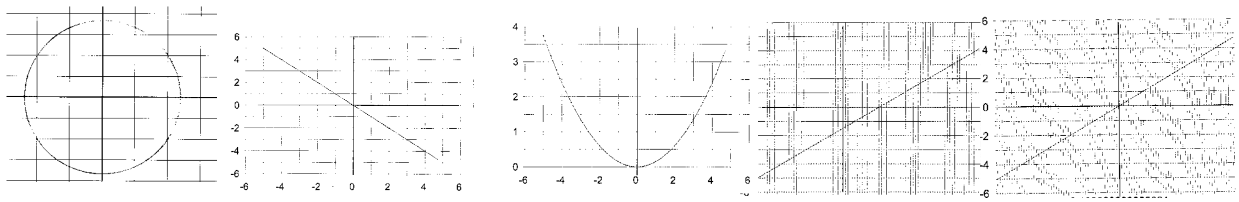
Figure 3.4


a. $(p\cdot q)v(p\cdot\bar{q})v(\bar{p}\cdot q)$
b. $\bar{p}\cdot\bar{q}$
c. $p\cdot q$
d. $(p\cdot\bar{q})v(\bar{p}\cdot q)v(\bar{p}\cdot\bar{q})$
e. $(p\cdot q)v(\bar{p}\cdot q)v(\bar{p}\cdot\bar{q})$
f. $p\cdot\bar{q}$
g. $(p\cdot q)v(p\cdot\bar{q})v(\bar{p}\cdot\bar{q})$
h. $\bar{p}\cdot q$
i. $(p\cdot q)v(\bar{p}\cdot\bar{q})$
j. $(p\cdot\bar{q})v(\bar{p}\cdot q)$
k. $(p\cdot q)v(p\cdot\bar{q})$
l. $(\bar{p}\cdot q)v(\bar{p}\cdot\bar{q})$
m. $(p\cdot q)v(\bar{p}\cdot q)$
n. $(p\cdot\bar{q})v(\bar{p}\cdot\bar{q})$
o. $(p\cdot q)v(p\cdot\bar{q})v(\bar{p}\cdot q)v(\bar{p}\cdot\bar{q})$

the “If and only if ...” examples.

3.6 Proportionality

Which of the graphs in the figure below show x proportional to y?



a *b* *c* *d* *e*

[a. Circle, centered about origin. b. Straight line with negative slope passing through origin. c. Parabola “sitting on” point (0,0). d. Straight line with positive slope passing near origin. e. Straight line with positive slope of about 0.05 passing through origin.]
Figure 3.5

Many introductory science textbooks give false or misleading statements concerning proportionality. For example, some confuse proportionality with linearity and some experienced teachers believe proportionality can be represented by a variety of shapes of graphs.

Perhaps the easiest explanation of a proportional relationship is that

if y is proportional to x, then doubling x will double y; tripling x will triple y; *etc.*
If y is proportional to x, then x is proportional to y.

To avoid minor ambiguities, we may assume that y is a single-valued function of x and x is a single-valued function of y. Then one and only one value of y corresponds to each value of x, and one and only one value of x corresponds to each value of y. (That would *not* be true, for example, if we had $y = x^2$ or $y = \sin x$, but these are not proportional relationships, anyway.)

3.7 Spatial Visualization

Because physics is concerned with objects in space, it is often necessary to envision how an object would appear if conditions were changed, including how an object would appear if you were observing it from some other location. The process is called *spatial visualization*. It is an acquired skill.

An example of spatial visualization that can be practiced, and useful, outside of class is interpreting the appearance of the Moon in terms of the geometry of Earth, Moon, and Sun. Start with the representation of the Moon as it might appear, as shown in Figure 3.6. Then answer the following questions. When the Moon appears as shown:

Figure 3.6.

1. What time of day is it?
2. What direction are you looking?

3. Is the Moon rising or setting?
4. Will the Moon appear in the same location in the sky at the same time, or earlier, or later, next day?
5. Is the Moon waxing (getting larger, in appearance) or waning (getting smaller)?

Other, more subtle, differences will appear as a consequence of your latitude, but even with a basic understanding you can tell something about directions and time of day anytime you can see the Moon, even in a strange area.

There are two keys to understanding phases of the Moon. First (as you could discover yourself by making observations over a few days), the (apparent) monthly rotation of the Moon about the Earth causes the Moon to rise later tomorrow than tonight..

*****

Example 3.1

A. Is the angular motion of the Moon about the Earth in the same direction, or opposite direction, to the rotation of the Earth?

B. A lunar month is about 29 days (depending on the reference frame for definition). How much later will the Moon rise tomorrow than today?

*****

Figure 3.7 a. Conventional Sun-Earth-Moon Diagram b. Classroom "Experiment"

The second key to understanding phases of the Moon is to grasp the significance of the position of the Sun in relation to the Moon and to the observer. Suspend a light-colored ball (such as a white polystyrene Christmas ball) in a semi-darkened room. Let a flashlight directed toward the ball represent the Sun. (Usually one student would hold the flashlight and other students would be observers. However, it is possible to mount the flashlight and ball in fixed positions for your own "checkout" observations.) Then, without moving the "Sun", move around the ball (looking *at the ball*) to discover where you will see a full Moon, where you will see a half Moon, and where you will see a crescent Moon. (A "new Moon" occurs when the Sun is directly opposite the observer, "behind" the Moon, and thus the new Moon is completely dark except for reflected light from the Earth, and is also obscured because of the bright Sun behind it.)

One caveat: When you believe you have located the necessary location of the Sun, while observing the Moon at night (*e.g.*, a half Moon), it may be difficult to fully appreciate the importance of the enormous distance to the Sun in determining angles. Textbook diagrams, such as Figure 3.6 (Earth, Sun, and Moon in four phases), may be helpful at this point in recognizing angles and directions.

3.8 Methods of Physics

Before we can apply logical analysis to a physical problem, we must know a great deal about the problem. Yet most problem descriptions, whether "real life" or "textbook" problems, are quite brief. Understanding the information that is (apparently) *not* there is perhaps the most important part of knowing how to apply physics to the solution of problems.

The universe is astonishingly complex. There are uncountable stellar-type objects, as well as unknown numbers of planets and smaller objects. Each of these has its own properties and peculiarities. Even on the Earth, there are uncounted numbers of living creatures. The atmosphere is subject to so many influences that its behavior cannot be predicted more than a short time ahead, and realistically, we cannot even describe the atmosphere at a given instant except in terms of averages over substantial areas at selected heights. We could continue indefinitely describing why we never really understand what is going on around us. There are too many objects and too many ways the objects can interact to allow an analysis of a significant fraction of the possibilities.

3.81. *Models*. How, then, can physicists expect to "explain" or "understand" the universe and its processes? The key is to break it down into simpler parts. We don't try to describe real objects. That would be too time consuming. It would take much too long to discover everything that might be known about any one object and the influences it experiences, and the answers would not be very helpful when we then wanted to look at some other real object. We choose, rather, to deal with simpler objects, or idealizations, which are *models* of objects in the real universe.

A model is a simplified representation of reality. Thus we represent the Sun and the planets as physical particles for initial analysis of orbits. We represent a wall as an impenetrable, immovable surface when we bounce a ball, represented as an elastic particle, off the wall. You may, if you wish, think of this as an elaborate code. It permits a very brief description of the situation that is to be analyzed, but more importantly, it tells the expert that a certain "box full" (or "chunk") of information already at hand may be applied to this particular problem, including geometric descriptions of the motion and algebraic formulas that describe the motion.

For example, springs are represented as being perfectly elastic (exactly following Hooke's law, $f = - k x$), without internal friction. A physicist looking at a mass attached to a spring automatically includes in the problem description such assumptions as the mass rigidly attached to one end of the spring, the other end of the spring rigidly attached to a supporting rigid surface, the mass acting like a (physical) particle, the spring acting like a perfect spring, and the combination moving without any friction with the air or nearby objects. Typically, the original model is analyzed first (substituting specific values for properties of the mass and the spring). Only then, and only when necessary, is the model modified to include additional details (such as the mass of the spring, internal friction that decreases the energy of motion, friction with the air, and so forth). The problem is described, among physicists, as "a mass on a spring", adding refinements when necessary.

Much of the process of learning physics is learning the models that are implied (and often never quite explicitly discussed). We will attempt to present these models explicitly, especially in the early chapters, so that you may gain practice in seeing how the models are set up and applied. This will make some parts of the discussion seem longer than necessary, but in the long run, if you learn to go through the process yourself, it will soon become automatic and you will be better able to explain the process to students. This may save much time, and anguish, later. The key is to *understand what you are trying to do before you attempt to do it.* In other words, an important part of the overall model of the universe employed by physicists is to describe

Answers 3.1

A. The Moon rotates in the same direction as the Earth, so the Earth must go slightly more than one revolution to catch up to the angular position of the Moon.

B. A mean lunar month is 29.53 d. It takes about 29 days for the Moon to go around the Earth once, so in one day it goes 1/29 of a circle. It takes 24 h for the Earth to rotate once, so $1/24^{th}$ of a circle would represent 1 h. Calculating more accurately, 1/29.5324 h x 60 min/h = 49 min approximately for the Earth to "catch up" with the orbital motion of the Moon. Thus the Moon rises about 50 minutes later each day.

(The length of the month depends on whether you include the extra circuit of the Earth to compensate for diurnal rotation. Because Earth and Moon are moving about the Sun, the Moon never actually makes a loop about the Earth.)

*****

certain *objects*[10], and certain *agents* that act on those objects. You must understand the objects, and agents, which means you must be able to visualize the model under investigation.

3.82. *Descriptors.* Because we follow the logical development proposed by Newton, the *agents* will initially always be *forces*, pushes or pulls exerted by some other object (perhaps unspecified) on the object of interest. For example, if the motion of a ball (the object) is altered by a collision with a wall, we may simply describe the wall by the force that it exerts on the ball to prevent the ball from continuing in its original direction with its original speed. If an object, such as a baseball, is thrown, it will be acted upon by an agent which is the force exerted by the Earth pulling the ball downward (which we will describe as the ball acted on by a *gravitational field*).

Each object, and each agent, must be specified before we can adequately predict their behavior. That is, we must know the *variables*, and hence the *properties* of the objects, and agents. Consider, for now, only the objects. We can think of this as a sub-model, or a model within our overall model. Is the object a point particle? That is quite unlikely, in physics, as we will see (especially in section 7.1). Is it a physical particle, for which any internal structure is ignorable? Or is it an elastic particle, or a liquid, or a gas, or an electric charge carried by a physical particle? Do we need to know the mass, or the shape, or the size, or the hardness, or other properties of the object? These may be called *static variables*. They change from one object to another, but typically they remain constant for any one object during the process we are examining.

In addition, the object will generally be described by variables that do change. Such properties as the position of the object, its speed, and its direction of motion would be typical of descriptors

---

[10] Recall that it will be convenient to replace *objects* with "systems", reflecting the fact that what is acted upon may, itself, be complex, consisting of more than one part, each of the parts having distinctive properties. For the moment, the term *object* will be sufficient and probably more familiar.

we call *state variables*. They describe the *state* of the object, including its *state of motion.*

To help you ease into the methods of physics, we will begin with descriptions of motion in which we ignore the agents causing the changes in state variables. In other words, we start with the (partial) answers in terms of how the object moves, and examine only how we describe the motion and relate each of the state variables to other state variables. Then, in chapter 5, we will introduce the agents that cause the changes of motion. If you practice identifying the variables, or descriptors, applying to moving objects in the early chapters, you will be prepared for the next step of describing the agents that cause changes in the state of the object. In every type of situation, you should look for what model is being considered, and how that model can be represented geometrically (*e.g.*, by a graph of position against time, or some other space-time relationship) and how it can be represented algebraically, with a mathematical equation.

It is always more fun to be "on the inside". If you learn the codes for describing the models, you will find physics to be much easier than if you attempt to memorize formulas and match formulas to problems. When the problem is understood, then you will be prepared to apply logical analysis to the problem. *If* the conditions are such and so, *then* the objects will behave in such a way. (Which of course is *not* the same as saying that if the object behaves in a certain way, then the conditions are necessarily such and so. More than one set of conditions *may* lead to the same observed behavior.)

Chapter Summary

Some studies of problem solving have shown that one of the most important requirements for solving problems is to be aware of your own thought processes in the steps of problem solution.

Understanding "implication" (*e.g.*, "If ... then ..." statements) is a necessary step toward understanding science and mathematics.

A recurring principle in all of science and mathematics is proportionality, which has the algebraic form $y = a\,x$, but is often confused with linearity or direction of change.

It is dangerous to assume students are familiar with implication, necessary and sufficient conditions, and proportionality until you have confirmed their state of understanding.

Spatial visualization is an important skill, gained through practice.

For necessary economy of effort, descriptions are sought for selected models, that include physical objects and the agents that act on them.