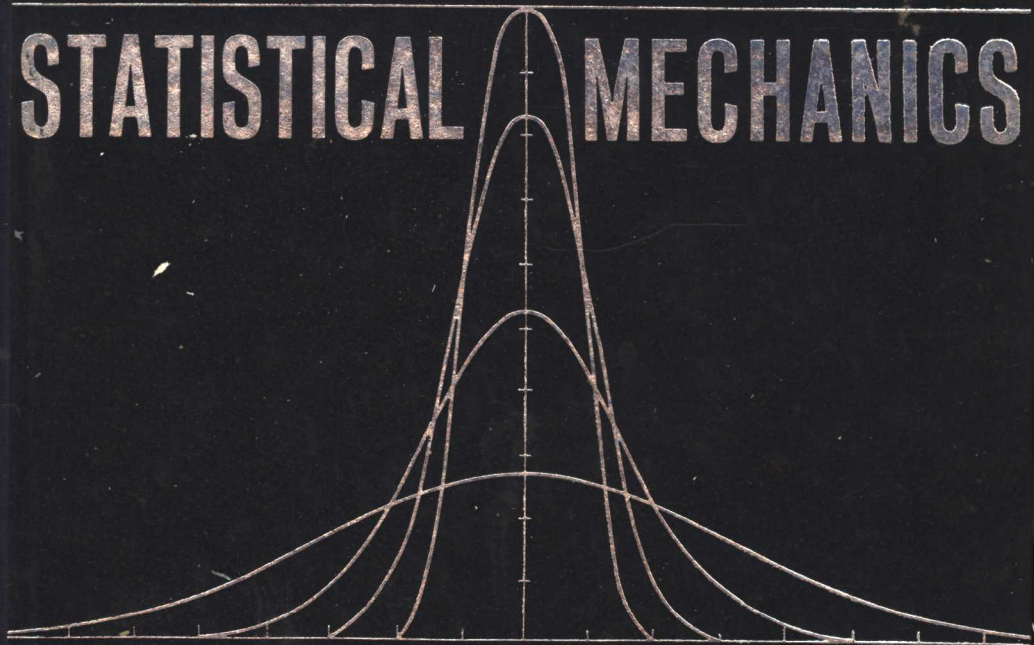


# STATISTICAL MECHANICS



NORMAN DAVIDSON

# **STATISTICAL MECHANICS**

**NORMAN DAVIDSON**

*Professor of Chemistry  
California Institute of Technology*

**McGRAW-HILL BOOK COMPANY, INC.** 1962

*New York San Francisco Toronto London*

## Preface

The statistical mechanics of dilute systems of independent particles at equilibrium is a subject which is essentially fully developed. The practicing chemist should be able to apply this theory with assurance and accuracy to calculate the thermodynamic properties of substances in the ideal-gas state from molecular structure data.

My first objectives have been to develop this theory in a simple, logical, and understandable way, and to describe the applications in sufficient detail, with illustrative problems, so that the reader can acquire the desired skill in this elementary part of the subject.

In the latter part of the book, I have tried to give a simple but rigorous account of the theories of the canonical ensemble and of the grand ensemble for treating systems of interacting particles. Several rather elementary but important examples are then studied, to illustrate the applications of these general theories. The most prominent and important manifestation of fluctuation phenomena for the experimental scientist is noise. I have therefore included a discussion of this topic in the chapter on fluctuations. However, I have not included a serious discussion of practical theories of liquids and dense gases. This is an important topic in statistical mechanics but it is difficult and complicated. The theoretical calculations do not as yet give good agreement with experiment. There are several excellent treatises by experts in the field, and I prefer to refer the interested student to this literature. I hope that I have carried the theory of systems of interacting particles far enough to provide a good foundation for such further study.

In addition, the latter half of the book presents discussions of some topics—paramagnetism, dielectrics, and ionic solutions—essentially from the point of view of the statistical mechanics of independent particles—but with at least an indication of how the theories of the canonical ensemble and the grand ensemble can be applied when interparticle interactions are important.

This text has developed from my lecture notes for a one-year graduate course at the California Institute. Many of the students are first-year graduate students (with an occasional bright senior) who have not had

a serious course in quantum mechanics, although they are usually taking such a course simultaneously. I have therefore included a chapter which is an introductory discussion of the relevant parts of quantum mechanics. Fortunately, the principles of equilibrium statistical mechanics can be developed very satisfactorily on the basis of an elementary and unsophisticated formulation of quantum mechanics.

I hope that the problems are instructive and/or interesting. They do not however cover the principles and their application so well that the ability to do them demonstrates a mastery of statistical mechanics as expounded in the text. Therefore, in examinations and homework assignments, in addition to problems I often ask for derivations which are given in the text. Some problems are inserted in the body of the text because I consider it desirable that they be done before proceeding further.

At the end of the course, I usually ask for a short paper on a subject of the student's own choosing. There is a tendency for the student to bite off more than he can chew; nevertheless, the results are generally healthy. Occasionally, I ask my students to make up some new problems and, if possible, to solve them. The results of this assignment are interesting and informative. Several of the problems in the text were obtained in this way.

I have solicited and received help and advice from so many colleagues that it would be unwise to attempt individual acknowledgments. The students in my course have contributed much to this book by their conscious criticisms and by my observations of their natural reactions (including some yawning and sleeping). The cheerful cooperation and painstaking care of Mrs. Ruth Hanson, Della Brown, and Allene Luke of the departmental secretarial staff are deeply appreciated.

I do have a special debt to Dr. Verner Schomaker, Dr. Robert Mazo, and the late Dr. William Moffitt. Each, in his own way, has added greatly to my understanding by answering many questions and discussing many problems with me.

I am dedicating this book to William Moffitt as an expression of my admiration and affection. It was from my conversations with him during my one-year stay at Harvard that I first gained confidence in the validity of my approach to statistical mechanics. Without this confidence, I would not have had the courage to write this book. In his own work, Bill Moffitt was a theorist with a passion for elegance and generality; but he insisted that the function of the theorist was to be useful. I hope that, were he still alive, he would think this book useful.

*Norman Davidson*

# Contents

<b>PREFACE</b> . . . . .	vii
<b>1. Introduction</b> . . . . .	1
<b>2. Classical Mechanics</b> . . . . .	4
<b>3. Quantum Mechanics</b> . . . . .	22
<b>4. Thermodynamics</b> . . . . .	43
<b>5. Mathematics</b> . . . . .	65
<b>6. The Statistical Mechanics of a System of Independent Particles</b> . . . . .	75
<b>7. Statistical Mechanics and Chemical Equilibrium</b> . . . . .	97
<b>8. Distribution Laws, Partition Functions, and Thermodynamic Functions for Atoms and Diatomic Molecules</b> . . . . .	105
<b>9. Nuclear Spin Statistics. Isotope Effects.</b> . . . . .	130
<b>10. Classical Statistical Mechanics. The Kinetic Theory of Gases.</b> . . . . .	146
<b>11. Polyatomic Molecules</b> . . . . .	169
<b>12. Black-body Radiation</b> . . . . .	211
<b>13. Canonical and Grand Canonical Ensembles</b> . . . . .	240
<b>14. Fluctuations, Noise, Brownian Motion</b> . . . . .	265
<b>15. Real Gases</b> . . . . .	310
<b>16. Lattice Vibrations and the Thermodynamic Properties of Crystals. Order-Disorder Phenomena</b> . . . . .	351
<b>17. A Digression on Electricity, Magnetism, and Units.</b> . . . . .	394
<b>18. Dielectric Phenomena</b> . . . . .	402
<b>19. Magnetic Phenomena</b> . . . . .	428
<b>20. Distribution Functions and the Theory of Dense Fluids</b> . . . . .	468
<b>21. Solutions of Electrolytes. The Debye-Hückel Theory</b> . . . . .	485
<b>GENERAL REFERENCES</b> . . . . .	528
<b>Appendix 1. Physical Constants and Conversion Factors.</b> . . . . .	530
<b>Appendix 2. Thermodynamic Functions for the Harmonic Oscillator</b> . . . . .	532
<b>INDEX</b> . . . . .	535

# 1

## Introduction

A macroscopic system at equilibrium has certain properties, for example, the energy, heat capacity, entropy, volume, pressure, and coefficient of expansion, which are of particular interest in thermodynamics. From the standpoint of thermodynamics, some of these quantities must be determined by experiment; the laws of thermodynamics provide relations by which it is then possible to calculate other quantities. Thus, if we know from experiment the equation of state of a substance in the form  $V = V(T, P)$  and we know the entropy at one pressure,  $P$ , we can calculate the entropy at any other pressure and the same temperature from the thermodynamic relation  $(\partial S / \partial P)_T = -(\partial V / \partial T)_P$ .

Our intuition tells us that it should not be necessary to measure the macroscopic properties of a system but that it should be possible to calculate them if the properties of the constituent molecules and the laws of force (the intermolecular interactions) between the molecules are known.

Statistical mechanics is a method, and in practice the method, for calculating the properties of macroscopic systems from the properties of the constituent molecules. Quantum mechanics provides the fundamental laws for calculating the properties of individual molecules and their intermolecular interactions. Statistical mechanics starts with these results and introduces a statistical hypothesis about the behavior of systems containing a large number of molecules. (The hypothesis that we shall use is actually that of equal a priori probabilities of individual quantum states; but its exact nature does not concern us right now.) It is then possible to predict many of the important properties of macroscopic systems.

This is not the only conceivable method for making predictions about macroscopic systems. One could, in principle, resort to a straightforward mechanical calculation. Consider, for example, a gas containing  $10^{24}$  atoms. Suppose that, in this case, classical mechanics is a satisfactory approximation and that we need not use quantum mechanics. It is necessary to know, at some initial time  $t_0$ , the  $3 \times 10^{24}$  position

coordinates and the  $3 \times 10^{24}$  velocity coordinates of all the particles. As we shall see in the next chapter, one can then, in principle, solve the equations of motion and calculate the positions and velocities of all the particles at all future times.

There are two difficulties with this direct approach. In the first place, the calculation is far too complex and cannot actually be performed. But suppose that, with a fantastically effective computer, it were possible to calculate the trajectories of all the particles. The results might be a gigantic data sheet giving the  $3 \times 10^{24}$  position coordinates of the atoms every  $10^{-11}$  sec. (This interval of time is chosen as reasonable because a typical atom at room temperature and atmospheric pressure undergoes a collision about every  $10^{-10}$  sec.) The history of the system for 1 sec would require  $3 \times 10^{35}$  entries. Such an enumeration of the data would be quite indigestible. We would look for a statistical summary of the data, and we would calculate certain statistical functions: the number of atoms with velocities in a certain interval, the number of collisions between atoms per second, the average number of atoms that are within a given distance of another atom at any particular time, the momentum exchange with the walls in any time interval, etc.

We shall see that by the methods of statistical mechanics it is possible to calculate these functions directly, without first calculating the detailed behavior of the system. A knowledge of these statistical functions is usually sufficient for the calculation of the macroscopic properties of a system. The statistical mechanical calculation is not just a cowardly expedient that we resort to because of our inability to make a complete calculation (although it is that, too); for most problems it contains all the information that we want about the system without going into unnecessary detail.\*

Equilibrium statistical mechanics treats the properties of systems at equilibrium. The calculation of the time-varying properties of a system which is not at equilibrium is more difficult. We are then interested in such properties as viscosity, heat conductivity, paramagnetic relaxation times, and chemical reaction rates. This is, in general, nonequilibrium statistical mechanics. It is not yet nearly so well developed a subject as equilibrium statistical mechanics. We shall be principally, but not exclusively, concerned with topics in equilibrium statistical mechanics.

Statistical mechanics is firmly based on quantum mechanics. The

\* It is interesting to note, however, that some difficult statistical mechanical problems are now being studied by detailed calculations of the mechanical behavior of small prototype systems. In one such calculation, the behavior of a system of hard spheres containing 32 particles was computed through 7,000 total collisions in an hour with an IBM-704 computer [B. J. Alder and T. E. Wainwright, *J. Chem. Phys.*, **27**: 1208 (1957)].

usual presentations of quantum mechanics presume a prior knowledge of classical mechanics; furthermore, classical mechanics is directly useful for many problems in statistical mechanics. We shall see that thermodynamics is closely related to statistical mechanics. Therefore, in the next three chapters, we review classical mechanics, quantum mechanics, and thermodynamics.



# 2

## Classical Mechanics\*

**2-1. Introduction.** We begin with a brief review of classical mechanics. We shall derive the equations of motion in Hamiltonian form from the more familiar Newtonian equations and shall introduce the concept of phase space.

**2-2. Mathematical Prelude.** Our object in this section is to illustrate, for a simple case, some of the mathematical operations needed in transforming the equations of motion from one system of coordinates to another. We shall treat the same problem for the general case in the next section.

Consider a single particle constrained to move in the  $xy$  plane. Newton's equations of motion are

$$m \frac{d^2x}{dt^2} = m\ddot{x} = F_x \quad m \frac{d^2y}{dt^2} = m\ddot{y} = F_y \quad (2-1)$$

where  $F_x$  and  $F_y$  are the forces in the  $x$  and  $y$  directions on the particle. We assume that the forces are derivable from a potential energy  $U(x,y)$ ,

$$F_x = - \frac{\partial U(x,y)}{\partial x} \quad F_y = - \frac{\partial U(x,y)}{\partial y} \quad (2-2)$$

and that the potential energy is a function of the position coordinates of the particle, but not an explicit function of time or of the velocities  $\dot{x}$  and  $\dot{y}$ . If these conditions are satisfied, the system is said to be conservative.

Suppose, for example, that the potential function is

$$U(x,y) = \frac{1}{2}ax^2 + bxy + \frac{1}{2}cy^2 \quad (2-3)$$

The equations of motion then are

$$\begin{aligned} m\ddot{x} &= -ax - by \\ m\ddot{y} &= -bx - cy \end{aligned} \quad (2-4)$$

\* A word of apology is in order. As regards mathematical content, this beginning chapter is one of the more difficult ones in the text. This may be bad pedagogy, but it is advantageous to base our further studies on a more general formulation of classical mechanics than Newton's laws of motion as presented in elementary physics classes.

The equations of motion (2-4) can be integrated to give a solution

$$x = x(t) \quad y = y(t) \quad (2-5)$$

which we call a trajectory. The particular trajectory depends upon the initial conditions, for example, the positions and velocities at  $t = 0$ , and there is a family of solutions for different initial conditions.

The kinetic energy  $K$  is

$$K(x, y) = \frac{1}{2}m \left( \frac{dx}{dt} \right)^2 + \frac{1}{2}m \left( \frac{dy}{dt} \right)^2 = \frac{1}{2}m\dot{x}^2 + \frac{1}{2}m\dot{y}^2 \quad (2-6)$$

The Lagrangian function  $L$  is defined as

$$L(x, y, \dot{x}, \dot{y}) = K - U = \frac{1}{2}m\dot{x}^2 + \frac{1}{2}m\dot{y}^2 - \frac{1}{2}ax^2 - bxy - \frac{1}{2}cy^2 \quad (2-7)$$

If we transform to polar coordinates,

$$\begin{aligned} x &= r \cos \phi \\ y &= r \sin \phi \end{aligned} \quad (2-8a)$$

then

$$\begin{aligned} \dot{x} &= \dot{r} \cos \phi - r \sin \phi \dot{\phi} \\ \dot{y} &= \dot{r} \sin \phi + r \cos \phi \dot{\phi} \end{aligned} \quad (2-8b)$$

By substitution in (2-6) and (2-3), we find

$$K = \frac{1}{2}m\dot{r}^2 + \frac{1}{2}mr^2\dot{\phi}^2 \quad (2-9a)$$

$$U = \frac{1}{2}ar^2 \cos^2 \phi + br^2 \sin \phi \cos \phi + \frac{1}{2}cr^2 \sin^2 \phi \quad (2-9b)$$

so that

$$\begin{aligned} L(r, \phi, \dot{r}, \dot{\phi}) &= \frac{1}{2}m\dot{r}^2 + \frac{1}{2}mr^2\dot{\phi}^2 - \frac{1}{2}ar^2 \cos^2 \phi \\ &\quad - br^2 \sin \phi \cos \phi - \frac{1}{2}cr^2 \sin^2 \phi \end{aligned} \quad (2-9c)$$

We notice that the position variable  $r$  enters into the expression for  $K$  and that the expression is quadratic in the velocities.

In Eq. (2-7), we can regard the Lagrangian  $L$  as a function of the independent variables  $x$ ,  $y$ ,  $\dot{x}$ , and  $\dot{y}$ . We can then write

$$\begin{aligned} \frac{\partial L(x, y, \dot{x}, \dot{y})}{\partial \dot{x}} &= m\dot{x} & \frac{\partial L}{\partial \dot{y}} &= m\dot{y} \\ \frac{\partial L(x, y, \dot{x}, \dot{y})}{\partial x} &= -ax - by & \frac{\partial L}{\partial y} &= -bx - cy \end{aligned}$$

For  $L$  as a function of polar coordinates  $r$ ,  $\phi$ ,  $\dot{r}$ ,  $\dot{\phi}$ ,

$$\frac{\partial L(r, \phi, \dot{r}, \dot{\phi})}{\partial \dot{\phi}} = mr^2\dot{\phi}$$

$$\frac{\partial L(r, \phi, \dot{r}, \dot{\phi})}{\partial r} = m\dot{\phi}^2 - ar \cos^2 \phi - 2br \sin \phi \cos \phi - cr \sin^2 \phi$$

and similar equations for  $\partial L / \partial \dot{r}$ ,  $\partial L / \partial \phi$ .

Sometimes the variables which are held fixed during partial differentiation will be indicated by the notation

$$\left(\frac{\partial L}{\partial \phi}\right)_{r, \phi, \dot{r}}$$

This can also be indicated by writing

$$\frac{\partial L(r, \phi, \dot{r}, \phi)}{\partial \phi}$$

Where the context makes the meaning clear, we shall often write simply

$$\frac{\partial L}{\partial \phi}$$

[Incidentally, note the difference between the ways in which a natural scientist and a mathematician regard function notation. When we write  $L(x, y, \dot{x}, \dot{y})$  we mean the physical quantity, the Lagrangian function,  $K - U$ , expressed in cartesian coordinates, as in Eq. (2-7), and  $L(r, \phi, \dot{r}, \dot{\phi})$  means the same physical quantity,  $K - U$ , expressed in polar coordinates. To a mathematician, the function  $L(x, y, \dot{x}, \dot{y})$  means the functional form

$$L(x, y, \dot{x}, \dot{y}) = \frac{m}{2} \dot{x}^2 + \frac{m}{2} \dot{y}^2 - \frac{1}{2} a x^2 - b x y - \frac{1}{2} c y^2$$

so that  $L(r, \phi, \dot{r}, \dot{\phi})$  would be the same form with  $r$  in place of  $x$ ,  $\phi$  in place of  $y$ , etc.; i.e.,

$$L(r, \phi, \dot{r}, \dot{\phi}) = \frac{m}{2} \dot{r}^2 + \frac{m}{2} \dot{\phi}^2 - \frac{1}{2} a r^2 - b r \phi - \frac{1}{2} c \phi^2$$

For correct mathematical use of function notation, if  $L(x, y, \dot{x}, \dot{y})$  is defined by (2-7), then the transformation (2-8) would transform  $L$  to a new function  $M(r, \phi, \dot{r}, \dot{\phi})$ , with

$$\begin{aligned} L[x(r, \phi), y(r, \phi), \dot{x}(r, \phi, \dot{r}, \dot{\phi}), \dot{y}(r, \phi, \dot{r}, \dot{\phi})] &= M(r, \phi, \dot{r}, \dot{\phi}) \\ M(r, \phi, \dot{r}, \dot{\phi}) &= \frac{1}{2} m \dot{r}^2 + \frac{1}{2} m r^2 \dot{\phi}^2 - \frac{1}{2} a r^2 \cos^2 \phi \\ &\quad - b r^2 \sin \phi \cos \phi - \frac{1}{2} c r^2 \sin^2 \phi \end{aligned}$$

so that the function  $M$  is the function which, in physical language, we called  $L(r, \phi, \dot{r}, \dot{\phi})$  in (2-9c).]

For arbitrary variations in  $x$ ,  $y$ ,  $\dot{x}$ , and  $\dot{y}$ , the variation in  $L$  is given by

$$dL = \frac{\partial L}{\partial x} dx + \frac{\partial L}{\partial y} dy + \frac{\partial L}{\partial \dot{x}} d\dot{x} + \frac{\partial L}{\partial \dot{y}} d\dot{y}$$

For a particular trajectory,  $x$ ,  $y$ ,  $\dot{x}$ , and  $\dot{y}$  are known functions of  $t$ . The variation of  $L$  with time can then be calculated from the equation

$$\frac{dL}{dt} = \frac{\partial L}{\partial x} \frac{dx}{dt} + \frac{\partial L}{\partial y} \frac{dy}{dt} + \frac{\partial L}{\partial \dot{x}} \frac{d\dot{x}}{dt} + \frac{\partial L}{\partial \dot{y}} \frac{d\dot{y}}{dt} = \frac{\partial L}{\partial x} \dot{x} + \frac{\partial L}{\partial y} \dot{y} + \frac{\partial L}{\partial \dot{x}} \ddot{x} + \frac{\partial L}{\partial \dot{y}} \ddot{y}$$

**Problem 2-1.** For a system of two particles moving in two dimensions, with masses  $m_1$  and  $m_2$  and cartesian coordinates  $x_1, y_1, x_2, y_2$ , the kinetic energy is

$$K = \frac{1}{2}m_1(\dot{x}_1^2 + \dot{y}_1^2) + \frac{1}{2}m_2(\dot{x}_2^2 + \dot{y}_2^2)$$

We now replace  $x_1, y_1, x_2, y_2$  by four new variables,  $X, Y, x_{12}, y_{12}$ , where

$$\begin{aligned} X &= \frac{m_1x_1 + m_2x_2}{m_1 + m_2} & Y &= \frac{m_1y_1 + m_2y_2}{m_1 + m_2} \\ x_{12} &= x_2 - x_1 & y_{12} &= y_2 - y_1 \end{aligned}$$

$X$  and  $Y$  are the coordinates of the center of gravity;  $x_{12}$  and  $y_{12}$  are the relative coordinates, which give the position of the second particle with respect to the first. Express  $K$  in terms of the velocities  $\dot{X}, \dot{Y}, \dot{x}_{12}, \dot{y}_{12}$  in the new system of variables. Explain the significance of this calculation.

By analogy, you can now write the corresponding expressions for the kinetic energy for a system of two particles in three dimensions.

**Problem 2-2.** The transformation between spherical polar coordinates and cartesian coordinates, as illustrated in Fig. 2-1, is

$$\begin{aligned} x &= r \sin \theta \cos \phi \\ y &= r \sin \theta \sin \phi \\ z &= r \cos \theta \end{aligned}$$

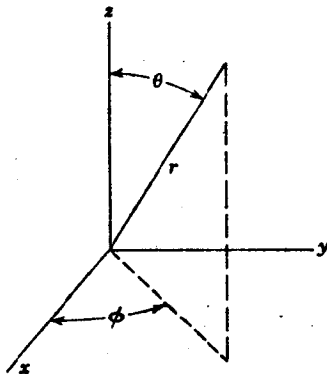


Fig. 2-1. Relation between cartesian coordinates and spherical polar coordinates.

Express the kinetic energy of a single particle in terms of its spherical coordinates  $r, \theta, \phi$ , and  $\dot{r}, \dot{\theta}, \dot{\phi}$ .

**2-3. The Lagrangian Equations of Motion.\*** There are several formulations of the laws of mechanics that are more general than Newton's

\* The subject matter of the next few sections is discussed in innumerable texts on classical mechanics. There are also clear expositions in Pauling and Wilson, "Introduction to Quantum Mechanics" [21], and Eyring, Walter, and Kimball, "Quantum Chemistry" [20].

equations. The two that we shall consider are the equations of motion in the Lagrangian form and in the Hamiltonian form. These formulations are easier to apply than Newton's equations for a number of problems in mechanics—especially when the most suitable coordinate system is not cartesian coordinates. However, our principal purpose is to derive Hamilton's equations, because these equations are used in the formulation of quantum mechanics and because they play a central role in statistical mechanics. In our treatment, Hamilton's equations of motion will be derived from Lagrange's equations, which will be derived from Newton's equations.

Consider a system composed of  $n$  particles, with masses  $m_1, \dots, m_i, \dots, m_n$ , and cartesian coordinates  $x_1, y_1, z_1, \dots, x_n, y_n, z_n$ .

The kinetic energy is given by

$$K = \frac{1}{2} \sum_{i=1}^n m_i (\dot{x}_i^2 + \dot{y}_i^2 + \dot{z}_i^2) \quad (2-10)$$

We assume that there is a potential,  $U(x_1, \dots, z_n)$ , which is a function of the position coordinates only.

Newton's equations of motion are

$$\begin{aligned} m_i \ddot{x}_i &= - \frac{\partial U}{\partial x_i} \\ m_i \ddot{y}_i &= - \frac{\partial U}{\partial y_i} \\ m_i \ddot{z}_i &= - \frac{\partial U}{\partial z_i} \end{aligned} \quad i = 1, \dots, n \quad (2-11)$$

From (2-10), regarding  $K$  as a function of  $\dot{x}_i, \dot{y}_i$ , and  $\dot{z}_i$ , we have

$$\frac{\partial K}{\partial \dot{x}_i} = m_i \dot{x}_i$$

The Lagrangian function  $L$  is defined by

$$L(x_1, \dots, z_n; \dot{x}_1, \dots, \dot{z}_n) = K(\dot{x}_1, \dots, \dot{z}_n) - U(x_1, \dots, z_n) \quad (2-12)$$

and we see that

$$\frac{\partial L}{\partial \dot{x}_i} = \frac{\partial K}{\partial \dot{x}_i} \quad \frac{\partial L}{\partial x_i} = - \frac{\partial U}{\partial x_i} \quad (2-13)$$

so that Eqs. (2-11) can be rewritten as

$$\begin{aligned}
 \frac{d}{dt} \left( \frac{\partial L}{\partial \dot{x}_i} \right) &= \frac{\partial L}{\partial x_i} \\
 \frac{d}{dt} \left( \frac{\partial L}{\partial \dot{y}_i} \right) &= \frac{\partial L}{\partial y_i} \\
 \frac{d}{dt} \left( \frac{\partial L}{\partial \dot{z}_i} \right) &= \frac{\partial L}{\partial z_i}
 \end{aligned}
 \tag{2-14}$$

These are the equations of motion in Lagrangian form for cartesian coordinates. Let us see what happens to these equations under a transformation of coordinates. Let there be  $3n$  generalized coordinates  $q_1, \dots, q_{3n}$  (center of mass plus internal coordinates, spherical polar coordinates, elliptical coordinates, or any other suitable coordinates for the problem at hand), which are related to the cartesian coordinates of the individual particles by the  $3n$  transformation equations

$$\begin{aligned}
 x_1 &= x_1(q_1, \dots, q_{3n}) \\
 y_1 &= y_1(q_1, \dots, q_{3n}) \\
 &\dots \dots \dots \\
 x_i &= x_i(q_1, \dots, q_{3n}) \\
 &\dots \dots \dots \\
 z_n &= z_n(q_1, \dots, q_{3n})
 \end{aligned}
 \tag{2-15a}$$

The velocities are then given by

$$\begin{aligned}
 \dot{x}_1 &= \sum_{j=1}^{3n} \frac{\partial x_1}{\partial q_j} \dot{q}_j \\
 \dot{y}_i &= \sum_{j=1}^{3n} \frac{\partial y_i}{\partial q_j} \dot{q}_j \quad i = 1, \dots, n \\
 \dot{z}_n &= \sum_{j=1}^{3n} \frac{\partial z_n}{\partial q_j} \dot{q}_j
 \end{aligned}
 \tag{2-15b}$$

We now want to regard  $x_1, \dots, z_n; \dot{x}_1, \dots, \dot{z}_n$  as  $6n$  independent variables for expressing  $L$  and  $q_1, \dots, q_{3n}; \dot{q}_1, \dots, \dot{q}_{3n}$  as another set of  $6n$  variables for expressing  $L$ . The relations between the two sets of variables are given by the  $6n$  equations (2-15). According to Eq. (2-15a), the functional relationships between the position coordinates do not contain the velocities, so that

$$\frac{\partial x_i}{\partial \dot{q}_j} = \frac{\partial y_i}{\partial \dot{q}_j} = \frac{\partial z_i}{\partial \dot{q}_j} = 0 \quad i = 1, \dots, n; j = 1, \dots, 3n \tag{2-16}$$

Furthermore, according to (2-15b), the velocity  $\dot{y}_i$  is a linear function of

all the  $\dot{q}_j$ , and

$$\frac{\partial \dot{y}_i(q_1, \dots, q_{3n}; \dot{q}_1, \dots, \dot{q}_{3n})}{\partial \dot{q}_j} = \frac{\partial y_i}{\partial q_j} \quad (2-17)$$

We can now express  $\partial L / \partial q_k$  and  $\partial L / \partial \dot{q}_k$  in terms of the derivatives  $\partial L / \partial x_i$  and  $\partial L / \partial \dot{x}_i$ , using the transformation relations (2-15).

We have

$$\begin{aligned} \frac{\partial L}{\partial q_k} &= \sum_{i=1}^n \left( \frac{\partial L}{\partial x_i} \frac{\partial x_i}{\partial q_k} + \frac{\partial L}{\partial y_i} \frac{\partial y_i}{\partial q_k} + \frac{\partial L}{\partial z_i} \frac{\partial z_i}{\partial q_k} \right) \\ &\quad + \sum_{i=1}^n \left( \frac{\partial L}{\partial \dot{x}_i} \frac{\partial \dot{x}_i}{\partial q_k} + \frac{\partial L}{\partial \dot{y}_i} \frac{\partial \dot{y}_i}{\partial q_k} + \frac{\partial L}{\partial \dot{z}_i} \frac{\partial \dot{z}_i}{\partial q_k} \right) \quad k = 1, \dots, 3n \end{aligned} \quad (2-18)$$

$$\begin{aligned} \frac{\partial L}{\partial \dot{q}_k} &= \sum_{i=1}^n \left( \frac{\partial L}{\partial x_i} \frac{\partial x_i}{\partial \dot{q}_k} + \frac{\partial L}{\partial y_i} \frac{\partial y_i}{\partial \dot{q}_k} + \frac{\partial L}{\partial z_i} \frac{\partial z_i}{\partial \dot{q}_k} \right) \\ &\quad + \sum_{i=1}^n \left( \frac{\partial L}{\partial \dot{x}_i} \frac{\partial \dot{x}_i}{\partial \dot{q}_k} + \frac{\partial L}{\partial \dot{y}_i} \frac{\partial \dot{y}_i}{\partial \dot{q}_k} + \frac{\partial L}{\partial \dot{z}_i} \frac{\partial \dot{z}_i}{\partial \dot{q}_k} \right) \end{aligned} \quad (2-19)$$

The first sum of terms in (2-19) is zero because  $\partial x_i / \partial \dot{q}_k = 0$  [Eq. (2-16)]. In the second sum, we substitute  $\partial x_i / \partial q_k$  for  $\partial \dot{x}_i / \partial \dot{q}_k$ , which is justified by Eq. (2-17). Then take the total derivative of both sides of (2-19) with respect to time:

$$\begin{aligned} \frac{d}{dt} \left( \frac{\partial L}{\partial \dot{q}_k} \right) &= \sum_{i=1}^n \frac{\partial x_i}{\partial q_k} \frac{d}{dt} \left( \frac{\partial L}{\partial \dot{x}_i} \right) + \frac{\partial y_i}{\partial q_k} \frac{d}{dt} \left( \frac{\partial L}{\partial \dot{y}_i} \right) + \frac{\partial z_i}{\partial q_k} \frac{d}{dt} \left( \frac{\partial L}{\partial \dot{z}_i} \right) \\ &\quad + \sum_{i=1}^n \frac{\partial L}{\partial \dot{x}_i} \frac{d}{dt} \left( \frac{\partial x_i}{\partial q_k} \right) + \frac{\partial L}{\partial \dot{y}_i} \frac{d}{dt} \left( \frac{\partial y_i}{\partial q_k} \right) + \frac{\partial L}{\partial \dot{z}_i} \frac{d}{dt} \left( \frac{\partial z_i}{\partial q_k} \right) \end{aligned} \quad (2-20)$$

In the first sum of (2-20), we set

$$\frac{d}{dt} \left( \frac{\partial L}{\partial \dot{x}_i} \right) = \frac{\partial L}{\partial x_i}$$

by Lagrange's equation. For the second sum, we have

$$\frac{d}{dt} \left( \frac{\partial x_i}{\partial q_k} \right) = \frac{\partial \dot{x}_i}{\partial q_k}$$

since the order of differentiation is immaterial. With these substitutions, (2-20) becomes

$$\frac{d}{dt} \left( \frac{\partial L}{\partial \dot{q}_k} \right) = \sum_{i=1}^n \left( \frac{\partial L}{\partial x_i} \frac{\partial x_i}{\partial q_k} + \frac{\partial L}{\partial y_i} \frac{\partial y_i}{\partial q_k} + \frac{\partial L}{\partial z_i} \frac{\partial z_i}{\partial q_k} \right) + \sum_{i=1}^n \left( \frac{\partial L}{\partial \dot{x}_i} \frac{\partial \dot{x}_i}{\partial \dot{q}_k} + \frac{\partial L}{\partial \dot{y}_i} \frac{\partial \dot{y}_i}{\partial \dot{q}_k} + \frac{\partial L}{\partial \dot{z}_i} \frac{\partial \dot{z}_i}{\partial \dot{q}_k} \right) \quad k = 1, \dots, 3n \quad (2-21)$$

We recognize from (2-18) that the right-hand side of (2-21) is just  $\partial L / \partial q_k$ . Therefore

$$\frac{d}{dt} \left( \frac{\partial L}{\partial \dot{q}_k} \right) = \frac{\partial L}{\partial q_k} \quad k = 1, \dots, 3n \quad (2-22)$$

We have thus proved the important theorem that Lagrange's equations of motion are the same in all coordinate systems. If by one means or another we can express the Lagrangian  $L = K - U$  in terms of the coordinates  $q, \dot{q}$ , we can immediately write the equations of motion (2-22) in the  $q$  coordinate system.

**Problem 2-3.** For a particle moving in three dimensions under the influence of a spherically symmetrical potential  $U = U(r)$ , use the results of Prob. 2-2 to write out the Lagrangian and the equations of motion in spherical coordinates.

**Problem 2-4.** Given a system of two particles as in Prob. 2-1, with the only potential their potential of interaction, which is a function of the relative coordinates only,

$$U = U(x_2 - x_1, y_2 - y_1, z_2 - z_1) = U(x_{12}, y_{12}, z_{12})$$

Use Lagrange's equations to write the equations of motion for the system in terms of the coordinates  $X, Y, Z$  and  $x_{12}, y_{12}, z_{12}$ . Describe the significance of the results in words.

**2-4. Hamilton's Equations.** The Lagrangian function  $L$  is regarded as a function of the  $6n$  coordinates  $q_1, \dots, q_{3n}; \dot{q}_1, \dots, \dot{q}_{3n}$ . We shall eliminate the  $\dot{q}$ 's as independent variables by introducing  $3n$  new variables by the equations

$$p_k = \frac{\partial L(q_1, \dots, q_{3n}; \dot{q}_1, \dots, \dot{q}_{3n})}{\partial \dot{q}_k} \quad k = 1, \dots, 3n \quad (2-23)$$

By solving the  $3n$  equations (2-23), we can express each  $\dot{q}$  as a function of  $q_1, \dots, q_{3n}; p_1, \dots, p_{3n}$ :

$$\dot{q}_j = \dot{q}_j(q_1, \dots, q_{3n}; p_1, \dots, p_{3n}) \quad j = 1, \dots, 3n \quad (2-24)$$

The  $p$  variables are called the generalized momenta corresponding to the generalized coordinates  $q$ . The two variables  $q_i$  and  $p_i$  are spoken of as being conjugate variables.

We shall define Hamilton's function  $H$  as being the total energy  $K + U$ :

$$H(q_1, \dots, q_{3n}; p_1, \dots, p_{3n}) = K + U = 2K - L \quad (2-25)$$



In (2-25), we can think that  $K$  and  $L$  were initially given as functions of the  $q$ 's and  $\dot{q}$ 's; by using Eq. (2-24), they can be expressed as functions of the  $q$ 's and  $p$ 's.

It is essential to realize that the kinetic energy is always a quadratic function of the velocities  $\dot{q}_i$ , that is,

$$K(q_1, \dots, q_{3n}; \dot{q}_1, \dots, \dot{q}_{3n}) = \sum_{i=1}^{3n} \sum_{j=1}^{3n} \dot{q}_i a_{ij} \dot{q}_j \\ = a_{11}\dot{q}_1^2 + 2a_{12}\dot{q}_1\dot{q}_2 + \dots + 2a_{3n-1, 3n}\dot{q}_{3n-1}\dot{q}_{3n} + a_{3n, 3n}\dot{q}_{3n}^2 \quad (2-26)$$

The coefficients  $a_{ij}$  in (2-26) are in general functions of the  $q$ 's as in (2-9a), where the coefficient of  $\dot{\phi}^2$  is  $\frac{1}{2}mr^2$ . We can always choose the coefficients  $a_{ij} = a_{ji}$  in the double sum of (2-26), so that we write, for example,  $2a_{12}\dot{q}_1\dot{q}_2$  instead of  $a_{12}\dot{q}_1\dot{q}_2 + a_{21}\dot{q}_2\dot{q}_1$ . Then, since  $U$  is not a function of the velocities,

$$p_k = \frac{\partial L}{\partial \dot{q}_k} = \frac{\partial(K - U)}{\partial \dot{q}_k} = \frac{\partial K(q_1, \dots, q_{3n}; \dot{q}_1, \dots, \dot{q}_{3n})}{\partial \dot{q}_k} \\ = 2 \sum_{j=1}^{3n} a_{kj} \dot{q}_j \quad (2-27)$$

[Readers who are not experienced with general summation notation can verify (2-27) by differentiating the expression for  $K$  as displayed in detail in the right-hand expression of (2-26).]

It follows from (2-27) that

$$\sum_{k=1}^{3n} \dot{q}_k p_k = \sum_{k=1}^{3n} \sum_{j=1}^{3n} 2\dot{q}_k a_{kj} \dot{q}_j = 2K \quad (2-28)$$

We can therefore rewrite (2-25) as

$$H(q_1, \dots, q_{3n}; p_1, \dots, p_{3n}) \\ = \sum_{k=1}^{3n} p_k \dot{q}_k - L(q_1, \dots, q_{3n}; \dot{q}_1, \dots, \dot{q}_{3n}) \quad (2-29)$$

where it is understood that the  $\dot{q}$ 's can be expressed in terms of  $q$ 's and  $p$ 's by (2-24). [We may remark that Eq. (2-29) is commonly taken as the definition of the Hamiltonian function; it has the advantage of being applicable for the more general case of nonconservative systems. For our purposes, the definitions (2-29) and (2-25) are equivalent.]

For an arbitrary variation in the  $q$ 's and  $p$ 's, we have, from (2-29),

$$dH = \sum \dot{q}_k dp_k + \sum p_k d\dot{q}_k - \sum_{k=1}^{3n} \frac{\partial L(q_1, \dots, q_{3n}; \dot{q}_1, \dots, \dot{q}_{3n})}{\partial \dot{q}_k} d\dot{q}_k \\ - \sum_{k=1}^{3n} \frac{\partial L(q_1, \dots, q_{3n}; \dot{q}_1, \dots, \dot{q}_{3n})}{\partial q_k} dq_k \quad (2-30)$$