Rayleigh - Jean Formula

Electromagnetic wave

Standing Wave

One dimension

Wave equation

$$\frac{\partial^2 E}{\partial x^2} = \frac{1}{c^2}\frac{\partial^2 E}{\partial t^2}$$

Boundary condition

$E = 0$ at $x = 0$ and $x = a$

$$E(x,t) = E_0 \sin kx \sin \omega t$$

$$k^2 = \frac{1}{c^2}\omega^2 \qquad \nu = \frac{\omega}{2\pi}, \quad k = \frac{2\pi}{\lambda}, \quad \lambda\omega = c$$

Three dimension

Wave Equation :

$$\frac{\partial^2 E}{\partial x^2} + \frac{\partial^2 E}{\partial y^2} + \frac{\partial^2 E}{\partial z^2} = \frac{1}{c^2}\frac{\partial^2 E}{\partial t^2}$$

$$E(x,t) = E_0 \sin k_x x \sin k_y y \sin k_z z \sin \omega t$$

Boundary condition

$x = 0,\ x = a \quad E = 0$
$y = 0,\ y = a \Rightarrow E = 0$
$z = 0,\ z = a \quad E = 0$

$k_x a = n_x \pi$
$k_y a = n_y \pi$
$k_z a = n_z \pi$

n_x, n_y, n_z are integers

$$k_x^2 + k_y^2 + k_z^2 = \frac{1}{c^2}\omega^2$$

$$\Rightarrow \nu = \frac{c}{2a}\sqrt{n_x^2 + n_y^2 + n_z^2}$$

(n_x, n_y, n_z) positive integers $\longrightarrow$ one mode

$n_x \to -n_x$
$\sin k_x x = -\sin k_x x$
linearly depend on each other

$\Rightarrow$ Number of modes

$\downarrow$

physcially each mode is a standing wave of ν

分類：
編號：3 - 6
總號：

Number of modes with frequency between ν and $d\nu$

$= \frac{8\pi\nu^2}{c^3} d\nu \cdot V$ See Eisberg and Resnick P.8

Electromagatic wave (standing) in one dimension box ($x: 0 \to a$)

$$E(x,t) = E_0 \sin kx \sin \omega t$$

Boundary condition $E = 0$ at $x = 0$ and a

$ka = n\pi \leftrightarrow \frac{2\pi}{\lambda} = \frac{n\pi}{a}$ n = integers (positive)

$$\nu = \frac{c}{\lambda} = \frac{cn}{2a}$$

Generalize to three dimensional case

$$\nu = \frac{c}{2a}\sqrt{n_x^2 + n_y^2 + n_z^2}$$

n_x, n_y, n_z are positive integers

Define $r^2 = n_x^2 + n_y^2 + n_z^2$

$$\nu = \frac{c}{2a} r$$

$N(\nu)\, d\nu$ = number of modes with frequency in the range between ν and $d\nu$

= number of lattice points enclosed by the shell r and $r+dr$ in the (n_x, n_y, n_z) space (with n_x, n_y, n_z positive)

Number of lattice point in the $\vec{n}$ space = volume in $\vec{n}$ space $= 4\pi r^2 dr$

Construct cuts at n_x, n_y, n_z at half-integers

It is obvious that there will be one lattice point in each unit volume cell.

$$N(\nu)\, d\nu = \frac{1}{8} 4\pi r^2 dr$$

$$= \frac{1}{8} 4\pi \cdot \left(\frac{2a\nu}{c}\right)^2 \frac{2a}{c} d\nu$$

$$= \frac{4\pi\nu^2}{c^3} V d\nu$$

For electromagnetic wave, for each mode, there are two possible states of polarization (vector field $\vec{E} \perp \vec{k}$)

$$\Rightarrow \quad N(\nu)\, d\nu = \frac{8\pi\nu^2}{c^3} V d\nu$$

Equipartition of Energy

$$\bar{\epsilon} = \frac{1}{2} kT$$

Example: Free particle, Maxwell-Boltzmann distribution $\langle K.E. \rangle = \frac{1}{2} kT$

$\bar{\epsilon} = kT$ for simple harmonic oscillator

since $\langle$kinetic energy$\rangle$ $= \langle$potential energy$\rangle$

Normal mode of frequency ν $\longleftrightarrow$ harmonic oscillator of frequency ν

This lead to the Rayleigh-Jean formula.

Note: integration is involved.

↓

Now we need to discuss the Boltzmann factor

Degrees of freedom (physics and chemistry)

From Wikipedia, the free encyclopedia

A **degree of freedom** is an independent physical parameter, often called a dimension, in the formal description of the state of a physical system. The set of all dimensions of a system is known as a phase space.

Contents

Definition

In physics, a degree of freedom of a system is a formal description of a parameter that contributes to the state of a physical system.

It can also be defined as the minimum number of coordinates required to specify the position of a particle or system of particles.

In mechanics, a point particle can move independently in the three directions of space. Thus, the momentum of a particle consists of three components, each called a *degree of freedom*. A system of N independent particles, therefore, has the total of $3N$ degrees of freedom.

Similarly in statistical mechanics, a degree of freedom is a single scalar number describing the microstate of a system. The specification of all microstates of a system is a point in the system's phase space.

A degree of freedom may be any useful property that is not dependent on other variables. For example, in the 3D ideal chain model, two angles are necessary to describe each monomer's orientation.

Example: diatomic gas

In three-dimensional space, three degrees of freedom are associated with the movement of a

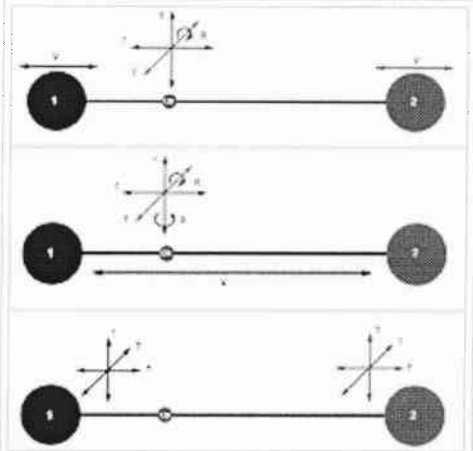
mechanical particle. A diatomic gas molecule thus has 6 degrees of freedom. This set may be decomposed in terms of translations, rotations, and vibrations of the molecule. The center of mass motion of the entire molecule accounts for 3 degrees of freedom. In addition, the molecule has one vibrational mode and two rotational degrees of motion. The rotations occur around the two axes perpendicular to the line between the two atoms. The rotation around the atom-atom bond is not counted. This yields, for a diatomic molecule, a decomposition of:

$$3N = 6 = 3 + 1 + 2.$$

For a general molecule with $N > 2$ atoms, all 3 rotational degrees of freedom are considered, resulting in the decomposition:

$$3\ N = 3 + 3 + (3\ N - 6)$$

which means that an N-atom molecule has 3N - 6 vibrational degrees of freedom for $N > 2$.



Different ways of visualizing the 3 degrees of freedom of a dumbbell-shaped diatomic molecule. (CM: center of mass of the system, T: translational motion, R: rotational motion, V: vibrational motion.)

As defined above one can also count degrees of freedom using the minimum number of coordinates required to specify a position. This is done as follows: 1. For a single particle we need 2 coordinates in a 2-D plane to specify its position and 3 coordinates in 3-D plane. Thus its degree of freedom in a 3-D plane is 3. 2. For a body consisting of 2 particles (ex. a diatomic molecule) in a 3-D plane with constant distance between them (let's say d) we can show (below) its degree of freedom to be 5. Let's say one particle in this body has coordinates (x_1,y_1,z_1) and the other has x-coordinate(x_2) and y-coordinate(y_2). Application of the formula for distance between two coordinates ($d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2}$) results in one equation with one unknown, in which we can solve for z_2. (Note:Here any one of x_1, x_2, y_1, y_2, z_1, or z_2 can be unknown.)

Contrary to the classical equipartition theorem, at room temperature, the vibrational motion of molecules typically makes negligible contributions to the heat capacity. This is because these degrees of freedom are *frozen* because the spacing between the energy eigenvalues exceeds the energy corresponding to ambient temperatures (kT). In the following table such degrees of freedom are disregarded because of their low effect on total energy. However, at very high temperatures they cannot be neglected.

	Monatomic	Linear molecules	Non-Linear molecules
Position (x, y and z)	3	3	3
Rotation (x, y and z)	0	2	3
Vibration	0	3N - 5	3N - 6
Total	**3**	**3N**	**3N**

Independent degrees of freedom

The set of degrees of freedom $X_1, \ldots, X_N$ of a system is independent if the energy associated with the set can be written in the following form:

$$E = \sum_{i=1}^{N} E_i(X_i),$$

where E_i is a function of the sole variable X_i.

example: if X_1 and X_2 are two degrees of freedom, and E is the associated energy:

- If $E = X_1^4 + X_2^4$, then the two degrees of freedom are independent.
- If $E = X_1^4 + X_1 X_2 + X_2^4$, then the two degrees of freedom are *not* independent. The term involving the product of X_1 and X_2 is a coupling term, that describes an interaction between the two degrees of freedom.

At thermodynamic equilibrium, $X_1, \ldots, X_n$ are all statistically independent of each other.

For i from 1 to N, the value of the ith degree of freedom X_i is distributed according to the Boltzmann distribution. Its probability density function is the following:

$$p_i(X_i) = \frac{e^{-\frac{E_i}{k_B T}}}{\int dX_i \, e^{-\frac{E_i}{k_B T}}},$$

In this section, and throughout the article the brackets $\langle\rangle$ denote the mean of the quantity they enclose.

The internal energy of the system is the sum of the average energies associated to each of the degrees of freedom:

$$\langle E \rangle = \sum_{i=1}^{N} \langle E_i \rangle.$$

Demonstrations

A system exchanges energy in the form of heat with its surroundings and the number of

particles in the system remains fixed. This corresponds to studying the system in the canonical ensemble. Note that in statistical mechanics, a result that is demonstrated for a system in a particular ensemble remains true for this system at the thermodynamic limit in any ensemble. In the canonical ensemble, at thermodynamic equilibrium, the state of the system is distributed among all micro-states according to the Boltzmann distribution. If T is the system's temperature and k_B is Boltzmann's constant, then the probability density function associated to each micro-state is the following:

$$P(X_1, \ldots, X_N) = \frac{e^{-\frac{E}{k_B T}}}{\int dX_1\, dX_2 \ldots dX_N e^{-\frac{E}{k_B T}}},$$

The denominator in the above expression plays an important role.[1] This expression immediately breaks down into a product of terms depending of a single degree of freedom:

$$P(X_1, \ldots, X_N) = p_1(X_1) \ldots p_N(X_N)$$

The existence of such a breakdown of the multidimensional probability density function into a product of functions of one variable is enough by itself to demonstrate that $X_1 \ldots X_N$ are statistically independent from each other.

Since each function p_i is normalized, it follows immediately that p_i is the probability density function of the degree of freedom X_i, for i from 1 to N.

Finally, the internal energy of the system is its mean energy. The energy of a degree of freedom E_i is a function of the sole variable X_i. Since $X_1, \ldots, X_N$ are independent from each other, the energies $E_1(X_1), \ldots, E_N(X_N)$ are also statistically independent from each other. The total internal energy of the system can thus be written as:

$$U = \langle E \rangle = \langle \sum_{i=1}^{N} E_i \rangle = \sum_{i=1}^{N} \langle E_i \rangle$$

Quadratic degrees of freedom

A degree of freedom X_i is quadratic if the energy terms associated to this degree of freedom can be written as

$$E = \alpha_i\, X_i^2 + \beta_i\, X_i Y,$$

where Y is a linear combination of other quadratic degrees of freedom.

example: if X_1 and X_2 are two degrees of freedom, and E is the associated energy:

- If $E = X_1^4 + X_1^3 X_2 + X_2^4$, then the two degrees of freedom are not independent and non-quadratic.
- If $E = X_1^4 + X_2^4$, then the two degrees of freedom are independent and non-

quadratic.

- If $E = X_1^2 + X_1 X_2 + 2X_2^2$, then the two degrees of freedom are not independent but are quadratic.
- If $E = X_1^2 + 2X_2^2$, then the two degrees of freedom are independent and quadratic.

For example, in Newtonian mechanics, the dynamics of a system of quadratic degrees of freedom are controlled by a set of homogeneous linear differential equations with constant coefficients.

Quadratic and independent degree of freedom

$X_1, \ldots, X_N$ are quadratic and independent degrees of freedom if the energy associated to a microstate of the system they represent can be written as:

$$E = \sum_{i=1}^{N} \alpha_i X_i^2$$

Equipartition theorem

In the classical limit of statistical mechanics, at thermodynamic equilibrium, the internal energy of a system of *N* quadratic and independent degrees of freedom is:

$$U = \langle E \rangle = N \frac{k_B T}{2}$$

Here, the mean energy associated with a degree of freedom is:

$$\langle E_i \rangle = \int dX_i \; \alpha_i X_i^2 \; p_i(X_i) = \frac{\int dX_i \; \alpha_i X_i^2 \; e^{-\frac{\alpha_i X_i^2}{k_B T}}}{\int dX_i \; e^{-\frac{\alpha_i X_i^2}{k_B T}}}$$

$$\langle E_i \rangle = \frac{k_B T}{2} \frac{\int dx \; x^2 \; e^{-\frac{x^2}{2}}}{\int dx \; e^{-\frac{x^2}{2}}} = \frac{k_B T}{2}$$

Since the degrees of freedom are independent, the internal energy of the system is equal to the sum of the mean energy associated with each degree of freedom, which demonstrates the result.

References

1. ^ "Configuration integral (statistical mechanics)"

(http://clesm.mae.ufl.edu/wiki.pub/index.php/Configuration_integral_%28statistical_mechanics%29#Thermodynamic_properties) .
http://clesm.mae.ufl.edu/wiki.pub/index.php/Configuration_integral_%28statistical_mechanics%29#Thermodynamic_properties.

Retrieved from "http://en.wikipedia.org/w/index.php?title=Degrees_of_freedom_(physics_and_chemistry)&oldid=504126422"
Categories: Concepts in physics | Dimension

分類：
編號：3. - 7
總號：

The law of equipartition of energy: classical kinetic theory

A system of gas molecules in thermal equilibrium at temperature T, the average kinetic energy of a molecule per degree of freedom is $\frac{1}{2}kT$

↳ Boltzmann constant

↓

Also applicable to any classical system containing, in equilibrium, a large number of entitle of the same kind.

↓

average kinetic energy of the sinusoildally standing wave is $\frac{1}{2}kT$

The average total energy of the sinusoilally standing wave is twice its average kinetic energy

↓

common property of physical systems, with one degree of freedom, that execute simple harmonic oscillation in time

$$\bar{E} = kT$$

Note: it is independent of ν.

Classical statistical physics is based on the following fundamental postulates:

(i) The particles of the system are identical but distinguishable

(ii) There is no restriction on the number of particles that may occupy a particular energy state

(iii) At thermal equilibrium, the distribution of particles among the accessible energy states is the most probable distribution consistent with prescribed constraints such as total energy and total number of particles

(iv) Every microstate of the system has equal a priori probability.

Boltzmann Factor

Want to find $\{N_i\}$ that makes $W_{\{N_i\}}$ maximum under the restriction

$$\sum_{i=1}^{k} N_i = N \quad , \quad \sum_{i=1}^{k} N_i \epsilon_i = E$$

$$W_{\{N_i\}} = \frac{N!}{N_1! N_2! \cdots N_k!}$$

It is useful to take

$$\ln W = \ln N! - \sum_{i=1}^{k} \ln N_i!$$

Stirling approximation.

$$\ln N! \simeq N \ln N - N$$

$$\ln W = N \ln N - N - \left(\sum_{i=1}^{k} N_i \ln N_i - \sum_{i=1}^{k} N_i \right)$$

$$= N \ln N - \sum_{i=1}^{k} N_i \ln N_i \qquad \left(\sum_{i=1}^{k} N_i = N \right)$$

$$\delta \ln W = - \sum_{i=1}^{k} \left(\ln N_i \, \delta N_i + N_i \frac{\delta N_i}{N_i} \right) = 0$$

Constraint $\sum_{i=1}^{k} N_i = N \Rightarrow \sum_{i=1}^{k} \delta N_i = 0$

$$\sum_{i=1}^{k} N_i \epsilon_i = E \Rightarrow \sum_{i=1}^{k} \epsilon_i \delta N_i = 0$$

Introduce the Lagrange multiplier λ, μ

$$\sum_{i=1}^{k} (\ln N_i + \lambda + \mu \epsilon_i) \delta N_i = 0$$

With the presence of Lagrange multiplier λ, μ

δN_i can be treated as independent

$$\Rightarrow \quad \ln N_i + \lambda + \mu \epsilon_i = 0$$

$$N_i = e^{-\lambda - \mu \epsilon_i}$$

$$\downarrow$$

$$f_{MB}(\epsilon_i) = e^{-\lambda - \mu \epsilon_i} = A e^{-\mu \epsilon_i} \qquad (\beta = \mu)$$

$$\downarrow$$

Maxwell - Boltzmann distribution

BINOMIAL AND MULTINOMIAL DISTRIBUTION

During the course of our discussion of the canonical ensemble, we shall encounter the problem of determining how many ways it is possible to divide N distinguishable systems into groups such that there are n_1 systems in the first group, n_2 systems in the second group, and so on, and such that $n_1 + n_2 + \cdots = N$, that is, all the systems are accounted for. This is actually one of the easiest problems in combinatorial analysis. To solve this, we first calculate the number of permutations of N distinguishable objects, that is, the number of possible different arrangements or ways to order N distinguishable objects. Let us choose one of the N objects and place it in the first position, one of the $N-1$ remaining objects and place it in the second position, and so on, until all N objects are ordered. Clearly there are N choices for the first position, $N-1$ choices for the second position, and so on, until finally there is only one object left for the Nth position. The total number of ways of doing this is then the product of all the choices,

$$N(N-1)(N-2)\cdots(2)(1) \equiv N! \qquad \text{(distinguishable objects)}$$

Next we calculate the number of ways of dividing N distinguishable objects into two groups, one group containing N_1 objects, say, and the other containing the remaining $N - N_1$. There are $N(N-1)\cdots(N-N_1+1)$ ways to form the first group, and $N_2! = (N-N_1)!$ ways to form the second group. The total number is, then, the product

$$N(N-1)\cdots(N-N_1+1)\times(N-N_1)! = \frac{N!}{(N-N_1)!}\times(N-N_1)! = N!$$

But this has overcounted the situation drastically, since the order in which we place N_1 members in the first group and N_2 in the second group is immaterial to the problem as stated. All $N_1!$ orders of the first group and $N_2!$ orders of the second group correspond to just one division of N objects into N_1 objects and N_2 objects. Therefore the desired result is

$$\frac{N!}{N_1!(N-N_1)!} = \frac{N!}{N_1!N_2!} \tag{1–75}$$

Since the combination of factorials in Eq. (1–75) occurs in the binomial expansion,

$$(x+y)^N = \sum_{N_1=0}^{N} \frac{N!x^{N-N_1}y^{N_1}}{N_1!(N-N_1)!} = \sum_{N_1N_2}^{*} \frac{N!x^{N_1}y^{N_2}}{N_1!N_2!} \tag{1–76}$$

$N!/N_1!(N-N_1)!$ is called a binomial coefficient. The asterisk on the second summation in Eq. (1–76) signifies the restriction $N_1 + N_2 = N$.

The generalization of Eq. (1–75) to the division of N into r groups, the first containing N_1, and so on, is easily seen to be

$$\frac{N!}{N_1!N_2!\cdots N_r!} = \frac{N!}{\prod_{j=1}^{r} N_j!} \tag{1–77}$$

where $N_1 + N_2 + \cdots + N_r = N$. This is known as a multinomial coefficient, since it occurs in the expansion

$$(x_1 + x_2 + \cdots + x_r)^N = \sum_{N_1=0}^{N}\sum_{N_2=0}^{N}\cdots\sum_{N_r=0}^{N*} \frac{N!x_1^{N_1}\cdots x_r^{N_r}}{\prod_{j=1}^{r} N_j!} \tag{1–78}$$

where this time the asterisk signifies the restriction $N_1 + N_2 + \cdots + N_r = N$.

There are a number of other combinatorial formulas that are useful in statistical thermodynamics, but Eq. (1–77) is the most useful for our purposes. Combinatorial formulas can become rather demanding to derive. We refer to Appendix AVII of Mayer and Mayer* which contains a collection of formulas.

STIRLING'S APPROXIMATION

In statistical thermodynamics we often encounter factorials of very large numbers, such as Avogadro's number. The calculation and mathematical manipulation of factorials become awkward for large N. Therefore it is desirable to find an approximation for $N!$ for large N. Problems of this sort occur often in mathematics and are called asymptotic approximations, that is, an approximation to a function which improves as the argument of that function increases. Since $N!$ is actually a product, it is convenient to deal with $\ln N!$ because this is a sum. The asymptotic approximation to $\ln N!$ is called Stirling's approximation, which we now derive.

Since $N! = N(N-1)(N-2)\cdots(2)(1)$, $\ln N!$ is

$$\ln N! = \sum_{m=1}^{N} \ln m \tag{1–73}$$

Figure 1–5 shows $\ln x$ plotted versus x. The sum of the areas under these rectangles up to N is $\ln N!$. Figure 1–5 also shows the continuous curve $\ln x$ plotted on the same graph. Thus $\ln x$ is seen to form an envelope to the rectangles, and this envelope becomes a steadily smoother approximation to the rectangles as x increases. We can approximate the area under these rectangles by the integral of $\ln x$. The area under $\ln x$ will poorly approximate the rectangles only in the beginning. If N is large enough (we are deriving an asymptotic expansion), this area will make a negligible contribution to the total area. We may write, then,

$$\ln N! = \sum_{m=1}^{N} \ln m \approx \int_{1}^{N} \ln x\, dx = N \ln N - N \qquad (N \text{ large}) \tag{1–74}$$

which is Stirling's approximation to $\ln N!$. The lower limit could just as well have been taken as 0 in Eq. (1–74), since N is large. (Remember that $x \ln x \to 0$ as $x \to 0$.)

A more refined derivation of Stirling's approximation gives $\ln N! \approx N \ln N - N + \ln(2\pi N)^{1/2}$, but this additional term is seldom necessary. (See Problem 1–59.)

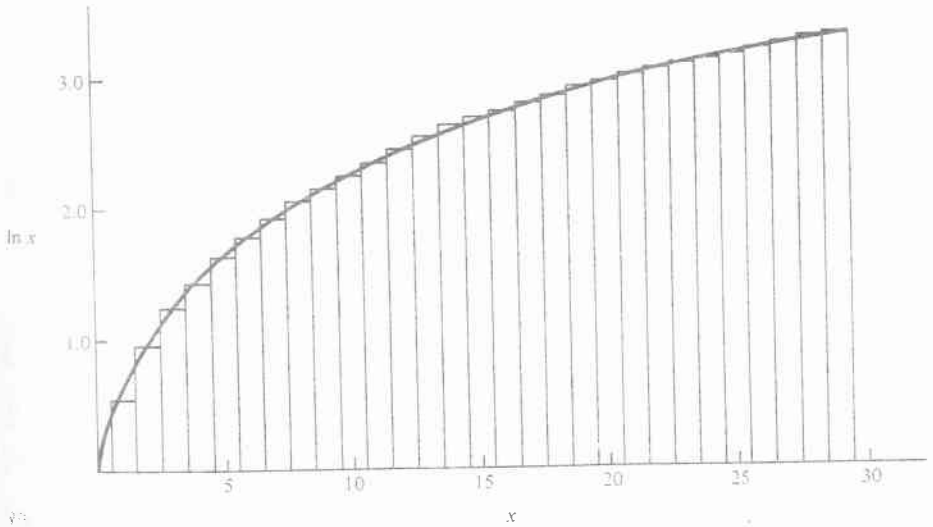


Figure 1–5. **A plot of $\ln x$ versus x, showing how the summation of $\ln m$ can be approximated by the integral of $\ln x$.**

METHOD OF LAGRANGE MULTIPLIERS

It will be necessary, later, to maximize Eq. (1–77) with the constraint $N_1 + N_2 + \cdots + N_r = \text{constant}$. This brings us to the mathematical problem of maximizing a function of several (or many) variables $f(x_1, x_2, \ldots, x_r)$ when the variables are connected by other equations, say $g_1(x_1, \ldots, x_r) = 0$, $g_2(x_1, \ldots, x_r) = 0$, and so on. This type of problem is readily handled by the method of Lagrange undetermined multipliers.

If it were not for the constraints, $g_j(x_1, x_2, \ldots, x_r) = 0$, the maximum of $f(x_1, \ldots, x_r)$ would be given by

$$\delta f = \sum_{j=1}^{r} \left(\frac{\partial f}{\partial x_j}\right)_0 \delta x_j = 0 \tag{1–79}$$

where the zero subscript indicates that this equation equals zero only when the r partial derivatives are evaluated at the maximum (or minimum) of f. Denote these values of x_j by x_j^0. If there were no constraints, each of the δx_j would be able to be varied independently and arbitrarily, and so we would conclude that $(\partial f/\partial x_j) = 0$ for every j, since δf must equal zero. This would give r equations from which the values of the r x_j^0 could be obtained.

On the other hand, if there is some other relation between the x's, such as $g(x_1, x_2, \ldots, x_r) = 0$, we have the additional equation

$$\delta g = \sum_{j=1}^{r} \left(\frac{\partial g}{\partial x_j}\right)_0 \delta x_j = 0 \tag{1–80}$$

This equation serves as a constraint that the δx_j must satisfy, thus making one of them depend upon the other $r - 1$. In the Lagrange method, one multiplies Eq. (1–80) by some parameter, say λ, and adds the result to Eq. (1–79) to get

$$\sum_{j=1}^{r} \left(\frac{\partial f}{\partial x_j} - \lambda \frac{\partial g}{\partial x_j}\right)_0 \delta x_j = 0 \tag{1–81}$$

The δx_j are still not independent, because of Eq. (1–80), and so they cannot be varied independently. Equation (1–80), however, can be treated as an equation giving one of the δx_j in terms of the other $r - 1$ independent ones. Pick any one of the r δx_j as the dependent one. Let this be δx_μ.

The trick now is that we have not specified λ yet. We set it equal to $(\partial f/\partial x_\mu)_0/(\partial g/\partial x_\mu)_0$, making the coefficient of δx_μ in Eq. (1–81) vanish. The subscript zero here indicates that $(\partial f/\partial x_\mu)$ and $(\partial g/\partial x_\mu)$ are to be evaluated at values of the x_j such that f is at its maximum (or minimum) under the constraint of Eq. (1–80). Of course, we do not know these values of x_j yet, but we can nevertheless formally define λ in this manner. This leaves a sum of terms in Eq. (1–81) involving only the independent δx_j, which can be varied independently, yielding that

$$\left(\frac{\partial f}{\partial x_j}\right)_0 - \lambda\left(\frac{\partial g}{\partial x_j}\right)_0 = 0 \qquad j = 1, 2, \ldots, \mu - 1, \mu + 1, \ldots, r$$

* See Mayer and Mayer, *Statistical Mechanics* (New York: Wiley, 1940).

If we combine these $r - 1$ equations with our choice for λ, we have

$$\left(\frac{\partial f}{\partial x_j}\right)_0 - \lambda\left(\frac{\partial g}{\partial x_j}\right)_0 = 0 \tag{1–82}$$

for all j.

As we said above, the choice of λ here is certainly formal, since both $(\partial f/\partial x_\mu)_0$ and $(\partial g/\partial x_\mu)_0$ must be evaluated at these values of x_j which maximizes f, but these are known from Eq. (1–82) only in terms of λ. But this presents no difficulty, since in practice λ is determined by physical requirements. Examples of this will occur in the next two chapters.

Lagrange's method becomes no more difficult in the case in which there are several constraints. Let $g_1(x_1, \ldots, x_r)$, $g_2(x_1, \ldots, x_r)$, ... be a set of constraints. We introduce a Lagrange multiplier for each $g_i(x_1, \ldots, x_r)$ and proceed as above to get

$$\frac{\partial f}{\partial x_j} - \lambda_1 \frac{\partial g_1}{\partial x_j} - \lambda_2 \frac{\partial g_2}{\partial x_j} - \cdots = 0 \tag{1–83}$$

$N = 6$

Each particles can occupy states of energy $E = n\epsilon$

Total energy of the system is 9ϵ

Macrostates and Microstates of Six Identical Particles with a Total Energy of 9ϵ

Macrostate	Number of Particles with Energy $n\epsilon$										Number of Microstates
	0ϵ	1ϵ	2ϵ	3ϵ	4ϵ	5ϵ	6ϵ	7ϵ	8ϵ	9ϵ	
A	5	0	0	0	0	0	0	0	0	1	6
B	4	1	0	0	0	0	0	0	1	0	30
C	4	0	1	0	0	0	0	1	0	0	30
D	3	2	0	0	0	0	0	1	0	0	60
E	4	0	0	1	0	0	1	0	0	0	30
F	3	1	1	0	0	0	1	0	0	0	120
G	2	3	0	0	0	0	1	0	0	0	60
H	4	0	0	0	1	1	0	0	0	0	30
I	3	1	0	1	0	1	0	0	0	0	120
J	3	0	2	0	0	1	0	0	0	0	60
K (F)	2	2	1	0	0	1	0	0	0	0	180
L	1	4	0	0	0	1	0	0	0	0	30
M	3	1	0	0	2	0	0	0	0	0	60
N	3	0	1	1	1	0	0	0	0	0	120
O (F)	2	1	2	0	1	0	0	0	0	0	180
P (F)	2	2	0	1	1	0	0	0	0	0	180
Q	1	3	1	0	1	0	0	0	0	0	120
R	0	5	0	0	1	0	0	0	0	0	6
S	3	0	0	3	0	0	0	0	0	0	20
T (F)	2	1	1	2	0	0	0	0	0	0	180
U	1	3	0	2	0	0	0	0	0	0	60
V	2	0	3	1	0	0	0	0	0	0	60
W (F)	1	2	2	1	0	0	0	0	0	0	180
X	0	4	1	1	0	0	0	0	0	0	30
Y	1	1	4	0	0	0	0	0	0	0	30
Z	0	3	3	0	0	0	0	0	0	0	20
Total Number of Microstates											2002

Macrostate B contoins 30 microstates

6 different particles that could be in the energy 8ϵ
Having chosen the particles in that energy state, there are then 5 particles that can have the energy ϵ
$\Rightarrow$ $6 \times 5 = 30$ different and distinct microstates

In general, the number of microstates contained within a microstate is given by

$$\mathcal{N} = \frac{N!}{n_1!\, n_2!\, n_3! \cdots}$$

N = total number of particles

n_j number of particles in the jth level

$n_i = \sum_j n_{ij} P_j$ → average number of particles in of energy E_i

n_{ij} = number of particles of energy E_i in the mocrostate j

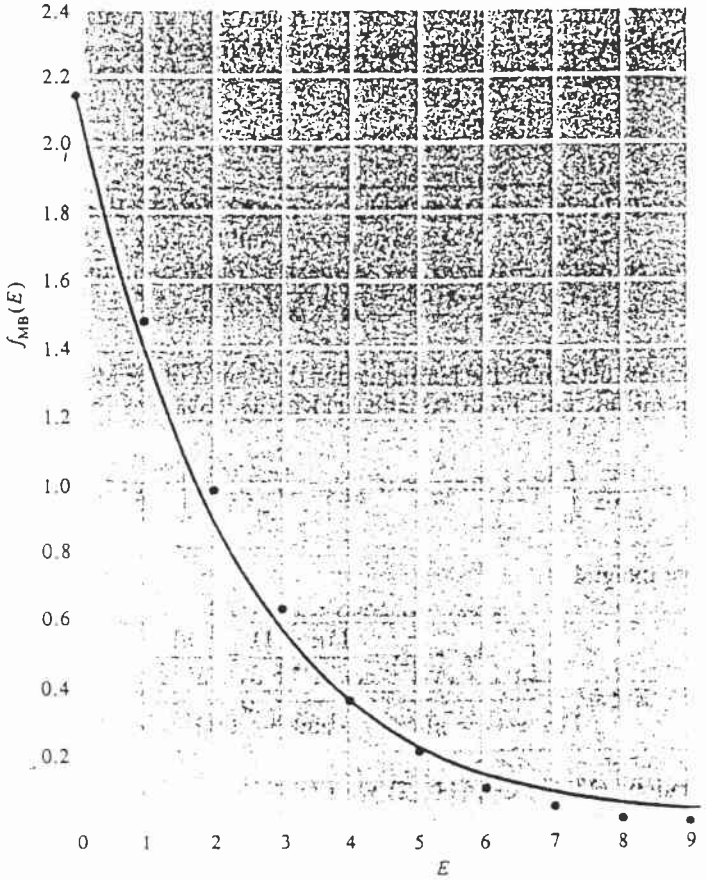
P_j is the probability of that macrostate.

[$n_0 = (5\times6 + 4\times30 + 3\times60 + 4\times30 + 3\times120 + 2\times60$
$+ 4\times30 + 3\times120 + 3\times60 + 2\times180 + 1\times30 + 3\times60 + 3\times120$
$+ 2\times180 + 2\times180 + 1\times120 + 0\times6 + 3\times20 + 2\times180 + 1\times60$
$+ 2\times60 + 1\times180 + 0\times30 + 1\times30 + 0\times20)/2002$
$= 2.143$]

Similarly,

$n_1 = 1.484$, $n_2 = 0.989$, $n_3 = 0.629$, $n_4 = 0.378$

$n_5 =$



: The distribution function for a classical gas of six noninteracting identical particles of total energy 9ϵ. Also shown is the exponential function $f_{MB}(E) = Ae^{-E/kT}$.

When N increases, the most probable distribution approaches the distribution of the average.

Now we shall use the calculus of variation technique to derive the Maxwell-Boltzmann distribution.

$$P(\varepsilon) = \frac{e^{-\varepsilon/kT}}{kT} \qquad (3)$$

This is a proper density distribution (the integral from zero to infinity equals 1), and the mean value of ε with this distribution over the range from zero to infinity is kT, regardless of the frequency of the mode, so this implies the equipartition of energy (which we know leads to non-sensical results). This distribution can be derived classically in several different ways, but in general it corresponds to the proposition that states with higher energy are less probable (or less populated), and the weight factors of any two states are inversely proportional to the exponentials of their energy levels. In turn this is related to the idea that the weight factor for a given region on the constant energy surface in phase space equals the volume swept out by that region as the total energy changes incrementally from ε to $\varepsilon+d\varepsilon$. For a certain mode with high energy, the change in the corresponding coordinate and momenta necessary to increment the energy by de is less than for modes with low energy – just as the increase in speed needed to give a certain increase in kinetic energy is less for a particle that is already moving rapidly than for one that is moving slowly (because the kinetic energy is quadratic in speed). Therefore, the volume of phase space swept out near a high-energy mode on the energy surface is less than near a low-energy mode, so the weight factors are correspondingly less.

Classically it was assumed that every energy mode is capable of possessing any amount of energy, from zero to infinity, so the phase space was continuous and had no natural scale, which presents some subtle problems when trying to decide how to count states. However, following Planck, we could hypothesize that energy modes are actually capable only of possessing integer multiples of a certain fundamental irreducible quantum of energy, and we could suppose that this quantity depends on the frequency of the energy mode. The simplest supposition is that the energy ε of a mode with frequency ν can only take on one of the discrete values

$$\varepsilon = nh\nu \quad , \quad n = 0,1,2,... \qquad (4)$$

where h is a fundamental constant of nature (now called Planck's constant). We still assert that the weight factors to be assigned to the energy levels are related according to the exponential formula (3), but we simply restrict the values of ε to the appropriate set of discrete values depending on the frequency.

The task now is to determine the mean value of energy for a mode with frequency n. Recall that if energy is treated continuously with the distribution (3) we get a mean value of kT for the energy, regardless of frequency. However, using the same exponential relation for the weight factors, but restricting the energy levels to the discrete values given by (4), the mean value of energy is

$$\overline{\varepsilon_\nu} = \frac{\sum_{n=0}^{\infty} nh\nu \frac{e^{-nh\nu/kT}}{kT}}{\sum_{n=0}^{\infty} \frac{e^{-nh\nu/kT}}{kT}}$$

If ν is small, this approaches the continuous case, so the mean energy approaches kT, but for larger ν the weight factor for the n = 0 term (which is constant) begins to predominate over the

weight factors for n > 0. The denominator can never be smaller than 1/kT, whereas the numerator goes to zero. Making use of the geometric series identities

$$\sum_{n=0}^{\infty} x^n = \frac{1}{1-x} \qquad \sum_{n=0}^{\infty} nx^n = \frac{x}{(1-x)^2}$$

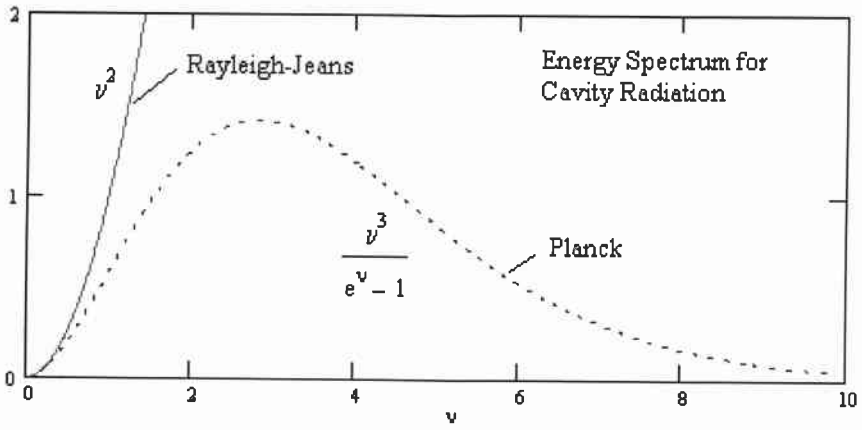
we can evaluate the summations to give the mean energy level for the frequency ν

$$\overline{\varepsilon_\nu} = \frac{h\nu}{e^{h\nu/kT} - 1} \tag{5}$$

Multiplying this by the spectral density (1) gives Planck's formula for the energy density per unit volume of cavity radiation as a function of frequency

$$\rho(\nu) = 8\pi h \frac{\nu^3}{c^3} \frac{1}{e^{h\nu/kT} - 1} \tag{6}$$

This replaces the classical Rayleigh-Jeans law given by equation (2). A normalized plot of this function is shown below.



The usual derivation of Planck's law, as described above, simply assumes that the Boltzmann distribution (3) gives the correct weight factors for discrete energy levels, even though that distribution was derived classically from the dynamics of a continuous distribution. There are actually several different classical derivations of Boltzmann's distribution for various circumstances, but they all involve continuously distributed energy levels, and some of them explicitly rely on the continuity. For a discrete set of energy levels one might have expected, a priori, something like a

Difficulty with Rayleigh-Jean formula

$U(T) = \int u(\nu, T)\, d\nu \to \infty \Rightarrow$ ultraviolet catastrophe

total radiation energy per unit volume

1900 Dec 14 Planck's paper → beginning of quantum theory

2. Planck's quantum theory (1900)

Thermodynamic consideration and interpolation

See T. Y. Wu, P. 32

$\Rightarrow$ Planck's formula

$$u(\nu, T) = \frac{8\pi\nu^2}{c^3} \frac{h\nu}{e^{h\nu/kT} - 1}$$

h is an adjustable parameter with dimension of energy · time

Large ν limit

$$u(\nu, T) \longrightarrow \frac{8\pi h}{c^3} \nu^3 e^{-h\nu/kT}$$ Wien's formula

Small ν limit

$$u(\nu, T) \longrightarrow \frac{8\pi\nu^2}{c^3} \frac{h\nu}{1 + \frac{h\nu}{kT} + \cdots - 1}$$

$$\to \frac{8\pi\nu^2}{c^3} kT$$ Rayleigh-Jean formula

$$U(T) = \frac{8\pi h}{c^3} \int_0^\infty \frac{\nu^3}{e^{h\nu/kT} - 1} d\nu$$

$$= \frac{8\pi h}{c^3} \left(\frac{kT}{h}\right)^4 \int_0^\infty \frac{x^3 dx}{e^x - 1} \qquad x = \frac{h\nu}{kT}, \quad \nu = \frac{kT}{h} x, \quad d\nu = \frac{kT}{h} dx$$

$$= aT^4$$ Stefan-Boltzmann law

Claim. $\int_0^\infty \frac{x^3 dx}{e^x - 1} = \frac{\pi^4}{15}$

Proof.

$$\int_0^\infty \frac{x^3 dx}{e^x - 1}$$

$$= \int_0^\infty x^3 dx \sum_{n=1}^\infty e^{-nx}$$

$$= \sum_{n=1}^\infty \int_0^\infty x^3 dx\, e^{-nx} \qquad y = nx$$

$$= \sum_{n=1}^\infty \frac{1}{n^4} \int_0^\infty dy\, y^3 e^{-y}$$

$$= 6 \sum_{n=1}^\infty \frac{1}{n^4} = \frac{\pi^4}{15}$$

Derivation of Planck's formula

Postulate: each of these standing waves in the box cannot take on all possible energies, as classical physics implies, but can only take on only discrete energies $E_n = nh\nu$ ($n = 1, 2, \cdots$)

$$p(E_n) = \frac{e^{-E_n/kT}}{\sum e^{-E_n/kT}}$$

derivation of the equation will be given in the appendix

$$\bar{E} = \sum E_n P(E_n)$$

$$= \frac{\sum nh\nu\, e^{-nh\nu/kT}}{\sum e^{-nh\nu/kT}}$$

Let $x = \frac{h\nu}{kT}$

$$\bar{E} = \frac{xkT \sum n e^{-nx}}{\sum e^{-nx}} = \frac{-xkT \frac{d}{dx} \sum e^{-nx}}{\sum e^{-nx}}$$

Claim $\sum e^{-nx} = \frac{1}{1-e^{-x}}$

Proof $(1-y)^{-1} = 1 + y + y^2 + \cdots$

Take $y = e^{-x}$

$$\Rightarrow \bar{E} = \frac{-xkT \frac{d}{dx}\left(\frac{1}{1-e^{-x}}\right)}{\left(\frac{1}{1-e^{-x}}\right)}$$

$$= \frac{xkT\, e^{-x}}{1-e^{-x}}$$

$$= \frac{h\nu}{e^{h\nu/kT} - 1}$$

$$\Rightarrow u(\nu, T)\, d\nu = \frac{8\pi\nu^2}{c^3} d\nu \frac{h\nu}{e^{h\nu/kT} - 1}$$ Planck's formula.

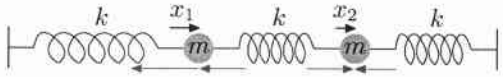
From experiments $\Rightarrow$ h = Planck's constant $= 6.63 \cdot 10^{-27}$ erg·sec

Planck: for some unknown reasons the atoms in the walls of the cavity emitted radiation in "quanta" with energy $nh\nu$ ($n = 1, 2, 3, \cdots$)

$E = nh\nu$

A Simple Problem simply done

Consider two equal masses m connected by springs of equal strength k connected as shown.



The Newton's equations are

$$m\ddot{x}_1 = -kx_1 - k(x_1 - x_2)\,,$$
$$m\ddot{x}_2 = k(x_1 - x_2) - kx_2\,.$$

Or

$$m\ddot{x}_1 = -2kx_1 + kx_2\,,$$
$$m\ddot{x}_2 = kx_1 - 2kx_2\,.$$

These are two coupled differential equations and call for some ingenuity to make them separable. Adding and subtracting, we get

$$m(\ddot{x}_1 + \ddot{x}_2) = -k(x_1 + x_2)\,,$$
$$m(\ddot{x}_1 - \ddot{x}_2) = -3k(x_1 - x_2)\,.$$

So by using a good combination of the co-ordinates, the equations of motion separate into two uncoupled oscillations with their specific frequencies.

$$x_1 + x_2 = A\cos(\omega_- t + \delta_1)\,,$$
$$x_1 - x_2 = B\cos(\omega_+ t + \delta_2)\,,$$

where $\omega_-^2 = k/m$ and $\omega_+^2 = 3k/m$.

Chapter 6

HEAT CAPACITIES

The importance of heat capacities in thermodynamic discussions can scarcely be over-estimated. At constant volume, the heat capacity measures quantitatively the ability of the system to take energy into its internal degrees of freedom. These are, in turn, intimately related to the atomic and molecular characteristics of the particular system. The heat capacity can thus provide an important link between the observed, macroscopic behaviour of a system and its detailed atomic or molecular structure.

Before we consider specific examples, let us look at the general way in which the energy-level ladder of an aspect of the system influences the ability of that aspect to take in energy, i.e. its heat capacity. When the levels are closely spaced and the energy gap between them is small, application of the distribution law in the form

$$n_i = n_0 p_i \exp(-\varepsilon_i / kT)$$

shows that it will be easy for particles to leave the ground state. In order to raise the temperature it is necessary to promote many particles to upper energy levels. There will be a large intake of energy and a heat capacity near the classical value which was deduced without considering quantization of energy. Conversely, when the energy levels are widely spaced and the energy gaps large, a rise in temperature promotes rather few particles and the heat capacity is low. Now the description 'large' or 'small' for the spacings of the energy levels are of necessity comparative; any quantity or object is

only large or small compared with something else. In this discussion, the comparison is with thermal energy, kT. We can illustrate this with two important physical situations:

$\varepsilon \ll kT$ (when the heat capacity has its classical value)

and

$\varepsilon \gg kT$ (when the heat capacity tends to zero)

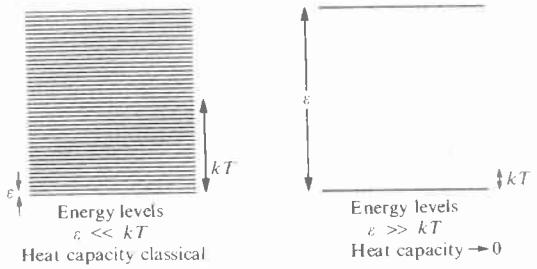


Fig. 6.1. Effect of energy-level spacing on heat capacities.

as shown in Fig. 6.1. Put another way, when the energy separation is much less than kT it is as though the effect of quantization were not noticed by the thermal energy. The classical result, which is based on a continuous distribution of energy levels, is then obtained. But once the energy becomes of the order of or greater than kT, quantization effects become important. We shall now consider two important systems, the study of which can help greatly in understanding the effect of quantization of energy on the thermodynamic properties of substance.

Heat Capacities of Gases

When a simple gas is heated at constant volume the energy is taken up by the modes of the molecule: translation, rotation, and vibration. Although translation energy is quantized, the quanta are so small ($\varepsilon_{\text{trans}} \ll kT$) that under all experimentally accessible conditions the translational partition function can be evaluated explicitly by integration. As we saw in Chapter 4, the energy of translation motion calculated from the partition function, is:

$$U_{\text{trans}} = \frac{3}{2}RT\,.$$

Therefore

$$C_{\text{trans}} = \frac{\mathrm{d}U_{\text{trans}}}{\mathrm{d}T} = \frac{3}{2}R\,.$$

No deviations from this formula due to quantization effects have been detected.

The energy levels of a linear rotating molecule are given by the formula discussed in Chapter 4

$$E_{\text{rot}} = \frac{h^2}{8\pi^2 I}J(J+1)$$

where J is a positive integer or zero. From this we deduced in Chapter 4 that, for diatomic molecules other than the isotopes of hydrogen, to a very good approximation:

$$q_{\text{rot}} = \frac{8\pi^2 IkT}{\sigma h^2}\,.$$

For a linear molecule:

$$U_{\text{rot}} = RT^2\frac{\mathrm{d}\,\ln q_{\text{rot}}}{\mathrm{d}T} = RT.$$

Therefore,

$$C_{\text{rot}} = \frac{\mathrm{d}U_{\text{rot}}}{\mathrm{d}T} = R.$$

(This result is identical with the classical formula in which the two degrees of rotational freedom of a linear molecule each contribute $\frac{1}{2}R$ to C_{rot}.) In practice we find that for all diatomic gases, except hydrogen and its isotopes, the experimental value is indeed R and quantization does not affect the rotational heat capacity. That quantization of rotation has the greatest effect for hydrogen should not be surprising. The spacing of rotational energy levels is inversely proportional to the moment of inertia, with the result that the energy separation of $J = 0$ and $J = 1$ is greatest for hydrogen. For example, this spacing is 30 times as large in hydrogen as it is in nitrogen. Thus, whilst quantization effects are noticeable for hydrogen between 20 K and 300 K, the temperature would need to be lowered by a factor of 30 to produce similar effects in nitrogen. The highest temperature at which non-classical behaviour could be expected from nitrogen is thus 10 K, at which temperature it is frozen.

The vibrational energy levels of a diatomic simple harmonic oscillator are given the formula:

$$E_{\text{vib}} = h\nu(v + \frac{1}{2})$$

where ν is the frequency of the vibration and v is the vibrational quantum number, a position integer or zero. As we saw in Chapter 4 this leads to the vibrational partition function

$$q_{\text{vib}} = (1 - \mathrm{e}^{-h\nu/kT})^{-1}$$

$$\ln q_{\text{vib}} = -\ln(1 - \mathrm{e}^{-h\nu/kT})$$

$$\frac{\mathrm{d}(\ln q_{\text{vib}})}{\mathrm{d}T} = \frac{h\nu}{kT^2} \cdot \frac{\mathrm{e}^{-h\nu/kT}}{1 - \mathrm{e}^{-h\nu/kT}}$$

$$U_{\text{vib}} = RT^2 \frac{\mathrm{d}(\ln q_{\text{vib}})}{\mathrm{d}T} = RT \cdot \frac{h\nu}{kT} \cdot \frac{\mathrm{e}^{-h\nu/kT}}{1 - \mathrm{e}^{-h\nu/kT}}.$$

When the substitution $u = h\nu/kT$ is made, this equation is of the form quoted in Appendix 3. If this equation is rearranged we can obtain:

$$U_{\text{vib}} = \frac{Rh\nu}{k} \cdot \frac{1}{\mathrm{e}^{(h\nu/kT)} - 1}$$

$$C_{\text{vib}} = R\left(\frac{h\nu}{kT}\right)^2 \frac{\mathrm{e}^{h\nu/kT}}{(\mathrm{e}^{h\nu/kT} - 1)^2}. \tag{6.1}$$

With the substitution $u = h\nu/kT$, this equation is also quoted in Appendix 3.

The vibrational heat capacity of a diatomic molecule, calculated from Eq. (6.1), can vary between zero and R, depending upon the value of the ratio $h\nu/kT$ (see Appendix 3). At high temperatures when $h\nu/kT$ is small, we can expand $(1 - \mathrm{e}^{-h\nu/kT})^{-1}$ as a power series and neglect terms in $(h\nu/kT)^2$ and beyond. We then get:

$$q_{\text{vib}} = \frac{kT}{h\nu} (\text{high temperature})$$

$$C_{\text{vib}} = \frac{\mathrm{d}}{\mathrm{d}T}\left(RT^2 \frac{\mathrm{d}\ln q_{\text{vib}}}{\mathrm{d}T}\right),$$

whence

$$C_{\text{vib}} = R.$$

At low temperatures where $h\nu/kT$ is large, $\mathrm{e}^{-h\nu/kT} \to 0$ and $q_{\text{vib}} \to 1$. Hence

$$C_{\text{vib}} \to 0 \text{ as } T \to 0.$$

The frequency ν is related to the mass and bond strength of the molecule by the equation $\nu = \frac{1}{2\pi}\sqrt{(k/\mu)}$, where k is the force-constant of the bond and μ the reduced mass of the molecule. Thus light (low μ) strongly bonded (high k) molecules have high vibration frequencies and low vibrational heat capacities at room temperature. Conversely, heavy, weakly-bonded molecules have near-classical vibration heat capacities, e.g. at 300 K C_{vib} ~ 0 J K^{-1}mol^{-1} for H_2; ~ 4.2 J K^{-1}mol^{-1} for Cl_2, and ~ 8.4 J K^{-1}mol^{-1} for I_2.

It is interesting to note that a ten-fold change in the ratio $h\nu/kT$ from a low-temperature value of, say 6 to a high-temperature value of 0.6 has the effect of changing the heat capacity from $0.09R$ to $0.97R$.

The temperature dependence of the heat capacity of a diatomic gas can be illustrated schematically (Fig. 6.2).

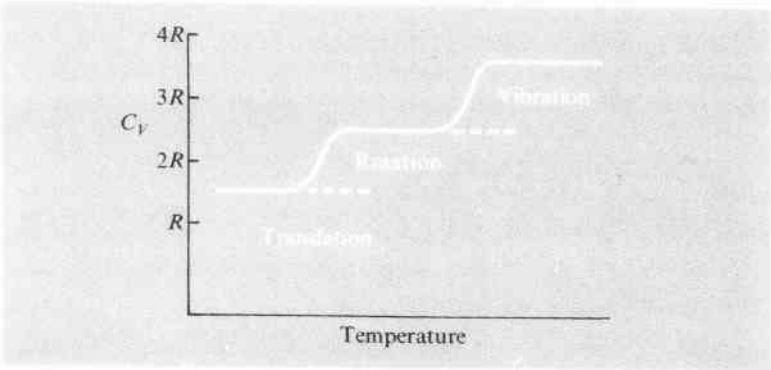


Fig. 6.2. Variation of heat capacity of a diatomic gas with temperature (only hydrogen isotopes give purely translation heat capacities).

Heat Capacity of Simple Solids

One consequence of the third law is a prediction that the heat capacity of a solid tends to zero at very low temperatures. We can see how this happens if we note that the only way in which a monatomic solid can take in heat is by increasing the vibrational excitation of its constituent atoms. This excitation naturally increases the entropy of the system. When the temperature is lowered sufficiently all the particles fall back to the lowest available level, thus reducing the entropy to zero. As we have seen for the diatomic molecule, when there is no vibrational excitation ($h\nu \ll kT$)

the vibrational heat capacity tends to zero. This result can be expressed mathematically for the heat capacity at constant pressure as follows:

$$C_P = \left(\frac{\partial H}{\partial T}\right)_P \text{ by definition}$$

$$= \left(\frac{\partial q}{\partial T}\right)_P \text{ (since only } P - V \text{ work is done by the system)}$$

$$= \left(T\frac{\partial S}{\partial T}\right)_P \text{ (a reversible process)}$$

$$= \left(\frac{\partial S}{\partial \ln T}\right)_P .$$

Now as T tends to zero, the third law predicts that the entropy tends to zero. Since $\ln T$ tends to minus infinity as $T \to 0$, the bracket tends to zero, i.e. $C_P \to 0$. Similarly $C_V \to 0$ as $\to T \to 0$. The temperature dependence of C_V of for some monatomic solids is shown in Fig. 6.3.

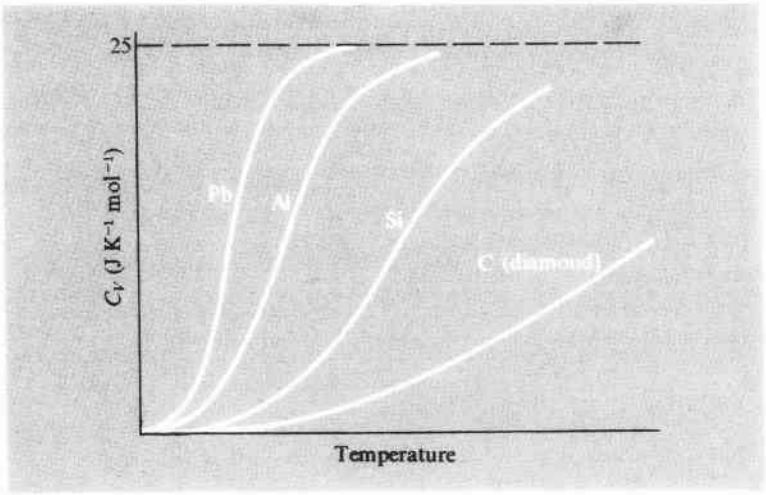


Fig. 6.3. Heat capacities of some monatomic solids.

The high temperature limiting value of C_V for a monatomic solid is $3R$, about 25 J K^{-1}mol^{-1}. This is, of course, Dulong and Petit's Law. Classically it was derived by assuming each atom could vibrate in three directions, with each mode of vibration contributing R to the heat capacity.

This approach, however, could provide no explanation for a heat capacity which goes to zero.

In order to understand the behaviour of solids, we need to find a model for the system which seems physically reasonable and whose thermodynamic properties are in accord with the experimental results. There have been two particularly important approaches to this problem, first by Einstein and then by Debye.

The Einstein Theory of Heat Capacities

In this treatment, the model is of a solid which consists of N independent simple harmonic oscillators whose energy is quantized according to the equation $E = nh\nu$, where ν is the fundamental frequency of the oscillators and n is a positive integer or zero, i.e. an oscillator may only have discrete energies $E, 2E, 3E \ldots$ above the zero-point level. It is at this point that the model differs from the classical treatment, in which the oscillating atoms can have any frequency and, therefore, any energy. When there is no constraint on the vibrations the solid has a temperature-independent heat capacity of $3R$.

The problem we now face is one of finding the energy of a system of solid-state simple harmonic oscillators. It can be solved by the same mathematical procedure as that we have already used earlier in this Chapter for the vibrations of a diatomic molecule. The only point of difference arises because a gaseous diatomic molecule has but a single mode of vibration, along the internuclear axis, while the motion of a solid-state oscillator can be resolved into three components, one along each of the Cartesian coordinates. Thus from our earlier equation for a vibrating diatomic molecule:

$$U_{\text{vib}} = \frac{Rh\nu}{k} \cdot \frac{1}{e^{h\nu/kT} - 1}$$

we can calculate the average energy of an oscillator $\bar{E}$ using the relation

$$\bar{E} = \frac{U}{N}.$$

Hence

$$\bar{E} = \frac{h\nu}{e^{h\nu/kT} - 1} \qquad (R = kN).$$

Since each solid-state oscillator has 3 directions of vibration or 'degrees of freedom' the total vibrational energy $3N\bar{E}$. We can now reach the Einstein value of the heat capacity of differentiation:

$$U_{\mathrm{E}} = 3N\bar{E} = 3N\frac{h\nu}{\mathrm{e}^{h\nu/kT}-1}$$

$$C_{\mathrm{E}} = \left(\frac{\partial U}{\partial T}\right)_V = \frac{3Nk\left(\frac{h\nu}{kT}\right)^2 \mathrm{e}^{h\nu/kT}}{\left(\mathrm{e}^{h\nu/kT}-1\right)^2}\,.$$

At high temperature ($h\nu \ll kT$) the upper limit of C_{E} can be obtained by the same method as the one used for a gas. The result is

$$(C_{\mathrm{E}})_{\mathrm{High}\ T} = 3R\,.$$

At low temperature ($h\nu \gg /kT$) the energy is:

$$U_{\mathrm{E}} = 3Nh\nu \mathrm{e}^{-h\nu/kT}.$$

Therefore

$$(C_{\mathrm{E}})_{\mathrm{Low}T} = 3Nk\left(\frac{h\nu}{kT}\right)^2 \mathrm{e}^{-h\nu/kT}.$$

As $T \to 0$ the exponential term controls the equation and $C_V \to 0$. This again is the correct limit. However, the exponential decrease at the lowest temperature is more rapid than the experimental results, which generally have a limiting T^3 dependence.

In the equation for C_V the only unknown parameter is the frequency of the oscillators, ν. Thus, by fitting the equation as well as possible to the experimental results, a value of this characteristic frequency of any particular solid can be obtained. It is sometimes more convenient to use temperature than frequency as the characteristic parameter. The two are related by the equation

$$h\nu_{\mathrm{E}} = k\theta_{\mathrm{E}}$$

where θ_{E} is the Einstein temperature.

For many common inorganic crystals $\theta_{\mathrm{E}} \sim 200K$, so that $\nu_{\mathrm{E}} \sim 4 \times 10^{12}\mathrm{Hz}$. It is usually possible to get good agreement between theory and experiment down to a temperature of about $0.2\theta_{\mathrm{E}}$. Below this temperature the exponential factor causes too rapid a fall in the heat capacity.

It is useful to consider the shortcoming of the Einstein model. The postulate of a single vibration frequency for the particles appears to be a good approximation at intermediate and high temperatures ($T > 0.2\theta_{\mathrm{E}}$). But at low temperatures the spacing of the lowest energy levels, which is $h\nu_{\mathrm{E}}$, is too large. The result is that the heat capacity falls too rapidly. This conclusion illustrates our general observation that widely spaced energy levels, or large quanta, lead to a small heat capacity. In choosing a new model, therefore, we should seek one in which there are some, but not too many, low-lying energy levels, which can be populated at low temperatures. The improvement at low temperatures which the Debye treatment offers over the Einstein model arises from this feature.

The Debye theory of heat capacities

Instead of treating the particles as though the motions were independent of one another, the Debye approach recognizes that the particles do interact. In this model the solid is treated as a homogeneous continuum, and the allowed vibrational energy levels become those of the crystal as a whole. The vibrational motion can be thought of as the three dimensional analogue of the familiar vibrations of a violin string. The quantization of energy is then governed by the condition that the vibrations have a node at the edge of the solid, i.e. $l = n\lambda/2$ where l is the length of the side of the solid and λ the wavelength of the vibration. n is a positive integer.

In order to calculate the total energy of vibration we need to know how many oscillations there are at each allowed frequency. We will call the mathematical function which gives us this result $f(\nu)$. The elastic theory of solids can be used to calculate $f(\nu)$, with the result $f(\nu) \propto \nu^2$. Over any small range of frequency, at any particular frequency ν_i, the number of vibrations is $f(\nu_i)\delta\nu$. Since we still have N atoms, the overall number of allowed vibrations remains at $3N$, i.e.

$$\sum_{\nu} f(\nu)\delta\nu = 3N\,.$$

An important consequence of this equation is that $f(\nu)$ and, therefore, ν cannot go on increasing indefinitely but must reach a limiting value. This limiting frequency is called the *Debye cut-off frequency*, ν_{D}. In this model the energy levels of the vibrating solid are closely spaced, so that no serious mathematical error is introduced if we integrate instead of summating the

equation for $f(\nu)$, i.e.

$$\int_0^{\nu_D} a\nu^2 d\nu = 3N$$

where a is a constant of proportionality.

Hence we get

$$a = 9N/\nu_D^3, \quad f(\nu) = 9N\nu^2/\nu_D^3.$$

We are now ready to write down the total vibrational energy of the crystal. This energy is the product of the average energy of an oscillator of frequency ν and the number of oscillators at that frequency, summed or integrated over all the allowed frequencies. This latter range is effectively 0 to ν_D. The average energy of an oscillator we have calculated previously in the Einstein treatment; it is

$$\bar{E} = \frac{h\nu}{e^{h\nu/kT} - 1}.$$

Thus

$$U = \int_0^{\nu_D} \bar{E} f(\nu) d\nu$$
$$= \int_0^{\nu_D} \frac{h\nu}{e^{h\nu/kT} - 1} \frac{9N\nu^2 d\nu}{\nu_D^3}.$$

Once again we can express the characteristic frequency ν_D as a temperature, θ_D, by means of the equation $h\nu_D = k\theta_D$. If we let $u = h\nu/kT$ and differentiate U to get C_V the result is:

$$C_D = 9R\left(\frac{T}{\theta_D}\right)^3 \int_0^{\theta_D/T} \frac{u^4 e^u du}{(e^u - 1)^2}$$

This integral can be calculated and is tabulated in standard works; at low temperatures the limit θ_D/T on the integral may be replaced by ∞, the value of the integral becomes independent of T, and so the heat capacity varies as T^3/θ_D^3. Once again C_V depends on the ratio of the temperature of a characteristic temperature, this time θ_D. This latter quantity is the only parameter specific to any particular monatomic solid. It follows that if $\theta_D$is chosen correctly and C_V is plotted against T/θ_D all solids should fall on the same curve. This should also be true of the Einstein model if C_V is plotted against T/θ_E. The experimental results show that this expectation is rather accurately fulfilled but that, as we have seen, the Einstein heat

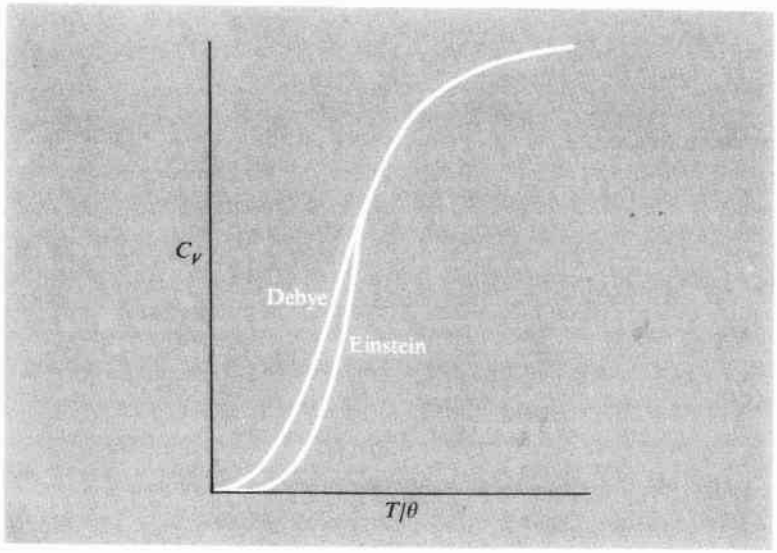


Fig. 6.4. Comparison of Einstein and Debye theories of heat capacities.

capacity falls off too fast at low temperatures. The Debye and Einstein curves are shown in Fig. 6.4.

Having obtained θ_D from the curve it is possible to make a further test of the Debye theory, because the elastic theory of solids allows a quite independent calculation of θ_D to be made, using only the independently measured elastic properties of the solid. The relation is:

$$\theta_D = \frac{h}{k} \cdot V_0 \left(6\pi^2 \frac{N}{V}\right)^{\frac{1}{3}}$$

where V_0 is the velocity of sound in the solid of volume V. There is good, though not perfect, agreement between the two results. The discrepancies arise because the distribution of allowed vibration frequencies chosen in the Debye treatment is something of a simplification. A more elaborate treatment gives better agreement but a less useful formula.

A particularly valuable feature of the Debye equation is the simple form of the low-temperature region. This T^3 temperature dependence can be used to extrapolate the experimental results from the lowest accessible temperature to the absolute zero. As we shall see, this extrapolation is required for the experimental determination of the third-law entropies.

Effects of Electron on Heat Capacities

One distinctive property of metals is their electrical conductivity. This property results from the ability of the conduction electrons to move freely through the lattice. It might have been expected, therefore, that these electrons would make a significant contribution to the heat capacity of the metal. However, at room temperature the difference between the heat capacity of a metal and of a dielectric is hardly significant. By contrast, at very low temperatures, where the lattice heat capacity is low and falling rapidly (Debye T^3 region), the contribution of the electrons does become important and it can be measured. The equation for the total heat capacity in the Debye T^3 region is

$$C_{\text{metal}} = \underset{\text{(lattice)}}{\alpha T^3} + \underset{\text{(electrons)}}{\gamma T} .$$

Thus at low temperatures the electrons make an appreciable contribution to the heat capacity. This is most clearly seen plotting C/T against T^2 for a metal and an insulator as in Fig. 6.5.

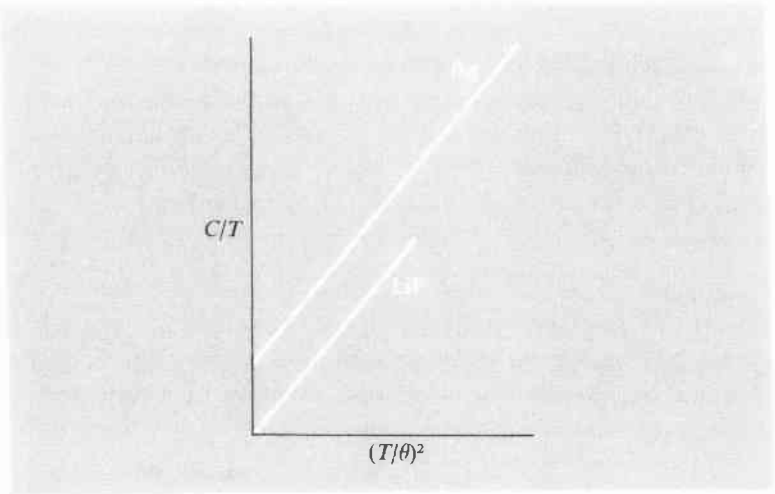


Fig. 6.5. Comparison of low-temperature heat capacities of a metal and an insulator.

The reason why the electrons have a heat capacity which is so much less that the classical value can be understood from our earlier discussions of

the effect of the quantization of energy levels on heat capacities. We have seen that when the characteristic energy is large compared with kT the associated heat capacity is small. The elementary band theory of metals can be used to show how it comes about that the energy of the electrons is indeed much greater than kT.

If we imagine that the metal consists of a regular crystalline array of atoms with the electrons free to move throughout the lattice then we find that there are groups of contiguous energy levels, 'bands', and that these may be separated from neighbouring bands by an energy gap. Within each band, the electrons (being Fermi-Dirac particles) occupy the lowest energy levels consistent with the Pauli principle, i.e. only two electrons, with opposed spins, can occupy each energy level. Thus, as electrons are fed in to the lattice, higher and higher energy states must be occupied. The energy-level diagram for this process usually drawn by plotting the number of states with a particular energy, $N(E)$, against energy. The resulting curve is the band. Electrons then occupy the $N_e/2$ lowest energy levels, where N_e is the number of conduction electrons. The resulting situation at absolute zero is shown in Fig. 6.6. The area shown shaded represents energy levels occupied by electrons, the clear area represents empty levels and the vertical line E_F marks the boundary. E_F is called the *Fermi level.*

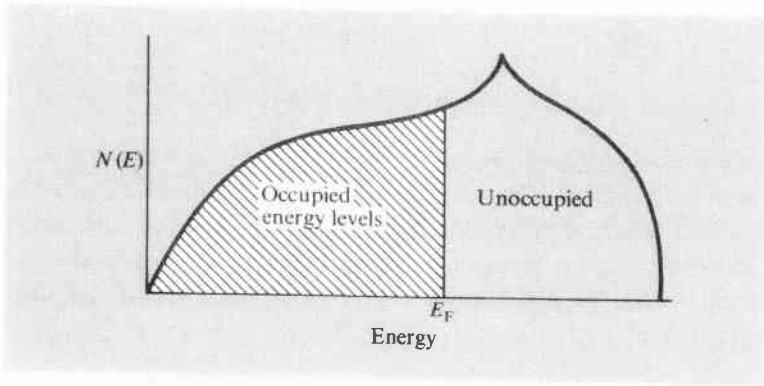


Fig. 6.6. A simple band diagram for a metal.

As the diagram shows, it is the electrons at the Fermi level and these alone which have unoccupied energy levels adjacent to them. Excitation to these empty levels may be achieved thermally or by the application of

Specific Heat of Solid

Consider monatomic molecules

A mole of solid consists of N molecules ← Avogadro's number

Each molecule has three oscillating degree of freedom

$$U = 3N\underbrace{k}_{R}T$$

R ← gas constant

U: internal energy of a mole of solid

$$\Rightarrow \quad C_V = 3R$$ law of Dulong and Petit

At lower temperature, the molar heat capacities vary.
In fact $C_V \to T^3$ as $T \to 0$

1907 Einstein treat the solid as a set of harmonic oscillators. all have the same frequency ν_0, and calculate the average internal energy on the assumption that these oscillators can have only the discrete energies proposed by Planck, i.e., $nh\nu_0$

$$U = 3N \frac{h\nu_0}{e^{h\nu_0/RT} - 1}$$

$\frac{h\nu_0}{e^{h\nu_0/RT}-1}$: average energy of oscillation per degree of freedom

Application of quantum concept to solid state physics

all oscillating with same ν

$$C_V = \frac{dU}{dT} = 3Nk \left(\frac{h\nu_0}{RT}\right)^2 \left[e^{h\nu_0/RT} - 1\right]^{-2} e^{h\nu_0/RT}$$

reasonably agreement with experiments at reasonably low temperature

ν_0 is a parameter

Debye's theory (1912)

$$N(\nu)\,d\nu = \frac{4\pi V}{v^3}\nu^2 d\nu$$

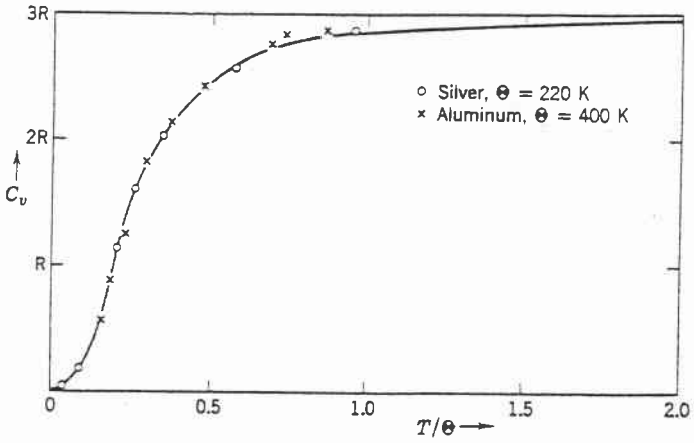
v: velocity of sound

$$\int_0^{\nu_m} N(\nu)\,d\nu = 3N$$

$$\Rightarrow \quad \nu_m = v\left(\frac{9N}{4\pi V}\right)^{1/3}$$

$$\bar{E} = \frac{h\nu}{e^{h\nu/RT} - 1}$$

$$\Rightarrow \quad U = \int_0^{\nu_m} \frac{4\pi V}{v^3}\nu^2 d\nu \frac{h\nu}{e^{h\nu/RT} - 1}$$



Heat capacity at constant volume as a function of temperature. The solid curve is the Debye function (eq. 8-17). The curve was fitted to the data points for each metal in order to determine the Debye temperature Θ for the metal, and then the data were replotted as a function of T/Θ. (From F. Seitz, *Modern Theory of Solids*, McGraw-Hill, New York, 1940.)

分類：
編號： 3 -11
總號：

Define $x = \frac{h\nu}{kT}$, $x_m = \frac{h\nu_m}{kT} \equiv \frac{\Theta}{T}$ ← Debye temperature

$$\Rightarrow \quad U = 3RT \frac{3}{x_m^3} \int_0^{x_m} \frac{x^3 dx}{e^x - 1}$$

$$= 9R \frac{T^4}{\Theta^3} \int_0^{\Theta/T} \frac{x^3 dx}{e^x - 1} \qquad T \to 0 \qquad \int_0^{\infty} \frac{x^3 dx}{e^x - 1} \to \text{constant}$$

$$\Rightarrow \quad C_V = \frac{dU}{dT} = 9R \left[4 \left(\frac{T}{\Theta} \right)^3 \int_0^{\Theta/T} \frac{x^3}{e^x - 1} dx - \frac{\Theta}{T} \frac{1}{e^{\Theta/T} - 1} \right]$$

This is Debye's theory of specific heat.

$T >> \Theta$ $\qquad U \cong 9R \frac{T^4}{\Theta^3} \int_0^{\Theta/T} x^2 dx$

$$= 9R \frac{T^4}{\Theta^3} \cdot \frac{\Theta^3}{3T^3} = 3RT$$

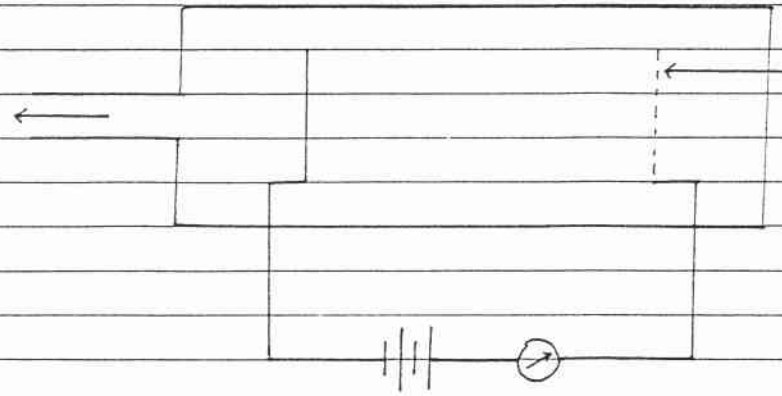
$\Rightarrow \quad C_V = 3R$ $\qquad$ law of Dulong and Petit

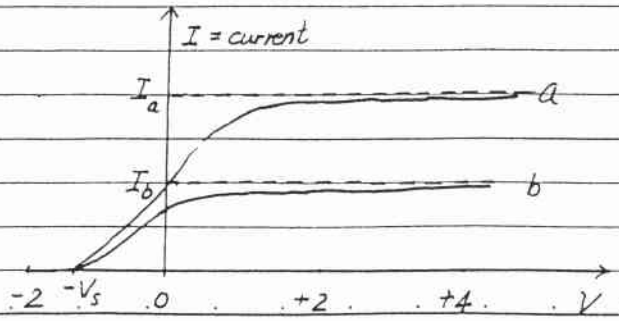
Furthermore $\qquad C_V \propto T^3$ $\qquad$ as $T \to 0$

Photoelectric Effect

Definition : the release of electrons from matter under the influence of electromagnetic radiation

Discovered by Hertz in 1887





V = potential of the second electrode with respect to photocathode

Modern Physics

Michael Fowler, University of Virginia

The Photoelectric Effect

Michael Fowler
University of Virginia

Hertz Finds Maxwell's Waves: and Something Else

The most dramatic prediction of Maxwell's theory of electromagnetism, published in 1865, was the existence of electromagnetic waves moving at the speed of light, and the conclusion that light itself was just such a wave. This challenged experimentalists to generate and detect electromagnetic radiation using some form of electrical apparatus. The first clearly successful attempt was by Heinrich Hertz in 1886. He used a high voltage induction coil to cause a spark discharge between two pieces of brass, to quote him, "*Imagine a cylindrical brass body, 3 cm in diameter and 26 cm long, interrupted midway along its length by a spark gap whose poles on either side are formed by spheres of 2 cm radius.*" The idea was that once a spark formed a conducting path between the two brass conductors, charge would rapidly oscillate back and forth, emitting electromagnetic radiation of a wavelength similar to the size of the conductors themselves.

To prove there really was radiation emitted, it had to be detected. Hertz used a piece of copper wire 1 mm thick bent into a circle of diameter 7.5 cms, with a small brass sphere on one end, and the other end of the wire was pointed, with the point near the sphere. He added a screw mechanism so that the point could be moved very close to the sphere in a controlled fashion. This "receiver" was designed so that current oscillating back and forth in the wire would have a natural period close to that of the "transmitter" described above. The presence of oscillating charge in the receiver would be signaled by a spark across the (tiny) gap between the point and the sphere (typically, this gap was hundredths of a millimeter). (It was suggested to Hertz that this spark gap could be replaced as a detector by a suitably prepared frog's leg, but that apparently didn't work.)

The experiment was very successful - Hertz was able to detect the radiation up to fifty feet away, and in a series of ingenious experiments established that the radiation was reflected and refracted as expected, and that it was polarized. The main problem - the limiting factor in detection -- was being able to *see* the tiny spark in the receiver. In trying to improve the spark's visibility, he came upon something very mysterious. To quote from Hertz again (he called the transmitter spark *A*, the receiver *B*): "*I occasionally enclosed the spark B in a dark case so as to more easily make the observations; and in so doing I observed that*

the maximum spark-length became decidedly smaller in the case than it was before. On removing in succession the various parts of the case, it was seen that the only portion of it which exercised this prejudicial effect was that which screened the spark B from the spark A. The partition on that side exhibited this effect, not only when it was in the immediate neighbourhood of the spark B, but also when it was interposed at greater distances from B between A and B. A phenomenon so remarkable called for closer investigation."

Hertz then embarked on a very thorough investigation. He found that the small receiver spark was more vigorous if it was exposed to ultraviolet light from the transmitter spark. It took a long time to figure this out - he first checked for some kind of electromagnetic effect, but found a sheet of glass effectively shielded the spark. He then found a slab of quartz did not shield the spark, whereupon he used a quartz prism to break up the light from the big spark into its components, and discovered that the wavelength which made the little spark more powerful was beyond the visible, in the ultraviolet.

In 1887, Hertz concluded what must have been months of investigation: "... *I confine myself at present to communicating the results obtained, without attempting any theory respecting the manner in which the observed phenomena are brought about.*"

Hallwachs' Simpler Approach

The next year, 1888, another German physicist, Wilhelm Hallwachs, in Dresden, wrote:

"*In a recent publication Hertz has described investigations on the dependence of the maximum length of an induction spark on the radiation received by it from another induction spark. He proved that the phenomenon observed is an action of the ultraviolet light. No further light on the nature of the phenomenon could be obtained, because of the complicated conditions of the research in which it appeared. I have endeavored to obtain related phenomena which would occur under simpler conditions, in order to make the explanation of the phenomena easier. Success was obtained by investigating the action of the electric light on electrically charged bodies.*"

He then describes his very simple experiment: a clean circular plate of zinc was mounted on an insulating stand and attached by a wire to a gold leaf electroscope, which was then charged negatively. The electroscope lost its charge very slowly. However, if the zinc plate was exposed to ultraviolet light from an arc lamp, or from burning magnesium, charge leaked away quickly. If the plate was positively charged, there was no fast charge leakage. (We showed this as a lecture demo, using a UV lamp as source.)

Questions for the reader: Could it be that the ultraviolet light somehow spoiled the insulating properties of the stand the zinc plate was on? Could it be that electric or magnetic effects from the large current in the arc lamp somehow

caused the charge leakage?

Although Hallwach's experiment certainly clarified the situation, he did not offer any theory of what was going on.

J.J. Thomson Identifies the Particles

In fact, the situation remained unclear until 1899, when Thomson established that the ultraviolet light caused *electrons* to be emitted, the same particles found in cathode rays. His method was to enclose the metallic surface to be exposed to radiation in a vacuum tube, in other words to make it the cathode in a cathode ray tube. The new feature was that electrons were to be ejected from the cathode by the radiation, rather than by the strong electric field used previously.

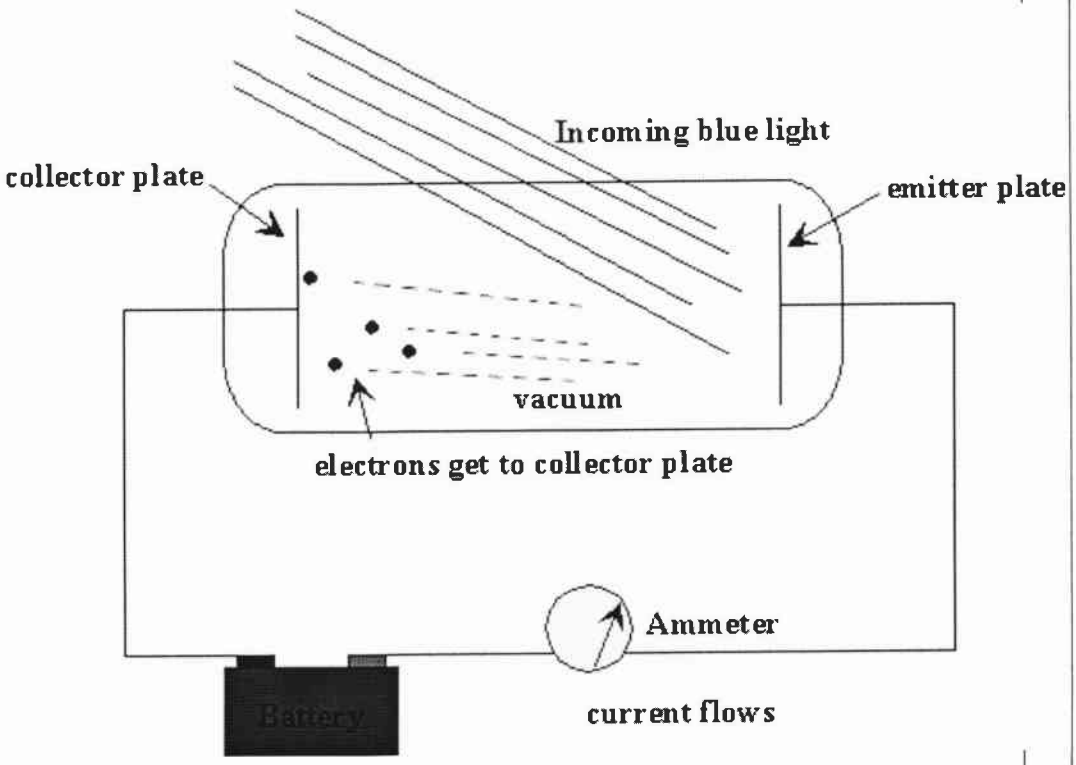
By this time, there was a plausible picture of what was going on. Atoms in the cathode contained electrons, which were shaken and caused to vibrate by the oscillating electric field of the incident radiation. Eventually some of them would be shaken loose, and would be ejected from the cathode. It is worthwhile considering carefully how the *number* and *speed* of electrons emitted would be expected to vary with the *intensity* and *color* of the incident radiation. Increasing the intensity of radiation would shake the electrons more violently, so one would expect more to be emitted, and they would shoot out at greater speed, on average. Increasing the frequency of the radiation would shake the electrons faster, so might cause the electrons to come out faster. For very dim light, it would take some time for an electron to work up to a sufficient amplitude of vibration to shake loose.

Lenard Finds Some Surprises

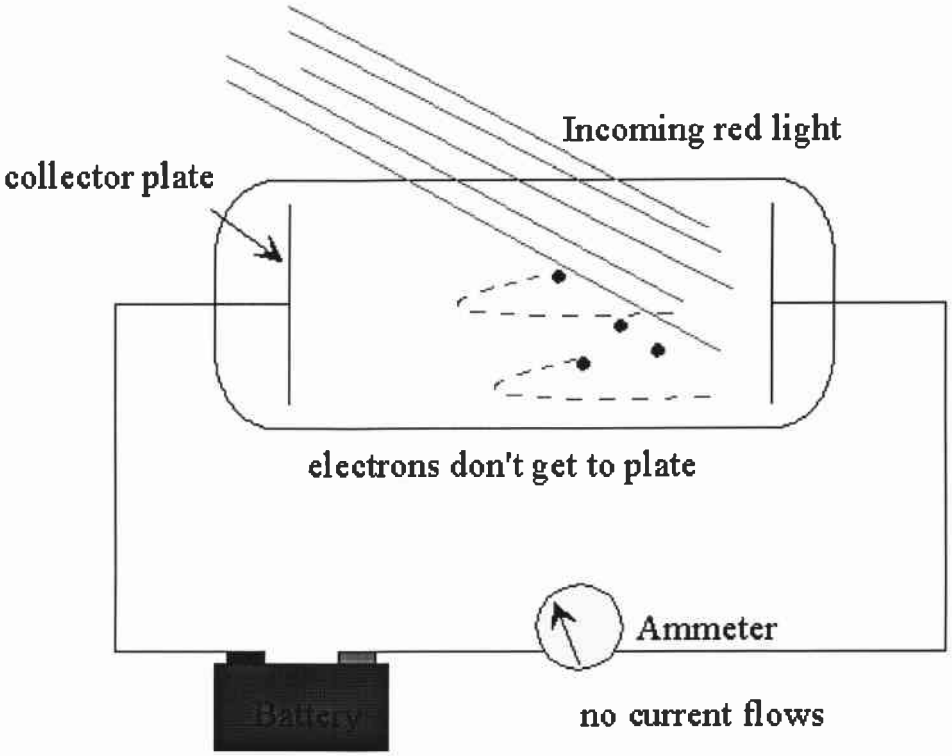
In 1902, Lenard studied how the energy of the emitted photoelectrons varied with the intensity of the light. He used a carbon arc light, and could increase the intensity a thousand-fold. The ejected electrons hit another metal plate, the collector, which was connected to the cathode by a wire with a sensitive ammeter, to measure the current produced by the illumination. To measure the energy of the ejected electrons, Lenard charged the collector plate negatively, to repel the electrons coming towards it. Thus, only electrons ejected with enough kinetic energy to get up this potential hill would contribute to the current. Lenard discovered that there was a well defined minimum voltage that stopped any electrons getting through, we'll call it V_{stop}. To his surprise, he found that V_{stop} did not depend at all on the intensity of the light! Doubling the light intensity doubled the *number* of electrons emitted, but did not affect the *energies* of the emitted electrons. The more powerful oscillating field ejected more electrons, but the maximum individual energy of the ejected electrons was the same as for the weaker field.

But Lenard did something else. With his very powerful arc lamp, there was sufficient intensity to separate out the colors and check the photoelectric effect using light of different colors. He found that the maximum energy of the ejected electrons *did* depend on the color --- the shorter wavelength, higher frequency light caused electrons to be ejected with more energy. This was, however, a

fairly qualitative conclusion --- the energy measurements were not very reproducible, because they were extremely sensitive to the condition of the surface, in particular its state of partial oxidation. In the best vacua available at that time, significant oxidation of a fresh surface took place in tens of minutes. (The details of the surface are crucial because the fastest electrons emitted are those from right at the surface, and their binding to the solid depends strongly on the nature of the surface --- is it pure metal or a mixture of metal and oxygen atoms?)



Question: In the above figure, the battery represents the potential Lenard used to charge the collector plate negatively, which would actually be a variable voltage source. Since the electrons ejected by the blue light are getting to the collector plate, evidently the potential supplied by the battery is less than V_{stop} for blue light. Show with an arrow on the wire the direction of the electric current in the wire.



Einstein Suggests an Explanation

In 1905 Einstein gave a very simple interpretation of Lenard's results. He just assumed that the incoming radiation should be thought of as quanta of frequency hf, with f the frequency. In photoemission, one such quantum is absorbed by one electron. If the electron is some distance into the material of the cathode, some energy will be lost as it moves towards the surface. There will always be some electrostatic cost as the electron leaves the surface, this is usually called the work function, W. The most energetic electrons emitted will be those very close to the surface, and they will leave the cathode with kinetic energy

$$E = hf - W.$$

On cranking up the negative voltage on the collector plate until the current just stops, that is, to V_{stop}, the highest kinetic energy electrons must have had energy eV_{stop} on leaving the cathode. Thus,

$$eV_{stop} = hf - W$$

Thus Einstein's theory makes a very definite quantitative prediction: if the frequency of the incident light is varied, and V_{stop} plotted as a function of frequency, the slope of the line should be h/e.

It is also clear that there is a minimum light frequency for a given metal, that for which the quantum of energy is equal to the work function. Light below that frequency , no matter how bright, will not cause photoemission.

Millikan's Attempts to Disprove Einstein's Theory

If we accept Einstein's theory, then, this is a completely different way to measure

Planck's constant. The American experimental physicist Robert Millikan, who did not accept Einstein's theory, which he saw as an attack on the wave theory of light, worked for ten years, until 1916, on the photoelectric effect. He even devised techniques for scraping clean the metal surfaces inside the vacuum tube. For all his efforts he found disappointing results: he confirmed Einstein's theory, measuring Planck's constant to within 0.5% by this method. One consolation was that he did get a Nobel prize for this series of experiments.

References

'*Subtle is the Lord...' The Science and Life of Albert Einstein*, Abraham Pais, Oxford 1982.

Inward Bound, Abraham Pais, Oxford, 1986

The Project Physics Course, Text, Holt, Rinehart, Winston, 1970

Life of Lenard

Life of Millikan

previous home next

Photoelectric effect

1887 H. Hertz

Shining light on metal surface ⇒ spark
↓
UV

1899 J. J. Thomson — 1906

Shining UV light → electron
↓
photoelectric effect

1902 Lenard — 1905

Electron energy showed not the slightest dependence on the light intensity

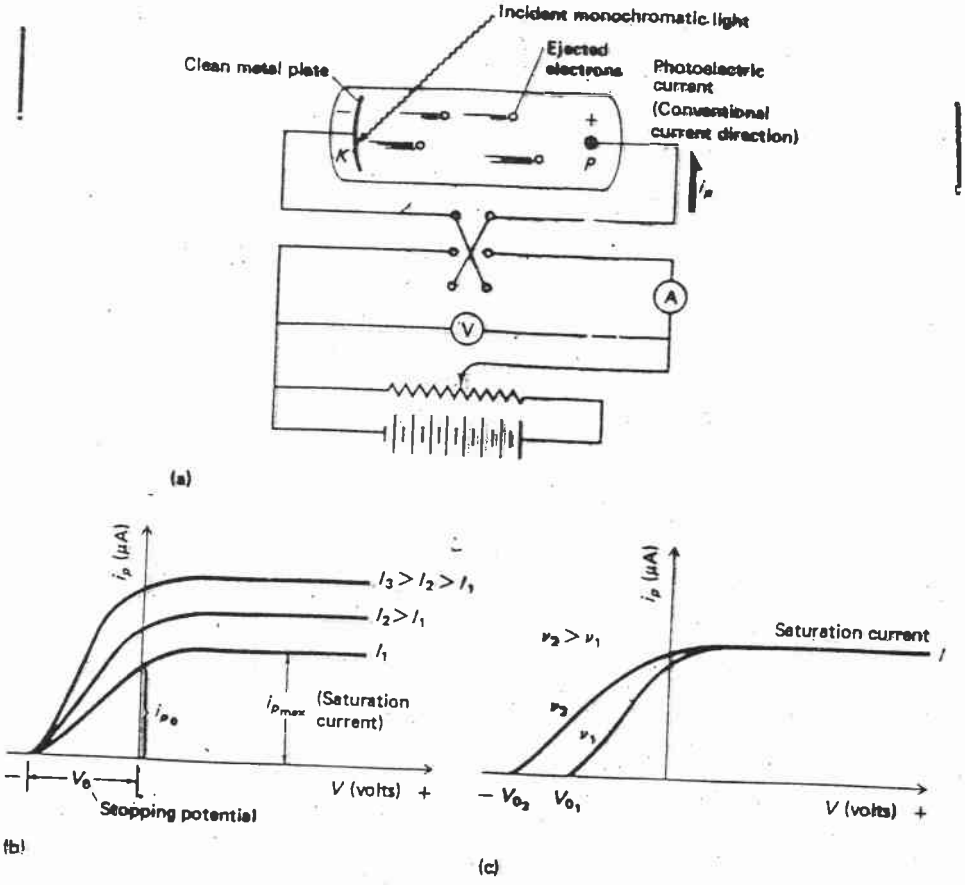
1905 Einstein — 1921

Light - quanta hypothesis

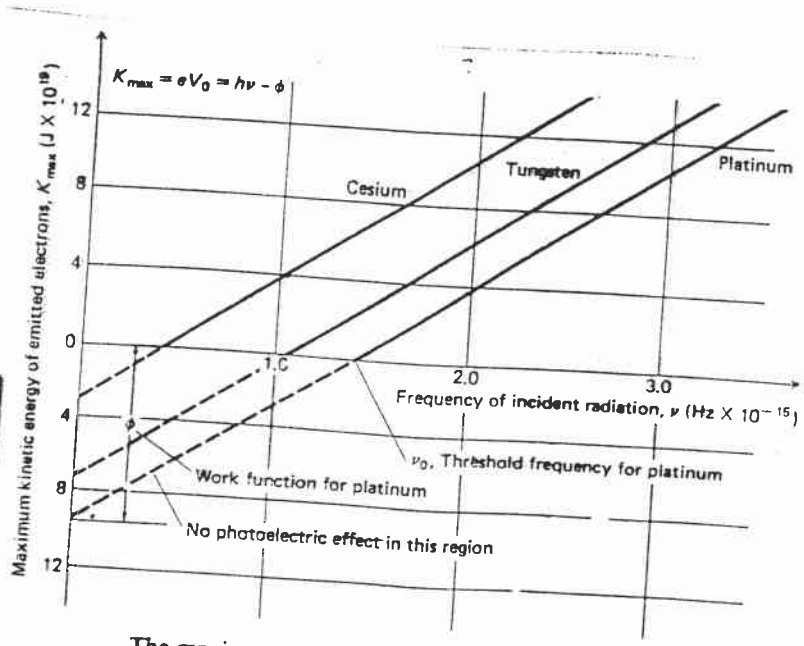
$E_{max} = h\nu - W$
↓ ↳ work function
Einstein equation.

1915 Millikan — 1923

Experimental verification of the Einstein equation



(a) Schematic for photoelectric experiment. (b) Photoelectric current versus the accelerating potential V for incident monochromatic light of wavelength λ. (c) Photoelectric current versus accelerating potential to show frequency dependence.



The maximum kinetic energy of photoelectrons K_{max} $(= eV_0)$ versus the frequency of incident radiation.

Experimental results.

(i) The effect persists even the tube is evulated to a very low pressure ⇒ gaseous ions are not the carrier of the current.

(ii) In curve b the intensity of incident light has been reduced to $\frac{1}{2}$ that of curve a
⇒ the magnitude of the current, when it exists, is proportional to the intensity of the light source

(iii) Some current still reaches the second electrode when $V<0$
⇒ photoelectrons are ejected from the photocathode with non-negligible kinetic energy
Well defined end point ⇒ well defined kinetic energy

$$(K.E.)_{max} = eV_s$$

Photoelectrons of maximum kinetic energy ↔ electrons emitted from the surfaces

∴

Photoelectron with lower energy ↔ originated inside the surface and lost kinetic energy in the process of reaching the surface

$(K.E.)_{max}$ ↔ energy given to an electron in the photoelectric process

(iv) V_s in independent of light intensity ⇒ E_{max} is independent of the intensity of the incident light.

(v) Whether the electrons are emitted depend on the frequency of the light. In general, there will be a threshold that varies from metal to metal: only light with a frequency greater than a given threshold frequency will produce photoelectric effect.

Classical theory

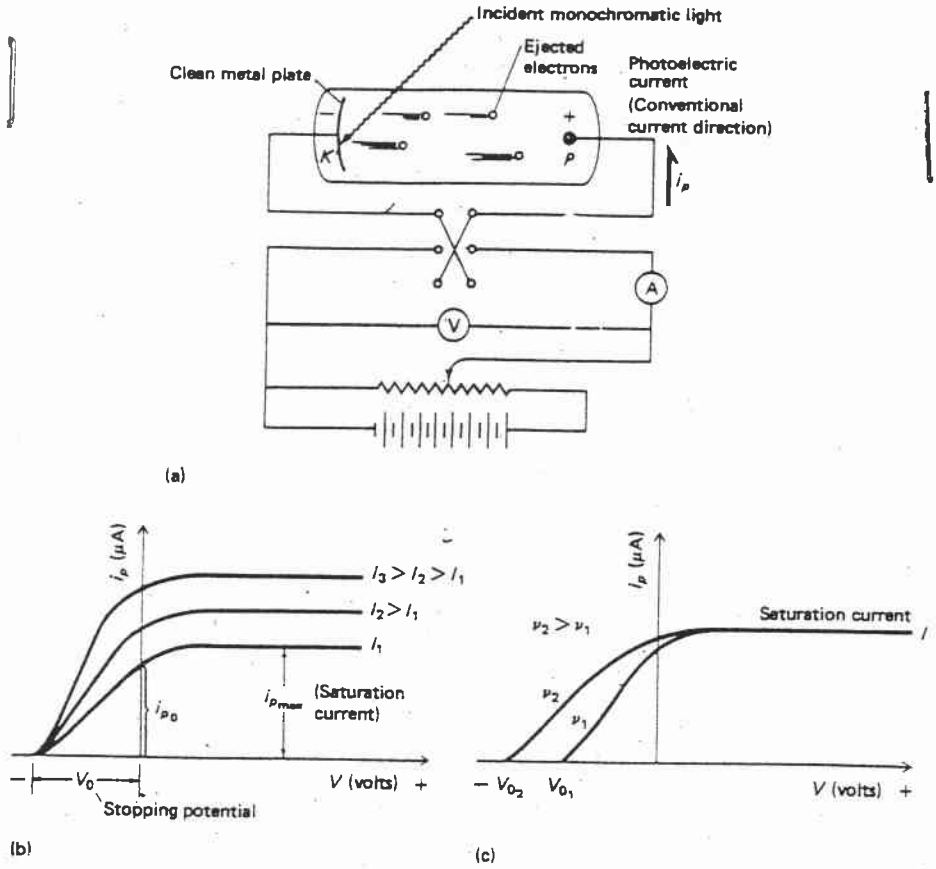
Energy carried by an EM wave ∝ intensity of the light

↓

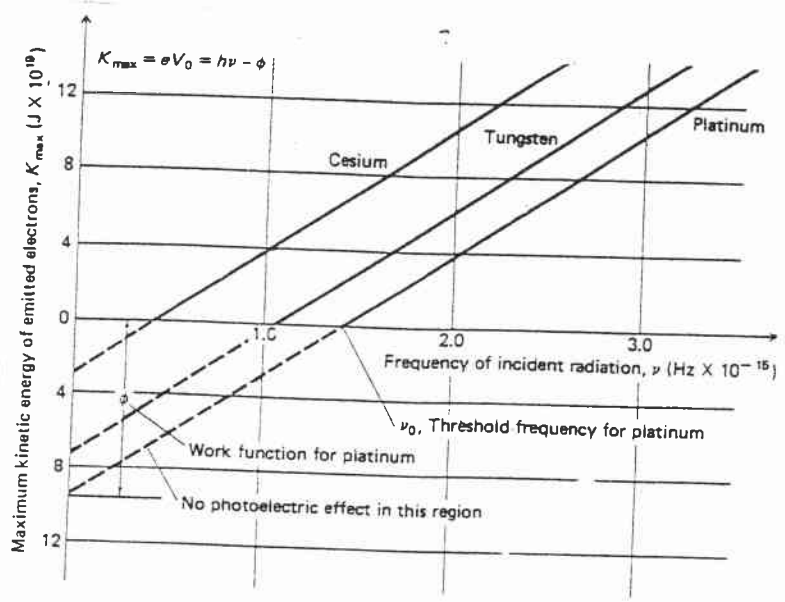
kinetic energy of the emitted electrons

↓

inconsistent with the experimental result



(a) Schematic for photoelectric experiment. (b) Photoelectric current versus the accelerating potential V for incident monochromatic light of wavelength λ. (c) Photoelectric current versus accelerating potential to show frequency dependence.



The maximum kinetic energy of photoelectrons K_{max} ($= eV_0$) versus the frequency of incident radiation.

Time delay problem

In classical theory of light, energy is uniformly distributed over the wave front

Assume the target is placed at 3 m from a weak light source whose power is 1 watt

Atomic radius $\sim 10^{-10}$ m

Power fall on the target $\sim$ 1 watt $\frac{\pi(10^{-10}\,m)^2}{4\pi(3\,m)^2} \sim 28\times10^{-23}$ J/sec

Time required to absorbed 1 eV of energy

$$\tau = \frac{1.6\cdot10^{-19}\,J}{28\cdot10^{-23}\,J/sec} \sim 572\ sec$$

↓

there should be a time lag of this order between the impinging of light on the surface and the ejection of photoelectrons

↓

No such delay (time) was observed

1905 Einstein's quantum theory of the photoelectric effect (Nobel prize, 1921)

Einstein: radiation consists of a collection of quanta (photons) with $E = h\nu$ (Einstein relation) which are absorbed individually in photoelectric process

$$K.E = h\nu - W$$

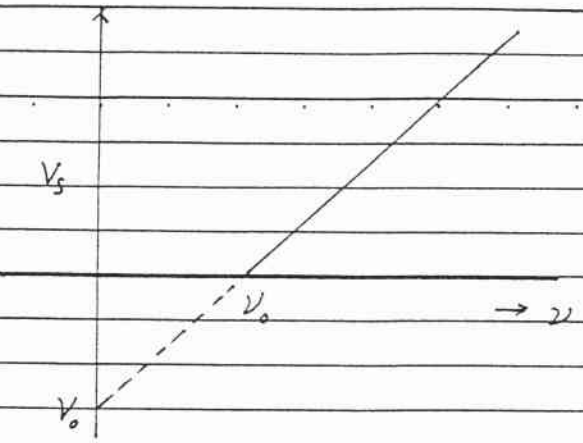
W: work required to remove the electrons from the metal

$$K.E_{max} = h\nu - W_0$$

W_0: minimum work required to remove the electron from the metal

$$K.E_{max} = eV_s$$

$$h\nu = eV_s + W_0 \quad (W_0 = eV_0) \quad \Rightarrow \quad V_s = \frac{h}{e}\nu - V_0$$



V_s vs ν plot should be given as above

(i) Intercept → ν_0
↳ threshold frequency

(ii) Extrapolation of the line to $\nu = 0$ axis ⇒ V_0 → work function

(iii) Slope of the line ⇒ $\frac{h}{e}$ ⇒ h
↓
agree with the value determined from black-body radiation.

These predictions were verified by Millikan in 1916 (Nobel prize, 1923)

A more intense source emits more photons and these photons in turn can liberate more photoelectrons ⇒ (ii)

Compton Effect (1923)

Monochromatic X-ray are scattered by a suitable scatter, the scattered radiation consists of two components
original wavelength λ_0
longer wavelength λ_1

$\lambda_1 - \lambda_0$ is a function of scattering angle only
independent of wavelength of the incident radiation, and the scattering material.

Classical picture: electron will oscillate under the influence of the electric field of the incident wave
⇒ will radiate a wave of the same wavelength (frequency)

$\gamma + "e^-" \rightarrow \gamma + e^-$

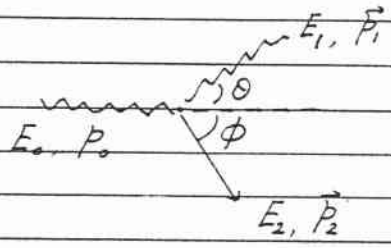
If the energy transfer to the electron is >> original binding energy ⇒ electrons can be treated approximately as "free electron"

Photon

$E = h\nu$

$p = \frac{E}{c} = \frac{h\nu}{c} = \frac{h}{\lambda}$

↳ since photon has zero rest mass



Momentum conservation

$P_0 = |\vec{P}_1| \cos\theta + |\vec{P}_2| \cos\phi$

$|\vec{P}_1| \sin\theta = |\vec{P}_2| \sin\phi$

$|\vec{P}_2|^2 = (P_0 - |\vec{P}_1|\cos\theta)^2 + |\vec{P}_1|^2 \sin^2\theta$

$= (P_0 - |\vec{P}_1|)^2 + 2P_0|\vec{P}_1|(1-\cos\theta)$ (1)

Energy conservation

$E_0 + mc^2 = E_1 + E_2$

$E_2 = E_0 - E_1 + mc^2$

$= c(P_0 - |\vec{P}_1|) + mc^2$

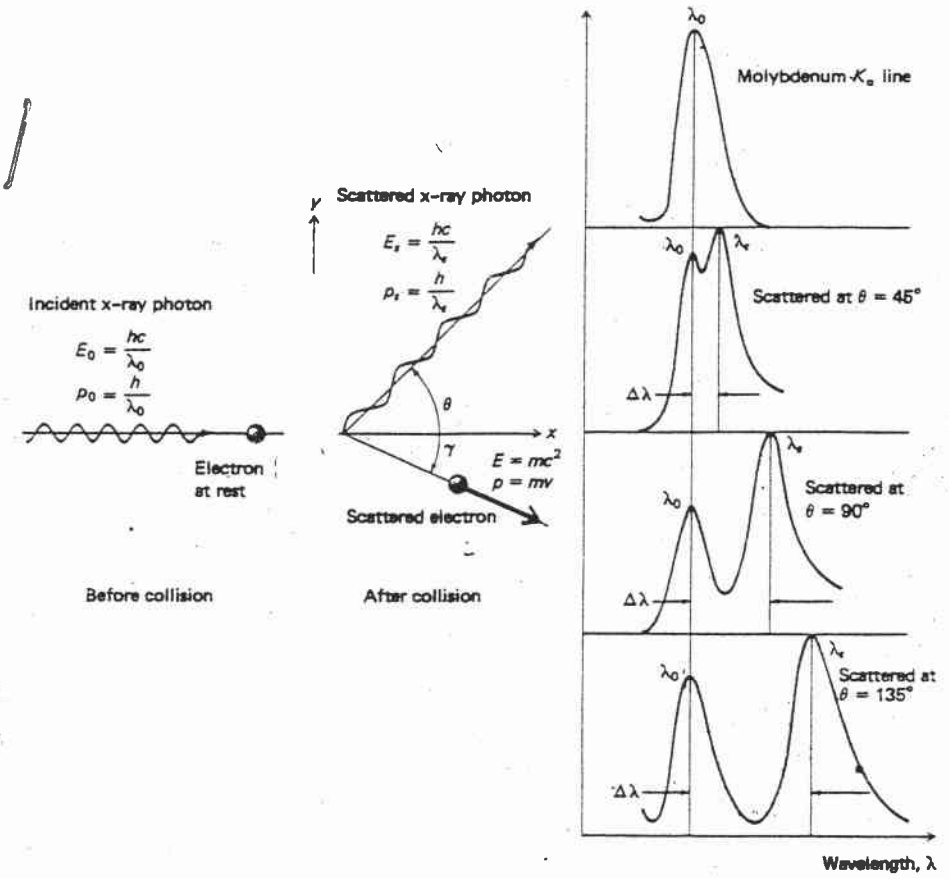
$|\vec{P}_2|^2 = \frac{1}{c^2}[E_2^2 - m^2c^4]$

$= \frac{1}{c^2}[c^2(P_0 - |\vec{P}_1|)^2 + 2mc^3(P_0 - |\vec{P}_1|) + m^2c^4 - m^2c^4]$

$= (P_0 - |\vec{P}_1|)^2 + 2mc(P_0 - |\vec{P}_1|)$ (2)

Combine (1), (2)

$\Rightarrow \quad mc(P_0 - |\vec{P}_1|) = P_0|\vec{P}_1|(1-\cos\theta)$



Compton scattering of a photon from an electron at rest. The graphs at the right show the shift in the K_α radiation from molybdenum scattered from carbon.

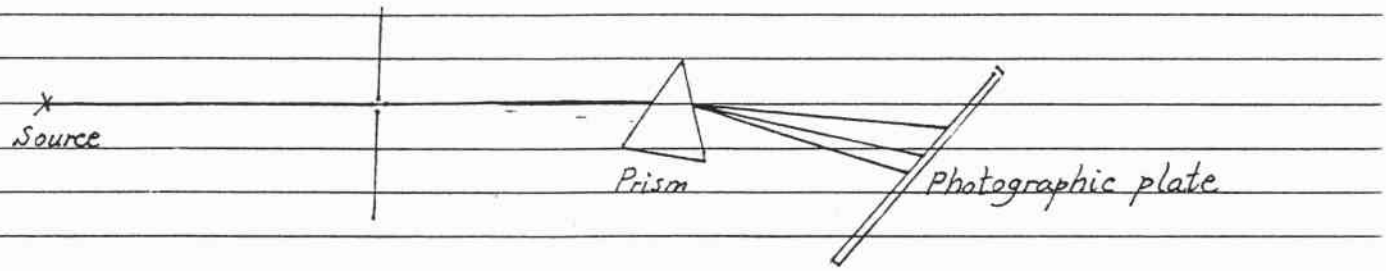
$$\frac{1}{|\vec{P}_1|} - \frac{1}{P_0} = \frac{1}{mc}(1-\cos\theta) \qquad \cdot h$$

$$\Rightarrow \lambda_1 - \lambda_0 = \frac{h}{mc}(1-\cos\theta) \leftarrow \text{Compton formula}$$

$\frac{h}{mc}$ ↓ $0.0243\ \text{Å}$ ↑ Compton wavelength of the electron

3. Bohr Model of the Atom

Atomic spectra



Source consists of electric discharge passing through a region containing a monatomic gas.
Collisions with electrons and with each other
⇒ some of the atoms become excited
From excited state → normal state
⇒ electromagnetic radiation

Radiation:
(i) collimated by slit
(ii) passes through a prism → break up into spectra
(iii) recorded on photographic plate

Results
(i) the electromagnetic radiations by free atoms are concentrated at discrete wavelengths
(ii) each discrete wavelength ↔ line
⇒ emission line spectra

Every element ↔ unique line spectra
Spectroscopy is a useful tool for analyzing the composition of unknown substances

Furthermore, wavelengths fall into definite set ⇒ spectral series
⇒ the wavelength in each series are specified by empirical formula

Hydrogen spectrum

1885 Balmer series

$$\lambda = 3646\ \text{Å}\ \frac{n^2}{n^2-4}$$

1890 Rydberg

$$\frac{1}{\lambda} = R_H\left(\frac{1}{2^2} - \frac{1}{n^2}\right) \qquad n = 3, 4, 5, \cdots \quad \text{Balmer series}$$

↳ Rydberg constant

$$\frac{1}{\lambda} = R_H\left(\frac{1}{1^2} - \frac{1}{n^2}\right) \qquad n = 2, 3, 4, \cdots \quad \text{Lyman series}$$

$$\frac{1}{\lambda} = R_H\left(\frac{1}{m^2} - \frac{1}{n^2}\right) \qquad n > m$$

Bohr model (1913)

(i) Atoms exist in "stationary states" of definite energy, in which states they do not radiate and are stable.

(ii) Atoms emit or absorb radiation only when atom goes from one stationary state to another

$$\Delta E = h\nu$$

(iii) Correspondence principle: quantum theory should give the same results as classical theory in the limit of large system
[Similar to relativity they at $\frac{v}{c} \to 0 \Rightarrow$ Newtonian theory]

Hydrogen atom case

Correspondence principle ⇒ in the limit of large system, where the allowed energies approach a continuum the quantum radiation condition must yield the same result as classical calculation

If one pictures an electron orbiting about a nucleus ⇒ for very large orbital radii, such that the atom has macroscopic size, the frequency of the radiation emitted by the hydrogen should be the same as frequency of revolution of the electron

$$\frac{1}{\lambda_{nm}} = R_H\left(\frac{1}{m^2} - \frac{1}{n^2}\right)$$

$h\frac{c}{\lambda} = hcR_H\left(\frac{1}{m^2} - \frac{1}{n^2}\right)$ $\quad n \to m \quad n > m$

$\| $

$h\nu_{nm}$

$\| $

$E_n - E_m$

$\Rightarrow \quad E_n = -\frac{hcR_H}{n^2}$

For orbiting electron

$E_n = -\frac{e^2}{8\pi\epsilon_0 r_n}$

$\| $

$-\frac{hcR_H}{n^2}$

$n \to \infty, \quad r_n \to \infty$

As $n \to \infty$, the seperation between energy levels become smaller and smaller

$\Rightarrow$ $n \to \infty$ correspond to the classical limit

Classical frequency of revolution

$$\frac{e^2}{4\pi\epsilon_0 r_n^2} = m\frac{v_n^2}{r_n}$$

$$\Rightarrow \quad v_n = \frac{e}{\sqrt{4\pi\epsilon_0 m r_n}} \qquad E_n = -\frac{e^2}{8\pi\epsilon_0 r_n}$$

$$\nu_n = \frac{v_n}{2\pi r_n}$$

$$= \frac{e}{\sqrt{4\pi\epsilon_0 m r_n}\cdot r_n}$$

$$= \frac{4\epsilon_0}{e^2}\sqrt{\frac{2}{m}}\,|E_n|^{3/2}$$

$$= \frac{4\epsilon_0}{e^2}\sqrt{\frac{2}{m}}\,\frac{(hcR_H)^{3/2}}{n^3} \qquad \left[\text{using } E_n = -\frac{hcR_H}{n^2}\right]$$

Quantum theory for large n

$$\nu_{if} = cR_H\left(\frac{1}{n_f^2} - \frac{1}{n_i^2}\right) \qquad n \to n-1$$

$$\to cR_H\left(\frac{1}{(n-1)^2} - \frac{1}{n^2}\right)$$

$$= cR_H\,\frac{n^2-(n-1)^2}{n^2(n-1)^2}$$

$$\xrightarrow[n\to\infty]{} cR_H\,\frac{2}{n^3}$$

Correspondence principle

$$\frac{4\epsilon_0}{e^2}\sqrt{\frac{2}{m}}\frac{(hcR_H)^{3/2}}{n^3} = cR_H\frac{2}{n^3} \quad \text{as } n\to\infty$$

$$\Rightarrow \quad R_H = \frac{me^4}{8\epsilon_0^2 h^3 c}$$

↓

Rydberg constant is determined from correspondence principle

$$\Rightarrow \quad E_n = -\frac{hcR_H}{n^2}$$

$$= -\frac{me^4}{8\epsilon_0^2 h^2}\frac{1}{n^2} = -13.6\,eV\cdot\frac{1}{n^2}$$

↓

allowed energy levels for hydrogen atom

$$E_n = -\frac{e^2}{8\pi\epsilon_0 r_n}$$

$$\Rightarrow \quad r_n = \frac{4\pi\epsilon_0\hbar^2}{me^2}n^2$$

$$v_n = \frac{e}{\sqrt{4\pi\epsilon_0 m r_n}}$$

$$= \frac{e^2}{4\pi\epsilon_0\hbar}\frac{1}{n} = \frac{1}{137}c\cdot\frac{1}{n}$$

Quantization of angular momentum

$$L_n = m r_n v_n$$

$$= m\,\frac{4\pi\epsilon_0\hbar^2}{me^2}n^2\cdot\frac{e^2}{4\pi\epsilon_0\hbar}\frac{1}{n}$$

$$= n\hbar$$

⇒ the orbital angular momentum of the electron is quantized, taking on only integral multiples of $\hbar$

↓

adopt this as the third postulate (replacing the correspondence principle)

With the circular orbit assumption

↓

Bohr's model of hydrogen described in mast textbook

分類：
編號： 3-20
總號：

The Frank-Hertz experiment gives a direct and striking demonstration of the discrete stationary state postulated by Bohr.
The Frank-Hertz experiment was first performed in 1914.

where, as a result of the periodic conditions, we have

$$k_x = \frac{2\pi n}{L}; \quad n = 0, \pm 1, \pm 2, \ldots$$
$$k_y = \frac{2\pi m}{L}; \quad m = 0, \pm 1, \pm 2, \ldots$$
$$k_z = \frac{2\pi l}{L}; \quad l = 0, \pm 1, \pm 2, \ldots \tag{1.25}$$

The wave solutions fill up the entire space but once again the k-vectors are quantized due to the boundary conditions.

The quantization aspect discussed above is a feature quite unique to waves. As we noted in the previous section, it does not occur when we discuss the classical behavior of particles. Another unique feature of waves is discussed in the next section.

1.6 WAVES, WAVEPACKETS, AND UNCERTAINTY

In classical physics when we deal with wave phenomena we are aware of a little "fuzziness" in the description of certain features of the wave. For example, let us imagine a child creating a wave on the surface of a pond by throwing a stone into the pond. We know from experience that a "wavepacket" or wave "pulse" intially localized around the point where the stone hit the water surface is produced. This wave then propagates toward the edges of the pond. Can we, at any time, *precisely* define the location of the wave and its wavelength? We know from experience and from classical physics that this is not possible. If we try to create a wavepacket highly localized in space we lose the knowledge of the wave's wavelength. On the other hand, if we try to create a "plane wave" with a well-defined wavelength, we lose the knowledge regarding the spatial position of the wave.

The uncertainty described above is not of concern in classical physics when we deal with particles. For example, we have no problem defining *precisely* any combinations of physical observables of a particle. However, a wave description will inevitably bring in an uncertainty in the precision with which we can simultaneously define certain physical observables. To see how this occurs, we examine the uncertainty arising in the wavelength or, for convenience, the wavevector k ($k = 2\pi/\lambda$), and the position of waves.

To describe the wavepacket, let us begin from a plane wave given by

$$\psi_k(x) = e^{ikx} \tag{1.26}$$

The position of the wave is completely undefined (as shown in Fig. 1.7a). To create a wavepacket localized at some point x_0 in space, we have to combine several plane waves. One example is to use an equal amplitude combination of waves centered around k_0 with a spread $\pm\Delta k$. The resulting function, say k_0, but from a spread $\pm\Delta k$, then the function

$$F(x, x_0) = \int_{k_0-\Delta k}^{k_0+\Delta k} dk\, e^{ik(x-x_0)}$$

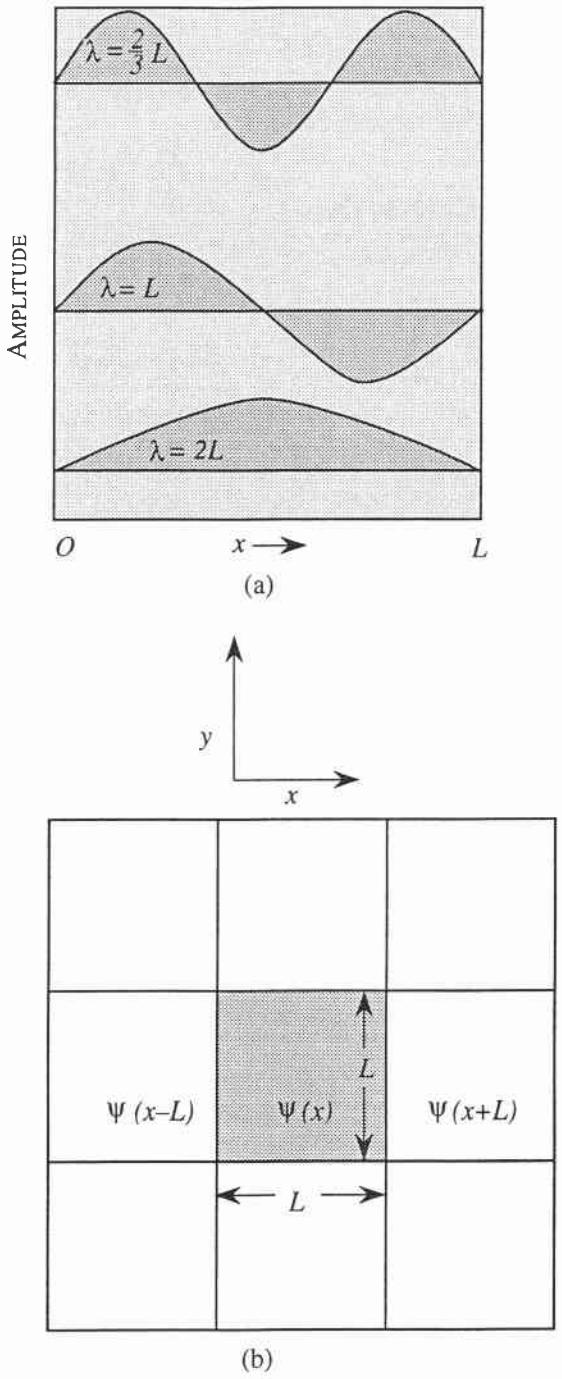


Figure 1.6: A schematic showing (a) the stationary boundary conditions applicable to a string clamped at $x = 0$ and $x = L$; and (b) periodic boundary conditions leading to plane wave solutions.

$$= \frac{2\sin(\Delta k\ (x - x_0))}{x - x_0}\ e^{ik_0(x-x_0)} \tag{1.27}$$

is centered around the point x_0 and the probability ($|F|^2$) decays from its maximum value at x_0 to a very small value within a distance $\pi/\Delta k$, as shown in Fig. 1.7b.

A more useful wavepacket is constructed by multiplying the integrand in the wavepacket by a Gaussian weighting factor:

$$f(k - k_0) = \exp\left[-\frac{(k - k_0)^2}{2(\Delta k)^2}\right] \tag{1.28}$$

and extending the range of integration from $-\infty$ to $+\infty$. This wave packet has the form

$$\begin{aligned}
\psi(x, x_0) &= \int_{-\infty}^{\infty} \exp\left[-\frac{(k - k_0)^2}{2(\Delta k)^2}\right] \exp\left[ik(x - x_0)\right]\ dk \\
&= \exp\left[ik_0(x - x_0) - \frac{(x - x_0)^2}{2}(\Delta k)^2\right] \\
&\quad \times \int_{-\infty}^{\infty} \exp\left[-\frac{(k - k_0)^2}{2(\Delta k)^2} + i(k - k_0)(x - x_0) + \frac{(x - x_0)^2}{2}(\Delta k)^2\right] \\
&= \sqrt{2\pi\Delta k}\ \exp\left[ik_0(x - x_0) - \frac{1}{2}(x - x_0)^2(\Delta k)^2\right]
\end{aligned} \tag{1.29}$$

$\psi(x, x_0)$ represents a Gaussian wavepacket in space which decays rapidly away from x_0. We note that when we considered the original state $\exp(ik_0x)$, the wave was spread infinitely in space, but has a precise k-value. By constructing a wavepacket, we sacrificed its precision in k-space by Δk and gained a precision Δx in real space. In general, the width of the wavepacket in real and k-space can be seen to have the relation

$$\Delta k\ \Delta x \approx 1 \tag{1.30}$$

This "uncertainty relation" which exists for waves is quite important. In the next chapter we will see that under some conditions, particles behave as waves. Similar uncertainty relations then exist in their properties. These relations are then called *Heisenberg uncertainty relations.*

1.6.1 Propagation of a Wavepacket

In classical wave propagation we are often interested in the question: How does a wave signal or pulse propagate in a medium? For example, we may create an optical pulse by switching a laser and this pulse may move down an optical fiber. We may be interested in the velocity at which the pulse moves and whether it distorts as it moves.

The simplest solution to wave equations have the form

$$\psi_{\mathbf{k}}(\mathbf{r}) \sim \mathbf{A}e^{\mathbf{ik}\cdot\mathbf{r}} \tag{1.31}$$

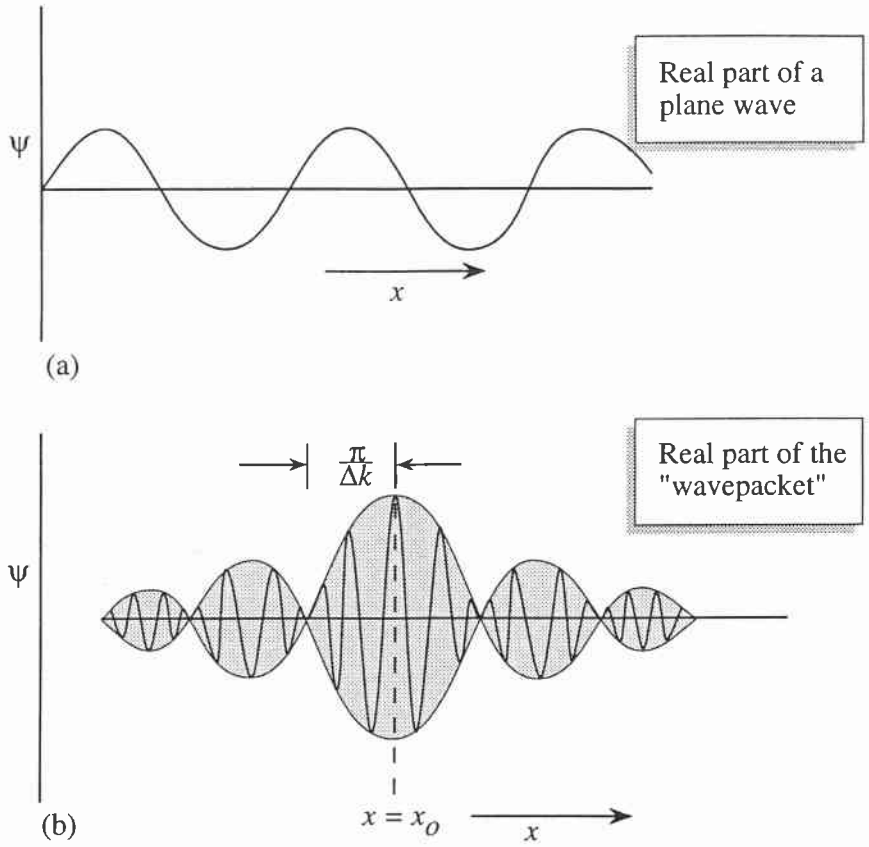


Figure 1.7: (a) A schematic description of a one-dimensional wave $e^{ik\cdot x}$ which is extended over all space; (b) a wavepacket produced by combining several waves produces a packet that is localized in space with a finite spread. The wavepacket is shown centered at x_o and having a spread Δx. The spread is such that $\Delta k \cdot \Delta x \sim 1$. This is an "uncertainty relation" in classical physics for waves. No such uncertainty exists in classical physics for particles.

In such a plane wave the probability density of the the wave, $\psi\psi^*$, is the same in all regions of space. Such a description is not useful if one wants to discuss transport of an optical pulse or of a particle from one point to another. For example, in describing electron transport we wish to describe an electron which moves from one point to another. Thus the wavefunction must be peaked at a particular place in space for such a description. This physical picture is realized by constructing a wavepacket picture.

Construction of a Wavepacket

Let us examine a one-dimensional plane wave state with a wave vector k_0

$$\psi_{k_0}(x) = e^{ik_0 x} \tag{1.32}$$

We note that if a state was constructed not from a single k_0 component, but from a spread $\pm\Delta k$, then the function

$$F(x, x_0) = \int_{k_0-\Delta k}^{k_0+\Delta k} dk\, e^{ik(x-x_0)} = \frac{2\sin(\Delta k\,(x-x_0))}{(x-x_0)}\, e^{ik_0(x-x_0)} \tag{1.33}$$

is centered around the point x_0 and the probability ($|F|^2$) decays from its maximum value at x_0 to a very small value within a distance $\pi/\Delta k$.

If Δk is small, this new "wavepacket" has essentially the same properties as ψ at k_0, but is localized in space and is thus very useful to describe motion of the particle. A more useful wavepacket is constructed by multiplying the integrand in the wavepacket by a Gaussian weighting factor

$$f(k-k_0) = \exp\left[-\frac{(k-k_0)^2}{2(\Delta k)^2}\right] \tag{1.34}$$

$$\begin{aligned}\psi(x, x_0) &= \int_{-\infty}^{\infty} \exp\left[-\frac{(k-k_0)^2}{2(\Delta k)^2} + ik(x-x_0)\right] dk \\ &= \exp\left[ik_0(x-x_0) - \frac{(x-x_0)^2}{2}(\Delta k)^2\right] \\ &\times \int_{-\infty}^{\infty} \exp\left[-\frac{(k-k_0)^2}{2(\Delta k)^2} + i(k-k_0)(x-x_0) + \frac{(x-x_0)^2}{2}(\Delta k)^2\right] \\ &= \sqrt{2\pi\Delta k}\, \exp\left[ik_0(x-x_0) - \frac{1}{2}(x-x_0)^2(\Delta k)^2\right]\end{aligned} \tag{1.35}$$

$\psi(x, x_0)$ represents a Gaussian wavepacket in space which decays rapidly away from x_0. We note that when we considered the original state $\exp(ik_0x)$, the wave was spread infinitely in space, but has a precise k-value. By constructing a wavepacket, we sacrificed its precision in k-space by Δk and gained a precision Δx in real space. In general, the width of the wavepacket in real and k-space can be seen to have the relation

$$\Delta k\ \Delta x \approx 1 \tag{1.36}$$

We can repeat this procedure for a wave of the form

$$\psi \sim e^{i\omega t} \tag{1.37}$$

and also obtain a wavepacket which is localized in time and frequency, the widths again being related by

$$\Delta\omega\ \Delta t \approx 1 \tag{1.38}$$

Let us now consider how a wavepacket moves through space and time. For this we need to bring in the time dependence of the wavefunction, i.e., the term $\exp(-iEt/\hbar)$ or $\exp(-i\omega t)$.

$$\psi(x, t) = \int_{-\infty}^{\infty} f(k-k_0)\ \exp\{i[k(x-x_0) - \omega t]\}\ dk \tag{1.39}$$

If ω has a simple dependence on k

$$\omega = ck \tag{1.40}$$

we can write

$$\psi(x,t) = \int_{-\infty}^{\infty} f(k-k_0)\ \exp\left[ik(x-x_0-ct)\right]\ dk \tag{1.41}$$

which means that the wavepacket simply moves with its center at

$$x - x_0 = ct \tag{1.42}$$

and its shape is unchanged with time. If, however, we have a dispersive media and the ω vs. k relation is more complex, we can, in general, write

$$\omega(k) = \omega(k_0) + \left.\frac{\partial\omega}{\partial k}\right|_{k=k_0} \cdot (k-k_0) + \frac{1}{2}\left.\frac{\partial^2\omega}{\partial k^2}\right|_{k=k_0} (k-k_0)^2 + \cdots \tag{1.43}$$

Setting

$$\begin{aligned} \omega(k_0) &= \omega_0 \\ \left.\frac{\partial\omega}{\partial k}\right|_{k=k_0} &= v_g \\ \left.\frac{\partial^2\omega}{\partial k^2}\right|_{k=k_0} &= \alpha \end{aligned} \tag{1.44}$$

we get

$$\begin{aligned} \psi(x,t) = \exp\left[i(k_0(x-x_0)-\omega_0 t)\right] &\int_{-\infty}^{\infty} f(k-k_0) \\ \times \exp\left[i(k-k_0)(x-x_0-v_g t) - \frac{i\alpha}{2}(k-k_0)^2 t\right]\ &dk \end{aligned} \tag{1.45}$$

If α were zero, the wavepacket would move with its peak centered at

$$x - x_0 = v_g t \tag{1.46}$$

i.e., with a velocity

$$v_g = \left.\frac{\partial\omega}{\partial k}\right|_{k=k_0} \tag{1.47}$$

However, for nonzero α, we show that the shape of the wavepacket also changes. To see this, let us again assume that

$$\begin{aligned} f(k-k_0) &= f(k') \\ &= \exp\left(\frac{-k'^2}{2\Delta k^2}\right) \end{aligned} \tag{1.48}$$

Then

$$\psi(x,t) = \exp\{i[k_0(x-x_0)-\omega_0 t]\} \times \int_{-\infty}^{\infty} \exp\left[ik'(x-x_0-v_g t) - \frac{k'^2}{2}\left(i\alpha t + \frac{1}{(\Delta k)^2}\right)\right] dk' \tag{1.49}$$

To evaluate this integral we complete the square in the integrand by adding and subtracting terms

$$\psi(x,t) = \exp\left\{i[k_0(x-x_0)-\omega_0 t] - \frac{(x-x_0-v_g t)^2(\Delta k)^2}{2[1+i\alpha t(\Delta k)^2]}\right\} \times \int_{-\infty}^{\infty} \exp\left\{\frac{-1}{2}\left[\frac{1+i\alpha t\,(\Delta k)^2}{(\Delta k)^2}\right] \times \left[k' - i\frac{(x-x_0-v_g t)\,(\Delta k)^2}{1+i\alpha t\,(\Delta k)^2}\right]^2\right\} dk' \tag{1.50}$$

The integral has a value

$$\sqrt{\frac{2\pi\,(\Delta k)^2}{1+i\alpha t\,(\Delta k)^2}}$$

Further multiplying and dividing the right-hand side exponent by $(1 - i\alpha t\,(\Delta k)^2)$ we get

$$\psi(x,t) = \exp\{i[k_0(x-x_0)-\omega_0 t]\}\sqrt{\frac{2\pi\,(\Delta k)^2}{1+i\alpha t\,(\Delta k)^2}} \times \exp\left[-\frac{(\Delta k)^2}{2}\frac{(x-x_0-v_g t)^2}{1+t^2\,(\Delta k)^4\alpha^2}\right] \times \exp\left[\frac{i\alpha t}{2}\frac{(\Delta k)^4\,(x-x_0-v_g t)^2}{1+t^2\alpha^2\,(\Delta k)^4}\right] \tag{1.51}$$

The probability $|\psi|^2$ has the dependence on space and time given by

$$|\psi(x,t)|^2 = \exp\left[-\frac{(\Delta k)^2\,(x-x_0-v_g t)^2}{1+\alpha^2 t^2\,(\Delta k)^4}\right] \tag{1.52}$$

This is a Gaussian distribution centered around $x = x_0 + v_g t$ and the mean width in real space is given by

$$\delta x = \frac{1}{\Delta k}\sqrt{1+\alpha^2 t^2\,(\Delta k)^4} = \delta x(t=0)\sqrt{1+\frac{\alpha^2 t^2}{[\delta x(t=0)]^4}} \tag{1.53}$$

For short times such that

$$\alpha^2 t^2 \, (\Delta k)^4 \ll 1 \tag{1.54}$$

the width does not change appreciably from its starting value, but as time passes, if $\alpha \neq 0$, the wavepacket will start spreading.

1.7 SYSTEMS WITH LARGE NUMBER OF PARTICLES

We have seen in Section 1.4 that the principles of mechanics given by Newton's equations are capable of describing the behavior of particles. What happens when the number of particles, all interacting with each other via collisions and mutual interactions, starts to increase? If the particle number is small, say less than a hundred or so, it is possible to use a powerful computer to find how each particle will behave in time and space. However, as the particle number increases, it becomes impossible to use Newtonian mechanics to describe how the system will behave. For example, if we were to examine the air in a room, we would find a mixture of oxygen, nitrogen, and carbon dioxide molecules. These molecules are bouncing off the wall and interacting with each other. The number of the particles and their densities are so large that it is simply not possible to apply the principles of mechanics to follow their trajectories. Fortunately, to describe the *measurable properties* of this and other such systems, we don't need to know the precise trajectories of the individual molecules. Such systems containing large number of particles are described by *statistical averages*. The measurable properties we are referring to are pressure, temperature, volume, etc.

Systems containing a large number of particles include: gases (air in a room, gases in a combustion engine, etc.); liquids (particles suspended in a liquid, chemical reagents, etc.); and solids (atoms and molecules in a piece of solid). Properties of these systems are described by the field of thermodynamics.

Let us consider the molecules in the air inside a room. While we cannot describe the individual trajectories of the molecules, we can ask and answer the following questions:

(i) What is the relation between the pressure of the gas and its volume? How is this relation dependent on the density of molecules?

(ii) What is the average kinetic energy of the molecules?

(iii) What is the probability that the molecules have an energy E?

(iv) Is there a difference between the average energy of oxygen molecules and nitrogen molecules?

The field of thermodynamics gives us answers to such questions. We will now state the important concepts of thermodynamics. The essence of thermodynamics is contained in the laws of thermodynamics outlined in Fig. 1.8. We will briefly review some important issues.

Thermal Equilibrium

Let us examine a system consisting of a large number of particles. If we were to observe the system on a microscopic level, we will see a lot of activity going on. Molecules are moving helter-skelter, sometimes suffering a collision and changing their

directions and speeds. However, under thermodynamic equilibrium the macroscopic properties of the system will stay constant in time. There will be no net transfer of energy between the system under observation and the rest of the universe.

An important outcome of thermodynamics is that if two systems are in thermal equilibrium with a third system, then they are also in equilibrium with each other. This is known as the zeroeth law of thermodynamics.

If we examine a system on a microscopic level, each particle can be assigned certain *degrees of freedom*. For example, an electron moving in space has three degrees of freedom since it can move in the x, y, and z directions. If we have a molecule with r atoms, there are $3r$ degrees of freedom of which 3 correspond to the motion of the molecule (center of mass motion) and $3(r-1)$ correspond to the internal motion (vibration and rotation).

Consider a system at thermal equilibrium containing a collection of different species of masses m_1, m_2, m_3, and so on. Under equilibrium we have the following equality:

$$\left\langle \frac{m_1 v_1^2}{2} \right\rangle = \left\langle \frac{m_2 v_2^2}{2} \right\rangle = \left\langle \frac{m_3 v_3^2}{2} \right\rangle \tag{1.55}$$

i.e., the average kinetic energy of the species is equal. Thus the mean kinetic energy is not a function of the particle masses but is a property of the system. *This allows us to use the mean kinetic energy of a particle in a system with a large number of particles to define the temperature of the material.*

The mean kinetic energy per degree of freedom k has a value (see Example 1.7)

$$\langle E_k \rangle = \frac{1}{2} k_B T$$

In 3-dimensional space we have

$$\langle E \rangle = \frac{3}{2} k_B T \tag{1.56}$$

The definition of temperature given here creates what is known as the absolute temperature and T is measured in the units of Kelvin (K).

Internal Energy, Free Energy, and Entropy

An important question that thermodynamics answers for us is the following: If we have a system in thermodynamic equilibrium at a temperature T, what is the probability that a particle has an energy E? The equilibrium state is given by the state in which the free energy of the system is minimum. The free energy is different from the internal energy which is simply given by

$$U = \Sigma_i E_i \tag{1.57}$$

where E_i is the energy of the i^{th} particle. The difference between the free energy and the internal energy arises from the quantity known as entropy of the system. A thermodynamic system is described not only by its internal energy but also by another property related to the heat contained in the system and the temperature.

A macroscopic definition of the entropy S is

$$S = \frac{Q}{T} \tag{1.58}$$

where Q is the heat contained in the system and T is the temperature. The entropy of the system is zero at $T = 0$. (This is the third law of thermodynamics.)

There is another definition of entropy that is very useful in defining the distribution function, i.e., the function that gives us the probability of finding a particle at an energy E. The entropy is defined as

$$S = k_B ln W \tag{1.59}$$

where W is the degeneracy of the system, i.e., the number of different ways in which particles can be arranged in the system to create the same total energy.

The free energy F of a system is given by

$$F = U - TS \tag{1.60}$$

As noted above, the equilibrium state is that where the free energy of the system is a minimum. Note that if we simply minimize the internal energy of the system, all the particles would occupy the lowest energy of the system. In this case, the degeneracy of the system is just unity and $S = 0$. At finite temperature there is a competition between the internal energy and entropy. The entropy term $-TS$, decreases as the particles are arranged in higher energy states where the degeneracy is large, but the internal energy term increases. At any given temperature one has to find the minima of F to find the actual distribution of particles. We will carry out this exercise in Chapter 3.

An important manifestation of the classical thermodynamics is the distribution function, which is called the Maxwell-Boltzmann distribution. Consider an ideal gas consisting of non-interacting particles of density N, mass m at equilibrium at a temperature T. The distribution of these particles in energy E or speed w is given by the following expressions:

$$\frac{dN}{dw} = \frac{4N}{\sqrt{\pi}} \left(\frac{m}{2k_B T}\right)^{3/2} w^2 \exp\left(-\frac{1}{2}\frac{mw^2}{k_B T}\right) \tag{1.61}$$

$$\frac{dN}{dE} = \frac{2N}{\sqrt{\pi}} E^{1/2} (k_B T)^{-3/2} \exp\left(-\frac{E}{k_B T}\right) \tag{1.62}$$

Fig. 1.9 shows the distribution of particles as a function of speed at various temperatures. As the temperature increases, the average particle speed increases, as one expects intuitively.

EXAMPLE 1.7 Show that the average of the square of the molecular speeds in an ideal gas is $3k_B T/m$.

The average of the square of the speeds is given by

$$\langle \omega^2 \rangle = \frac{4}{\sqrt{\pi}} \left(\frac{m}{2k_B T}\right)^{3/2} \int_0^\infty \omega^4 \exp\left(-\frac{1}{2}\frac{m\omega^2}{k_B T}\right) d\omega$$

LAWS OF THERMODYNAMICS

ZEROTH LAW: Two systems in thermal equilibrium with a third system are in thermal equilibrium with each other.

FIRST LAW: Represents the law of conservation of energy $\Longrightarrow$ Heat put into a system, dQ + Work done on a system, dW = Increase in internal energy of the system, dU:

$$dQ + dW = dU$$

SECOND LAW: A process whose only effect on all systems is to take heat from a reservoir and convert it to work is impossible $\Longrightarrow$ This law in its mathematical form places limits on the most efficient heat engine, i.e., it says that no heat engine taking heat θ_1 from a system at temperature T_1 and delivering heat θ_2 to a system at T_2 can do more work than a reversible engine for which

$$W = \theta_1 - \theta_2 = \theta_1 \left(\frac{T_1 - T_2}{T_1} \right)$$

The second law is also stated in terms of entropy: The entropy of the universe is always increasing. In a reversible process the entropy is unchanged.

THIRD LAW: The entropy of a system is zero at $T = 0$

Figure 1.8: Laws of thermodynamics.

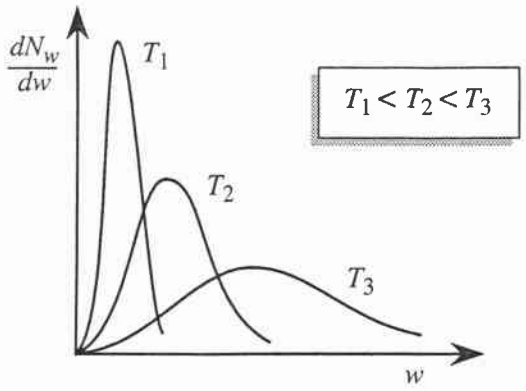


Figure 1.9: Distribution of particle speeds for different temperatures.

The integral has a value $\frac{3}{8}\sqrt{\pi}(2k_BT/m)^{5/2}$. Thus

$$< \omega^2 >= \frac{3k_BT}{m}$$

This is, of course, consistent with the observations that the average energy is

$$< E >= \frac{1}{2}m < \omega^2 >= \frac{3}{2}k_BT$$

1.8 CHAPTER SUMMARY

Summary table 1.1 covers key findings and topics studied in this chapter.

1.9 PROBLEMS

Problem 1.1 A child is swinging a 100 gm mass attached to a 1 m string at a rate of two swings per second. Express the angular momentum of the mass in units of $\hbar = 1.05 \times 10^{-34}$ J.s. How small does the string have to be for the angular momentum to be 100 $\hbar$?

Problem 1.2 Consider a pendulum made from a string of length 1 m and mass 100 gm. What is the amplitude of the pendulum (displacement from equilibrium) if the energy of the pendulum is to be 10 $\hbar\omega$ where $\hbar = 1.05 \times 10^{-34}$ J.s and ω is the angular frequency of the pendulum.

Problem 1.3 Calculate the speed of a satellite so that it orbits the earth. How does this speed compare to the speed at which a rocket must be fired in order to leave the earth completely?

Problem 1.4 The separation of the nuclei in a silicon crystal is 2.35 Å. Calculate the gravitational potential energy due to the attraction between the nuclei. Assume that a Si atom is surrounded by four neighbors.

Problem 1.5 In Bohr's model of the hydrogen atom an electron moves in a circular orbit of radius 0.53 Å with an angular momentum of 1.05×10^{-34} J.s. Calculate the

TOPICS STUDIED	KEY OBSERVATIONS
Scientific method	The scientific method depends upon observation, theoretical models, and predictions. Experimental evidence is the final arbiter of all theories.
Classical physics	Classical physics consists of: • Newtonian mechanics for particles • Wave equations for waves • Statistical thermodynamics for collection of large number of particles
Newtonian mechanics	Newtonian mechanics depends upon Newton's equations of motion. The position-momentum trajectories of particles can be exactly predicted if the forces are known.
Wave theory	• Wave equations describe properties of waves in media. • Maxwell equations describe the most important wave phenomena—that of electromagnetic radiation. • All waves have an uncertainty relation between their wavevector and position.
Classical thermodynamics	Properties of systems with very large number of particles are described by thermodynamics. The laws of thermodynamics describe the properties of such systems. Thermodynamic equilibrium is determined by the minimization of the free energy which is equal to the total internal energy—entropy times temperature.

Table 1.1: Summary table

speed of the electron.

Problem 1.6 The optical power density impinging upon a detector has a value of 10^{-6} mW/cm^2. Calculate the electric field amplitude associated with this power.

Problem 1.7 A detector is designed to detect a minimum electric field of 2.0 mV/cm. Calculate the minimum power density level this detector can detect.

Problem 1.8 A microwave oven is designed to produce a maximum electric field (rms) of 1 kV/cm. Calculate the maximum electromagnetic power density produced by this oven.

Problem 1.9 A typical silicon MOSFET "breaks down" when the electric field reaches 2×10^5 V/cm. Calculate the optical power density needed to cause breakdown. Assume that the field needed for breakdown is the rms field of the radiation.

Problem 1.10 A calibrated tuning fork is used to determine the frequency of an instrument by observing the beat frequencies. How long will the observation time have to be to have an accuracy of 0.01 Hz?

Problem 1.11 Consider a diffraction grating with n lines. Show that the resolving power for an m^{th} order beam is

$$\frac{\Delta\lambda}{\lambda} = \frac{1}{mn}$$

Problem 1.12 A gas cylinder contains n molecules of mass 4.7×10^{-26} kg. The temperature is changed from 273 K to 200 K. Calculate the change in the average kinetic energy of the molecules. Also calculate how much the height of the cylinder will have to be altered with respect to the earth's surface to produce the same change in potential energy.

Problem 1.13 Estimate the average speeds of oxygen, nitrogen, and carbon dioxide molecules in air at room temperature.

Problem 1.14 Calculate the probability of the density of atoms of a species in air as a function of height from earth's surface.

Problem 1.15 Calculate the ratio of the density of oxygen molecules at earth's surface to their density at 40 km from the surface. Repeat this problem for hydrogen molecules.

CHAPTER 2

QUANTUM MECHANICS AND THE UNIVERSE

CHAPTER AT A GLANCE