

Understanding Physical Chemistry

Dor Ben-Amotz

Purdue University
West Lafayette, IN

January 9, 2008

Chapter 1

The Basic Ideas

This chapter contains a summary of some of the most important ideas that underly all of physical chemistry. In other words, it could be subtitled *Ingredients in a Physical Chemists' Cookbook* or *Tools in a Physical Chemists' Workshop*. These ideas are ones that physical chemists frequently refer to when they are having conversations with each other. So, you could think of this chapter as a Quick-Start guide to thinking, talking and walking like a physical chemist. Having these basic ideas in mind can help make physical chemistry less confusing by providing a broad overview of how various pieces of nature's puzzle fit together to produce a beautiful big picture.

1.1 Things to Keep in Mind

Physical Chemistry is a Conversation

Science is sometimes incorrectly envisioned as a static and impersonal body of knowledge – in fact it is much more like an interesting conversation which evolves in endlessly surprising ways. This multi-faceted conversation often takes place between good friends, over lunch or coffee (or some other beverage), or while taking a break in the lab, or during a walk in the woods. It also often includes people who live in very different places (and times), *via* email, over the phone, at scientific meetings, or in journal articles, both in the latest issues and in archives extending back many years, spanning centuries, and drawing on memories that reach deeply into the foggy depths of recorded history, and beyond.

A classroom is one of the main places in which scientific conversations happen. A classroom, of some kind or other, is where every single scientist throughout history has come to find out more about the most interesting discussions and realizations that other scientists have had. The best classroom experiences are themselves conversations in which students and teachers struggle to improve their individual and collective understandings by working hard to clearly communicate and think in new ways.

Like any good conversation, scientific progress requires a free exchange of ideas and an open-minded attitude. Obviously, having a conversation also requires speaking the same language and sharing a common body of knowledge and experience. However, the preconceptions that inevitably come along with any body of knowledge can also be among the greatest impediments to scientific progress, or, for that matter, any other kind of productive exchange of ideas. So, the feeling of confusion or disorientation that may at times overtake you while you are struggling to learn physical chemistry is not necessarily a bad thing – that is often how it feels when an interesting conversation is on the verge of a breakthrough.

Longing for Equilibrium

All changes in the world appear to be driven by an irresistible longing for equilibrium. Although this longing is not the same as a subjective feeling of longing, the effect can be much the same. Any change in the world clearly implies the existence of an underlying driving force. Moreover, our experience suggests that some changes can and do often occur spontaneously while others are highly improbable or even impossible. These ideas are best illustrated by some simple examples.

Consider a boulder situated very comfortably up on the side of a mountain. Although this boulder may remain in more-or-less the same spot for many years, if the ground holding it gives way, the boulder will spontaneously careen down into the valley below – dramatically converting its potential energy into a great burst of kinetic energy. However, our experience also tells us that under no circumstances would the boulder ever spontaneously roll uphill, unless significant work were expended to push it.

This same tendency is also obviously responsible for the fact that rivers invariably flow downhill, rather than uphill. Understanding this universal tendency can be of great practical value. For example, one can build a waterwheel or a hydroelectric generator in order to use the tendency of water to

flow downhill in order to perform useful work, such as mechanically grinding wheat into flour or generating horsepower in the form of electricity. As another example, the tendency of electrons to flow downhill in potential energy from one chemical compound to another may be used to produce batteries and fuel cells, as well as to flex muscles and create brain storms.

The sun is another good example of the importance of disequilibrium. Hydrogen atoms are just like boulders sitting high up on a hill, where they can remain in a very stable state for many years. However, given the right circumstances (such as the very high pressures and temperatures inside the sun) hydrogen atoms can be dislodged to undergo fusion reactions such as, $4\text{H} \rightarrow \text{He} + 2e^- + 2e^+$, releasing a great bust of energy in the form of sunlight.¹

The longing for equilibrium is of keen interest not only to physical chemists but also to engineers and mathematicians – whose research expenses are often subsidized by investors anxious to capitalize on nature's tendencies. Although the above examples make it obvious how some spontaneous processes may be converted to useful work, the general analysis of nature's proclivity for equilibration is a deep and complex subject which motivated the development of thermodynamics.

Among the most remarkable results of thermodynamics is the discovery of a function, called entropy, which expresses the longing of all systems for equilibrium in rigorous mathematical terms. This function may be used to predict whether a given process can or cannot occur spontaneously. Even more importantly, entropy can be used to predict the maximum amount of work which can be obtained from any spontaneous process (or conversely, the work required to drive a non-spontaneous process). We will return to revisit these issues more closely in Chapter 2.

Invariants, Constraints and Symmetry

A recurring theme underlying all of physical chemistry (and other branches of science) is the search for universal principles, or fundamental quantities, which give rise to all observed phenomena. The search for such invariant

¹Note that the reaction of four H (${}^1_1\text{H}$) atoms to form a He (${}^4_2\text{He}$) atom makes use of the fact that a proton may decompose into a neutron plus a positron. There are also other lower order reactions which can produce helium from heavy isotopes of hydrogen (deuterium and tritium), such as ${}^2_1\text{H} + {}^3_1\text{H} \rightarrow {}^4_2\text{He} + {}^1_0\text{n}$, which is among the processes that may some day form the basis of environmentally safe nuclear fusion power plants on earth.

properties of nature has ancient roots, tracing back at least to the Ionian school of Greek philosophy, which thrived in the 6th century BC, and whose adherents, including Thales and Anaximander, postulated that all things are composed of a single elemental substance. This school of thought also influenced a young Ionian named Democretus, who proposed that everything in the world is composed of atoms which are too small to be visible with the naked eye.

The idea that some quantities are conserved in the course of chemical processes seems pretty obvious. For example, although a chemical reaction may produce dramatical changes in color, texture and other measurable properties, one would naturally expect the products of a reaction to weigh the same amount as the reactants. Careful experimental measurements demonstrate that mass is indeed conserved during chemical reactions, to within the accuracy of a typical analytical balance. However, it turns out that mass is not in fact perfectly conserved! This failure proves to be linked to a deeper principle of energy conservation, as we will see.

The invariant properties of a system are also intimately linked to the constraints and symmetries which characterize the system. These deep interconnections underly some of the most amazing scientific discoveries. Because these ideas are so profound, they are often reserved for more advanced (graduate level) courses in chemistry, physics and mathematics. However, there is no harm in learning something about them, even before we fully understand where they come from. The truth is not always easy to fathom, but it is always worth the effort.

A good example of the connection between invariants, constraints and symmetry emerges from considering the motion of an object in a central force field – such as the earth moving in the central gravitational force field of the sun, or an electron moving the central coulombic (electrostatic) force field of a proton. In the 17th century Johannes Kepler demonstrated that a planet which is constrained to move under the influence of the sun's gravitational force must sweep out a constant (invariant) area per unit time. This is a special case of the principle of conservation of angular momentum, which also holds for an electron in a quantum mechanical orbital around an atomic nucleus.

The conservation of angular momentum is a necessary consequence of the spherical symmetry of a central-force constraint. Similar connections between invariants, constraints and symmetry underly the conservation of linear momentum in a system with translational invariance, such as objects moving

in free space, or billiard balls rolling on a pool table. Thinking about these connections led Einstein to develop the special theory of relativity, which is a consequence of the experimentally observed invariance of the speed of light, independent of the relative velocity of an observer (and the corresponding invariance of Maxwell's electromagnetic equations). The theory of relativity leads to all sorts of surprising predictions about the interrelations between light, space, time, energy and mass. Among these is a prediction that mass cannot be perfectly conserved in any chemical reaction which either releases or absorbs energy (as further discussed in section 1.2).

Constraints and symmetries also play an important role in thermodynamics. For example, as first clearly demonstrated by Joseph Black, an 18th century Scottish professor of medicine and chemistry, two chemical systems which are constrained in such a way that they cannot physically mix but can exchange heat (e.g. because they are separated by partition made from copper or some other good thermal conductor), invariably evolve to an equilibrium state of the same temperature. Similarly, two systems which are separated by a constraint which can freely translate, such as a movable piston, will evolve to an equilibrium state of the same pressure. As another example, if we remove a constraint that separated two different kinds of gases (i.e. by opening a stopcock between the containers that hold the two gases) then they will evolve to a state in which the concentration of each gas is the same everywhere.

A common theme underlying the above examples is that systems tend to evolve toward a state of maximum symmetry, to the extent allowed by the constraints imposed on the system. Note that in the gas mixing example we implicitly assumed that the two gases don't react with each other and that there is no difference in potential energy between one part of the system and another. However, even when molecules can react or when there are any sort of complicated potential energy differences within the system, Willard Gibb's brilliantly demonstrated that one can nevertheless identify a quantity called the chemical potential whose invariance (i.e. uniformity throughout the system) is assured at equilibrium (as we will see in Chapter 2).

1.2 Why is Energy so Important?

Conservation of Energy

The principle of energy conservation, which is closely related to the first law of thermodynamics, identifies energy as the one quantity which is invariably conserved during any process, chemical or otherwise. This principle also leads to the recognition of different forms of energy, including kinetic and potential energy, as well as work and heat, which represent means of exchanging energy between a system and its surroundings.

The connection between kinetic, K , and potential, V , energies may be illustrated by considering an apple falling off of a tree. If the apple of mass, m , is initially hanging at a height, h , then its potential energy (relative to the ground) is $V = mgh$ (where $g = 9.8 \text{ m/s}^2$ is the acceleration due to gravity). Once the apple hits the ground all of its potential energy will be converted to kinetic energy, $K = \frac{1}{2}mv^2 = mgh = V$ (where v is the velocity of the apple). Similar expressions could be used to obtain the increase in kinetic energy of any object which results when it freely falls through a given potential energy drop, or to calculate how high up an object could go if it is launched with a particular initial value of kinetic energy. In the next section we will see why it is that these two kinds of energy have the functional forms that they do.

Kinetic and potential energies can also be related to work, W , which is defined as the product of the force, F , experienced by an object times the distance, x over which that force is imposed. Thus, the work associated with an infinitesimal displacement is

$$dW = Fdx \quad (1.1)$$

and so the total work exchanged during a given process is

$$W = \int dW = \int Fdx \quad (1.2)$$

where both integrals are performed from the starting point to the end point of the path of interest. For example, if a constant force is used to accelerate a object over a distance Δx then a total work of $W = F\Delta x$ will be performed. Moreover, the work that is done will produce an increase in kinetic and/or potential energy which is exactly equal to W . We can also use eq 1.2 to calculate the work associated with many other types of processes, such as

compressing a gas, or moving an electron through a voltage gradient, or breaking a chemical bond.

The way in which heat is related to other forms of energy is a subtle and interesting issue. From a macroscopic perspective the heat exchanged between a system and its surroundings may be *defined* as any change in the energy of the system, other than that due to the performance of work on the system (by the surroundings). From a molecular perspective one may identify heat exchange with changes in the kinetic and/or potential energy of molecules (as opposed to macroscopic objects). The flow of energy from the macroscopic to the molecular scale is an interesting subject with practically important ramifications.

The irreversible loss of useful macroscopically organized energy into less useful random molecular energy is intimately linked to the concept of entropy. Given that this concept has required centuries to develop, we should not be surprised if it takes us some time and effort to fully apprehend its significance and implications. One of the primary aims of physical chemistry is to attain such an understanding by revisiting this and the other key ideas from various different perspectives. Just as the different perspectives provided by our two eyes are required to produce a three dimensional image of the world, so too are different perspectives required in order to better visualize the world of physical chemistry.

The Hamiltonian

The significance of energy conservation may be further illuminated by considering the interactions of particles moving on a flat potential surface (such as billiard balls colliding on a pool table) or objects which move under the influence of external forces (such as those produced by magnetic, electric or gravitational fields). In the late 17th Century, Isaac Newton formulated his famous second law, which applies to all such processes

$$F = ma \tag{1.3}$$

where F is the force acting on an object of mass, m , and $a = \frac{dv}{dt}$ is the acceleration it experiences as a result.

The force on an object may also be related to the slope of a potential energy function. For example, consider a car parked on a hill. If you release the brakes, then the car will tend to accelerated down the hill. The force

which produces this motion is proportional to the slope of the hill. More specifically, if some object is moving along the x -direction under the influence of a potential energy function, $V(x)$, then the force it experiences is

$$F = - \left[\frac{dV(x)}{dx} \right]. \quad (1.4)$$

The minus sign simply indicates that a potential function (hill) which goes up when you move forward will exert of force which pushes you back (down the hill), as illustrate in Fig. 1.1.

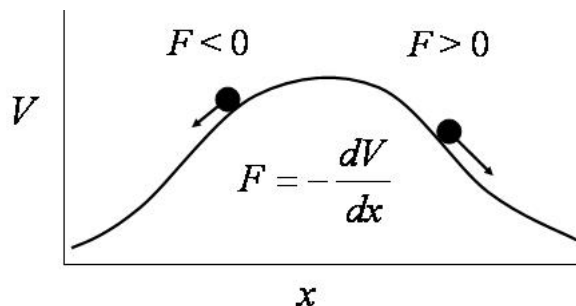


Figure 1.1: A ball on a hill feels a force that is opposite in sign to the slope of the hill. In other words, when the slope is positive, the ball is pushed backwards, while when the slope is negative the ball is pushed forward.

The total energy of any system is defined as the sum of its kinetic, K , and potential, V , energies, and is also called the Hamiltonian function, H .

$$H = K + V = \text{Total Energy} \quad (1.5)$$

This function is named after William Rowan Hamilton, a leading 19th century physicists.²

²Hamilton was born in Ireland in 1805, and demonstrated an early brilliance by learning more than ten languages by the age of 12. His interest in mathematics apparently began around that same time, when he met an American named Zerah Colburn who could mentally calculate the solutions of equations involving large numbers. Soon after than he began avidly reading all the mathematical physics books he could get his hands on, included Newton's *Principia* and Laplace's *Celestial Mechanics*, in which he uncovered a key error. He published a brilliant and highly influential paper on optics while he was still an undergraduate, and was appointed a professor of Astronomy at Trinity College at the age of 21.

The usefulness of the Hamiltonian function can be illustrated by considering a particle moving in the x -direction under the influence of a potential energy function, $V(x)$. The kinetic energy of the particle is $K = \frac{1}{2}mv^2$, and so the Hamiltonian of such a system is

$$H = \frac{1}{2}mv^2 + V(x). \quad (1.6)$$

If we take the time-derivative of both sides we discover a very interesting property of the Hamiltonian function.

$$\begin{aligned} \frac{dH}{dt} &= \frac{d}{dt} \left[\frac{1}{2}mv^2 + V(x) \right] \\ &= \frac{1}{2}m \left[2v \left(\frac{dv}{dt} \right) \right] + \left[\frac{dV(x)}{dt} \right] \\ &= mv \left(\frac{dv}{dt} \right) + \left[\frac{dV(x)}{dx} \right] \left(\frac{dx}{dt} \right) \\ &= mva - Fv \\ &= (ma - F)v = (0)v \\ &= 0 \end{aligned}$$

This result clearly implies that H is time-independent, and so the total energy of the system is conserved! In other words, we have shown that Newton's law implies the conservation of energy. However, while Newton formulated classical mechanics in terms of forces which may have a complicated time-dependence, the Hamiltonian formulation of classical mechanics is founded on a time-independent (conserved) property – the total energy.³ Notice that the above derivation also demonstrates why we define $K \equiv \frac{1}{2}mv^2$, as this is the quantity which combines with potential energy to form a Hamiltonian that is time-independent.

³Hamilton also demonstrated that all of classical mechanics could be obtained from what is now called Hamilton's principle, which states that $\delta \int (K - V)dt = 0$. In other words, he demonstrated that the path followed by any mechanical system is one which minimizes the time integral of the difference between its kinetic and potential energies. This principle is closely related to Fermat's principle of least time, which applies to the path followed by light in a medium of varying refractive index. Hamilton's principle also played a central role in both Schrodinger's and Feynman's 20th century contributions to the development of quantum mechanics.

Although the above derivation only considered a single particle moving in the x -direction, the result can be generalized to show that the Hamiltonian of any *isolated* system, no matter how complicated, must also be time-independent. Note that an isolated system is defined as one from which nothing can leave (or enter). Thus, the entire universe is one example of an isolated system, which implies that the energy of the universe must be conserved!

Our experience tells us that the energy of some sub-systems within the universe may *not* be conserved. For example, a car dissipates energy when it drives, and so it is clearly not an isolated mechanical system. This is also why a car is valuable, because it can use chemical energy to drive up hills and speed along a highway for many miles at a steady clip, in spite of frictional drag and wind resistance.

One of the simplest examples of a non-isolated system is an object that experiences a frictional force which is inversely proportional to its velocity, $F_{\text{friction}} = -fv$. Thus, the total force on such an object can be expressed as the sum of this frictional force plus the force arising from its potential energy.

$$F = ma = -\left(\frac{dV}{dx}\right) - fv \quad (1.7)$$

Notice that this equation can also be obtained by equating the time-derivative of H with $-fv^2$.

$$\begin{aligned} \frac{dH}{dt} &= \frac{d}{dt} \left[\frac{1}{2}mv^2 + V(x) \right] = -fv^2 \\ \left[ma + \left(\frac{dV}{dx} \right) \right] v &= -(fv)v \\ ma + \left(\frac{dV}{dx} \right) &= -fv \\ ma &= -\left(\frac{dV}{dx} \right) - fv \end{aligned}$$

This indicates that the Hamiltonian of such a system is *not* constant, $\frac{dH}{dt} = -fv^2$. In other words, a frictional force has the effect of dissipating the total energy of the system at a rate of $-fv^2$. But where does this energy go? Our experience tells us that friction is often accompanied by an increase in temperature, such as that which you can feel when you rub your hand

together rapidly. So, heat must be related in some simple way to dissipative energy loss, as we will further explore in Chapter 2.

In summary, the Hamiltonian (total energy) of any isolated system is time-independent (conserved), while that of a non-isolated system may not be. However, since the entire universe (which includes the system and all of its surroundings) is itself isolated, the universe must have a fixed amount of total energy. Thus, any energy which leaves a system is not lost but simply goes into some other part of the universe. We will further investigate such energy exchange processes in Chapter 2, and will also re-encounter the Hamiltonian when we investigate study the internal energy-level structures of atoms and molecules in Section 1.4 (and then in greater detail in Chapter 3).

Relation Between Energy and Mass

A remarkable extension of the principle of energy conservation was discovered in the early 20th century by Albert Einstein, whose theory of relativity made it clear that there is an intimate connection between the conservation of energy and mass. Einstein first reported this monumental finding in a short note entitled *Does the Inertia of a Body Depend on its Energy Content?* which he published in 1905 as an afterthought to his famous first paper about relativity. In this note he analyzed the implications of relativity when applied to processes involving the emission of light by atoms. This analysis suggested that the measured mass of an atom must decrease when it loses energy. Thus, Einstein obtained what may well be the most famous equation in all of science.

$$E = mc^2 \tag{1.8}$$

This states that mass, m , and energy, E , are not independent variables, but are related to each other by a constant of proportionality that is equal to the square of the velocity of light, c^2 . Einstein actually wrote the equation as $m = E/c^2$, which better emphasizes the fact that the mass of an object depends on how much energy it has, and so *any* energy change must be accompanied by a change in mass.

For example, the combustion of methane, $\text{CH}_4 + 2 \text{O}_2 \rightarrow \text{CO}_2 + 2 \text{H}_2\text{O}$, is accompanied by the release of -604.5 kJ of heat (per mole of methane). Equation 1.8 implies that this change in energy must also be accompanied by a decrease in mass of about about 6.7 ng (6.7×10^{-9} g). Although such

a change in mass is too small to be readily measurable, it does clearly imply that mass is not strictly conserved during chemical reactions.

A more dramatic demonstration of the validity of eq 1.8 is the experimentally observed annihilation of an electron and a positron to form two high energy (gamma ray) photons, $e^- + e^+ \rightarrow 2\gamma$, in which the entire mass of the electron and positron is converted into energy (in the form of two photons with no rest mass). This process also implies that the energy released in the nuclear reaction, $4\text{H} \rightarrow \text{He} + 2e^- + 2e^+$, is exactly equivalent to the difference in mass between one helium atom and four hydrogen atoms, which is 0.029 g or 2.6×10^9 kJ! (per mole of He).

Comparison of the above chemical and nuclear reactions makes it clear why nuclear fusion might some day prove to be an attractive alternative to fossil fuels as a source of energy, although we have not yet devised a practical means of performing such reactions in a safe and controlled way. Alternatively, future generations may decide that the safest place to carry out nuclear reactions is in the sun, where they already occur naturally, and thus focus research efforts on improving the efficiency with which the sun's highly reliable and freely distributed supply of energy may best be harvested.

1.3 Quantization is Everywhere

Given that atoms and molecules are over 1000 times smaller than the thickness of this page, it should not be too surprising that the way the world looks and behaves on such very short length-scales is quite different from the macroscopic world of our everyday experience. So, although some of the things we learn about the atomic world can seem kind of strange, this is largely due to the quite natural difficulties associated projecting our macroscopic experiences onto the sub-microscopic scale. One of the most persistently troubling examples of such difficulties are those associated with a blurring of the lines between what we perceive as the wave-like and particle-like properties of objects.

In our everyday experience, we have little trouble distinguishing the difference between a wave on the ocean and a ball bouncing on the beach, or between the vibration of a guitar string and a bullet shooting out of a gun. That is because our macroscopic experience tells us that waves and particles are very different sort of objects. However, on the atomic scale such distinctions are not so clear, as the same object can sometimes behave like

a wave and at other times behave like a particle. This phenomena is also closely related to the quantization of energy, and even more generally to the quantization of *action* – a product of momentum and position whose units are the same as those of angular momentum and Planck’s quantum of action h , as we will see.

Much of our everyday experience suggests that energy is a continuous function. For example, when we are driving a car we are able to continuously accelerate from a state of zero kinetic energy up to a dangerously high kinetic energy. The same is true of the kinetic energy of a baseball or a billiard ball. Moreover, we expect a pendulum or a ball on a spring to be capable of oscillating over a continuous range of amplitudes, and thus to have a continuously variable energy.

However, on the atomic scale energies are usually quantized, in the sense that they have discrete rather than continuous values. The energy spacing between quantum states depends on the nature of the motion involved. When a given degree of freedom has a quantum state spacing that is small compared to the ambient thermal energy then it will behave classically, while when the spacing is larger than the available thermal energy it will behave non-classically, as further discussed in Section 1.4.

The Quantization of Light

The early development of quantum mechanics was marked by over two decades of bold speculation, aimed at repairing glaring disagreements between classical predictions and experimental measurements. The ensuing debate generated a fascinating plethora of proposals regarding the fundamental constituents underlying macroscopically observable phenomena.

The most famous failure of classical electrodynamics and thermodynamics pertains to the spectra of so-called “black-bodies” – which in fact closely resemble coals glowing in a campfire and the light emitted by stars overhead. Classical theory predicted that the intensity of the light radiated by such bodies should increase with increasing frequency, while experiments invariably showed intensities decreasing to zero at the highest frequencies. Planck resolved the discrepancy in 1900, by postulating that the energy emitted at each black-body frequency, ν , is quantized in packets of $h\nu$, with a universal constant of proportionality, h , which now bears his name. However, it was initially far from clear whether the required quantization should be attributed to light or to the material from which the glowing body is composed,

or both.

An important clarification of the above question was suggested by Einstein in the first of his three famous papers written in 1905, in which he presented various arguments all leading to the conclusion that light itself is quantized in packets of energy $h\nu$, now known as photons.⁴ The following are his own words (in translation) from the introduction to that paper.

It seems to me that the observations associated with blackbody radiation, fluorescence, the production of cathode rays by ultraviolet light, and other related phenomena connected with the emission or transformation of light are more readily understood if one assumes that the energy of light is discontinuously distributed in space. In accordance with the assumption to be considered here, the energy of a light ray spreading out from a point source is not continuously distributed over an increasing space but consists of a finite number of energy quanta which are localized at points in space, which move without dividing, and which can only be produced and absorbed as complete units.

At the end of the above paper Einstein noted that the quantization of light could explain the so-called “photo-electric effect”, in which electrons are ejected when a metal surface is irradiated with light. The problematic feature of the associated experimental observations was that the kinetic energies of the ejected electrons were found to be proportional to the frequency of the light, rather than its intensity. Einstein pointed out that this apparently paradoxical phenomena can readily be understood if light is composed of particles (photons) with energy $h\nu$. These speculations were not widely embraced for over a decade, until Millikan reported the results of additional key experiments. The following extended quotation from the introduction of Millikan’s 1916 paper, entitled *A Direct Photoelectric Determination of Planck’s “h”*, provides an interesting glimpse into the prevailing view of Einstein’s photon postulate.

⁴Although Planck and Einstein developed our current understanding of photons, it is an interesting and little known fact that the term “photon” was first introduced in a short note submitted to the journal *Nature* in 1926 by a prominent physical chemist named Gilbert Newton Lewis – the same G. N. Lewis who created the Lewis dot-structure representation of chemical bonds, and the concept of Lewis acids and bases, as well as many other important ideas pertaining to the thermodynamics of chemical processes.

Quantum theory was not originally developed for the sake of interpreting photoelectric phenomena. It was solely a theory as to the mechanism of absorption and emission of electromagnetic waves by resonators of atomic or subatomic dimensions. It had nothing to say about the energy of an escaping electron or about the conditions under which such an electron could make its escape, and up to this day the form of the theory developed by its author has not been able to account satisfactorily for the photoelectric facts presented herewith. We are confronted, however, by the astonishing situation that these facts were correctly and exactly predicted nine years ago by a form of quantum theory which has now been pretty generally abandoned.

It was in 1905 that Einstein made the first coupling of photo effects with any form of quantum theory by bringing forward the bold, not to say reckless, hypothesis of an electro-magnetic light copuscle of energy $h\nu$, which energy was transferred upon absorption to an electron. This hypothesis may well be called reckless first because an electro-magnetic disturbance which remains localized in space seems a violation of the very conception of an electromagnetic disturbance, and second because it flies in the face of the thoroughly established facts of interference. The hypothesis was apparently made solely because it furnished a ready explanation of one of the most remarkable facts brought to light by recent investigations, viz., that the energy with which an electron is thrown out of a metal by ultra-violet light or X-rays is independent of the intensity of the light while it depends on its frequency. This fact alone seems to demand some modification of classical theory or, at any rate, it has not yet been interpreted satisfactorily in terms of classical theory.

Even after Millikan's paper, and after Einstein received a Nobel prize "for his service to theoretical physics and particularly for his discovery of the law of the photo-electric effect", the subject of photon quantization remained, and continues to be, an active and interesting area of research, all the results of which are entirely consistent with Einstein's original proposal. However, Einstein himself apparently retained some concerns about the photon concept, as illustrated by the following quotation from the end of his 1917 paper entitled *On the Quantum Theory of Radiation* (which is most famous for

predicting stimulated emission, long before the development of lasers).

These properties of elementary particles... make the formulation of a proper quantum theory of radiation appear almost unavoidable. The weakness of the theory lies on the one hand in the fact that it does not get us any closer to making the connection with wave theory; on the other, that it leaves the duration and direction of the elementary processes to ‘chance’. Nevertheless I am fully confident that the approach chosen here is a reliable one.

Wave-Particles and Particles-Waves

The photo-electric effect is also closely related to the photo-ionization of atoms and molecules by light. In both cases the energy of the emitted electron is proportional to the frequency of the light. Also, in both cases no electrons are emitted when the photon energy $h\nu$ is too small. This makes sense, because some energy is required in order to overcome the binding energy of the electron to the material. So, for both the emission of photo-electrons from a metal and the photo-ionization of molecules,

$$K = h\nu - \Phi \tag{1.9}$$

where K is the kinetic energy of the ejected electron and Φ is the binding energy of the electron (which is a constant that is different for different materials).

Other examples of the particle-like properties of light include phenomena known as Compton scattering and Raman scattering. These both involve the inelastic scattering of light by a chemical substance (either a solid or a molecule). In other words, the energy of a photon is changed as it either gives up or gains energy from an object with which it collides.

Compton scattering involves the collision of a high energy (x-ray) photon with a free electron.⁵ So, the energy and momentum of the photon and the electron both change when they collide. Raman scattering, on the other hand, involves the interaction of a photon from with the vibrational modes of a molecule, so a Raman scattered photon either gains or loses energy

⁵The electron in a Compton scattering experiment often starts out inside a solid material. However, since the binding energy of the electron is much smaller than the energy of an x-ray photon, the electron behaves essentially as if it has no binding energy.

when it collides with a vibrating molecule.⁶ Both processes are similar to what might happen if you were to throw a baseball at a mattress spring, and the ball bounced back with less energy because it lost energy to the mattress springs.

Compton scattering is named after Arthur H. Compton, whose experiments performed in 1923 revealed the particle-like properties of photons in remarkable details. Compton's experiments showed that when a photon hits an electron, the two particles bounce off each other just like billiard balls on a pool table. The angles of the outgoing electron and photon are exactly those required in order to conserve both the energy and momentum of the two particles.⁷

Energy and momentum are clearly closely related to each other. For example, the kinetic energy of a free particle is $E = \frac{1}{2}mv^2 = p^2/2m$. Moreover, Einstein's theory of relativity implies that $E = mc^2$ for any particle. Since $p = mv$, eq 1.8 implies that $p = Ev/c^2$. The energy of a photon is

$$E = h\nu \quad (1.10)$$

while the velocity and frequency of light are $v = c$ and $\nu = c/\lambda$, respectively (where λ is the wavelength light). These identities may be combined to

⁶As we will see in Section 1.4, the vibrational energy of molecules is also quantized with an energy spacing of $h\nu_V$, where ν_V is the vibrational frequency of the molecule (which is equal to the frequency of light which is resonant with the molecular vibration).

⁷More specifically, if the input photon energy is $h\nu_{\text{in}}$ and the initially stationary electron is kicked out with an energy of $\Delta\varepsilon = \frac{1}{2}mv^2 = h(\nu_{\text{in}} - \nu_{\text{out}})$ and a momentum of $\Delta p = mv = h\left(\frac{1}{\lambda_{\text{in}}} - \frac{1}{\lambda_{\text{out}}}\right)$, then the observed deflection angles of the outgoing electron, ϕ , and photon, θ , will be related as follows.

$$\cos\theta = \frac{\Delta\varepsilon}{c\Delta p} \cos\phi$$

where $\cos\theta$ is also related the change in the photon wavelength $\Delta\lambda = \lambda_{\text{out}} - \lambda_{\text{in}}$. These scattering predictions are obtained by equating the energy and momentum of the ingoing and outgoing photon and electron, as described for example in Appendix XVII A in a book entitled *Light* by R. W. Ditchburn.

$$\Delta\lambda = \frac{h}{mc} (1 - \cos\theta)$$

obtain the following expression for the momentum of a photon.

$$p = \frac{h\nu c}{c^2} = \frac{h\nu}{c} = \frac{h\left(\frac{c}{\lambda}\right)}{c} = \frac{h}{\lambda} \quad (1.11)$$

The momentum of light can be observed experimentally by measuring the “radiation pressure” exerted by light as it reflects off the surface of a mirror. This pressure is exactly consistent with the particle-like properties of photons. But, quite remarkably, the pressure and energy density of light can also be correctly predicted from purely classical electromagnetic theory.

Photo-electric and Compton scattering experiments show that both photons and electrons have particle-like properties. However, the appearance of ν and λ in the equations for the energy and momentum of a photon indicates that the particle and wave properties of photons are inextricably linked. Such observations led a graduate student named Luis de Broglie to propose in 1924 that particles such as electrons, protons and atoms may also have wave-like properties. This astonishing prediction was beautifully confirmed in experiments which clearly show that electron and atoms do indeed have wave-like properties.

In the late 1920’s Otto Stern and co-workers set out to systematically test de Broglie’s hypothesis by conducting experiments in which a beams of various kinds of atoms and molecules were directed at salt crystals. Their results showed, for example, that a beam of He atoms undergoes diffraction when it is scattered off of salt crystals. The observed diffraction fringe spacing is related to the momentum of the He atoms and the lattice spacing of the rock salt crystal, exactly as predicted by de Broglie.⁸

Thus, not only photons but all other particles appear to have a wavelength which is related to their momentum.

$$p = \frac{h}{\lambda} \quad (1.12)$$

The apparently universal validity of this expression is one of the clearest pieces of evidence for the blurred distinction between particles and waves. All waves have particle-like properties and all particles have wave-like properties, but particles with very large momenta have very small wavelengths and waves with very long wavelengths have very small momenta. Objects with large

⁸The diffraction of beams of He and H₂ molecules produced by crystals of NaCl and LiF were reported by Otto Stern and coworkers in 1930.

momentum (such as macroscopic billiard balls) are more readily observable as particles while those with small momentum (such as photons) are more readily observable as waves.

1.4 Thermal Energies and Populations

Relation Between Energy and Probability

Energy plays a key role in determining the probability of finding a system in a given state. Not surprisingly, states of lower energy have a higher probability than those of higher energy. This is, for example, why the density of the atmosphere decreases with increasing altitude (i.e. with increasing gravitational potential energy). It is also why the density of a vapor is lower than that of the liquid with which it is at equilibrium (because molecules in the liquid experience a greater, more negative, cohesive interaction energy than they do in the vapor). The quantitative connection between energy and probability was worked out by Maxwell, Boltzmann and Gibbs, whose insights led to the following tremendously important, and yet remarkably simple, proportionality.⁹

$$P(\varepsilon) \propto e^{-\beta\varepsilon} \quad (1.13)$$

$P(\varepsilon)$ is the probability of finding a system in a state of a given energy, ε , and $\beta = 1/k_B T$, where $k_B T$ is a measure of thermal energy (equal to Boltzmann's constant, k_B , times the absolute temperature, T). So, eq 1.13 indicates that the probability of occupying a state not only decreases with increasing energy but also increases with temperature, which again makes good sense. In other words, the probability of observing any system in a state of energy ε is proportional to the Boltzmann factor $e^{-\beta\varepsilon}$, and this in turn only depends on the ratio $\varepsilon/k_B T = \beta\varepsilon$.

When Boltzmann's constant is multiplied by Avogadro's number it becomes equivalent to the gas constant, $R = N_A k_B$. So, if we choose to express energies in molar units then we should identify $\beta = 1/RT$. In other words, $k_B T$ and RT are essentially equivalent, and so physical chemists tend to

⁹This relation was first derived by Maxwell and Boltzmann for gas phase systems, and later generalized to any system by Gibbs. A nice summary of Gibbs' method may be found in Chapter 9 of a book entitled *Introduction to Theoretical Physical Chemistry* by S. Golden (Addison-Wesley, Reading MA, 1961)

switch back and forth between expressing thermal energy as $k_B T$ or RT , depending on whether the context calls for using molecular or molar units.

Notice that eq 1.13 is reminiscent of the well known relation between chemical equilibrium constants and reaction free energies, $K = e^{-\Delta G/RT}$, where K represents the ratio of the concentrations (probabilities) of the product and reactant molecules. This similarity is certainly no accident, as we will learn in Chapters 2 and 4.

The proportionality in eq 1.13 may be turned into an equality by making use of the fact that the total probability of observing a system in any state must be equal to 1. In other words, we require that $\sum_i P(\varepsilon_i) = 1$, where the sum is carried out over all the energies (quantum states) of the system. This also implies that the constant of proportionality that is missing in eq 1.13 is $1/\sum_i P(\varepsilon_i)$ and so $P(\varepsilon)$ is exactly given by the following expression.

$$P(\varepsilon_i) = \frac{e^{-\beta\varepsilon_i}}{\sum_i e^{-\beta\varepsilon_i}} \quad (1.14)$$

Note that the sum of $P(\varepsilon_i)$ over all states is $\sum_i e^{-\beta\varepsilon_i} / \sum_i e^{-\beta\varepsilon_i} = 1$, as expected.

The denominator in eq 1.14 plays a surprisingly important role in chemical thermodynamics – it is called the partition function, and is often represented by the letter q .¹⁰

$$q \equiv \sum_i e^{-\beta\varepsilon_i}. \quad (1.15)$$

One of the interesting facts about q is that it is equivalent to the number of thermally populated quantum states, in a given system at a given temperature, as further discussed below.

In order to see how we can use eq 1.14 to obtain practical predictions, it is useful to consider a system which has evenly spaced ladder of “quantum states”, $\varepsilon_n = n\Delta\varepsilon$, where $\Delta\varepsilon$ is a constant energy spacing and n is any integer ($0, 1, 2, 3, \dots$, so $\varepsilon_n = 0, \Delta\varepsilon, 2\Delta\varepsilon, 3\Delta\varepsilon, \dots$). Such an energy level structure arises in many situations, including molecular vibrations (described as harmonic oscillators) as well as light (which also consists of harmonic electromagnetic oscillations). Both experimental observations and quantum

¹⁰When describing a macroscopic system composed of many molecules (maintained at constant temperature and volume) the partition function is often designated as Q , or sometimes by other letters such as Z , but its definition is always the same.

mechanical predictions agree that such systems behave just as if they have an evenly spaced ladder of energy states.

Light of a given frequency (color) is composed of photons of energy $h\nu$ (where h is Planck's constant and ν is the frequency of the light). Thus, a beam of light can only have an energy of $n\Delta\varepsilon = nh\nu$, where n is the number of photons in the beam. Similarly, molecular vibrations can also only have energies of $n\Delta\varepsilon = nh\nu$ where ν is the frequency of the molecular vibration. When a molecule becomes macroscopically large, we call it a solid, and $h\nu$ becomes the energy of each vibrational "phonon" of the solid.

For any system with an evenly spaced ladder of energy states, one may use the following nifty mathematical trick to transform the above probability formula into a simple closed form. This begins by suggesting a change of variables, $x = e^{-\beta\Delta\varepsilon}$. Thus, the partition function (denominator) in Eq. 1.14 may be written as,¹¹

$$q = \sum_{n=0}^{\infty} x^n = 1 + x + x^2 + x^3 + \dots, \quad (1.16)$$

Since $\beta\Delta\varepsilon$ is invariably positive it follows that $0 < x < 1$. Under such conditions, the series converges exactly to $q = 1/(1-x) = 1/(1-e^{-\beta\Delta\varepsilon})$ when summed over all (the infinite number) of n values. Thus, the probability in eq 1.14 may be expressed simply as

$$P(\varepsilon_n) = \frac{e^{-\beta n\Delta\varepsilon}}{q} = e^{-\beta n\Delta\varepsilon} (1 - e^{-\beta\Delta\varepsilon}) . \quad (1.17)$$

This expression indicates that the probability of observing a quantized oscillator decreases exponentially with increasing n (at any fixed temperature). In other words, the oscillator is less likely to occupy higher energy states, as expected. Moreover, as the temperature approaches absolute zero, so does the Boltzmann factor $e^{-\beta\Delta\varepsilon}$. As a result, the only state that has a non-zero probability at very low temperature is the ground state, for which $P(\varepsilon_0) = e^{-0} = 1$, while for all other states $P(\varepsilon_n) = e^{-\beta n\Delta\varepsilon} = 0$. At high temperature, on the other hand, all the states for which $\beta\varepsilon_n \ll 1$ will have a finite probability while states for which $\beta\varepsilon_n \gg 1$ will have essentially zero probability. These results again make sense, as they indicate that the

¹¹We refer to this partition function as q simply as a reminder that it pertains to a system with n evenly spaced quantum state.

temperature of a system determines how many states will be significantly populated. At low temperature only the very lowest state will be populated, while at higher temperature only states for which $\varepsilon_n < k_B T$ will be significantly populated. So, as the temperature increases, so does the population of higher energy states.

The quantitative connection between q and the number of states with a significant thermal population can be inferred from the following considerations. The value of q is given by eq 1.15, which consists of a sum of terms each of which are equal to a number between 0 and 1. The lowest energy terms in the series are each approximately equal to one (since $e^{-\beta\varepsilon_i} \approx 1$ whenever $\varepsilon_i < k_B T$), while the high energy terms are approximately zero (since $e^{-\beta\varepsilon_i} \approx 0$ whenever $\varepsilon_i > k_B T$). Thus, q represents the average number of terms in the sum which have a value near one, which in turn represents the number of states which are significantly thermally populated.

The Boltzmann probabilities, $P(\varepsilon)$, may be used to calculate the weighted average of any property of the system. For example, the average energy of any system is given by the following weighted average.

$$\langle \varepsilon \rangle = \sum_i \varepsilon_i P(\varepsilon_i) = \frac{\sum_i \varepsilon_i e^{-\beta\varepsilon_i}}{\sum_i e^{-\beta\varepsilon_i}} \quad (1.18)$$

For a system with a ladder of evenly spaced energies the above sum reduced to the following more compact expression.

$$\langle \varepsilon \rangle = \frac{\sum \varepsilon_n e^{-\beta\varepsilon_n}}{\sum e^{-\beta\varepsilon_n}} = \frac{\Delta\varepsilon e^{-\beta\Delta\varepsilon}}{1 - e^{-\beta\Delta\varepsilon}} = \frac{\Delta\varepsilon}{e^{+\beta\Delta\varepsilon} - 1}. \quad (1.19)$$

This result was obtained by noting that the derivative of q with respect to β is $\sum -\varepsilon_n e^{-\beta\varepsilon_n}$, and so $\langle \varepsilon \rangle = \frac{\sum \varepsilon_n e^{-\beta\varepsilon_n}}{q} = \left(\frac{-1}{q}\right) \left(\frac{dq}{d\beta}\right)$.¹²

At very low temperature, the denominator of the above expression blows up, and so $\langle \varepsilon_n \rangle = 0$ (which makes sense since in this case all the population goes to the ground state of energy $\varepsilon_0 = 0$). On the other hand, at high temperature, $\Delta\varepsilon \ll k_B T$ and so $\beta\Delta\varepsilon \ll 1$ and $e^{\beta\Delta\varepsilon} \approx 1 + \beta\Delta\varepsilon$, which implies that $\langle \varepsilon_n \rangle \approx \Delta\varepsilon/(\beta\Delta\varepsilon) = 1/\beta = k_B T$. So, at high temperature the ladder of states is populated up to an energy equal to $k_B T$ (on the average).

¹²The required derivative is $dq_n/d\beta = -\Delta\varepsilon e^{-\beta\Delta\varepsilon}/(1 - e^{-\beta\Delta\varepsilon})^2$, which leads to eq 1.19 when divided by $-q$ (and the last identity in eq 1.19 is obtained by multiplying both the numerator and denominator by $e^{+\beta\Delta\varepsilon}$).

Note that this simple result pertains not only to very high temperatures but also to a system of lower temperature which has very closely spaced energy levels. This is because $\Delta\varepsilon/k_B T$ becomes small, and so $e^{\beta\Delta\varepsilon} \approx 1 + \beta\Delta\varepsilon$, either at high temperature or when the energy spacings are very small (compared to $k_B T$). So, *the high temperature limit is equivalent to the classical limit in which energies are essentially continuous.*

There is a deeper – not so obvious, but pretty interesting – reason that the average energy of an oscillator approaches $k_B T$ at high temperature. A hint as to the reason is provided by the fact that a similar analysis of systems with different sorts of energy level structures reveals that these often have average energies that are proportional, but not exactly equal to, $k_B T$. For example, a particle that is free to translate in the x -direction turns out to have an average kinetic energy of $\frac{1}{2}k_B T$, while a particle that is free to translate in three dimensions has an average kinetic energy of $\frac{3}{2}k_B T$. So, each additional dimension, or translational degree of freedom, increases the average kinetic energy by $\frac{1}{2}k_B T$.

The derivation of the above results also reveals that the constant of proportionality between the average energy and $k_B T$ depends in a simple way on the number of quadratic terms that appear in the Hamiltonian of the system (as shown in Section 1.5). So, for example, a free particle that is moving in the x -direction has a kinetic energy (and no potential energy) so its Hamiltonian is $H_x = \frac{1}{2}mv_x^2$, which is a quadratic function of the particles velocity (in the x -direction), and produces an average energy of $\frac{1}{2}k_B T$. The Hamiltonian of a particle that is free to move in three dimensions has three quadratic terms, $H_{xyz} = \frac{1}{2}mv_x^2 + \frac{1}{2}mv_y^2 + \frac{1}{2}mv_z^2$, and so has an average energy that is three times larger, $3(\frac{1}{2}k_B T) = \frac{3}{2}k_B T$. In other words, each quadratic term appearing in the Hamiltonian contributes $\frac{1}{2}k_B T$ to the average energy per molecule, or $\frac{1}{2}RT$ to the average energy per mole.

The energy of a harmonic oscillator (such as a vibrating diatomic molecule) contains both kinetic and potential energy contributions, each of which again give rise to quadratic terms in the Hamiltonian, $H_{\text{vib}} = \frac{1}{2}\mu v^2 + \frac{1}{2}f\delta^2$. The two quadratic variables in this expression are the velocity v and displacement δ of the oscillator (while μ and f are the corresponding reduced mass and harmonic force constant).¹³ Although these variables represent different

¹³More specifically, for a diatomic harmonic oscillator $v = |\vec{v}_2 - \vec{v}_1| = dr/dt$ (where \vec{v}_1 and \vec{v}_2 are the velocity vectors of each atom) and $\delta = r - r_b$ (where δ is the diatomic's instantaneous bond length, and r_b is its average bond length). The other parameters in

kinds of coordinates, it can be shown that each of the quadratic terms again contribute an average energy of $\frac{1}{2}k_B T$. *This explains why the average energy of a harmonic oscillator is equal to $k_B T$, since $\langle \varepsilon_{\text{vib}} \rangle = \frac{1}{2}k_B T + \frac{1}{2}k_B T = k_B T$.* This result again pertains to high temperatures (or small quantum spacings). At very low temperature (or in a system with large quantum spacings) the energy gap between the ground and first excited state prevents any energy from being absorbed by the system (as further discussed in section 1.4).

Rotational degrees of freedom also contribute quadratically to the Hamiltonian. For example, a diatomic molecule rotating about a given axis has a rotational kinetic energy of $H_{\text{rot}} = \frac{1}{2}I\omega^2$, where $I = \mu r_b^2$ is the moment of inertia of the diatomic and ω is its angular frequency of rotation (in units of radians per second). Since a diatomic has two independent axes of rotation, there are two quadratic rotational terms in its Hamiltonian. These contribute an additional $k_B T$ to the average energy (whenever the temperature is sufficiently high to overcome the rotational quantum state spacing).

We may now combine the above results to predict the total energy of a diatomic, including all 7 of its quadratic Hamiltonian terms (2 vibrational, 2 rotation and 3 translational). Thus, a diatomic molecule is expected to have a average energy of $\frac{7}{2}k_B T$ in the classical (high temperature) limit. The experimentally observed energies of real diatomic molecules may deviate from this prediction at low temperatures, as the result of energy quantization, as described in the next section.

Transition from Quantum to Classical Behavior

One of the first experimental clues that something was amiss with the purely classical view of physical chemistry came from experimental measurements of the heat capacities of solids and gases. In both cases, experiments revealed that at low temperatures chemical systems have smaller heat capacities than expected based on classical predictions. For example, the anomalous low-temperature heat capacities of solids were not understood until Albert Einstein (1907) and later Peter Debye (1912) showed that these are consistent with the quantization of phonons (vibrational motions of solids). Similar non-classical behavior of the heat capacity of diatomic gases was also recognized prior to 1901, as poignantly attested by J. Willard Gibbs in the

eq 1.45 are the reduced mass $\mu = \frac{m_1 m_2}{m_1 + m_2}$ and vibrational force constant $f = \mu (2\pi\nu)^2$ of the diatomic (where ν is its harmonic vibrational frequency).

following quotation from the preface to his *Statistical Mechanics* (the first and still classic foundational textbook on this subject).

... we do not escape difficulties in as simple a matter as the number of degrees of freedom of a diatomic gas. It is well known that while theory would assign to the gas six degrees of freedom per molecule, in our experiments on specific heats we cannot account for more than five. Certainly one is building on an insecure foundation, who rests his work on hypotheses concerning the constitution of matter.

Difficulties of this kind have deterred the author from attempting to explain the mysteries of nature, and have forced him to be contented with the more modest aim of deducing some of the more obvious propositions relating to the statistical branch of mechanics...¹⁴

Gibbs (who died in 1903) did not live to see these difficulties elegantly resolved in a way that left his statistical mechanics essentially unscathed (except for the quantization of his phase space using Planck's constant), and revealed that molecular and electromagnetic (blackbody) harmonic oscillators are mathematically isomorphic.

Recall that in the classical limit the energy of a mole of gas molecules is $U = N_A \langle \varepsilon \rangle = \frac{D}{2} N_A k_B T = \frac{D}{2} RT$, where D represents the number of degrees of freedom or quadratic terms in the Hamiltonian (and $R = k_B N_A$ is the gas constant and N_A is Avogadro's number). In this limit, the molar heat capacity (at constant-volume) is easy to calculate, since it is equal to first derivative of the energy, $C_V = \left(\frac{\partial U}{\partial T}\right)_V = \frac{D}{2} R$. However, as temperature decreases, different degrees of freedom are expected to become inactive, as $k_B T$ becomes smaller than the corresponding quantum state spacing.

The experimental results in Fig. 1.2 clearly show the transition from quantum to classical behavior with increasing temperature for two diatomic

¹⁴The fact that Gibbs expected a diatomic molecule to have six degrees of freedom suggests that he viewed diatomics as rigid (non-vibrating) molecules with three translational and three rotational degrees of freedom. We now know that molecules have an additional vibrational degree of freedom. Also, the reason that a diatomic only has two active rotation degrees of freedom is because rotation about the bond axis is of such high frequency that it has an energy spacing that is typically much larger than $k_B T$. This latter degree of rotational freedom corresponds to the rotation of electrons, and so gives rise to electronic state quantization.

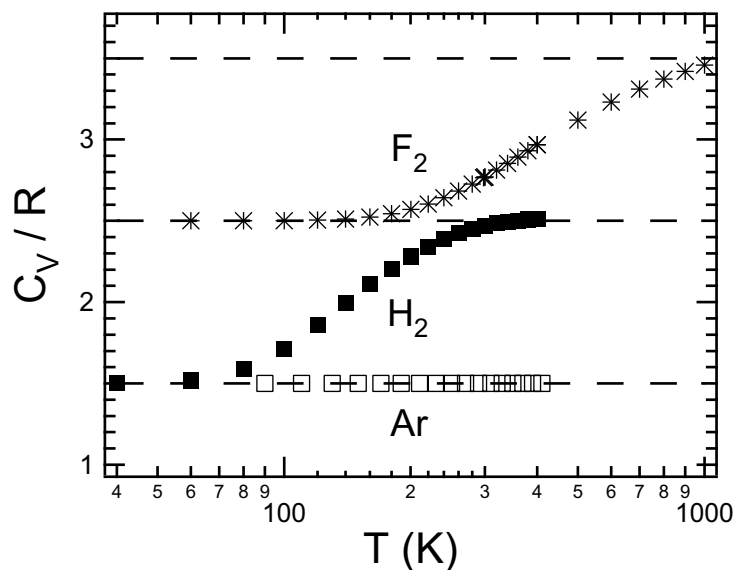


Figure 1.2: The experimental heat capacities (points) of argon and two diatomic molecules are plotted as a function of temperature. The horizontal dashed lines indicate the predicted heat capacity of a classical molecule with only translation (bottom), or only translation and rotation (middle), or translation, rotation and vibration (top) degrees of freedom.

molecules (H_2 and F_2), while argon behaves perfectly classically over this entire temperature range. Note that argon gas has three translation degrees of freedom and so it is predicted to have a classical translational heat capacity of $\frac{3}{2}R$, which is exactly what is observed experimentally. This clearly implies that the quantum spacing between translational states of argon is small compared to $k_B T$, even at temperatures below 100 K (in fact, translational quantum spacings are typically much smaller than $k_B T$ even at 1 K, and so translational kinetic energy can almost always be treated classically).

At low temperatures the hydrogen molecule H_2 also only appears to have three active degrees of freedom. This implies that the rotational quantum states of H_2 are sufficiently large that they only become active at temperatures well above 100 K. This is quite unusual, as most molecules have significantly smaller rotational quantum states spacings. The reason H_2 is exceptional is that it has an unusually high rotational frequency (because of its

small mass or moment of inertia). Hydrogen also has a very high vibrational frequency (again because of its small reduced mass). Thus both the rotational and vibrational quantum spacings of H_2 are significantly larger than those of other (heavier) diatomic molecules. So, only the three translations of H_2 are thermally active below about 100 K , and even at temperatures as high as 400 K the vibrational motion of H_2 is evidently not yet active. The heavier diatomic F_2 apparently has significantly smaller rotational and vibrational quantum state spacings than H_2 . Even at very low temperatures the two rotational degrees of freedom of F_2 are clearly already active. Moreover, the experimental results indicate that the vibrational quantum state spacing of F_2 has a magnitude of the order of $\varepsilon_V \approx k_B \times 1000K \approx 10^{-20}$ J or in molar units $R \times 1000K \approx 10$ kJ/mol.

The vibrational quantum spacing is equal to Planck's constant times the vibrational frequency, $\Delta\varepsilon = h\nu = hc\tilde{\nu} = hc/\lambda$, where $\tilde{\nu}$ is the frequency expressed in "wavenumber" units, which is equivalent to one over the wavelength of light with the same frequency $\tilde{\nu} = 1/\lambda$. In other words, λ is the wavelength of light whose frequency is resonant with that of the molecular vibration. Molecular vibrations typically have wavenumber frequencies somewhere between 100 cm^{-1} and 5000 cm^{-1} , which corresponds to wavelengths between 0.0002 cm and 0.01 cm (or 2 – 100 μm), in the infrared region of the electromagnetic spectrum.

The temperature dependence of the vibrational contribution to the heat capacity may be compared with the theoretical predictions of eq 1.19 (with $\Delta\varepsilon = hc\tilde{\nu}$). More specifically, the vibrational contribution to the heat capacity of a diatomic may be obtained by differentiating eq 1.19 with respect to temperature (or first differentiating with respect to β and noting that $d\beta/dT = -1/k_B T^2 = -k_B \beta^2$, which becomes $d\beta/dT = 1/RT^2$ when expressed in molar units).

$$C_V^{\text{vib}} = -\frac{1}{RT^2} \frac{d}{d\beta} \left[\frac{\Delta\varepsilon}{e^{\beta\Delta\varepsilon} - 1} \right] = R \frac{(\beta\Delta\varepsilon)^2 e^{\beta\Delta\varepsilon}}{[e^{\beta\Delta\varepsilon} - 1]^2} \quad (1.20)$$

Since translational and rotational energies typically have much smaller energy spacings, these degrees of freedom behave classically down to quite low temperatures. In other words, the sum of the translational and rotational contribution to the molar energy of a diatomic is $\frac{3}{2}RT + RT = \frac{5}{2}RT$, and so the corresponding contribution to the heat capacity is simply $\frac{5}{2}R$. Hence, the following expression is expected to accurately represent the total molar

heat capacity of a diatomic gas such as F_2 .

$$\frac{C_V}{R} = \frac{5}{2} + \frac{(\beta\Delta\varepsilon)^2 e^{\beta\Delta\varepsilon}}{[e^{\beta\Delta\varepsilon} - 1]^2} \quad (1.21)$$

Hydrogen (H_2) is exceptional, as previously noted, as its large rotational quantum state spacing gives rise to a rotational heat capacity that remains temperature dependent up to moderately high temperatures.

The spacings between a molecule's translational quantum states depend on the size of the box which contains the molecule. In a macroscopic container this energy spacing is so small that translational energies invariably behave classically. However, when a particle is confined to a container of nanometer (or smaller) dimensions then its translational quantum state spacing becomes much larger. For example, this is the case for an electron confined within an atom or molecule, whose translational quantization gives rise to electronic state spacings that are typically much larger than $k_B T$ (or RT in molar units), as we will see when we investigate the subject of quantum mechanics in greater detail (in Chapter 3).

1.5 Classical Energies and Hyper-Spheres

Calculating probabilities and average energies for degrees of freedom which are not quantized (or have very small quantum state spacings) requires replacing sums by integrals. For example, for systems with a continuum of energies eq 1.14 becomes

$$P(\varepsilon) = \frac{e^{-\beta\varepsilon} d\tau}{\int e^{-\beta\varepsilon} d\tau} \quad (1.22)$$

where τ represents any variable(s) on which the energy depends (such as the positions and velocities of each of the molecules in the system). In other words, the partition function is also now expressed as an integral.

$$q = \int e^{-\beta\varepsilon} d\tau \quad (1.23)$$

The average energy of any such classical degree of freedom is obtained from the following weighted integral of ε times the corresponding probability.

$$\langle \varepsilon \rangle = \int \varepsilon P(\varepsilon) = \frac{\int \varepsilon e^{-\beta\varepsilon} d\tau}{\int e^{-\beta\varepsilon} d\tau} \quad (1.24)$$

Since translational degrees of freedom typically behave classically (in any system of macroscopic volume), the above integral expressions may be used to calculate the associated probabilities and average energies. Lets first consider only the kinetic energy along the x -direction, for which the energy is $\varepsilon_x = \frac{1}{2}mv_x^2$. The variable τ in eqs 1.22–1.24 may in this case be taken to be equal to v_x . The corresponding partition function can again be evaluated analytically.¹⁵

$$q_x = \int e^{-\beta\varepsilon_x} d\tau = \int_{-\infty}^{\infty} e^{-mv_x^2/2k_B T} dv_x = \sqrt{\frac{2\pi k_B T}{m}} \quad (1.25)$$

Thus, the probability of observing a molecule with a velocity between v_x and $v_x + dv_x$ is,

$$P(v_x) = \sqrt{\frac{m}{2\pi k_B T}} e^{-mv_x^2/2k_B T} dv_x \quad (1.26)$$

The above expression may readily be extended to three dimensions by recalling that the probabilities of observing a combination of statistically independent events behave very much like the probabilities of tossing a sequence of coins, or rolling a pair of dice. For example, the probability of tossing two heads in a row is $\frac{1}{2}\frac{1}{2} = \frac{1}{4}$, while that of tossing three heads in a row is $\frac{1}{2}\frac{1}{2}\frac{1}{2} = \frac{1}{8}$. Similarly, the probability of throwing a snake eyes (two ones) with a pair of dice is $\frac{1}{6}\frac{1}{6} = \frac{1}{36}$. In other words, *the combined probabilities of statistically independent events always multiply*. This simple fact may be used to convert the probability associated with velocity in the x -direction, to the probability associated with the total velocity in three-dimensions, $v = \sqrt{v_x^2 + v_y^2 + v_z^2}$.

$$\begin{aligned} P(v) &= P(v_x) P(v_y) P(v_z) \\ &= \left(\frac{m}{2\pi k_B T}\right)^{3/2} e^{-m(v_x^2+v_y^2+v_z^2)/2k_B T} dv_x dv_y dv_z \end{aligned}$$

We may also transform the above three dimensional integral to a one dimensional integral with respect to the total velocity v .

$$P(v) = \left(\frac{m}{2\pi k_B T}\right)^{3/2} e^{-mv^2/2k_B T} 4\pi v^2 dv \quad (1.27)$$

¹⁵The last identity in eq 1.25 is obtained using $\int_{-\infty}^{\infty} e^{-ax^2} dx = 2 \int_0^{\infty} e^{-ax^2} dx = \sqrt{\frac{\pi}{a}}$ (from a table of integrals).

Notice that $4\pi v^2$ may be viewed as the “surface area” of a sphere of “radius” v , and so $4\pi v^2 dv$ is the “volume” of a shell of thickness dv . Thus, integration over $4\pi v^2 dv$ is equivalent to integration over $dv_x dv_y dv_z$.¹⁶

The total translational kinetic energy of a molecule (in three-dimensions) is $\varepsilon = \frac{1}{2}mv^2$ so we may also change the independent variable in eq 1.27 from v to ε (since $v = \sqrt{\frac{2\varepsilon}{m}}$ and so $dv = \frac{1}{\sqrt{2m\varepsilon}}d\varepsilon$),

$$P(\varepsilon) = \frac{2}{\sqrt{\pi}} \sqrt{\varepsilon^*} e^{-\varepsilon^*} d\varepsilon^* \quad (1.28)$$

where $\varepsilon^* = \varepsilon/k_B T$. This energy probability density can be used to calculate the average translational kinetic energy of a molecule (in three-dimensions).¹⁷

$$\langle \varepsilon \rangle = k_B T \frac{2}{\sqrt{\pi}} \int_0^\infty \varepsilon^* \sqrt{\varepsilon^*} e^{-\varepsilon^*} d\varepsilon^* = \frac{3}{2} k_B T \quad (1.29)$$

Thus, as expected, the average translational kinetic energy of a molecule in three dimensions is exactly three times its average energy in one dimension.

The above results clearly suggest that we may use a similar procedure to calculate probabilities and average energies for systems with any number of degrees of freedom which contribute quadratically to the total energy (Hamiltonian). In order to more easily see how we may generalize the above results, it is useful to note that the translational kinetic energy along each direction in space may be re-scaled to produce a new variable $r_i \equiv v_i \sqrt{m/2}$, so that

$$\varepsilon = r_x^2 + r_y^2 + r_z^2.$$

The right hand side of the above expression looks just like the square of the radius of a sphere, $\mathbf{r}^2 = r_x^2 + r_y^2 + r_z^2$, and so we may equate $\varepsilon = \mathbf{r}^2$. We can do the same thing for systems with any number of translational degrees of freedom, or any other degrees of freedom which contribute quadratic terms to the energy.

For example, the two quadratic variables which contribute to the energy of a classical harmonic oscillator may be re-scaled to express $\varepsilon_{\text{vib}} = r_1^2 + r_2^2 = \mathbf{r}^2$.

¹⁶The volume element $dv_x dv_y dv_z$ may be expressed as $v^2 \sin \phi d\phi d\theta dv$ in polar coordinates, and then integrated over all angles (i.e. from 0 to π for ϕ and 0 to 2π for θ) to obtain $4\pi v^2 dv$.

¹⁷The integral in eq 1.29 was evaluated using the fact that $\int_0^\infty x^b e^{-x} dx = \Gamma(b+1)$, for any real positive value of b . The Gamma function, Γ is further described in the paragraph following the eq 1.30.

If we add the two rotational degrees of freedom of a diatomic molecule we may express the total vibrational plus rotational energy of the diatomic as the sum of four re-scaled quadratic components, $\varepsilon_{\text{vib-rot}} = r_1^2 + r_2^2 + r_3^2 + r_4^2$.¹⁸ Note that this is again equivalent to the square of the radius of a four-dimensional hyper-sphere, $\mathbf{r}^2 = r_1^2 + r_2^2 + r_3^2 + r_4^2$ (or $\mathbf{r} = \sqrt{r_1^2 + r_2^2 + r_3^2 + r_4^2}$).

More generally, any system with D (quadratic) degrees of freedom can be represented as a D -dimensional hyper-sphere. In order to make use of this correspondence, we need to consider some interesting facts about hyper-spheres. The volume of a hyper-sphere may be expressed as the following function of its radius \mathbf{r} and dimension D .

$$V_D = \frac{\pi^{D/2}}{\Gamma(\frac{D}{2} + 1)} \mathbf{r}^D \quad (1.30)$$

The Gamma function, $\Gamma(b + 1)$, is closely related to the factorial function since $\Gamma(n + 1) = n! = n(n - 1)(n - 2) \dots 1$, when n is a positive integer, and $\Gamma(1) = 0! = 1$. For half-integer values of b the Gamma function may be evaluated by noting that $\Gamma(\frac{1}{2}) = \sqrt{\pi}$ and $\Gamma(b + 1) = b\Gamma(b)$. So, it is pretty easy to show that the above expression produces the correct “volumes” of $V_1 = 2\mathbf{r}$, $V_2 = \pi\mathbf{r}^2$ and $V_3 = \frac{4}{3}\pi\mathbf{r}^3$ for “spheres” in 1, 2, and 3 dimensions, respectively. The same formula also correctly predicts the volumes of hyper-spheres in higher dimensions, each of which are enclosed by a surface that is equidistant from a single point.

The surface area of a hyper-sphere may be obtained by differentiating the volume with respect to \mathbf{r} , and so¹⁹

$$A_D = \frac{dV_D}{d\mathbf{r}} = \frac{D\pi^{D/2}}{\Gamma(\frac{D}{2} + 1)} \mathbf{r}^{D-1} \quad (1.31)$$

This surface area may be used to obtain the volume of a hyper-spherical shell of radius \mathbf{r} and thickness $d\mathbf{r}$. The volume of such a shell is simply $A_D d\mathbf{r}$, and so the probability of observing a state with a value between \mathbf{r} and $\mathbf{r} + d\mathbf{r}$ in a system with D quadratic degrees of freedom is,

$$P_D(\mathbf{r}) = [P_1(r_1)]^D = \frac{D}{\Gamma(\frac{D}{2} + 1)} \mathbf{r}^{*D-1} e^{-\mathbf{r}^{*2}} d\mathbf{r}^*, \quad (1.32)$$

¹⁸More specifically, the rescaled variables are $r_1^2 = \frac{1}{2}\mu\Delta v^2$, $r_2^2 = \frac{1}{2}f\Delta r^2$, $r_3^2 = r_4^2 = \frac{1}{2}I\omega^2$.

¹⁹The derivative relation between volume and surface area follows from the fact that $V = \int_0^{\mathbf{r}} A d\mathbf{r}$ which implies that $A = dV/d\mathbf{r}$

where $P_1(r_1)$ is the probability associated with a single re-scaled quadratic degree of freedom and $\mathbf{r}^* = \mathbf{r}/\sqrt{k_B T}$.²⁰ By using $\varepsilon = \mathbf{r}^2$ we may transform to the corresponding energy probability density (with $\varepsilon^* = \varepsilon/k_B T$).

$$P_D(\varepsilon) = \frac{D}{2\Gamma(\frac{D}{2} + 1)} \varepsilon^{\frac{D}{2}-1} e^{-\varepsilon} d\varepsilon \quad (1.33)$$

The average value of any function of \mathbf{r} may be obtained using $P_D(\mathbf{r})$,

$$\langle f(\mathbf{r}) \rangle_D = \int f(\mathbf{r}) P_D(\mathbf{r}) = \frac{D}{\Gamma(\frac{D}{2} + 1)} \int_0^\infty f(\mathbf{r}) \mathbf{r}^{*D-1} e^{-\mathbf{r}^{*2}} d\mathbf{r}^* \quad (1.34)$$

while the corresponding average value of any function of ε may be obtained using $P_D(\varepsilon)$.

$$\langle f(\varepsilon) \rangle_D = \int f(\varepsilon) P_D(\varepsilon) = \frac{D}{2\Gamma(\frac{D}{2} + 1)} \int_0^\infty f(\varepsilon) \varepsilon^{\frac{D}{2}-1} e^{-\varepsilon} d\varepsilon \quad (1.35)$$

Thus, for example, we may use eq 1.35 (or eq 1.34) to predict the average energy of any system with D (quadratic) degrees of freedom.²¹

$$\langle \varepsilon \rangle_D = \frac{D}{2} k_B T \quad (1.36)$$

For a harmonic oscillator $D = 2$ and so eq 1.36 implies that the average vibrational energy of a diatomic molecule is $\langle \varepsilon \rangle = k_B T$, as expected. We may also use eq 1.34 to evaluate the average amplitude δ and velocity v of the vibrating diatomic. In this case, it is convenient to treat each quadratic term as a separate one-dimensional coordinate.

$$\langle \delta \rangle = \langle \mathbf{r}^* \rangle_1 \sqrt{\frac{2k_B T}{f}} = \sqrt{\frac{8k_B T}{\pi f}} \quad (1.37)$$

²⁰Equation 1.32 is obtained by noting that in one dimension the partition function associated with a single re-scaled coordinate r_1 is, $q = \int_{-\infty}^{\infty} e^{-r_1^2/k_B T} dr_1 = \sqrt{\pi k_B T}$, so $P_1(r_1) = \frac{e^{-r_1^2/k_B T} dr_1}{\sqrt{\pi k_B T}}$, and thus $[P_1(\mathbf{r}_1)]^D = \frac{e^{-\mathbf{r}^2/k_B T}}{(\pi k_B T)^{\frac{D}{2}}} A_D d\mathbf{x}$, where $\mathbf{r}^2 = \sum_{i=1}^D r_i^2$. Note that while r_1 extends over the entire real axis (including both positive and negative values), while $\mathbf{r} = \sqrt{\mathbf{r}^2} > 0$, as it represents the absolute “velocity” or “vibrational amplitude”, in the appropriately re-scaled units.

²¹The required integrals may be evaluated using one of the following standard integrals, $\int_0^\infty x^b e^{-ax} dx = \Gamma(b+1)/a^{b+1}$ or $\int_0^\infty x^b e^{-ax^2} dx = \Gamma(\frac{b+1}{2})/(2a^{\frac{b+1}{2}})$.

$$\langle v \rangle = \langle \mathbf{r}^* \rangle_1 \sqrt{\frac{2k_B T}{\mu}} = \sqrt{\frac{8k_B T}{\pi\mu}} \quad (1.38)$$

Thus, the average amplitude and velocity of the vibrating molecule are both predicted to increase with increasing temperature – which seems reasonable. They are also predicted to decrease with increasing force constant or reduced mass. Note that the square-root derives from the fact that the energy depends quadratically on δ and v .

Beyond the Quadratic Approximation

The above results are all restricted to systems whose Hamiltonians may be expressed as the sum of quadratic position and/or velocity (or momentum) variables. For translational kinetic energy this is exactly the case, while for vibrational and rotational motions the corresponding quadratic terms are approximations which ignore anharmonic and centrifugal effects that give rise to additional non-quadratic contributions to the Hamiltonian. However, the above procedure may readily be extended to treat such contributions.

As one example, the puckering motion of four-membered ring compounds, such as cyclobutanone and trimethylene oxide, have potential energies that scale approximately as the fourth power, rather than square, of the corresponding displacement, $\varepsilon = ax^4$. The average potential energy associated with this motion may again be obtained by applying eq 1.24.

$$\langle \varepsilon \rangle = \frac{\int ax^4 e^{-\beta ax^4} dx}{\int e^{-\beta ax^4} dx} = \frac{1}{4\beta} = \frac{1}{4} k_B T \quad (1.39)$$

So, this degree of freedom is expected to have an average potential energy that is smaller (by a factor of two) than that of a harmonic (quadratic) vibration, at high temperature.

More generally, any contribution to the energy which is proportional to the n^{th} -power of the corresponding variable gives rise to an average potential energy of $\frac{1}{n} k_B T$, in the classical limit. So, a particle in D -dimensions which is confined by a central force potential of the form $V(r) = cr^n$ has an average kinetic energy of $\langle K \rangle = \frac{D}{2} k_B T$ and an average potential energy of $\langle V \rangle = \frac{1}{n} k_B T$. This implies that the ratio of the potential and kinetic energies of such a particle is simply determined by the values of D and n .

$$\frac{\langle V \rangle}{\langle K \rangle} = \frac{2}{nD} \quad (1.40)$$

This important result is equivalent to that obtained from the Virial Theorem of Clausius. Note that for a 1-dimensional harmonic oscillator this predicts that $\langle V \rangle / \langle K \rangle = 1$, as expected. On the other hand, for a three dimensional particle confined by potential which scales as $1/r$ (such as an electron bound to a proton or a planet rotating around the sun), the ratio becomes $\langle V \rangle / \langle K \rangle = -2/3$, which is again consistent with experimental observations.