KINETICS

LibreTexts™

Kinetics - Davis

The LibreTexts libraries are Powered by NICE CXOne and are supported by the Department of Education Open Textbook Pilot Project, the UC Davis Office of the Provost, the UC Davis Library, the California State University Affordable Learning Solutions Program, and Merlot. This material is based upon work supported by the National Science Foundation under Grant No. 1246120, 1525057, and 1413739.

Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation nor the US Department of Education.

TABLE OF CONTENTS

6: The Distribution of Gas Velocities

7: Diffusion

8: Chemical Kinetics

9: Enzyme Kinetics

Numerical Solutions to Rate Laws

Licensing

*A detailed breakdown of this resource's licensing can be found in **Back Matter/Detailed Licensing**.*

CHAPTER OVERVIEW

1: Boltzmann

```
from matplotlib import pyplot as plt
import numpy as np
k=0.695 #In wavenumbers/K. We can change it to other units if we want to use $\vareps
Na=6.022045*10**23
h=6.262e-34
c=2.9979e10
def q(eps,T):
  sum=0
  for i in range(0,len(eps)):
    sum+=np.exp(-eps[i]/k/T)
  return sum
T=np.arange(1,1500,1)
#@title Default title text
upperlevelenergy =  500#@param {type:"integer"}
eps1=[0,upperlevelenergy] #energy levels at 0 and 50 cm^-1
plt.plot(T,q(eps1,T))
plt.ylim(1,2)
```

run | restart | restart & run all

```
y=3
print(y)
```

run | restart | restart & run all

CHAPTER OVERVIEW

2: Boltzmann Statistics

2.1: The Independent-Molecule Approximation

In Chapter 18, our survey of quantum mechanics introduces the idea that a molecule can have any of an infinite number of discrete energies, which we can put in order starting with the smallest. We now turn our attention to the properties of a system composed of a large number of molecules. This multi-molecule system must obey the laws of quantum mechanics. Therefore, there exists a Schrödinger equation, whose variables include all of the inter-nucleus, inter-electron, and electron-nucleus distance and potential terms in the entire multi-molecule system. The relevant boundary conditions apply at the physical boundaries of the macroscopic system. The solutions of this equation include a set of infinitely many wavefunctions, $\Psi_{i,j}$, each describing a quantum mechanical state of the entire multi-molecule system. In general, the collection of elementary particles that can be assembled into a particular multi-molecule system can also be assembled into many other multi-molecule systems. For example, an equimolar mixture of CO and H_2O can be reassembled into a system comprised of equimolar CO_2 and H_2, or into many other systems containing mixtures of CO, H_2O, CO_2, and H_2. Infinitely many quantum-mechanical states are available to each of these multi-molecule systems.

For every such multi-molecule wavefunction, $\psi_{i,j}$, there is a corresponding system energy, E_i. In general, the system energy, E_i, is Ω_i-fold degenerate; there are Ω_i wavefunctions, $\Psi_{i,1}$, $\Psi_{i,2}$, …, Ψ_{i,Ω_i}, whose energy is E_i. The wavefunctions include all of the interactions among the molecules of the system, and the energy levels of the system reflect all of these interactions. While generating and solving this multi-molecule Schrödinger equation is straightforward in principle, it is completely impossible in practice.

Fortunately, we can model multi-molecule systems in another way. The primary focus of chemistry is the study of the properties and reactions of molecules. Indeed, the science of chemistry exists, as we know it, only because the atoms comprising a molecule stick together more tenaciously than molecules stick to one another. (Where this is not true, we get macromolecular materials like metals, crystalline salts, *etc*.) This occurs because the energies that characterize the interactions of atoms within a molecule are much greater than the energies that characterize the interaction of one molecule with another. Consequently, the energy of the system can be viewed as the sum of two terms. One term is a sum of the energies that the component molecules would have if they were all infinitely far apart. The other term is a sum of the energies of all of the intermolecular interactions, which is the energy change that would occur if the molecules were brought from a state of infinite separation to the state of interest.

In principle, we can describe a multi-molecule system in this way with complete accuracy. This description has the advantage that it breaks a very large and complex problem into two smaller problems, one of which we have already solved: In Chapter 18, we see that we can approximate the quantum-mechanical description of a molecule and its energy levels by factoring molecular motions into translational, rotational, vibrational, and electronic components. It remains only to describe the intermolecular interactions. When intramolecular energies are much greater than intermolecular-interaction energies, it may be a good approximation to ignore the intermolecular interactions altogether. This occurs when we describe ideal gas molecules; in the limit that a gas behaves ideally, the force between any two of its molecules is nil.

In Chapter 23, we return to the idea of multi-molecule wavefunctions and energy levels. Meanwhile we assume that intermolecular interactions can be ignored. This is a poor approximation for many systems. However, it is a good approximation for many others, and it enables us to keep our description of the system simple while we use molecular properties in our development of the essential ideas of statistical thermodynamics.

We focus on developing a theory that gives the macroscopic thermodynamic properties of a pure substance in terms of the energy levels available to its individual molecules. To begin, we suppose that we solve the Schrödinger equation for an isolated molecule. In this Schrödinger equation, the variables include the inter-nucleus, inter-electron, and electron-nucleus distance and potential terms that are necessary to describe the molecule. The solutions are a set of infinitely many wavefunctions, $\psi_{i,j}$, each describing a different quantum-mechanical state of an isolated molecule. We refer to each of the possible wavefunctions as ***quantum state*** of the molecule. For every such wavefunction, there is a corresponding molecular energy, ϵ_i. Every unique molecular energy, ϵ_i, is called an ***energy level***. Several quantum states can have the same energy. When two or more quantum states have the same energy, we say that they belong to the same energy level, and the energy level is said to be ***degenerate***. In general, there are g_i quantum states that we can represent by the g_i wavefunctions, $\psi_{i,1}$, , $\psi_{i,2}$, ..., , ψ_{i,g_i}, each of whose energy is ϵ_i. The number of quantum states that have the same energy is called the ***degeneracy*** of the energy level. Figure 1 illustrates the terms we use to describe the quantum states and energy levels available to a molecule.

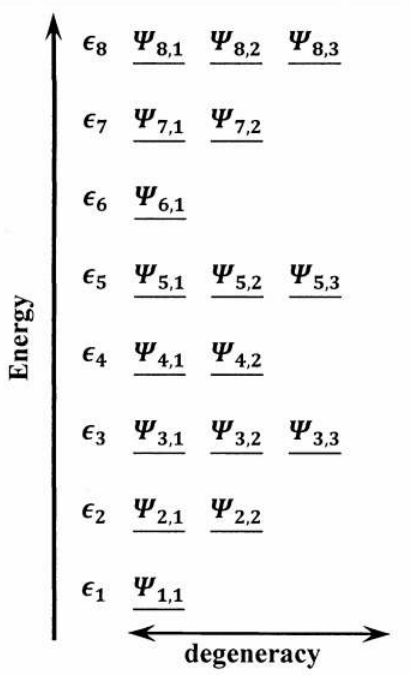


Figure 1. Quantum states and degenerate energy levels.

In our development of classical thermodynamics, we find it convenient to express the value of a thermodynamic property of a pure substance as the change that occurs during a formal process that forms one mole of the substance, in its standard state, from its unmixed constituent elements, in their standard states. In developing statistical thermodynamics, we find it convenient to express the value of a molecular energy, ϵ_i, as the change that occurs during a formal process that forms a molecule of the substance, in one of its quantum states, $\psi_{i,j}$, from its infinitely separated, stationary, constituent atoms. That is, we let the isolated constituent atoms be the reference state for the thermodynamic properties of a pure substance.

2.2: The Probability of An Energy Level at Constant N, V, and T

If only pressure–volume work is possible, the state of a closed, reversible system can be specified by specifying its volume and temperature. Since the system is closed, the number, N, of molecules is constant. Let us consider a closed, equilibrated, constant-volume, constant-temperature system in which the total number of molecules is very large. Let us imagine that we can monitor the quantum state of one particular molecule over a very long time. Eventually, we are able to calculate the fraction of the elapsed time that the molecule spends in each of the quantum states. We label the available quantum states with the wavefunction symbols, $\psi_{i,j}$.

We assume that the fraction of the time that a molecule spends in the quantum state $\psi_{i,j}$ is the same thing as the probability of finding the molecule in quantum state $\psi_{i,j}$. We denote this probability as $\rho\left(\psi_{i,j}\right)$. To develop the theory of statistical thermodynamics, we assume that this probability depends on the energy, and only on the energy, of the quantum state $\psi_{i,j}$. Consequently, any two quantum states whose energies are the same have the same probability, and the g_i-fold degenerate quantum states, $\psi_{i,j}$, whose energies are ϵ_i, all have the same probability. In our imaginary monitoring of the state of a particular molecule, we observe that the probabilities of two quantum states are the same if and only if their energies are the same; that is, we observe $\rho\left(\psi_{i,j}\right) = \rho\left(\psi_{k,m}\right)$ if and only if $i = k$.

The justification for this assumption is that the resulting theory successfully models experimental observations. We can ask, however, why we might be led to make this assumption in the first place. We can reason as follows: The fact that we observe a definite value for the energy of the macroscopic system implies that quantum states whose energies are much greater than the average molecular energy must be less probable than quantum states whose energies are smaller. Otherwise, the sum of the energies of high-energy molecules would exceed the energy of the system. Therefore, we can reasonably infer that the probability of a quantum state depends on its energy. On the other hand, we can think of no plausible reason for a given molecule to prefer one quantum state to another quantum state that has the same energy.

This assumption means that a single function suffices to specify the probability of finding a given molecule in any quantum state, $\psi_{i,j}$, and the only independent variable is the quantum-state energy, ϵ_i. We denote the probability of a single quantum state, $\psi_{i,j}$, whose energy is ϵ_i, as $\rho\left(\epsilon_i\right)$. Since this is the probability of each of the g_i-fold degenerate quantum states, $\psi_{i,j}$, that have energy ϵ_i, the probability of finding a given molecule in any energy level, ϵ_i, is $P\left(\epsilon_i\right) = g_i\rho\left(\epsilon_i\right)$. We find it convenient to introduce "P_i" to abbreviate this probability; that is, we let

$$P_i = \sum_{j=1}^{g_i} \rho\left(\psi_{i,j}\right) = P\left(\epsilon_i\right) = g_i\rho\left(\epsilon_i\right)$$

(the probability of energy level ϵ_i)

There is a P_i for every energy level ϵ_i. P_i must be the same for any molecule, since every molecule has the same properties. If the population set $\{N_1^\bullet, N_2^\bullet, \ldots, N_i^\bullet, \ldots\}$ characterizes the equilibrium system, the fraction of the molecules that have energy ϵ_i is $N_i^\bullet/N$. (Elsewhere, an energy-level population set is often called a "distribution." Since we define a distribution somewhat differently, we avoid this usage.) Since the fraction of the molecules in an energy level at any instant of time is the same as the fraction of the time that one molecule spends in that energy level, we have

$$P_i = P\left(\epsilon_i\right) = g_i\rho\left(\epsilon_i\right) = \frac{N_i^\bullet}{N}$$

As long as the system is at equilibrium, this fraction is constant. In Chapter 21, we find an explicit equation for the probability function, $\rho\left(\epsilon_i\right)$.

The energy levels, ϵ_i, depend on the properties of the molecules. In developing Boltzmann statistics for non-interacting molecules, we assume that the probability of finding a molecule in a particular energy level is independent of the number of molecules present in the system. While P_i and $\rho\left(\epsilon_i\right)$ depend on the energy level, ϵ_i, neither depends on the number of molecules, N. If we imagine inserting a barrier that converts an equilibrated collection of molecules into two half-size collections, each of the new collections is still at equilibrium. Each contains half as many molecules and has half the total energy of the original. In our model, the fraction of the molecules in any given energy level remains constant. Consequently, the probabilities associated with each energy level remain constant. (In Chapter 25, we introduce Fermi-Dirac and Bose-Einstein statistics. When we must use either of these models to describe the system, P_i is affected by rules for the number of molecules that can occupy an energy level.)

The number of molecules and the total energy are extensive properties and vary in direct proportion to the size of the system. The probability, P_i, is an intensive variable that is a characteristic property of the macroscopic system. P_i is a state function. P_i depends on ϵ_i. So long as the thermodynamic variables that determine the state of the system remain constant, the ϵ_i are constant. For a given macroscopic system in which only pressure–volume work is possible, the quantum mechanical energy levels, ϵ_i, are constant so long as the system volume and temperature are constant. However, the ϵ_i are quantum-mechanical quantities that depend on our specification of the molecule and on the boundary values in our specification of the system. If we change any molecular properties or the dimensions of the system, the probabilities, P_i, change.

2.3: The Population Sets of a System at Equilibrium at Constant N, V, and T

In developing Boltzmann statistics, we assume that we can tell different molecules of the same substance apart. We say that the molecules are ***distinguishable***. This assumption is valid for molecules that occupy lattice sites in a crystal. In a crystal, we can specify a particular molecule by specifying its position in the lattice. In other systems, we may be unable to distinguish between different molecules of the same substance. Most notably, we cannot distinguish between two molecules of the same substance in the gas phase. The fact that gas molecules are indistinguishable, while we assume otherwise in developing Boltzmann statistics, turns out to be a problem that is readily overcome. We discuss this in Section 24.2.

We want to model properties of a system that contains N, identical, distinguishable, non-interacting molecules. The solutions of the Schrödinger equation presume fixed boundary conditions. This means that the volume of this N-molecule system is constant. We assume also that the temperature of the N-molecule system is constant. Thus, our goal is a theory that predicts the properties of a system when N, V, and T are specified. When there are no intermolecular interactions, the energy of the system is just the sum of the energies of the individual molecules. If we know how the molecules are allocated among the energy levels, we can find the energy of the system. Letting N_i be the population of the energy level ϵ_i, any such allocation is a population set $\{N_1,\ N_2, \ldots, N_i, \ldots\}$. We have

$$N = \sum_{i=1}^{\infty} N_i$$

and the system energy is

$$E = \sum_{i=1}^{\infty} N_i \epsilon_i$$

Let us imagine that we can assemble a system with the molecules allocated among the energy levels in any way we please. Let $\{N_1^o,\ N_2^o, \ldots, N_i^o, \ldots\}$ represent an initial population set that describes a system that we assemble in this way. This population set corresponds to a well-defined system energy. We imagine immersing the container in a constant-temperature bath. Since the system can exchange energy with the bath, the molecules of the system gain or lose energy until the system attains the temperature of the bath in which it is immersed. As this occurs, the populations of the energy levels change. A series of different population sets characterizes the state of the system as it evolves toward thermal equilibrium. When the system reaches equilibrium, the population sets that characterize it are different from the initial one, $\{N_1^o,\ N_2^o, \ldots, N_i^o, \ldots\}$.

Evidently, the macroscopic properties of such a system also change with time. The changes in the macroscopic properties of the system parallel the changing energy-level populations. At thermal equilibrium, macroscopic properties of the system cease to undergo any further change. In Section 3.9, we introduce the idea that the most probable population set, which we denote as

$$\left\{N_1^{\bullet},\ N_2^{\bullet}, \ldots, N_i^{\bullet}, \ldots\right\}$$

or its proxy,

$$\{NP(\epsilon_1), NP(\epsilon_2), \ldots, NP(\epsilon_i), \ldots\}$$

(where $N = N_1^{\bullet} + N_2^{\bullet} + \ldots + N_i^{\bullet} + \ldots$), is the best prediction we can make about the outcomes in a future set of experiments in which we find the energy of each of N different molecules at a particular instant. We hypothesize that the most probable population set specifies all of the properties of the macroscopic system in its equilibrium state. When we develop the logical consequences of this hypothesis, we find a theory that expresses macroscopic thermodynamic properties in terms of the energy levels available to individual molecules. In the end, the justification of this hypothesis is that it enables us to calculate thermodynamic properties that agree with experimental measurements made on macroscopic systems.

Our hypothesis asserts that the properties of the equilibrium state are the same as the properties of the system when it is described by the most probable population set. Evidently, we can predict the system's equilibrium state if we can find the equilibrium $N_i^{\bullet}$ values, and *vice versa*. To within an arbitrary factor representing its size, an equilibrated system can be completely described by its intensive properties. In the present instance, the fractions $N_1^{\bullet}/N$, $N_2^{\bullet}/N$, ..., $N_i^{\bullet}/N$, ... describe the equilibrated system to within the factor, N, that specifies its size. Since we infer that $P_i = P(\epsilon_i) = N_i^{\bullet}/N$, the equilibrated system is also described by the probabilities $(P_1, P_2, \ldots,\ P_i, \ldots)$.

Our hypothesis does not assert that the most-probable population set is the only population set possible at equilibrium. A very large number of other population sets may describe an equilibrium system at different instants of time. However, when its state is specified by any such population set, the macroscopic properties of the system are indistinguishable from the macroscopic properties of the system when its state is specified by the most-probable population set. The most-probable population set characterizes the equilibrium state of the system in the sense that we can calculate the properties of the equilibrium state of the macroscopic system by using the single-molecule energy levels and the most probable population set—or its proxy. The relationship between a molecular energy level, ϵ_i, and its equilibrium population, $N_i^\bullet$, is called the ***Boltzmann equation***. From $P_i = N_i^\bullet / N$, we see that the Boltzmann equation specifies the probability of finding a given molecule in energy level ϵ_i.

Although we calculate thermodynamic properties from the most probable population set, the population set that describes the system can vary from instant to instant while the system remains at equilibrium. The central limit theorem enables us to characterize the amount of variation that can occur. When N is comparable to the number of molecules in a macroscopic system, the probability that variation among population sets can result in a macroscopically observable effect is vanishingly small. The hypothesis is successful because the most probable population set is an excellent proxy for any other population set that the equilibrium system is remotely likely to attain.

We develop the theory of statistical thermodynamics for N-molecule systems by considering the energy levels, ϵ_i, available to a single molecule that does not interact with other molecules. Thereafter, we develop a parallel set of statistical thermodynamic results by considering the energy levels, $\hat{E}_i$, available to a system of N molecules. These N-molecule-system energies can reflect the effects of any amount of intermolecular interaction. We can apply the same arguments to find that the Boltzmann equation also describes the equilibrium properties of systems in which intermolecular interactions are important. That is, the probability, $P_i\left(\hat{E}_i\right)$, that an N-molecule system has energy $\hat{E}_i$ is the same function of $\hat{E}_i$ as the molecular-energy probability, $P_i = P\left(\epsilon_i\right)$, is of ϵ_i.

When we finish our development based on single-molecule energy levels, we understand nearly all of the ideas that we need in order to complete the development for the energies of an N-molecule system. This development is an elegant augmentation of the basic argument called the ***ensemble treatment*** or the ***ensemble method***. The ensemble treatment is due to J. Willard Gibbs; we discuss it in Chapter 23. For now, we simply note that our approach involves no wasted effort. When we discuss the ensemble method, we use all of the ideas that we develop in this chapter and the next. The extension of these arguments that is required for the ensemble treatment is so straightforward as to be (almost) painless.

2.4: How can Infinitely Many Probabilities Sum to Unity?

There are an infinite number of successively greater energies for a quantum mechanical system. We infer that the probability that a given energy level is occupied is a property of the energy level. Each of the probabilities must be between 0 and 1. When we sum the fixed probabilities associated with the energy levels, the sum contains an infinite number of terms. By the nature of probability, the sum of this infinite number of terms must be one:

$$\begin{aligned} 1 &= P_1 + P_2 + \cdots + P_i + \ldots \\ &= P(\epsilon_1) + P(\epsilon_2) + \cdots + P(\epsilon_i) + \ldots \\ &= \sum_{i=1}^{\infty} P(\epsilon_i) \end{aligned}$$

That is, the sum of the probabilities is an infinite series, which must converge: The sum of all of the occupancy probabilities must be unity. This can happen only if all later members of the series are very small. In the remainder of this chapter, we explore some of the thermodynamic ramifications of these facts. In the next chapter, we use this relationship to find the functional dependence of the P_i on the energy levels, ϵ_i. To obtain these results, we need to think further about the probabilities associated with the various population sets that can occur. Also, we need to introduce a new fundamental postulate.

To focus on the implications of this sum of probabilities, let us review geometric series. A **geometric series** is a sum of terms, in which each successive term is a multiple of its predecessor. A geometric series is an infinite sum that can converge:

$$T = a + ar + ar^2 + \cdots + ar^i \cdots = a\left(1 + r + r^2 + \cdots + r^i + \ldots\right) = a + a\sum_{i=1}^{\infty} r^i$$

Successive terms approach zero if $|r| < 1$. If $|r| \geq 1$, successive terms do not become smaller, and the sum does not have a finite limit. If $|r| \geq 1$, we say that the infinite series **diverges**.

We can multiply an infinite geometric series by its constant factor to obtain

$$\begin{aligned} rT &= ar + ar^2 + ar^3 + \cdots + ar^i + \ldots \\ &= a\left(r + r^2 + r^3 + \cdots + r^i + \ldots\right) \\ &= a\sum_{i=1}^{\infty} r^i \end{aligned}$$

If $|r| < 1$, we can subtract and find the value of the infinite sum:

$$T - rT = a$$

so that

$$T = a/(1-r)$$

In a geometric series, the ratio of two successive terms is $r^{n+1}/r^n = r$ The condition of convergence for a geometric series can also be written as

$$\left|\frac{r^{n+1}}{r^n}\right| < 1$$

We might anticipate that any other series also converges if its successive terms become smaller at least as fast as those of a geometric series. In fact, this is true and is the basis for the ***ratio test*** for convergence of an infinite series. If we represent successive terms in an infinite series as t_i, their sum is

$$T = \sum_{i=0}^{\infty} t_i$$

The ratio test is a theorem which states that the series converges, and T has a finite value, if

$$\lim_{n\to\infty}\left|\frac{t_{n+1}}{t_n}\right| < 1$$

One of our goals is to discover the relationship between the energy, ϵ_i, of a quantum state and the probability that a molecule will occupy one of the quantum states that have this energy, $P_i = g_i \rho(\epsilon_i)$. When we do so, we find that the probabilities for all of the quantum mechanical systems that we discuss in Chapter 18 satisfy the ratio test.

2.5: The Total Probability Sum at Constant N, V, and T

In a collection of distinguishable independent molecules at constant N, V, and T, the probability that a randomly selected molecule has energy ϵ_i is P_i; we have $1 = P_1 + P_2 + \cdots + P_i + \ldots$. At any instant, every molecule in the N-molecule system has a specific energy, and the state of the system is described by a population set, $\{N_1,\ N_2, \ldots, N_i, \ldots.\}$, wherein N_i can have any value in the range $0 \leq N_i \leq N$, subject to the condition that

$$N = \sum_{i=1}^{\infty} N_i$$

The probabilities that we assume for this system of molecules have the properties we assume in Chapter 19 where we find the total probability sum by raising the sum of the energy-level probabilities to the N^{th} power.

$$1 = \ (P_1 + P_2 + \cdots + P_i + \ldots)^{\ N} = \sum_{\{N_i\}} \frac{N!}{N_1!N_2! \ldots N_i! \ldots} P_1^{N_1} P_2^{N_2} \ldots P_i^{N_i} \ldots$$

The total-probability sum is over all possible population sets, $\{N_1,\ N_2, \ldots, N_i, \ldots.\}$, which we abbreviate to $\{N_i\}$, in indicating the range of the summation. Each term in this sum represents the probability of the corresponding population set $\{N_1,\ N_2, \ldots, N_i, \ldots.\}$,. At any given instant, one of the possible population sets describes the way that the molecules of the physical system are apportioned among the energy levels. The corresponding term in the total probability sum represents the probability of this apportionment. It is not necessary that all of the energy levels be occupied. We can have $N_k = 0$, in which case $P_k^{N_k} = P_k^0 = 1$ and $N_k! = 1$. Energy levels that are not occupied have no effect on the probability of a population set. The unique population set

$$\{N_1^{\bullet},\ N_2^{\bullet}, \ldots, N_i^{\bullet}, \ldots.\}$$

that we conjecture to characterize the equilibrium state is represented by one of the terms in this total probability sum. We want to focus on the relationship between a term in the total probability sum and the corresponding state of the physical system.

Each term in the total probability sum includes a probability factor, $P_1^{N_1} P_2^{N_2} \ldots P_i^{N_i} \ldots$ This factor is the probability that N_i molecules occupy each of the energy levels ϵ_i. This term is not affected by our assumption that the molecules are distinguishable. The probability factor is multiplied by the polynomial coefficient

$$\frac{N!}{N_1!N_2! \ldots N_i! \ldots}$$

This factor is the number of combinations of distinguishable molecules that arise from the population set $\{N_1,\ N_2, \ldots, N_i, \ldots\}$. It is the number of ways that the N distinguishable molecules can be assigned to the available energy levels so that N_1 of them are in energy level, ϵ_1, *etc*.

The combinations for the population set {3,2} are shown in Figure 2.

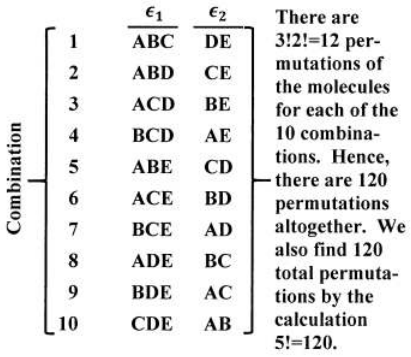


Figure 2. Combinations for the population set {3,2}.

The expression for the number of combinations takes the form it does only because the molecules can be distinguished from one another. To emphasize this point, let us find the number of combinations using the method we develop in Chapter 19. Briefly recapitulated, the argument is this:

1. We can permute the N molecules in $N!$ ways. If we were to distinguish (as different combinations) any two permutations of all of the molecules, this would also be the number of combinations.
2. In fact, however, we do not distinguish between different permutations of those molecules that are assigned to the same energy level. If the N_1 molecules assigned to the first energy level are B, C, Q,..., X, we do not distinguish the permutation $BCQ \ldots X$ from the permutation $CBQ \ldots X$ or from any other permutation of these N_1 molecules. Then the complete set of $N!$ permutations contains a subset of $N_1!$ permutations, all of which are equivalent because they have the same molecules in the first energy level. So the total number of permutations, $N!$, over-counts the number of combinations by a factor of $N_1!$ We can correct for this over-count by dividing by $N_1!$ That is, after correcting for the over-counting for the N_1 molecules in the first energy level, the number of combinations is $N!/N_1!$ (If all N of the molecules were in the first energy level, there would be only one combination. We would have $N = N_1$, and the number of combinations calculated from this formula would be $N!/N! = 1$, as required.)
3. The complete set of $N!$ permutations also includes $N_2!$ permutations of the N_2 molecules in the second energy level. In finding the number of combinations, we want to include only one of these permutations, so correcting for the over-counting due to both the N_1 molecules in the first energy level and the N_2 molecules in the second energy level gives

$$\frac{N!}{N_1!N_2!}$$

4. Continuing this argument through all of the occupied energy levels, we see that the total number of combinations is

$$C\left(N_1, N_2, \ldots, N_i, \ldots\right) = \frac{N!}{N_1!N_2! \ldots N_i! \ldots}$$

Because there are infinitely many energy levels and probabilities, P_i, there are infinitely many terms in the total-probability sum. Every energy available to the macroscopic system is represented by one or more terms in this total-probability sum. Since there is no restriction on the energy levels that can be occupied, there are an infinite number of such system energies. There is an enormously large number of terms each of which corresponds to an enormously large system energy. Nevertheless, the sum of all of these terms must be one. The P_i form a convergent series, and the total probability sum must sum to unity.

Just as the P_i series can converge only if the probabilities of high molecular energies become very small, so the total probability sum can converge only if the probabilities of high system energies become very small. If a population set has N_i molecules in the i^{th} energy level, the probability of that population set is proportional to $P_i^{N_i}$. We see therefore, that the probability of a population set in which there are many molecules in high energy levels must be very small. Terms in the total probability sum that correspond to population sets with many molecules in high energy levels must be negligible. Equivalently, at a particular temperature, macroscopic states in which the system energy is anomalously great must be exceedingly improbable.

What terms in the total probability sum do we need to consider? Evidently from among the infinitely many terms that occur, we can select a finite subset whose sum is very nearly one. If there are many terms that are small and nearly equal to one another, the number of terms in this finite subset could be large. Nevertheless, we can see that terms in this subset must involve the largest possible P_i values raised to the smallest possible powers, N_i, consistent with the requirement that the N_i sum to N.

If an equilibrium macroscopic system could have only one population set, the probability of that population set would be unity. Could an equilibrium system be characterized by two or more population sets for appreciable fractions of an observation period? Would this require that the macroscopic system change its properties with time as it jumps from one population set to another? Evidently, it would not, since our observations of macroscopic systems show that the equilibrium properties are unique. A system that wanders between two (or more) macroscopically distinguishable states cannot be at equilibrium. We are forced to the conclusion that, if a macroscopic equilibrium system has multiple population sets with non-negligible probabilities, the macroscopic properties associated with each of these population sets must be indistinguishably similar. (The alternative is to abandon the theory, which is useful only if its microscopic description of a system makes useful predictions about the system's macroscopic behavior.)

To be a bit more precise about this, we recognize that our theory also rests on another premise: Any intensive macroscopic property of many independent molecules depends on the energy levels available to an individual molecule and the fraction of the molecules that populate each energy level. The average energy is a prime example. For the population set $\{N_1, N_2, \ldots, N_i, \ldots.\}$, the average molecular energy is

$$\bar{\epsilon} = \sum_{i=1}^{\infty} \left(\frac{N_i}{N} \right) \epsilon_i$$

We recognize that many population sets may contribute to the total probability sum at equilibrium. If we calculate essentially the same $\bar{\epsilon}$ from each of these contributing population sets, then all of the contributing population sets correspond to indistinguishably different macroscopic energies. We see in the next section that the central limit theorem guarantees that this happens whenever N is as large as the number of molecules in a macroscopic system.

2.6: The Most Probable Population Set at Constant N, V, and T

We are imagining that we can examine a collection of N distinguishable molecules and determine the energy of each molecule in the collection at any particular instant. If we do so, we find the population set, $\{N_1,\ N_2, \ldots, N_i, \ldots .\}$, that characterizes the system at that instant. In Section 3.9, we introduce the idea that the most probable population set, $\{N_1^{\bullet}, N_2^{\bullet}, \ldots N_i^{\bullet}, ., ., ., \}$, or its proxy, $\{NP\left(\epsilon_1\right), NP\left(\epsilon_2\right), \ldots, NP\left(\epsilon_i\right), \ldots .\}$, is the best prediction we can make about the outcome of a future replication of this measurement. In Section 20.2, we hypothesize that the properties of the system when it is characterized by the most probable population set are indistinguishable from the properties of the system at equilibrium.

Now let us show that this hypothesis is implied by the central limit theorem. We suppose that the population set that characterizes the system varies from instant to instant and that we can find this population set at any given instant. The population set that we find at a particular instant comprises a random sample of N molecular energies. For this sample, we can find the average energy from

$$\bar{\epsilon} = \sum_{i=1}^{\infty} \left(\frac{N_i}{N} \right) \epsilon_i$$

The expected value of the molecular energy is

$$\langle \epsilon \rangle = \sum_{i=1}^{\infty} P_i \epsilon_i$$

It is important that we remember that $\bar{\epsilon}$ and $\langle \epsilon \rangle$ are not the same thing. There is a distribution of $\bar{\epsilon}$ values, one $\bar{\epsilon}$ value for each of the possible population sets $\{N_1,\ N_2, \ldots, N_i, \ldots .\}$. In contrast, when N, V, and T are fixed, the expected value, $\langle \epsilon \rangle$, is a constant; the value of $\langle \epsilon \rangle$ is completely determined by the values of the variables that determine the state of the system and fix the probabilities P_i. If our theory is to be useful, the value of $\langle \epsilon \rangle$ must be the per-molecule energy that we observe for the macroscopic system we are modeling.

According to the central limit theorem, the average energy of a randomly selected sample, $\bar{\epsilon}$, approaches the expected value for the distribution, $\langle \epsilon \rangle$, as the number of molecules in the sample becomes arbitrarily large. In the present instance, we hypothesize that the most probable population set, or its proxy, characterizes the equilibrium system. When N is sufficiently large, this hypothesis implies that the probability of the i^{th} energy level is given by $P_i = N_i^{\bullet}/N$. Then the expected value of a molecular energy is

$$\langle \epsilon \rangle = \sum_{i=1}^{\infty} P_i \epsilon_i = \sum_{i=1}^{\infty} \left(\frac{N_i^{\bullet}}{N} \right) \epsilon_i$$

Since the central limit theorem asserts that $\bar{\epsilon}$ approaches $\langle \epsilon \rangle$ as N becomes arbitrarily large:

$$0 = \lim_{N \to \infty} \left(\bar{\epsilon} - \langle \epsilon \rangle \right) = \lim_{N \to \infty} \sum_{i=1}^{\infty} \left(\frac{N_i}{N} - P_i \right) \varepsilon_i = \lim_{N \to \infty} \sum_{i=1}^{\infty} \left(\frac{N_i}{N} - \frac{N_i^{\bullet}}{N} \right) \epsilon_i$$

One way for the limit of this sum to be zero is for the limit of every individual term to be zero. If the ϵ_i were arbitrary, this would be the only way that the sum could always be zero. However, the ϵ_i and the P_i are related, so we might think that the sum is zero because of these relationships.

To see that the limit of every individual term must in fact be zero, we devise a new distribution. We assign a completely arbitrary number, X_i, to each energy level. Now the i^{th} energy level is associated with an X_i as well as an ϵ_i. We have an X distribution as well as an energy distribution. We can immediately calculate the expected value of X. It is

$$\langle X \rangle = \sum_{i=1}^{\infty} P_i X_i$$

When we find the population set $\{N_1,\ N_2, \ldots, N_i, \ldots .\}$, we can calculate the corresponding average value of X. It is

$$\overline{X} = \sum_{i=1}^{\infty} \left(\frac{N_i}{N} \right) X_i$$

The central limit theorem applies to any distribution. So, it certainly applies to the X distribution; the average value of X approaches the expected value of X as N becomes arbitrarily large:

$$0 = \lim_{N\to\infty}\left(\overline{X} - \langle X\rangle\right) = \lim_{N\to\infty}\sum_{i=1}^{\infty}\left(\frac{N_i}{N} - P_i\right)X_i = \lim_{N\to\infty}\sum_{i=1}^{\infty}\left(\frac{N_i}{N} - \frac{N_i^\bullet}{N}\right)X_i$$

Now, because the X_i can be chosen completely arbitrarily, the only way that the limit of this sum can always be zero is that every individual term becomes zero.

In the limit as $N \to \infty$, we find that

$$N_i/N \to N_i^\bullet/N$$

As the number of molecules in the equilibrium system becomes arbitrarily large, the fraction of the molecules in each energy level at an arbitrarily selected instant approaches the fraction in that energy level in the equilibrium-characterizing most-probable population set, $\{N_1^\bullet, N_2^\bullet, \ldots N_i^\bullet \ldots\}$. In other words, the only population sets that we have any significant chance of observing in a large equilibrium system are population sets whose occupation fractions, N_i/N, are all very close to those, $N_i^\bullet/N$, in the equilibrium-characterizing population set. Estimating P_i as the ratio N_i/N gives essentially the same result whichever of these population sets we use. Below, we see that the ϵ_i and the P_i determine the thermodynamic properties of the system. Consequently, when we calculate any observable property of the macroscopic system, each of these population sets gives the same result.

Since the only population sets that we have a significant chance of observing are those for which

$$N_i/N \approx N_i^\bullet/N$$

we frequently say that we can ignore all but the most probable population set. What we have in mind is that the most probable population set is the only one we need in order to calculate the macroscopic properties of the equilibrium system. We are incorrect, however, if we allow ourselves to think that the most probable population set is necessarily much more probable than any of the others. Nor does the fact that the N_i/N are all very close to the $N_i^\bullet/N$ mean that the N_i are all very close to the $N_i^\bullet$. Suppose that the difference between the two ratios is 10^{-10}. If $N = 10^{20}$, the difference between N_i and $N_i^\bullet$ is 10^{10}, which probably falls outside the range of values that we usually understand by the words “very close.”

We develop a theory that includes a mathematical model for the probability that a molecule has any one of its quantum-mechanically possible energies. It turns out that we are frequently interested in macroscopic systems in which the number of energy levels greatly exceeds the number of molecules. For such systems, we find $NP_i \ll 1$, and it is no longer possible to say that a single most-probable population set, $\{N_1^\bullet, N_2^\bullet, \ldots N_i^\bullet, \ldots\}$, describes the equilibrium state of the system. When it is very unlikely that any energy level is occupied by more than one molecule, the probability of any population set in which any N_i is greater than one becomes negligibly small. We can approximate the total probability sum as

$$1 = \left(P_1 + P_2 + \cdots + P_i + \ldots\right)^N \approx \sum_{\{N_i\}} N! P_1^{N_1} P_2^{N_2} \ldots P_i^{N_i} \ldots$$

However, the idea that the proxy, $\{NP(\epsilon_1), NP(\epsilon_2), \ldots, NP(\epsilon_i), \ldots.\}$, describes the equilibrium state of the system remains valid. In these circumstances, a great many population sets can have essentially identical properties; the properties calculated from any of these are indistinguishable from each other and indistinguishable from the properties calculated from the proxy. Since the equilibrium properties are fixed, the value of these extended products is fixed. For any of the population sets available to such a system at equilibrium, we have

$$P_1^{N_1} P_2^{N_2} \ldots P_i^{N_i} \cdots = P_1^{NP_1} P_2^{NP_2} \ldots P_i^{NP_i} \cdots = \text{constant}$$

It follows that, for some constant, c, we have

$$c = \sum_{i=1}^{\infty} NP_i \ln P_i = N\sum_{i=1}^{\infty} P_i \ln P_i$$

As it evolves, we see that the probability of finding a molecule in an energy level is the central feature of our theory.

2.7: The Microstates of a Given Population Set

Thus far, we have considered only the probabilities associated with the assignments of distinguishable molecules to the allowed energy levels. In Section 20.2, we introduce the hypothesis that all of the g_i degenerate quantum states with energy ϵ_i are equally probable, so that the probability that a molecule has energy ϵ_i is $P_i = P(\epsilon_i) = g_i \rho(\epsilon_i)$. Making this substitution, the total probability sum becomes

$$\begin{aligned} 1 = & \left(P_1 + P_2 + \cdots + P_i + \ldots\right)^N \\ & = \sum_{\{N_i\}} \frac{N!}{N_1! N_2! \ldots N_i! \ldots} P_1^{N_1} P_2^{N_2} \ldots P_i^{N_i} \ldots \\ & = \sum_{\{N_i\}} \frac{N! g_1^{N_1} g_2^{N_2} \ldots g_i^{N_i} \ldots}{N_1! N_2! \ldots N_i! \ldots} \rho(\epsilon_1)^{N_1} \rho(\epsilon_2)^{N_2} \ldots \rho(\epsilon_i)^{N_i} \ldots . \\ & = \sum_{\{N_i\}} N! \prod_{i=1}^{\infty} \left(\frac{g_i^{N_i}}{N_i!}\right) \rho(\epsilon_i)^{N_i} \\ & = \sum_{\{N_i\}} W \prod_{i=1}^{\infty} \rho(\epsilon_i)^{N_i} \end{aligned}$$

where we use the notation

$$a_1 \times a_2 \times \ldots a_i \times \ldots a_\omega \times = \prod_{i=1}^{\omega} a_i$$

for extended products and introduce the function

$$\begin{aligned} W & = W(N_i, g_i) \\ & = W(N_1, g_1, N_2, g_2, \ldots, N_i, g_i, \ldots .) \\ & = N! \prod_{i=1}^{\infty} \left(\frac{g_i^{N_i}}{N_i!}\right) \\ & = C(N_1, N_2, \ldots, N_i, \ldots) \prod_{i=1}^{\infty} g_i^{N_i} \end{aligned}$$

For reasons that become clear later, W is traditionally called the ***thermodynamic probability***. This name is somewhat unfortunate, because W is distinctly different from an ordinary probability.

In Section 20.5, we note that $P_1^{N_1} P_2^{N_2} \ldots P_i^{N_i}$ is the probability that N_i molecules occupy each of the energy levels ϵ_i and that $N!/(N_1! N_2! \ldots N_i! \ldots)$ is the number of combinations of distinguishable molecules that arise from the population set $\{N_1, N_2, \ldots, N_i, \ldots\}$. Now we observe that the extended product

$$\rho(\epsilon_1)^{N_1} \rho(\epsilon_2)^{N_2} \ldots \rho(\epsilon_i)^{N_i} \ldots .$$

is the probability of any one assignment of the distinguishable molecules to quantum states such that N_i molecules are in quantum states whose energies are ϵ_i. Since a given molecule of energy ϵ_i can be in any of the g_i degenerate quantum states, the probability that it is in the energy level ϵ_i is g_i-fold greater that the probability that it is in any one of these quantum states.

Microstates

We call a particular assignment of distinguishable molecules to the available quantum states a ***microstate***. For any population set, there are many combinations. When energy levels are degenerate, each combination gives rise to many microstates. The factor $\rho(\epsilon_1)^{N_1} \rho(\epsilon_2)^{N_2} \ldots \rho(\epsilon_i)^{N_i} \ldots .$ is the probability of any one microstate of the population set $\{N_1, N_2, \ldots, N_i, \ldots\}$. Evidently, the thermodynamic probability

$$W = N! \prod_{i=1}^{\infty} \left(\frac{g_i^{N_i}}{N_i!} \right) \tag{2.7.1}$$

is the total number of microstates of that population set.

To see directly that the number of microstates is dictated by Equation 2.7.1, let us consider the number of ways we can assign N distinguishable molecules to the quantum states when the population set is $\{N_1, N_2, \ldots, N_i, \ldots\}$ and energy level ϵ_i is g_i-fold degenerate. We begin by assigning the N_1 molecules in energy level ϵ_1. We can choose the first molecule from among any of the N distinguishable molecules and can choose to place it in any of the g_1 quantum states whose energy is ϵ_1. The number of ways we can make these choices is Ng_1. We can choose the second molecule from among the $N-1$ remaining distinguishable molecules. In Boltzmann statistics, we can place any number of molecules in any quantum state, so there are again g_1 quantum states in which we can place the second molecule. The total number of ways we can place the second molecule is $(N-1)\, g_1$.

The number of ways the first and second molecules can be chosen and placed is therefore $N\,(N-1)\, g_1^2$. We find the number of ways that successive molecules can be placed in the quantum states of energy ϵ_1 by the same argument. The last molecule whose energy is ϵ_1 can be chosen from among the $(N-N_1+1)$ remaining molecules and placed in any of the g_1 quantum states. The total number of ways of placing the N_1 molecules in energy level ϵ_1 is $N\,(N-1)\,(N-2)\ldots(N-N_1+1)\, g_1^{N_1}$.

This total includes all possible orders for placing every set of N_1 distinguishable molecules into every possible set of quantum states. However, the order doesn't matter; the only thing that affects the state of the system is which molecules go into which quantum state. (When we consider all of the ways our procedure puts all of the molecules into any of the quantum states, we find that any assignment of molecules A, B, and C to any particular set of quantum states occurs six times. Selections in the orders A, B,C; A,C,B; B,A,C; B,C,A; C,A,B; and C,B,A all put the same molecules in the same quantum states.) There are $N_1!$ orders in which our procedure chooses the N_1 molecules; to correct for this, we must divide by $N_1!$, so that the total number of assignments we want to include in our count is

$$N\,(N-1)\,(N-2)\ldots(N-N_1+1)\, g_1^{N_1}/N_1!$$

The first molecule that we assign to the second energy level can be chosen from among the $N-N_1$ remaining molecules and placed into any of the g_2 quantum states whose energy is ϵ_2. The last one can be chosen from among the remaining $(N-N_1-N_2+1)$ molecules. The number of assignments of the N_2 molecules to g_2-fold degenerate quantum states whose energy is ϵ_2 is

$$(N-N_1)\,(N-N_1-1)\ldots(N-N_1-N_2+1)\, g_2^{N_2}/N_2!$$

When we consider the number of assignments of molecules to quantum states with energies ϵ_1 and ϵ_2 we have

$$N\,(N-1)\ldots(N-N_1+1)\,(N-N_1)\,(N-N_1-1)\ldots$$
$$\times\,(N-N_1-N_2+1)\left(\frac{g_1^{N_1}}{N_1!}\right)\left(\frac{g_2^{N_2}}{N_2!}\right)$$

Let the last energy level to contain any molecules be ϵ_ω. The number of ways that the N_ω molecules can be assigned to the quantum states with energy ϵ_ω is $N_\omega\,(N_\omega-1)\ldots(1)\, g_\omega^{N_\omega}/N_\omega!$ The total number of microstates for the population set $\{N_1, N_2, \ldots, N_i, \ldots\}$ becomes

$$N\,(N-1)\ldots(N-N_1)\,(N-N_1-1)\ldots$$
$$\times\,(N_\omega)\,(N_\omega-1)\ldots(1)\prod_{i=1}^{\infty}\left(\frac{g_i^{N_i}}{N_i!}\right) = N!\prod_{i=1}^{\infty}\left(\frac{g_i^{N_i}}{N_i!}\right)$$

When we consider Fermi-Dirac and Bose-Einstein statistics, it is no longer true that the molecules are distinguishable. For Fermi-Dirac statistics, no more than one molecule can be assigned to a particular quantum state. For a given population set, Boltzmann, Fermi-Dirac, and Bose-Einstein statistics produce different numbers of microstates.

It is helpful to have notation that enables us to specify different combinations and different microstates. If ϵ_i is the energy associated with the wave equation that describes a particular molecule, it is convenient to say that the molecule is in energy level ϵ_i; that is, its quantum state is one of those that has energy ϵ_i. Using capital letters to represent molecules, we indicate that

molecule A is in energy level ϵ_i by writing $\epsilon_i\,(A)$. To indicate that A, B, and C are in ϵ_i, we write $\epsilon_i\,(A,B,C)$. Similarly, to indicate that molecules D and E are in ϵ_k, we write $\epsilon_k\,(D,E)$. For this system of five molecules, the assignment $\epsilon_i\,(A,B,C)\,\epsilon_k\,(D,E)$ represents one of the possible combinations. The order in which we present the molecules that have a given energy is immaterial: $\epsilon_i\,(A,B,C)\,\epsilon_k\,(D,E)$ and $\epsilon_i\,(C,B,A)\,\epsilon_k\,(E,D)$ represent the same combination. When any one molecule is distinguishable from others of the same substance, assignments in which a given molecule has different energies are physically different and represent different combinations. The assignments $\epsilon_i\,(A,B,C)\,\epsilon_k\,(D,E)$ and $\epsilon_i\,(D,B,C)\,\epsilon_k\,(A,E)$ represent different combinations. In Figure 2, we represent these assignments more schematically.

Any two assignments in which a particular molecule occupies different quantum states give rise to different microstates. If the i^{th} energy level is three-fold degenerate, a molecule in any of the quantum states $\psi_{i,1}$, $\psi_{i,2}$, or $\psi_{i,3}$ has energy ϵ_i. Let us write

$$\psi_{i,1}\,(A,B)\,\psi_{i,2}\,(C)\,\psi_{k,1}\,(DE)$$

to indicate the microstate arising from the combination $\epsilon_i\,(A,B,C)\,\epsilon_k\,(D,E)$ in which molecules A and B occupy $\psi_{i,1}$, molecule C occupies $\psi_{i,2}$, and molecules D and E occupy $\psi_{k,1}$. Then,

$$\psi_{i,1}\,(A,B)\,\psi_{i,2}\,(C)\,\psi_{k,1}\,(DE)$$

$$\psi_{i,1}\,(B,C)\,\psi_{i,2}\,(A)\,\psi_{k,1}\,(DE)$$

$$\psi_{i,1}\,(A)\,\psi_{i,2}\,(B,C)\,\psi_{k,1}\,(DE)$$

are three of the many microstates arising from the combination $\epsilon_i\,(A,B,C)\,\epsilon_k\,(D,E)$. Figure 3 shows all of the microstates possible for the population set $\{2,1\}$ when the quantum states of a molecule are $\psi_{1,1}$, $\psi_{1,2}$, and $\psi_{2,1}$.

Microstates		$\psi_{1,1}$	$\psi_{1,2}$	$\psi_{2,1}$
	1	AB		C
	2	A	B	C
	3	B	A	C
	4		AB	C
	5	AC		B
	6	A	C	B
	7	C	A	B
	8		AC	B
	9	BC		A
	10	B	C	A
	11	C	B	A
	12		BC	A

Figure 3. Microstates for {2,1} with quantum states $\Psi_{1,1}$, $\Psi_{1,2}$, and $\Psi_{2,1}$.

2.8: The Probabilities of Microstates that Have the Same Energy

In Section 20.2, we introduce the assumption that, for a molecule in a constant-N-V-T system, for which the g_i and ϵ_i are fixed, the probability of a quantum state, $\rho(\epsilon_i)$, depends only on its energy. It follows that two or more quantum states that have the same energy must have equal probabilities. We accept the idea that the probability depends only on energy primarily because we cannot see any reason for a molecule to prefer one state to another if both states have the same energy.

We extend this thinking to multi-molecule systems. If two microstates have the same energy, we cannot see any reason for the system to prefer one rather than the other. In a constant-N-V-T system, in which the total energy is not otherwise restricted, each microstate of $\{N_1,\ N_2, \ldots, N_i, \ldots\}$ occurs with probability $\rho(\epsilon_1)^{N_1}\rho(\epsilon_2)^{N_2}\ldots\rho(\epsilon_i)^{N_i}\ldots$, and each microstate of $\{N_1^{\#}, N_2^{\#},\ \ldots, N_I^{\#}, \ldots\}$ occurs with probability $\rho(\epsilon_1)^{N_1^{\#}}\rho(\epsilon_2)^{N_2^{\#}}\ldots\rho(\epsilon_i)^{N_i^{\#}}\ldots$. When the energies of these population sets are equal, we infer that these probabilities are equal, and their value is a constant of the system. That is,

$$\rho(\epsilon_1)^{N_1}\rho(\epsilon_2)^{N_2}\ldots\rho(\epsilon_i)^{N_i}\ldots$$

$$=\rho(\epsilon_1)^{N_1^{\#}}\rho(\epsilon_2)^{N_2^{\#}}\ldots\rho(\epsilon_i)^{N_i^{\#}}\ldots \quad =\rho_{MS,N,E}=\text{constant}$$

where we introduce $\rho_{MS,N,E}$ to represent the probability of a microstate of a system of N molecules that has total energy E. If $E=E^{\#}$, then $\rho_{MS,N,E}=\rho_{MS,N,E^{\#}}$.

When we think about it critically, the logical basis for this equal-probability idea is not very impressive. While the idea is plausible, it is not securely rooted in any particular empirical observation or prior postulate. The equal-probability idea is useful only if it leads us to theoretical models that successfully mirror the behavior of real macroscopic systems. This it does. Accordingly, we recognize that the equal-probability idea is really a fundamental postulate about the behavior of quantum-mechanical systems. It is often called the *principle of equal a priori probabilities*:

Definition: principle of equal *a priori* probabilities

For a particular system, all microstates that have the same energy have the same probability.

Our development of statistical thermodynamics relies on the principle of equal *a priori* probabilities. For now, let us summarize the important relationships that the principle of equal *a priori* probabilities imposes on our microscopic model for the probabilities of two population sets of a constant-N-V-T system that have the same energy:

- A given population set $\{N_1,\ N_2, \ldots, N_i, \ldots\}$ gives rise to $W(N_i, g_i)$ microstates, and each of these microstates has energy

$$E=\sum_{i=1}^{\infty}N_i\epsilon_i$$

- A second population set, $\{N_1^{\#}, N_2^{\#},\ \ldots, N_I^{\#}, \ldots\}$, that has the same energy need not—and usually will not—give rise to the same number of microstates. In general, for two such population sets,

$$W(N_i, g_i)\neq W\left(N_i^{\#}, g_i\right)$$

  However, because each microstate of either population set has the same energy, we have

$$E=\sum_{i=1}^{\infty}N_i\epsilon_i=\sum_{i=1}^{\infty}N_i^{\#}\epsilon_i$$

- The probability of a microstate of a given population set $\{N_1,\ N_2, \ldots, N_i, \ldots\}$ depends only on its energy:

$$\rho(\epsilon_1)^{N_1}\rho(\epsilon_2)^{N_2}\ldots\rho(\epsilon_i)^{N_i}\ldots=\rho_{MS,N,E}=\text{constant}$$

2.9: The Probabilities of the Population Sets of an Isolated System

In principle, the energy of an equilibrium system that is in contact with a constant-temperature heat reservoir can vary slightly with time. In contrast, the energy of an isolated system is constant. A more traditional and less general statement of the equal *a priori* probability principle focuses on isolated systems, for which all possible microstates necessarily have the same energy:

> *All microstates of an isolated (constant energy) system occur with equal probability.*

If we look at the fraction of the molecules of an isolated system that are in each microstate, we expect to find that these fractions are approximately equal. In consequence, for an isolated system, the probability of a population set, $\{N_1,\ N_2, \ldots, N_i, \ldots\}$, is proportional to the number of microstates, $W\left(N_i, g_i\right)$, to which that population set gives rise.

In principle, the population sets of a constant-N-V-T system can be significantly different from those of a constant-N-V-E system. That is, if we move an isolated system, whose temperature is T, into thermal contact with a heat reservoir at constant-temperature T, the population sets that characterize the system can change. In practice, however, for a system containing a large number of molecules, the population sets that contribute to the macroscopic properties of the system must be essentially the same.

The fact that the same population sets are important in both systems enables us to make two further assumptions that become important in our development. We assume that the proportionality between the probability of a population set and $W\left(N_i, g_i\right)$, which is strictly true only for a constant-N-V-E system, is also true for the corresponding constant-N-V-T system. We also assume that the probabilities of a quantum state, $\rho\left(\epsilon_i\right)$, and a microstate, $\rho_{MS,N,E}$, which we defined for the constant-N-V-T system, are the same for the corresponding constant-N-V-E system.

Let us see why we expect the same population sets to dominate the macroscopic properties of otherwise identical constant-energy and constant-temperature systems. Suppose that we isolate a constant-N-V-T system in such a way that the total energy, $E = \sum_{i=1}^{\infty} N_i \epsilon_i$, of the isolated system is exactly equal to the expected value, $\langle E \rangle = N \sum_{i=1}^{\infty} P_i \epsilon_i$, of the energy of the system when its temperature is constant. What we have in mind is a *gedanken* experiment, in which we monitor the energy of the thermostatted system as a function of time, waiting for an instant in which the system energy, $E = \sum_{i=1}^{\infty} N_i \epsilon_i$, is equal to the expected value of the system energy, $\langle E \rangle$. When this occurs, we instantaneously isolate the system.

We suppose that the isolation process is accomplished before any molecule can experience an energy change, so that the population set that characterizes the system immediately afterwards is the same as the one that characterizes it before. After isolation, of course, the molecules can exchange energy with one another, and many population sets may be available to the system.

Clearly, the value of every macroscopic property of the isolated system must be the same as its observable value in the original constant-temperature system. Our microscopic description of it is different. Every population set that is available to the isolated system has energy $E = \langle E \rangle$, and gives rise to

$$W\left(N_i, g_i\right) = N! \prod_{i=1}^{\infty} \left(\frac{g_i^{N_i}}{N_i!} \right)$$

microstates. At the same temperature, each of these microstates occurs with the same probability. Since the isolated-system energy is $\langle E \rangle$, this probability is $\rho_{MS,N,\langle E \rangle}$. The probability of an available population set is $W\left(N_i, g_i\right) \rho_{MS,N,\langle E \rangle}$.

Since the temperature can span a range of values centered on $\langle T \rangle$, where $\langle T \rangle$ is equal to the temperature of the original constant-N-V-T system, there is a range of $\rho_{MS,N,\langle E \rangle}$ values spanning the (small) range of temperatures available to the constant-energy system. Summing over all of the population sets that are available to the isolated system, we find

$$1 = \sum_{\{N_i\},\ E=\langle E \rangle, T=\langle T \rangle} W\left(N_i, g_i\right) \rho_{MS,N,\langle E \rangle} + \sum_{\{N_i\},\ E=\langle E \rangle, T \neq \langle T \rangle} W\left(N_i, g_i\right) \rho_{MS,N,\langle E \rangle}$$

The addition of "$E = \langle E \rangle$" beneath the summation sign emphasizes that the summation is to be carried out over the population sets that are consistent with both the molecule-number and total-energy constraints and no others. The total probability sum breaks into two terms, one spanning population sets whose temperature is exactly $\langle T \rangle$ and another spanning all of the other population sets. (Remember that the $\rho\left(\epsilon_i\right)$ are temperature dependent.)

The population sets available to the isolated system are slightly different from those available to the constant-temperature system. In our microscopic model, only population sets that have exactly the right total energy can occur in the isolated system. Only population sets that have exactly the right temperature can occur in the constant-temperature system.

Summing over all of the population sets that are available to the constant-temperature system, we partition the total probability sum into two terms:

$$1 = \sum_{\{N_i\}, E=\langle E\rangle, T=\langle T\rangle} W\left(N_i, g_i\right) \rho_{MS,N,\langle E\rangle} + \sum_{\{N_i\},\ E\neq\langle E\rangle, T=\langle T\rangle} W\left(N_i, g_i\right) \rho_{MS,N,\langle E\rangle}$$

From the central limit theorem, we expect the constant-energy system to have (relatively) few population that fail to meet the condition $E = \langle E\rangle$. Likewise, we expect the constant temperature system to have (relatively) few population sets that fail to meet the condition $T = \langle T\rangle$. The population sets that satisfy both of these criteria must dominate both sums. For the number of molecules in macroscopic systems, we expect the approximation to the total probability sum

$$1 = \sum_{\{N_i\}, E} W\left(N_i, g_i\right) \rho_{MS,N,\langle E\rangle} \approx \sum_{\{N_i\}, E=\langle E\rangle, T=\langle T\rangle} W\left(N_i, g_i\right) \rho_{MS,N,\langle E\rangle}$$

to be very good. The same population sets dominate both the constant-temperature and constant-energy systems. Each system must have a most probable population set, $\{N_1^\bullet, N_2^\bullet, \ldots N_i^\bullet, , , , \}$. If these are not identically the same set, they must be so close that the same macroscopic properties are calculated using either one.

Thus, the central limit theorem implies that the total probability sum, which we develop for the constant-temperature system, also describes the constant-energy system, so long as the number of molecules in the system is sufficiently large.

Now, two aspects of this development warrant elaboration. The first is that the probability of population sets that have energies and temperature that satisfy $E = \langle E\rangle$ and $T = \langle T\rangle$ *exactly* may actually be much less than one. The second is that constant-energy and constant-temperature systems are creatures of theory. No real system can actually have an *absolutely* constant energy or temperature.

Recognizing these facts, we see that when we stipulate $E = \langle E\rangle$ or $T = \langle T\rangle$, what we really mean is that $E = \langle E\rangle \pm \delta E$ and $T = \langle T\rangle \pm \delta T$, where the intervals $\pm\delta E$ and $\pm\delta T$ are vastly smaller than any differences we could actually measure experimentally. When we write $E \neq \langle E\rangle$ and $T \neq \langle T\rangle$, we really intend to specify energies and temperatures that fall outside the intervals $E = \langle E\rangle \pm \delta E$ and $T = \langle T\rangle \pm \delta T$. If the system contains sufficiently many molecules, the population sets whose energies and temperatures fall within the intervals $E = \langle E\rangle \pm \delta E$ and $T = \langle T\rangle \pm \delta T$ account for nearly all of the probability—no matter how small we choose δE and δT. All of the population sets whose energies and temperatures fall within the intervals $E = \langle E\rangle \pm \delta E$ and $T = \langle T\rangle \pm \delta T$ correspond to the same macroscopically observable properties.

2.10: Entropy and Equilibrium in an Isolated System

In an isolated system, the probability of population set $\{N_1,\ N_2, \ldots, N_i, \ldots\}$ is $W\left(N_i, g_i\right) \rho_{MS,N,\langle E\rangle}$, where $\rho_{MS,N,\langle E\rangle}$ is a constant. It follows that $W = W\left(N_i, g_i\right)$ is proportional to the probability that the system is in one of the microstates associated with the population set $\{N_1,\ N_2, \ldots, N_i, \ldots\}$. Likewise, $W^{\#} = W\left(N_i^{\#}, g_i\right)$ is proportional to the probability that the system is in one of the microstates associated with the population set $\{N_1^{\#}, N_2^{\#}, \ldots N_i^{\#}, \ldots\}$. Suppose that we observe the isolated system for a long time. Let F be the fraction of the time that the system is in microstates of population set $\{N_1,\ N_2, \ldots, N_i, \ldots\}$ and $F^{\#}$ be the fraction of the time that the system is in microstates of the population set $\{N_1^{\#}, N_2^{\#}, \ldots N_i^{\#}, \ldots\}$. The principle of equal *a priori* probabilities implies that we would find

$$\frac{F^{\#}}{F} = \frac{W^{\#}}{W}$$

Suppose that $W^{\#}$ is much larger than W. This means there are many more microstates for $\{N_1^{\#}, N_2^{\#}, \ldots N_i^{\#}, \ldots\}$ than there are for $\{N_1,\ N_2, \ldots, N_i, \ldots\}$. The fraction of the time that the population set $\{N_1^{\#}, N_2^{\#}, \ldots N_i^{\#}, \ldots\}$ characterizes the system is much greater than the fraction of the time $\{N_1,\ N_2, \ldots, N_i, \ldots\}$ characterizes it. Alternatively, if we examine the system at an arbitrary instant, we are much more likely to find the population set $\{N_1^{\#}, N_2^{\#}, \ldots N_i^{\#}, \ldots\}$ than the population set $\{N_1,\ N_2, \ldots, N_i, \ldots\}$. The larger $W\left(N_1, g_1,\ N_2, g_2, \ldots, N_i, g_i, \ldots\right)$, the more likely it is that the system will be in one of the microstates associated with the population set $\{N_1,\ N_2, \ldots, N_i, \ldots\}$. In short, W predicts the state of the system; it is a measure of the probability that the macroscopic properties of the system are those of the population set $\{N_1,\ N_2, \ldots, N_i, \ldots\}$.

If an isolated system can undergo change, and we re-examine it at after a few molecules have moved to different energy levels, we expect to find it in one of the microstates of a more-probable population set; that is, in one of the microstates of a population set for which W is larger. At still later times, we expect to see a more-or-less smooth progression: the system is in microstates of population sets for which the values of W are increasingly larger. This can continue only until the system occupies one of the microstates of the population set for which W is a maximum or a microstate of one of the population sets whose macroscopic properties are essentially the same as those of the constant-N-V-E population set for which W is a maximum.

Once this occurs, later inspection may find the system in other microstates, but it is overwhelmingly probable that the new microstate will still be one of those belonging to the largest-W population set or one of those that are macroscopically indistinguishable from it. Any of these microstates will belong to a population set for which W is very well approximated by $W\left(\ N_1^{\bullet}, g_1,\ N_2^{\bullet}, g_2, \ldots, N_i^{\bullet}, g_i, \ldots\right)$. Evidently, the largest-W population set characterizes the equilibrium state of the either the constant-N-V-T system or the constant–N-V-E system. Either system can undergo change until W reaches a maximum. Thereafter, it is at equilibrium and can undergo no further macroscopically observable change.

Boltzmann recognized this relationship between W, the thermodynamic probability, and equilibrium. He noted that the unidirectional behavior of W in an isolated system undergoing spontaneous change is like the behavior we found for the entropy function. Boltzmann proposed that, for an isolated (constant energy) system, S and W are related by the equation $S = k \ln W$, where k is Boltzmann's constant. This relationship associates an entropy value with every population set. For an isolated macroscopic system, equilibrium corresponds to a state of maximum entropy. In our microscopic model, equilibrium corresponds to the population set for which W is a maximum. By the argument we make in §6, this population set must be well approximated by the most probable population set, $\{N_1^{\bullet}, N_2^{\bullet}, \ldots N_i^{\bullet}, , , , , \}$. That is, the entropy of the equilibrium state of the macroscopic system is

$$\begin{aligned} S &= k \ln W_{max} \\ &= k \ln \frac{N!}{N_i^{\bullet}! N_i^{\bullet}! \ldots N_i^{\bullet}! \ldots} + k \sum_{i=1}^{\infty} N_i^{\bullet} \ln g_i \end{aligned}$$

This equation can be taken as the definition of entropy. Clearly, this definition is different from the thermochemical definition, $S = q^{rev}/T$. We can characterize—imperfectly—the situation by saying that the two definitions provide alternative scales for measuring the same physical property. As we see below, our statistical theory enables us to define entropy in still more ways, all of which prove to be functionally equivalent. Gibbs characterized these alternatives as "entropy analogues;" that is, functions whose properties parallel those of the thermochemically defined entropy.

We infer that the most probable population set characterizes the equilibrium state of either the constant-temperature or the constant-energy system. Since our procedure for isolating the constant-temperature system affects only the thermal interaction between the system and its surroundings, the entropy of the constant-temperature system must be the same as that of the constant-energy system. Using $N_i^* = NP_i = Ng_i\rho(\epsilon_i)$ and assuming that the approximation $\ln N_i^*! = N_i^* \ln N_i^* - N_i^*$ is adequate for all of the energy levels that make a significant contribution to S, substitution shows that the entropy of either system depends only on probabilities:

$$\begin{aligned} S &= kN\ln N - kN - k\sum_{i=1}^{\infty}[NP_i\ln(NP_i) - NP_i] + k\sum_{i=1}^{\infty} NP_i \ln g_i \\ &= kN\ln N - \mathrm{kN} - kN\sum_{i=1}^{\infty}[P_i\ln(N) + P_i\ln P_i - P_i - P_i\ln g_i] \\ &= k(N\ln N - N) - k(N\ln N - N)\sum_{i=1}^{\infty} P_i - kN\sum_{i=1}^{\infty} P_i[\ln P_i - \ln g_i] = -kN\sum_{i=1}^{\infty} P_i \ln\rho(\epsilon_i) \end{aligned}$$

The entropy per molecule, S/N, is proportional to the expected value of $\ln\rho(\epsilon_i)$; Boltzmann's constant is the proportionality constant. At constant temperature, $\rho(\epsilon_i)$ depends only on ϵ_i. The entropy per molecule depends only on the quantum state properties, g_i and ϵ_i.

2.11: Thermodynamic Probability and Equilibrium in an Isomerization Reaction

To relate these ideas to a change in a more specific macroscopic system, let us consider isomeric substances A and B. (We consider this example further in Chapter 21.) In principle, we can solve the Schrödinger equation for a molecule of isomer A and for a molecule of isomer B. We obtain all possible energy levels for a molecule of each isomer.[1] If we list these energy levels in order, beginning with the lowest, some of these levels belong to isomer A and the others belong to isomer B.

Now let us consider a mixture of N_A molecules of A and N_B molecules of B. We suppose that individual molecules are distinguishable and that intermolecular interactions can be ignored. Since a group of atoms that can form an A molecule can also form a B molecule, every energy level is accessible to this group of atoms; that is, we can view both sets of energy levels as being available to the atoms that make up the molecules. For a given system energy, there will be many population sets in which only the energy levels belonging to isomer A are occupied. For each of these population sets, there is a corresponding thermodynamic probability, W. Let W_A^{max} be the largest of these thermodynamic probabilities. Similarly, there will be many population sets in which only the energy levels corresponding to isomer B are occupied. Let W_B^{max} be the largest of the thermodynamic probabilities associated with these population sets. Finally, there will be many population sets in which the occupied energy levels belong to both isomer A and isomer B. Let $W_{A,B}^{max}$ be the largest of the thermodynamic probabilities associated with this group of population sets.

Now, W_A^{max} is a good approximation to the number of ways that the atoms of the system can come together to form isomer A. W_B^{max} is a good approximation to the the number of ways that the atoms of the system can come together to form isomer B. At equilibrium, therefore, we expect

$$K = \frac{N_B}{N_A} = \frac{W_B^{max}}{W_A^{max}}$$

If we consider the illustrative—if somewhat unrealistic—case of isomeric molecules whose energy levels all have the same degeneracy ($g_i = g$ for all i), we can readily see that the equilibrium system must contain some amount of each isomer. For a system containing N molecules, $N!g^N$ is the numerator in each of the thermodynamic probabilities W_A^{max}, W_B^{max}, and $W_{A,B}^{max}$. The denominators are different. The denominator of $W_{A,B}^{max}$ must contain terms, $N_i!$, for essentially all of the levels represented in the denominator of W_A^{max}. Likewise, it must contain terms, $N_j!$, for essentially all of the energy levels represented in the denominator of W_B^{max}. Then the denominator of $W_{A,B}^{max}$ is a product of $N_k!$ terms that are generally smaller than the corresponding factorial terms in the denominators of W_A^{max} and W_B^{max}. As a result, the denominators of W_A^{max} and W_B^{max} are larger than the denominator of $W_{A,B}^{max}$. In consequence, $W_{A,B}^{max} > W_A^{max}$ and $W_{A,B}^{max} > W_B^{max}$. (See problems 5 and 6.)

If we create the system as a collection of A molecules, or as a collection of B molecules, redistribution of the sets of atoms among all of the available energy levels must eventually produce a mixture of A molecules and B molecules. Viewed as a consequence of the principle of equal *a priori* probabilities, this occurs because there are necessarily more microstates of the same energy available to some mixture of A and B molecules than there are microstates available to either A molecules alone or B molecules alone. Viewed as a consequence of the tendency of the isolated system to attain the state of maximum entropy, this occurs because $k\ln W_{A,B}^{max} > k\ln W_A^{max}$ and $k\ln W_{A,B}^{max} > k\ln W_B^{max}$.

2.12: The Degeneracy of an Isolated System and Its Entropy

In Section 20.9, we find that the sum of the probabilities of the population sets of an isolated system is

$$1 = \sum_{\{N_i\},E} W\left(N_i, g_i\right) \rho_{MS,N,E}.$$

By the principle of equal *a priori* probabilities, $\rho_{MS,N,E}$ is a constant, and it can be factored out of the sum. We have

$$1 = \rho_{MS,N,E} \sum_{\{N_i\},E} W\left(N_i, g_i\right)$$

Moreover, the sum of the thermodynamic probabilities over all allowed population sets is just the number of microstates that have energy E. This sum is just the ***degeneracy of the system energy***, E. The symbol Ω_E is often given to this system-energy degeneracy. That is,

$$\Omega_E = \sum_{\{N_i\},E} W\left(N_i, g_i\right)$$

The sum of the probabilities of the population sets of an isolated system becomes

$$1 = \rho_{MS,N,E} \Omega_E$$

In Section 20.9, we infer that

$$\rho_{MS,N,E} = \prod_{i=1}^{\infty} \rho(\epsilon_i)^{N_i}$$

so we have

$$1 = \Omega_E \prod_{i=1}^{\infty} \rho(\epsilon_i)^{N_i}$$

2.13: The Degeneracy of an Isolated System and its Entropy

In Section 20.10, we observe that the entropy of an isolated equilibrium system can be defined as $S = k\ln W_{max}$. In Section 20.12, we see that the system-energy degeneracy is a sum of terms, one of which is $W_{max} = W\left(N_i^*, g_i\right)$. That is, we have

$$\Omega_E = W_{max} + \sum_{\{N_i\} \neq \left\{N_i^*\right\}, E_{total}} W\left(N_i, g_i\right)$$

where the last sum is taken over all energy-qualifying population sets other than the most-probable population set.

Let us now consider the relative magnitude of Ω_E and W_{max}. Clearly, $\Omega_E \geq W_{max}$. If only one population set is consistent with the total-molecule and total-energy constraints of the isolated system, then $\Omega_E = W_{max}$. In general, however, we must expect that there will be many, possibly an enormous number, of other population sets that meet the constraints. Ultimately, the relative magnitude of Ω_E and W_{max} depends on the energy levels available to the molecules and the number of molecules in the system and so could be almost anything. However, rather simple considerations lead us to expect that, for most macroscopic collections of molecules, the ratio $\alpha = \Omega_E / W_{max}$ will be much less than W_{max}. That is, although the value of α may be very large, for macroscopic systems we expect to find $\alpha \ll W_{max}$. If $\Omega_E = W_{max}$, then $\alpha = 1$, and $\ln \alpha = 0$.

Because W for any population set that contributes to Ω_E must be less than or equal to W_{max}, the maximum value of α must be less than the number of population sets which satisfy the system constraints. For macroscopic systems whose molecules have even a modest number of accessible energy levels, calculations show that W_{max} is a very large number indeed. Calculation of α for even a small collection of molecules is intractable unless the number of accessible molecular energy levels is small. Numerical experimentation on small systems, with small numbers of energy levels, shows that the number of qualifying population sets increases much less rapidly than W_{max} as the total number of molecules increases. Moreover, the contribution that most qualifying population sets make to Ω_E is much less than W_{max}.

For macroscopic systems, we can be confident that W_{max} is enormously greater than α. Hence Ω_E is enormously greater than α. When we substitute for W_{max} in the isolated-system entropy equation, we find

$$\begin{aligned} S &= k \ln W_{max} \\ &= k \ln(\Omega_E / \alpha) \\ &= k\ln\Omega_E - k \ln \alpha \\ &\approx k \ln \Omega_E \end{aligned}$$

where the last approximation is usually very good.

In many developments, the entropy of an isolated system is defined by the equation $S = k\ln \Omega_E$ rather than the equation we introduced first, $S = k\ln W_{max}$. From the considerations above, we expect the practical consequences to be the same. In Section 20.14, we see that the approximate equality of $\ln W_{max}$ and $\ln \Omega_E$ is a mathematical consequence of our other assumptions and approximations.

2.14: Effective Equivalence of the Isothermal and Constant-energy Conditions

In principle, an isolated system is different from a system with identical macroscopic properties that is in equilibrium with its surroundings. We emphasize this point, because this distinction is important in the logic of our development. However, our development also depends on the assumption that, when N is a number that approximates the number molecules in a macroscopic system, the constant-temperature and constant-energy systems are functionally equivalent.

In Section 20.9, we find that any calculation of macroscopic properties must produce the same result whether we consider the constant-temperature or the constant-energy system. The most probable population set, $\{N_1^\bullet, N_2^\bullet, \ldots N_i^\bullet, \ldots\}$, provides an adequate description of the macroscopic state of the constant-temperature system precisely because it is representative of all the population sets that contribute significantly to the total probability of the constant-temperature system. The effective equivalence of the constant-temperature and constant-energy systems ensures that the most probable population set is also representative of all the population sets that contribute significantly to the total probability of the constant-energy system.

In Section 20.12, we see that the essential equivalence of the isothermal and constant-energy systems means that we have

$$1 = \Omega_E \prod_{i=1}^{\infty} \rho(\epsilon_i)^{N_i^\bullet}$$

Taking logarithms of both sides, we find

$$\ln \Omega_E = -\sum_{i=1}^{\infty} N_i^\bullet \ln \rho\left(\epsilon_i\right)$$

From $S = k \ln \Omega_E$, it follows that

$$\mathrm{S} = -k \sum_{i=1}^{\infty} N_i^\bullet \ln \rho\left(\epsilon_i\right)$$

For the constant-temperature system, we have $N_i^\bullet = N P_i$. When we assume that the equilibrium constant-temperature and constant-energy systems are essentially equivalent, the entropy of the N-molecule system becomes

$$\begin{aligned} S &= -k \sum_{i=1}^{\infty} N_i^\bullet \ln \rho\left(\epsilon_i\right) \\ &= -kN \sum_{i=1}^{\infty} P_i \ln \rho\left(\epsilon_i\right) \end{aligned}$$

so that we obtain the same result from assuming that $S = k \ln \Omega_E$ as we do in Section 20.10 from assuming that $S = k \ln W_{max}$. Under the approximations we introduce, $\ln \Omega_E$ and $\ln W_{max}$ evaluate to the same thing.

2.15: Problems

1. Three non-degenerate energy levels are available to a set of five distinguishable molecules, $\{A,\ B,\ C,\ D,\ E\}$. The energies of these levels are 1, 2, and 3, in arbitrary units. Find all of the population sets that are possible in this system. For each population set, find the system energy, E, and the number of microstates, W. For each system energy, E, list the associated population sets and the total number of microstates. How many population sets are there? What is W_{max}? If this system is isolated with $E = 10$, how many population sets are possible? What is Ω_E for $E = 10$?

2. For the particle in a box, the allowed energies are proportional to the squares of the successive integers. What population sets are possible for the distinguishable molecules, $\{A,\ B,\ C,\ D,\ E\}$, if they can occupy three quantum states whose energies are 1, 4, and 9? For each population set, find the system energy, E, and the number of microstates. For each system energy, E, list the associated population sets and the total number of microstates. How many population sets are there? What is W_{max}? If this system is isolated with $E = 24$, how many population sets are possible? What is Ω_E for $E = 24$?

3. Consider the results you obtained in problem 2. In general, when the allowed energies are proportional to the squares of successive integers, how many population sets do you think will be associated with each system energy?

4.

(a) Compare W for the population set $\{3, 3, 3\}$ to W for the population set $\{2, 5, 2\}$. The energy levels are non-degenerate.

(b) Consider an N-molecule system that has a finite number, M, of quantum states. Show that W is (at least locally) a maximum when $N_1 = N_2 = \cdots = N_M = N/M$. (Hint: Let $U = N/M$, and assume that N can be chosen so that U is an integer. Let

$$W_U = N!/\left[U!U!\prod_{i=1}^{i=M-2} U!\right]$$

and let

$$W_O = N!/\left[(U+1)!\,(U-1)!\prod_{i=1}^{i=M-2} U!\right]$$

Show that $W_O/W_U < 1$.)

5. The energy levels available to isomer A are $\epsilon_0 = 1$, $\epsilon_2 = 2$, and $\epsilon_4 = 3$, in arbitrary units. The energy levels available to isomer B are $\epsilon_1 = 2$, $\epsilon_3 = 3$, and $\epsilon_5 = 4$. The energy levels are non-degenerate.

(a) A system contains five molecules. The energy of the system is 10. List the population sets that are consistent with $N = 5$ and $E = 10$. Find W for each of these population sets. What are $W_{A,B}^{max}$, W_A^{max}, and W_B^{max}? What is the total number of microstates, $= \Omega_{A,B}$, available to the system in all of the cases in which A and B molecules are present? What is the ratio $\Omega_{A,B}/W_{A,B}^{max}$?

(b) Repeat this analysis for a system that contains six molecules and whose energy is 12.

(c) Would the ratio $\Omega_{A,B}/W_{A,B}^{max}$ be larger or smaller for a system with $N = 50$ and $E = 100$?

(d) What would happen to this ratio if the number of molecules became very large, while the average energy per molecule remained the same?

6. In Section 20.11, we assume that all of the energy levels available to an isomeric pair of molecules have the same degeneracy. We then argue that the thermodynamic probabilities of a mixture of the isomers must be greater than the thermodynamic probability of either pure isomer: $W_{A,B}^{max} > W_A^{max}$ and $W_{A,B}^{max} > W_B^{max}$. Implicitly, we assume that many energy levels are multiply occupied: $N_i > 1$ for many energy levels ϵ_i. Now consider the case that $g_i > 1$ for most ϵ_i, but that nearly all energy levels are either unoccupied or contain only one molecule: $N_i = 0$ or $N_i = 1$. Show that under this assumption also, we must have $W_{A,B}^{max} > W_A^{max}$ and $W_{A,B}^{max} > W_B^{max}$.

Notes

[1]The statistical-mechanical procedures that have been developed for finding the energy levels available to a molecule express molecular energies as the difference between the molecule energy and the energy that its constituent atoms have when they are motionless. This is usually effected in two steps. The molecular energy levels are first expressed relative to the energy of the molecule's own lowest energy state. The energy released when the molecules is formed in its lowest energy state from the isolated

constituent atoms is then added. The energy of each level is then equal to the work done on the component atoms when they are brought together from infinite separation to form the molecule in that energy level. (Since energy is released in the formation of a stable molecule, the work done on the atoms and the energy of the resulting molecule are less than zero.) In our present discussion, we suppose that we can solve the Schrödinger equation to find the energies of the allowed quantum states. This corresponds to choosing the isolated constituent electrons and nuclei as the zero of energy for both isomers.

CHAPTER OVERVIEW

3: The Boltzmann Distribution Function

3.1: Finding the Boltzmann Equation

The probabilities of the energy levels of a constant-temperature system at equilibrium must depend only on the intensive variables that serve to characterize the equilibrium state. In Section 20.8, we introduce the principle of equal *a priori* probabilities, which asserts that any two microstates of an isolated system have the same probability. From the central limit theorem, we infer that an isolated system is functionally equivalent to a constant-temperature system when the system contains a sufficiently large number of molecules. From these ideas, we can now find the relationship between the energy values, ϵ_i, and the corresponding probabilities,

$$P_i = P(\epsilon_i) = g_i \rho(\epsilon_i).$$

Let us consider the microstates of an isolated system whose energy is $E^{\#}$. For any population set, $\{N_1,\ N_2, \ldots, N_i, \ldots\}$, that has energy $E^{\#}$, the following relationships apply.

1. The sum of the energy-level populations is the total number of molecules:

$$N = N_1 + N_2 + \cdots + N_i = \sum_{j=1}^{\infty} N_j$$

2. The energy of the system is the sum of the energies of its constituent molecules:

$$E^{\#} = N_1\epsilon_1 + N_2\epsilon_2 + \cdots + N_i\epsilon_i = \sum_{j=1}^{\infty} N_j\epsilon_j$$

3. The product of powers of quantum-state probabilities is a constant:

$$\rho(\epsilon_1)^{N_1}\rho(\epsilon_1)^{N_1}\ldots\rho(\epsilon_1)^{N_1}\cdots = \kappa$$

   or, equivalently,

$$\begin{aligned} N_1\ln\rho(\epsilon_1)\ + N_2\ln\rho(\epsilon_2)\ + \cdots + N_i\ln\rho(\epsilon_i)\ + \ldots\ &= \sum_{i=1}^{\infty} N_i\ln\rho(\epsilon_i) \\ &= \ln\kappa \end{aligned}$$

4. For the system at constant temperature, the sum of the energy-level probabilities is one. When we infer that the constant-temperature system and the isolated system are functionally equivalent, we assume that this is true also for the isolated system:

$$1 = P(\epsilon_1) + P(\epsilon_2) + \cdots + P(\epsilon_i) + \cdots = \sum_{j=1}^{\infty} P(\epsilon_j)$$

We want to find a function, $\rho(\epsilon)$, that satisfies all four of these conditions. One way is to keep trying functions that look like they might work until we find one that does. A slightly more sophisticated version of this approach is to try the most general possible version of each such function and see if any set of restrictions will make it work. We could even try an infinite series. Suppose that we are clever (or lucky) enough to try the series solution

$$\ln\rho(\epsilon)\ = c_0 + c_1\epsilon + \cdots + c_i\epsilon^i + \cdots = \sum_{k=0}^{\infty} c_k\epsilon^k$$

Then the third condition becomes

$$\begin{aligned}\ln \kappa &= \sum_{i=1}^{\infty} N_i \ln \rho\ (\epsilon_i) \\ &= \sum_{i=1}^{\infty} N_i \sum_{k=0}^{\infty} \left[c_k \epsilon_i^k\right] \\ &= \sum_{k=0}^{\infty} \sum_{i=1}^{\infty} c_k N_i \epsilon_i^k = c_0 \sum_{i=1}^{\infty} N_i \epsilon_i^0 + c_1 \sum_{i=1}^{\infty} N_i \epsilon_i^1 + \cdots + c_k \sum_{k=2}^{\infty} \sum_{i=1}^{\infty} N_i \epsilon_i^k + \ldots \\ &= c_0 N + c_1 E^{\#} + \cdots + c_k \sum_{k=2}^{\infty} \sum_{i=1}^{\infty} N_i \epsilon_i^k + \ldots\end{aligned}$$

We see that the coefficient of c_0 is N and the coefficient of c_1 is the total energy, $E^{\#}$. Therefore, the sum of the first two terms is a constant. We can make the trial function satisfy the third condition if we set $c_k = 0$ for all $k > 1$. We find

$$\ln \kappa = \sum_{i=1}^{\infty} N_i \ln \rho\ (\epsilon_i) = \sum_{i=1}^{\infty} N_i \left(c_0 + c_1 \epsilon_i\right)$$

The last equality is satisfied if, for each quantum state, we have

$$\ln \rho\ (\epsilon_i) = c_0 + c_1 \epsilon_i$$

or

$$\rho\left(\epsilon_i\right) = \alpha \exp\left(c_1 \epsilon_i\right)$$

where $\alpha = \exp\left(c_0\right)$. Since the ϵ_i are positive and the probabilities $\rho\left(\epsilon_i\right)$ lie in the interval $0 < \rho\left(\epsilon_i\right) < 1$, we must have $c_1 < 0$. Following custom, we let $c_1 = -\beta$, where β is a constant, and $\beta > 0$. Then,

$$\rho\left(\epsilon_i\right) = \alpha \exp\left(-\beta \epsilon_i\right)$$

and

$$P_i = g_i \rho\left(\epsilon_i\right) = \alpha g_i \exp\left(-\beta \epsilon_i\right)$$

The fourth condition is that the energy-level probabilities sum to one. Using this, we have

$$1 = \sum_{i=1}^{\infty} P\left(\epsilon_i\right) = \alpha \sum_{i=1}^{\infty} g_i \exp\left(-\beta \epsilon_i\right)$$

The sum of exponential terms is so important that it is given a name. It is called the ***molecular partition function***. It is often represented by the letter "z." Letting

$$z = \sum_{i=1}^{\infty} g_i \exp\left(-\beta \epsilon_i\right)$$

we have

$$\alpha = \frac{1}{\sum_{i=1}^{\infty} g_i \exp\left(-\beta \epsilon_i\right)} = z^{-1}$$

Thus, we have the Boltzmann probability:

$$\begin{aligned}P\left(\epsilon_i\right) &= g_i \rho\left(\epsilon_i\right) \\ &= \frac{g_i \exp\left(-\beta \epsilon_i\right)}{\sum_{i=1}^{\infty} g_i \exp\left(-\beta \epsilon_i\right)} \\ &= \frac{g_i}{z} \exp\left(-\beta \epsilon_i\right)\end{aligned}$$

The probability of an energy level depends only on its degeneracy, g_i, its energy, ϵ_i, and the constant β. Since the equilibrium-characterizing population set is determined by the probabilities, we have $P_i = N_i^\bullet / N$, and

$$\frac{N_i^\bullet}{N} = \frac{g_i}{z} \exp\left(-\beta \epsilon_i\right)$$

In Section 21.2, we develop **Lagrange's method of undetermined multipliers**. In Section 21.3, we develop the same result by applying Lagrange's method to our model for the probabilities of the microstates of an isolated system. That is, we find the Boltzmann probability equation by applying Lagrange's method to the entropy relationship,

$$S = -Nk \sum_{i=1}^{\infty} P_i \ln \rho\left(\epsilon_i\right)$$

that we first develop in §20-11. In §4, we find the Boltzmann probability equation by using Lagrange's method to find the values of $N_i^\bullet$ that produce the largest possible value for W_{max} in an isolated system. This argument requires us to assume that there is a very large number of molecules in each of the occupied energy levels of the most probable population set. Since our other arguments do not assume anything about the magnitude of the various $N_i^\bullet$, it is evident that some of the assumptions we make when we apply Lagrange's method to find the $N_i^\bullet$ are not inherent characteristics of our microscopic model.

3.2: Lagrange's Method of Undetermined Multipliers

Lagrange's method of undetermined multipliers is a method for finding the minimum or maximum value of a function subject to one or more constraints. A simple example serves to clarify the general problem. Consider the function

$$z = z_0 \exp\left(x^2 + y^2\right)$$

where z_0 is a constant. This function is a surface of revolution, which is tangent to the plane $z = z_0$ at $(0, 0, z_0)$. The point of tangency is the minimum value of z. At any other point in the xy-plane, $z(x, y)$ is greater than z_0. If either x or y becomes arbitrarily large, z does also. If we project a contour of constant z onto the xy-plane, the projection is a circle of radius

$$r = \left(x^2 + y^2\right)^{1/2}.$$

Suppose that we introduce an additional condition; we require $y = 1 - x$. Then we ask for the smallest value of z consistent with this constraint. In the xy-plane the constraint is a line of slope -1 and intercept 1. A plane that includes this line and is parallel to the z-axis intersects the function z. As sketched in Figure 1, this intersection is a curve. Far away from the origin, the value of z at which the intersection occurs is large. Nearer the origin, the value of z is smaller, and there is some (x, y) at which it is a minimum. Our objective is to find this minimum.

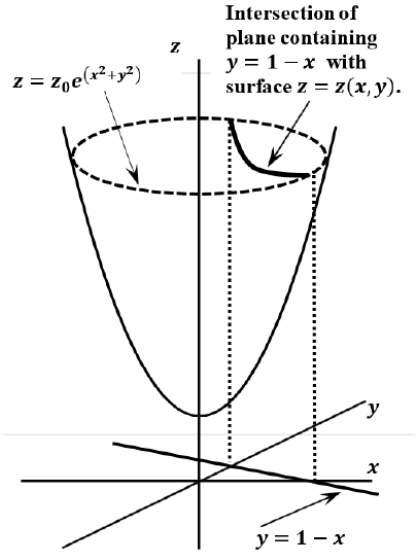


Figure 1: A surface and a constraint equation.

There is a straightforward solution of this problem; we can substitute the constraint equation for y into the equation for z, making z a function of only one variable, x. We have

$$\begin{aligned} z &= z_0 \exp\left(x^2 + (1 - x)^2\right) \\ &= z_0 \exp\left(2x^2 - 2x + 1\right) \end{aligned}$$

To find the minimum, we equate the derivative to zero, giving

$$0 = \frac{dz}{dx} = (4x - 2)\, z_0 \exp\left(2x^2 - 2x + 1\right)$$

so that the minimum occurs at $x = 1/2$, y$= 1/2$, and

$$z = z_0 \exp(1/2)$$

Solving such problems by elimination of variables can become difficult. Lagrange's method of undetermined multipliers is a general method, which is usually easy to apply and which is readily extended to cases in which there are multiple constraints. We can see how Lagrange's method arises by thinking further about our particular example. We can imagine that we "walk" along the constraint line in the xy-plane and measure the z that is directly overhead as we progress. The problem is to find the minimum

value of z that we encounter as we proceed along the line. This perspective highlights the central feature of the problem: While it is formally a problem in three dimensions (x, y, and z), the introduction of the constraint makes it a two-dimensional problem. We can think of one dimension as a displacement along the line $y = 1 - x$, from some arbitrary starting point on the line. The other dimension is the perpendicular distance from the xy-plane to the intersection with the surface z.

The relevant part of the xy-plane is just the one-dimensional constraint line. We can recognize this by parameterizing the line. Let t measure location on the line relative to some initial point at which $t = 0$. Then we have $x = x(t)$ and $y = y(t)$ and

$$z(x, y) = z(x(t), y(t)) = z(t).$$

The point we seek is the one at which $dz/dt = 0$.

Now let us examine a somewhat more general problem. We want a general way to find the values (x, y) that minimize (or maximize) a function $h = h(x, y)$ subject to a constraint of the form $c = g(x, y)$, where c is a constant. As in our example, this constraint requires a solution in which (x, y) are on a particular line. If we parameterize this problem, we have

$$h = h(x, y) = h(x(t), y(t)) = h(t)$$

and

$$c = g(x, y) = g(x(t), y(t)) = g(t)$$

Because c is a constant, $dc/dt = dg/dt = 0$. The solution we seek is the point at which h is an extremum. At this point, $dh/dt = 0$. Therefore, at the point we seek, we have

$$\frac{dh}{dt} = \left(\frac{\partial h}{\partial x}\right)_y \frac{dx}{dt} + \left(\frac{\partial h}{\partial y}\right)_x \frac{dy}{dt} = 0$$

and

$$\frac{dg}{dt} = \left(\frac{\partial g}{\partial x}\right)_y \frac{dx}{dt} + \left(\frac{\partial g}{\partial y}\right)_x \frac{dy}{dt} = 0$$

We can multiply either of these equations by any factor, and the product will be zero. We multiply dg/dt by λ (where $\lambda \neq 0$) and subtract the result from dh/dt. Then, at the point we seek,

$$0 = \frac{dh}{dt} - \lambda\frac{dg}{dt} = \left(\frac{\partial h}{\partial x} - \lambda\frac{\partial g}{\partial x}\right)_y \frac{dx}{dt} + \left(\frac{\partial h}{\partial y} - \lambda\frac{\partial g}{\partial y}\right)_x \frac{dy}{dt}$$

Since we can choose $x(t)$ and $y(t)$ any way we please, we can insure that $dx/dt \neq 0$ and $dy/dt \neq 0$ at the solution point. If we do so, the terms in parentheses must be zero at the solution point.

Conversely, setting

$$\left(\frac{\partial h}{\partial x} - \lambda\frac{\partial g}{\partial x}\right)_y = 0$$

and

$$\left(\frac{\partial h}{\partial y} - \lambda\frac{\partial g}{\partial y}\right)_x = 0$$

is sufficient to insure that

$$\frac{dh}{dt} = \lambda\frac{dg}{dt}$$

Since $dg/dt = 0$, these conditions insure that $dh/dt = 0$. This means that, if we can find a set $\{x, y, \lambda\}$ satisfying

$$\left(\frac{\partial h}{\partial x} - \lambda\frac{\partial g}{\partial x}\right)_y = 0$$

and

$$\left(\frac{\partial h}{\partial y}-\lambda\frac{\partial g}{\partial y}\right)_x=0$$

and

$$c-g(x,y)=0$$

then the values of x and y must be those make $h(x,y)$ an extremum, subject to the constraint that $c=g(x,y)$. We have not shown that the set $\{x,y,\lambda\}$ exists, but we have shown that if it exists, it is the desired solution.

A useful mnemonic simplifies the task of generating the family of equations that we need to use Lagrange's method. The mnemonic calls upon us to form a new function, which is a sum of the function whose extremum we seek and a series of additional terms. There is one additional term for each constraint equation. We generate this term by putting the constraint equation in the form $c-g(x,y)=0$ and multiplying by an undetermined parameter. For the case we just considered, the mnemonic function is

$$F_{mn}=h(x,y)+\lambda(c-g(x,y))$$

We can generate the set of equations that describe the solution set, $\{x,y,\lambda\}$, by equating the partial derivatives of F_{mn} with respect to x, y, and λ to zero. That is, the solution set satisfies the simultaneous equations

$$\frac{\partial F_{mn}}{\partial x}=0$$
$$\frac{\partial F_{mn}}{\partial y}=0$$

and

$$\frac{\partial F_{mn}}{\partial \lambda}=0$$

If there are multiple constraint equations, $c_\lambda-g_\lambda(x,y)=0$, $c_\alpha-g_\alpha(x,y)=0$, and $c_\beta-g_\beta(x,y)=0$, then the mnemonic function is

$$F_{mn}=h(x,y)+\lambda(c_\lambda-g_\lambda(x,y))+\alpha(c_\alpha-g_\alpha(x,y))+\beta(c_\beta-g_\beta(x,y))$$

and the simultaneous equations that represent the constrained extremum are

- $\partial F_{mn}/\partial x=0$,
- $\partial F_{mn}/\partial y=0$,
- $\partial F_{mn}/\partial \lambda=0$,
- $\partial F_{mn}/\partial \alpha=0$, and
- $\partial F_{mn}/\partial \beta=0$.

To illustrate the use of the mnemonic, let us return to the example with which we began. The mnemonic equation is

$$F_{mn}=z_0\exp\left(x^2+y^2\right)+\lambda(1-x-y)$$

so that

$$\frac{\partial F_{mn}}{\partial x}=2xz_0\exp\left(x^2+y^2\right)-\lambda=0,$$
$$\frac{\partial F_{mn}}{\partial y}=2yz_0\exp\left(x^2+y^2\right)-\lambda=0$$

and

$$\frac{\partial F_{mn}}{\partial \lambda}=1-x-y=0$$

which yield $x=1/2$, y$=1/2$, and $\lambda=z_0\exp(1/2)$.

3.3: Deriving the Boltzmann Equation I

In Sections 20-10 and 20-14, we develop the relationship between the system entropy and the probabilities of a microstate, $\rho(\epsilon_i)$, and an energy level, $P_i = g_i\rho(\epsilon_i)$, in our microscopic model. We find

$$S = -Nk\sum_{i=1}^{\infty} P_i \ln \rho(\epsilon_i)$$

$$= -Nk\sum_{i=1}^{\infty} g_i\rho(\epsilon_i) \ln \rho(\epsilon_i)$$

For an isolated system at equilibrium, the entropy must be a maximum, and hence

$$-\sum_{i=1}^{\infty} g_i\rho(\epsilon_i) \ln \rho(\epsilon_i) \tag{3.3.1}$$

must be a maximum. We can use Lagrange's method to find the dependence of the quantum-state probability on its energy. The $\rho(\epsilon_i)$ must be such as to maximize entropy (Equation 3.3.1) subject to the constraints

$$1 = \sum_{i=1}^{\infty} P_i = \sum_{i=1}^{\infty} g_i\rho(\epsilon_i)$$

and

$$\langle\epsilon\rangle = \sum_{i=1}^{\infty} P_i\epsilon_i = \sum_{i=1}^{\infty} g_i\varepsilon_i\rho(\epsilon_i)$$

where $\langle\epsilon\rangle$ is the expected value of the energy of one molecule. The mnemonic function becomes

$$F_{mn} = -\sum_{i=1}^{\infty} g_i\rho(\epsilon_i) \ln \rho(\epsilon_i) \; + \alpha^* \left(1 - \sum_{i=1}^{\infty} g_i\rho(\epsilon_i)\right) + \beta\left(\langle\epsilon\rangle - \sum_{i=1}^{\infty} g_i\varepsilon_i\rho(\epsilon_i)\right)$$

Equating the partial derivative with respect to $\rho(\epsilon_i)$ to zero,

$$\frac{\partial F_{mn}}{\partial \rho(\epsilon_i)} = -g_i \ln \rho(\epsilon_i) \; - g_i - \alpha^* g_i - \beta g_i\epsilon_i = 0$$

so that

$$\rho(\epsilon_i) = \exp(-\alpha^* - 1) \; \exp(-\beta\epsilon_i)$$

From

$$1 = \sum_{i=1}^{\infty} P_i = \sum_{i=1}^{\infty} g_i\rho(\epsilon_i)$$

the argument we use in Section 21.1 again leads to the partition function, z, and the Boltzmann equation

$$P_i = g_i\rho(\epsilon_i) = z^{-1} g_i \exp(-\beta\epsilon_i)$$

3.4: Deriving the Boltzmann Equation II

In Section 20-9, we find that the probability of the population set $\{N_1,\ N_2, \ldots, N_i, \ldots\}$ in an isolated system is

$$\rho_{MS,N,E} = N! \prod_{i=1}^{\infty} \frac{g_i^{N_i}}{N_i!}$$

The thermodynamic probability

$$W\left(N_i, g_i\right) = N! \prod_{i=1}^{\infty} \frac{g_i^{N_i}}{N_i!}$$

is the number of microstates of the population set. $\rho_{MS,N,E}$ is the constant probability of any one microstate. In consequence, as we see in Section 20.10, the probability of a population set is proportional to its thermodynamic probability, $W\left(N_i, g_i\right)$. It follows that the most probable population set is that for which $W\left(N_i, g_i\right)$ is a maximum. Our microscopic model asserts that the most probable population set, $\{N_1^{\bullet},\ N_2^{\bullet}, \ldots, N_i^{\bullet}, \ldots\}$, characterizes the equilibrium state, because the equilibrium system always occupies the either the most probable population set or another population set whose macroscopic properties are indistinguishable from those of the most probable one.

Evidently, the equilibrium-characterizing population set is the one for which $W\left(N_i, g_i\right)$, or $\ln W\left(N_i, g_i\right)$, is a maximum. Let us assume that the N_i are very large so that we can treat them as continuous variables, and we can use Stirling's approximation for $N_i!$. Then we can use Lagrange's method of undetermined multipliers to find the most probable population set by finding the set, $N_1,\ N_2, \ldots, N_i, \ldots$, for which $\ln W\left(N_i, g_i\right)$ is a maximum, subject to the constraints

$$N = \sum_{i=1}^{\infty} N_i$$

and

$$E = \sum_{i=1}^{\infty} N_i \epsilon_i .$$

From our definition of the system, both N and E are constant. The mnemonic function is

$$\begin{aligned} F_{mn} \ &= \ln\left(\frac{N! g_1^{N_1} g_2^{N_2} \ldots g_i^{N_i} \ldots}{N_1! N_2! \ldots N_i! \ldots}\right) \ + \alpha\left(N - \sum_{i=1}^{\infty} N_i\right) + \beta\left(E - \sum_{i=1}^{\infty} N_i \epsilon_i\right) \\ &\approx N \ln N - N - \sum_{i=1}^{\infty} N_i \ln N_i \ + \sum_{i=1}^{\infty} N_i + \sum_{i=1}^{\infty} N_i \ln g_i \ + \alpha\left(N - \sum_{i=1}^{\infty} N_i\right) + \beta\left(E - \sum_{i=1}^{\infty} N_i \epsilon_i\right) \end{aligned}$$

Taking the partial derivative with respect to N_i gives

$$\frac{\partial F_{mn}}{\partial N_i} = -N_i\left(\frac{1}{N_1}\right) - \ln N_i \ + 1 + \ln g_i \ - \alpha - \beta \epsilon_i = -\ln N_i \ + \ln g_i \ - \alpha - \beta \epsilon_i$$

from which we have, for the population set with the largest possible thermodynamic probability,

$$-\ln N_i^{\bullet} \ + \ln g_i \ - \alpha - \beta \epsilon_i = 0$$

or

$$N_i^{\bullet} = g_i \exp(-\alpha)\ \exp\left(-\beta \epsilon_i\right)$$

We can again make use of the constraint on the total number of molecules to find $\exp(-\alpha)$:

$$N = \sum_{i=1}^{\infty} N_i^{\bullet} = \exp(-\alpha)\ \sum_{i=1}^{\infty} g_i \exp\left(-\beta \epsilon_i\right)$$

so that $\exp(-\alpha) = Nz^{-1}$, where z is the partition function, $z = \sum_{i=1}^{\infty} g_i \exp(-\beta\epsilon_i)$. Therefore, in the most probable population set, the number of molecules having energy ϵ_i is

$$N_i^{\bullet} = Nz^{-1} g_i \exp(-\beta\epsilon_i)$$

The fraction with this energy is

$$\frac{N_i^{\bullet}}{N} = z^{-1} g_i \exp(-\beta\epsilon_i)$$

This fraction is also the probability of finding an arbitrary molecule in one of the quantum states whose energy is ϵ_i. When the isolated system and the corresponding constant-temperature system are functionally equivalent, this probability is P_i. As in the two previous analyses, we have

$$\begin{aligned} P_i &= g_i \rho(\epsilon_i) \\ &= z^{-1} g_i \exp(-\beta\epsilon_i). \end{aligned}$$

This derivation of Boltzmann's equation from W_{max} is the most common introductory treatment. It relies on the assumption that all of the N_i are large enough to justify treating them as continuous variables. This assumption proves to be invalid for many important systems. (For ideal gases, we find that $N_i = 0$ or $N_i = 1$ for nearly all of the very large number of energy levels that are available to a given molecule.) Nevertheless, the result obtained is clearly correct; not only is it the same as the result of our two previous arguments, but also it leads to satisfactory agreement between microscopic models and the macroscopic properties of a wide variety of systems.

3.5: Partition Functions and Equilibrium - Isomeric Molecules

In Section 20.11, we discuss chemical equilibrium between isomers from the perspective afforded by Boltzmann's definition of entropy. Now, let us consider equilibrium in this system from the perspective afforded by the energy-level probabilities. Let us assign even-integer labels to energy levels of isomer A and odd-integer labels to energy levels of isomer B. A group of atoms that can arrange itself into either a molecule of A or a molecule of B can occupy any of these energy levels. The partition function for this group of molecules to which all energy levels are available is

$$z_{A+B} = \sum_{i=1}^{\infty} g_i \exp\left(-\beta\epsilon_i\right)$$

The fraction of molecules in the first (odd) energy level associated with molecules of isomer B is

$$\frac{N_1^{\bullet}}{N_{A+B}} = g_1 \left(z_{A+B}\right)^{-1} \exp\left(-\beta\epsilon_1\right)$$

and the fraction in the next is

$$\frac{N_3^{\bullet}}{N_{A+B}} = g_3 \left(z_{A+B}\right)^{-1} \exp\left(-\beta\epsilon_3\right)$$

The total number of B molecules is

$$N_B^{\bullet} = \sum_{i\,odd} N_i$$

so that the fraction of all of the molecules that are B molecules is

$$\frac{N_B^{\bullet}}{N_{A+B}} = \left(z_{A+B}\right)^{-1} \sum_{i\,odd} g_i exp\left(-\beta\epsilon_i\right) \;\; = z_B / z_{A+B}$$

Likewise, the fraction that is A molecules is

$$\frac{N_A^{\bullet}}{N_{A+B}} = \left(z_{A+B}\right)^{-1} \sum_{i\,even} g_i exp\left(-\beta\epsilon_i\right) \;\; = z_A / z_{A+B}$$

The equilibrium constant for the equilibrium between A and B is

$$K_{eq} = \frac{N_B^{\bullet}}{N_A^{\bullet}} = \frac{z_B}{z_A}$$

We see that the equilibrium constant for the isomerization reaction is simply equal to the ratio of the partition functions of the isomers.

It is always true that the equilibrium constant is a product of partition functions for reaction-product molecules divided by a product of partition functions for reactant molecules. However, the partition functions for the various molecules must be expressed with a common zero of energy. Choosing the infinitely separated component atoms as the zero-energy state for every molecule assures that this is the case. However, it is often convenient to express the partition function for a molecule by measuring each molecular energy level, ϵ_i, relative to the lowest energy state of that isolated molecule. When we do this, the zero of energy is different for each molecule.

To adjust the energies in a molecule's partition function so that they are expressed relative to the energy of the molecule's infinitely separated atoms, we must add to each molecular energy the energy required to take the molecule from its lowest energy state to its isolated component atoms. If z is the partition function when the ϵ_i are measured relative to the lowest energy state of the isolated molecule, $\Delta\epsilon$ is the energy released when the isolated molecule is formed from its component atoms, and z^* is the partition function when the ϵ_i are measured relative to the molecule's separated atoms, we have $z^* = z \exp\left(+\Delta\epsilon / kT\right)$.

3.6: Finding ß and the Thermodynamic Functions for Distinguishable Molecules

All of a substance's thermodynamic functions can be derived from the molecular partition function. We begin with the entropy. We consider closed (constant N) systems of independent, distinguishable molecules in which only pressure–volume work is possible. In Sections 20.10 and 20.14, we find that two different approaches give the entropy of this system,

$$S = -Nk \sum_{i=1}^{\infty} P_i \ln \rho\left(\epsilon_i\right) \ .$$

In Sections 20.1, 20.3, and 20.4, we find that three different approaches give the Boltzmann equation,

$$P_i = g_i \rho\left(\epsilon_i\right) = z^{-1} g_i \exp\left(-\beta \epsilon_i\right).$$

We have

$$\ln \rho\left(\epsilon_i\right) = -\ln z - \beta \epsilon_i$$

Substituting, and recognizing that the energy of the N-molecule system is $E = N\left\langle \epsilon \right\rangle$, we find that the entropy of the system is

$$S = kN \sum_{i=1}^{\infty} P_i \left[\ln z + \beta \epsilon_i\right] = kN \ln z \sum_{i=1}^{\infty} P_i + k\beta N \sum_{i=1}^{\infty} P_i \epsilon_i = kN \ln z + k\beta E$$

In Section 10.1, we find that the fundamental equation implies that

$$\left(\frac{\partial E}{\partial S}\right)_V = T$$

Since the ϵ_i are fixed when the volume and temperature of the system are fixed, $\ln z$ is constant when the volume and temperature of the system are constant. Differentiating $S = kN \ln z + k\beta E$ with respect to S at constant V, we find

$$1 = k\beta \left(\frac{\partial E}{\partial S}\right)_V = k\beta T$$

so that

$$\beta = \frac{1}{kT}$$

This is an important result: Because we have now identified all of the parameters in our microscopic model, we can write the results we have found in forms that are more useful:

1. $$\underbrace{z = \sum_{i=1}^{\infty} g_i \exp\left(\frac{-\epsilon_i}{kT}\right)}_{\text{molecular partition function}}$$
2. $$\underbrace{P_i = g_i \rho\left(\epsilon_i\right) = z^{-1} g_i \exp\left(\frac{-\epsilon_i}{kT}\right)}_{\text{Boltzmann's equation}}$$
3. $$\underbrace{S = kN \ln z + \frac{E}{T}}_{\text{Entropy of an N-molecule system}}$$

To express the system energy in terms of the molecular partition function, we first observe that

$$E = N\left\langle \epsilon \right\rangle = N \sum_{i=1}^{\infty} P_i \epsilon_i = N z^{-1} \sum_{i=1}^{\infty} g_i \epsilon_i \exp\left(\frac{-\epsilon_i}{kT}\right)$$

Then we observe that

$$\left(\frac{\partial \ln z}{\partial T}\right)_V = z^{-1}\sum_{i=1}^{\infty} g_i\left(\frac{\epsilon_i}{kT^2}\right)\exp\left(\frac{-\epsilon_i}{kT}\right)$$
$$= \left(\frac{1}{NkT^2}\right) N z^{-1}\sum_{i=1}^{\infty} g_i\epsilon_i \exp\left(\frac{-\epsilon_i}{kT}\right)$$
$$= \frac{E}{NkT^2}$$

The system energy becomes

$$\underbrace{E = NkT^2\left(\frac{\partial \ln z}{\partial T}\right)_V}_{\text{energy of an N-molecule system}}$$

By definition, $A = E - TS$. Rearranging our entropy result, $S = kN\ln z + E/T$, we have $E - TS = -NkT\ln z$. Thus,

$$A = -NkT\ln z$$

(Helmholtz free energy of an N-molecule system)

From $dA = -SdT - PdV$, we have

$$\left(\frac{\partial A}{\partial V}\right)_T = -P$$

(Here, of course, P is the pressure of the system, not a probability.) Differentiating $A = -NkT\ln z$ with respect to V at constant T, we find

$$P = NkT\left(\frac{\partial \ln z}{\partial V}\right)_T$$

(pressure of an N-molecule system)

The pressure–volume product becomes

$$PV = NkTV\left(\frac{\partial \ln z}{\partial V}\right)_T$$

Substituting into $H = E + PV$, the enthalpy becomes

$$H = NkT\left[T\left(\frac{\partial \ln z}{\partial T}\right)_V + V\left(\frac{\partial \ln z}{\partial V}\right)_T\right]$$

(enthalpy of an N-molecule system)

The Gibbs free energy is given by $G = A + PV$. Substituting, we find

$$G = -NkT\ln z + NkTV\left(\frac{\partial \ln z}{\partial V}\right)_T$$

(Gibbs free energy of an N-molecule system)

The chemical potential can be found from

$$\mu = \left(\frac{\partial A}{\partial n}\right)_{V,T}$$

At constant volume and temperature, $kT\ln z$ is constant.

Substituting $N = n\overline{N}$ into $A = -NkT\ln z$ and taking the partial derivative, we find

$$\underbrace{\mu = -\overline{N}kT\ln z = -RT\ln z}_{\text{chemical potential of distinguishable molecules}}$$

In statistical thermodynamics we frequently express the chemical potential per molecule, rather than per mole; then,

$$\mu = \left(\frac{\partial A}{\partial N} \right)_{V,T}$$

and

$$\mu = -kT \ln z$$

(chemical potential per molecule)

3.7: The Microscopic Model for Reversible Change

Now let us return to the closed (constant-N) system to develop another perspective on the dependence of its macroscopic thermodynamic properties on the molecular energy levels and their probabilities. We undertake to describe the system using volume and temperature as the independent variables. In thinking about the energy-level probabilities, we stipulate that any parameters that affect the state of the system remain constant. Specifically, we mean that any parameters that appear in the Schrödinger equation remain constant. For example, the energy levels of a particle in a box depend on the mass of the particle and the length of the box. Any such parameter is called an ***exogenous*** variable. If we change an exogenous variable (say the length of the box) by a small amount, all of the energy levels change by a small amount, and all of the probabilities change by a small amount. The energy levels and their probabilities are smooth functions of the exogenous variable. If ξ is the exogenous variable, we have

$$P_i = P\left(\epsilon_i\right) = g_i \rho\left(\epsilon_i\left(\xi\right)\right)$$

A change in the exogenous variable corresponds to a reversible macroscopic process.

For a particle in a box, the successive ψ_i are functions that depend on the quantum number, i, and the length of the box, ℓ. When we change the length of the box, the wavefunction and its associated energy both change. Both are continuous functions of the length of the box. The energy is

$$\epsilon_i = \frac{i^2 h^2}{8m\ell^2}$$

Changing the length of the box is analogous to changing the volume of a system. A reversible volume change entails work. We see that changing the length of the box does work on the particle-in-a-box, just as changing the volume of a three-dimensional system does work on the system.

Temperature plays a central role in the description of equilibrium from the macroscopic perspective. We can see that temperature enters the description of equilibrium from the microscopic perspective through its effect on the probability factors. When we increase the temperature of a system, its energy increases. The average energy of its molecules increases. The probability of an energy level must depend on temperature. Evidently, the probabilities of energy levels that are higher than the original average energy increase when the temperature increases. The probabilities of energy levels that are lower than the original average energy decrease when the temperature increases. The effects of heat and work on the energy levels and their equilibrium populations are diagrammed in Figure 3.7.1.

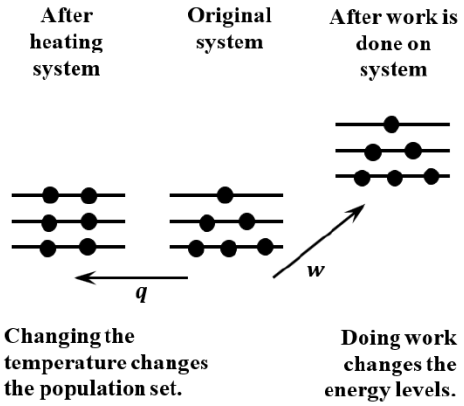


Figure 3.7.1: The effects of heat and work on energy levels and their populations.

If our theory is to be useful, the energy we measure for a macroscopic system must be indistinguishably close to the expected value of the system energy as calculated from our microscopic model:

$$E_{\text{experiment}} \approx \langle E \rangle = N \langle \epsilon \rangle = N \sum_{i=1}^{\infty} P_i \epsilon_i$$

We can use this equation to relate the probabilities, P_i, to other thermodynamic functions. Dropping the distinction between the experimental and expected energies, and assuming that the ϵ_i and the P_i are continuous variables, we find the total differential

$$dE = N\sum_{i=1}^{\infty}\epsilon_i dP_i + N\sum_{i=1}^{\infty}P_i d\epsilon_i$$

This equation is important because it describes a reversible macroscopic process in terms of the microscopic variables ϵ_i and P_i.

Let us consider the first term. Since N is a constant, we have from $N_i^{\bullet} = P_i N$ that $dN_i^{\bullet} = N dP_i$. Substituting, we have

$$(dE)_{\epsilon_i} = N\sum_{i=1}^{\infty}\epsilon_i dP_i = \sum_{i=1}^{\infty}\epsilon_i dN_i^{\bullet}$$

This asserts that the energy of the system changes if we redistribute the molecules among the various energy levels. If the redistribution takes molecules out of lower energy levels and puts them into higher energy levels, the energy of the system increases. This is our statistical-mechanical picture of the shift in the equilibrium position that occurs when we heat a system of independent molecules; the allocation of molecules among the available energy levels shifts to put more molecules in higher energy levels and fewer in lower ones. This corresponds to an increase in the temperature of the macroscopic system.

In terms of the macroscopic system, the first term represents an increment of heat added to the system in a reversible process; that is,

$$dq^{rev} = N\sum_{i=1}^{\infty}\epsilon_i dP_i$$

The second term, $N\sum_{i=1}^{\infty}P_i d\epsilon_i$, is a contribution to the change in the energy of the system from reversible changes in the energy of the various quantum states, while the number of molecules in each quantum state remains constant. This term corresponds to a process in which the quantum states (and their energies) evolve in a continuous way as the state of the system changes. The second term represents an increment of work done on the system in a reversible process; that is

$$dw^{rev} = N\sum_{i=1}^{\infty}P_i d\epsilon_i$$

Evidently, the total differential expression for dE is the fundamental equation of thermodynamics expressed in terms of the variables we use to characterize the molecular system. It enables us to relate the variables that characterize our microscopic model of the molecular system to the variables that characterize the macroscopic system.

For a system in which the reversible work is pressure–volume work, the energy levels depend on the volume. At constant temperature we have

$$dw^{rev} = -PdV = N\sum_{i=1}^{\infty}P_i d\epsilon_i = N\sum_{i=1}^{\infty}P_i\left(\frac{\partial\epsilon_i}{\partial V}\right)_T dV$$

so that the system pressure, P, is related to the energy-level probabilities, P_i, as

$$P = -N\sum_{i=1}^{\infty}P_i\left(\frac{\partial\epsilon_i}{\partial V}\right)_T$$

To evaluate the pressure, we must know how the energy levels depend on the volume of the system.

The first term relates the entropy to the energy-level probabilities. Since $dq^{rev} = TdS = N\sum_{i=1}^{\infty}\epsilon_i dP_i$, we have

$$dS = \frac{N}{T}\sum_{i=1}^{\infty}\epsilon_i dP_i$$

From the Boltzmann distribution function we have

$P_i = z^{-1}g_i\exp\left(-\epsilon_i/kT\right)$, or

$$\epsilon_i = -kT\ln P_i + kT\ln g_i - kT\ln z$$

Substituting into our expression for dS, we find

$$dS = -Nk\sum_{i=1}^{\infty}(\ln P_i)dP_i + Nk\sum_{i=1}^{\infty}(\ln g_i)dP_i - Nk(\ln z)\sum_{i=1}^{\infty}dP_i$$

Since $\sum_{i=1}^{\infty} P_i = 1$, we have $\sum_{i=1}^{\infty} dP_i = 0$, and the last term vanishes. Also,

$$\sum_{i=1}^{\infty} d(P_i \ln P_i) = \sum_{i=1}^{\infty}(\ln P_i)\, dP_i + \sum_{i=1}^{\infty} dP_i = \sum_{i=1}^{\infty}(\ln P_i)\, dP_i$$

so that

$$dS = -Nk\sum_{i=1}^{\infty} d(P_i \ln P_i) + Nk\sum_{i=1}^{\infty}(\ln g_i)dP_i$$

At any temperature, the probability ratio for any two successive energy levels is

$$\frac{P_{i+1}(T)}{P_i(T)} = \frac{P_{i+1}}{P_i} = \frac{g_{i+1}}{g_i}\exp\left(\frac{-(\epsilon_{i+1}-\epsilon_i)}{kT}\right)$$

In the limit as the temperature goes to zero,

$$\frac{P_{i+1}}{P_i} \to 0$$

It follows that $P_1(0) = 1$ and $P_i(0) = 0$ for $i > 1$. Integrating from $T = 0$ to T, the entropy of the system goes from $S(0) = S_0$ to $S(T)$, and the energy-level probabilities go from $P_i(0)$ to $P_i(T)$. We have

$$\int_{S_0}^{S(T)} dS = -Nk\sum_{i=1}^{\infty}\int_{P_i(0)}^{P_i(T)} d(P_i \ln P_i) + Nk\sum_{i=1}^{\infty}\int_{P_i(0)}^{P_i(T)}(\ln g_i)\, dP_i$$

so that

$$S(T) - S_0 = -Nk\sum_{i=1}^{\infty} P_i(T)\ln P_i(T) + NkP_1(0)\ln P_1(0) + Nk\sum_{i=1}^{\infty}(\ln g_i)\, P_i(T) - Nk(\ln g_1)\, P_1(0)$$

Since $P_1(0) = 1$, $\ln P_1(0)$ vanishes. The entropy change becomes

$$S(T) - S_0 = -Nk\sum_{i=1}^{\infty} P_i[\ln P_i - \ln g_i] - Nk\ln g_1 = -Nk\sum_{i=1}^{\infty} P_i \ln \rho(\epsilon_i) - Nk\ln g_1$$

We have $S_0 = Nk\ln g_1$. If $g_1 = 1$, the lowest energy level is non-degenerate, and $S_0 = 0$; then we have

$$S = -Nk\sum_{i=1}^{\infty} P_i \ln \rho(\epsilon_i)$$

This is the entropy of an N-molecule, constant-volume, constant-temperature system that is in thermal contact with its surroundings at the same temperature. We obtain this same result in Sections 20.10 and 20.14 by arguments in which we assume that the system is isolated. In all of these arguments, we assume that the constant-temperature system and its isolated counterpart are functionally equivalent; that is, a group of population sets that accounts for nearly all of the probability in one system also accounts for nearly all of the probability in the other.

Because we obtain this result by assuming that the system is composed of N, independent, non-interacting, distinguishable molecules, the entropy of this is system is N times the entropy contribution of an individual molecule. We can write

$$S_{\text{molecule}} = -k\sum_{i=1}^{\infty} P_i \ln \rho(\epsilon_i)$$

3.8: The Third Law of Thermodynamics

In Section 21.7, we obtain the entropy by a definite integration. We take the lower limits of integration, at $T = 0$, as $P_1(0) = 1$ and $P_i(0) = 0$, for $i > 1$. In doing so, we apply the third law of thermodynamics, which states that the entropy of a perfect crystal can be chosen to be zero when the temperature is at absolute zero. The idea behind the third law is that, at absolute zero, the molecules of a crystalline substance all are in the lowest energy level that is available to them. The probability that a molecule is in the lowest energy state is, therefore, $P_1 = 1$, and the probability that it is any higher energy level, $i > 1$, is $P_i = 0$.

While the fact is not relevant to the present development, we note in passing that the energy of a perfect crystal is not zero at absolute zero. While all of the constituent particles will be in their lowest vibrational energy levels at absolute zero, the energies of these lowest vibrational levels are not zero. In the harmonic oscillator approximation, the lowest energy possible for each oscillator is $h\nu/2$. (See Section 18.5).

By a perfect crystalline substance we mean one in which the lowest energy level is non-degenerate; that is, for which $g_1 = 1$. We see that our entropy equation conforms to the third law when we let

$$S_0 = Nk \ln g_1$$

so that $S_0 = 0$ when $g_1 = 1$.

Let us consider a crystalline substance in which the lowest energy level is degenerate; that is, one for which $g_1 > 1$. This substance is not a perfect crystal. In this case, the temperature-zero entropy is

$$S_0 = Nk \ln g_1 > 0$$

The question arises: How can we determine whether a crystalline substance is a perfect crystal? In Chapter 11, we discuss the use of the third law to determine the absolute entropy of substances at ordinary temperatures. If we assume that the substance is a perfect crystal at zero degrees when it is not, our theory predicts a value for the absolute entropy at higher temperatures that is too small, because it does not include the term $S_0 = Nk \ln g_1$. When we use this too-small absolute entropy value to calculate entropy changes for processes involving the substance, the results do not agree with experiment.

Absolute entropies based on the third law have been experimentally determined for many substances. As a rule, the resulting entropies are consistent with other experimentally observed entropy changes. In some cases, however, the assumption that the entropy is zero at absolute zero leads to absolute entropy values that are not consistent with other experiments. In these cases, the absolute entropies can be brought into agreement with other entropy measurements by assuming that, indeed, $g_1 > 1$ for such substances. In any particular case, the value of g_1 that must be used is readily reconciled with other information about the substance.

For example, the third law entropy for carbon monoxide must be calculated taking $g_1 = 2$ in order to obtain a value that is consistent with other entropy measurements. This observation is readily rationalized. In perfectly crystalline carbon monoxide, all of the carbon monoxide molecules point in the same direction, as sketched in Figure 11-2. However, the two ends of the carbon monoxide molecule are very similar, with the consequence that the carbon monoxide molecules in the crystal point randomly in either of two directions. Thus there are two (approximately) equally energetic states for a carbon monoxide molecule in a carbon monoxide crystal at absolute zero, and we can take $g_1 = 2$. (We are over-simplifying here. We explore this issue further in Section 22-7.)

3.9: The Partition Function for a System of N Molecules

At a given temperature, the Boltzmann equation gives the probability of finding a molecule in any of the energy levels that the molecule can occupy. Throughout our development, we assume that there are no energies of interaction among the molecules of the system. The molecular partition function contains information about the energy levels of only one molecule. We obtain equations for the thermodynamic functions of an N-molecule system in terms of this molecular partition function. However, since these results are based on assigning the same isolated-molecule energy levels to each of the molecules, they do not address the real-system situation in which intermolecular interactions make important contributions to the total energy of the system.

As we mention in Sections 20.1 and 20.3, the ensemble theory of statistical thermodynamics extends our arguments to express the thermodynamic properties of a macroscopic system in terms of all of the total energies that are available to the macroscopic system. The molecular origins of the energies of the system enter the ensemble treatment only indirectly. The theory deals with the relationships between the possible values of the energy of the system and its thermodynamic state. How molecular energy levels and intermolecular interactions give rise to these values of the system energy becomes a separate issue. Fortunately, ensemble theory just reuses—from a different perspective—all of the ideas we have just studied. The result is just the Boltzmann equation, again, but now the energies that appear in the partition function are the possible energies for the collection of N molecules, not the energies available to a single molecule.

3.10: Problems

1. Consider a system with three non-degenerate quantum states having energies $\epsilon_1 = 0.9\ kT$, $\epsilon_2 = 1.0\ kT$, and $\epsilon_3 = 1.1\ kT$. The system contains $N = 3 \times 10^{10}$ molecules. Calculate the partition function and the number of molecules in each quantum state when the system is at equilibrium. This is the equilibrium population set $\{N_1^*, N_2^*, N_3^*\}$. Let W_{mp} be the number of microstates associated with the equilibrium population set. Consider the population set when 10^{-5} of the molecules in ϵ_2 are moved to each of ϵ_1 and ϵ_3. This is the population set $\{N_1^* + 10^{-5} N_2^*,\quad N_2^* - 2 \times 10^{-5},\quad N_3^* + 10^{-5} N_2^*\}$. Let W be the number of microstates associated with this non-equilibrium population set.

(a) What percentage of the molecules are moved in converting the first population set into the second?

(b) How do the energies of these two populations sets differ from one another?

(c) Find W_{mp}/W. Use Stirling's approximation and carry as many significant figures as your calculator will allow. You need at least six.

(d) What does this calculation demonstrate?

2. Find the approximate number of energy levels for which $\epsilon < kt >$ for a molecule of molecular weight 40 in a box of volume 10^{-6} m^3 at 300 K.

3. The partition function plays a central role in relating the probability of finding a molecule in a particular quantum state to the energy of that state. The energy levels available to a particle in a one-dimensional box are

$$\epsilon_n = \frac{n^2 h^2}{8 m \ell^2}$$

where m is the mass of the particle and ℓ is the length of the box. For molecular masses and boxes of macroscopic lengths, the factor $h^2/8m\ell^2$ is a very small number. Consequently, the energy levels available to a molecule in such a box can be considered to be effectively continuous in the quantum number, n. That is, the partition function sum can be closely approximated by an integral in which the variable of integration, n, runs from 0 to ∞.

(a) Obtain a formula for the partition function of a particle in a one-dimensional box. Integral tables give

$$\int_0^\infty \exp\left(-an^2\right) dn = \sqrt{\pi/4a}$$

(b) The expected value of the energy of a molecule is given by

$$\langle \epsilon \rangle = kT^2 \left(\frac{\partial \ln z}{\partial T} \right)_V$$

What is $\langle \epsilon \rangle$ for a particle in a box?

(c) The relationship between the partition function and the per-molecule Helmholtz free energy is $A = -kT \ln z$. For a molecule in a one-dimensional box, we have $dA = -S dT - \rho \ell$, where ρ is the per-molecule "pressure" on the ends of the box and ℓ is the length of the box. (The increment of work associated with changing the length of the box is $dw = -\rho\, d\ell$. In this relationship, $d\ell$ is the incremental change in the length of the box and ρ is the one-dimensional "pressure" contribution from each molecule. ρ is, of course, just the force required to push the end of the box outward by a distance $d\ell$. $\rho d\ell$ is the one-dimensional analog of PdV.) For the one-dimensional system, it follows that

$$\rho = -\left(\frac{\partial A}{\partial \ell} \right)_T$$

Use this information to find ρ for a molecule in a one-dimensional box.

(d) We can find ρ for a molecule in a one-dimensional box in another way. The per-molecule contribution to the pressure of a three-dimensional system is related to the energy-level probabilities, P_i, by

$$P_{\text{molecule}}^{\text{system}} = -\sum_{n=1}^{\infty} P_n \left(\frac{\partial \epsilon_n}{\partial V} \right)_T$$

By the same argument we use for the three-dimensional case, we find that the per-molecule contribution to the “pressure” inside a one-dimensional box is

$$\rho = -\sum_{n=1}^{\infty} P_n \left(\frac{\partial \epsilon_n}{\partial \ell} \right)_T$$

From the equation for the energy levels of a particle in a one dimensional box, find an equation for

$$\left(\frac{\partial \epsilon_n}{\partial \ell} \right)_T$$

(Hint: We can express this derivative as a simple multiple of ϵ_n.)

(e) Using your result from part (d), show that the per molecule contribution, ρ, to the “one-dimensional pressure” of N molecules in a one-dimensional box is

$$\rho = 2 \langle \epsilon \rangle / \ell$$

(f) Use your results from parts (b) and (e) to express ρ as a function of k, T, and ℓ.

(g) Let Π be the pressure of a system of N molecules in a one-dimensional box. From your result in part (c) or part (f), give an equation for Π. Show how this equation is analogous to the ideal gas equation.

CHAPTER OVERVIEW

4: Some Basic Applications of Statistical Thermodynamics

4.1: Interpreting the Partition Function

When it is a good approximation to say that the energy of a molecule is the sum of translational, rotational, vibrational, and electronic components, we have

$$\epsilon_{i,j,k,m} = \epsilon_{t,i} + \epsilon_{r,j} + \epsilon_{v,k} + \epsilon_{e,m}$$

where the indices i, j, k, and m run over all possible translational, rotational, vibrational, and electronic quantum states, respectively. Then the partition function for the molecule can be expressed as a product of the individual partition functions z_t, z_r, z_v, and z_e; that is,

$$\begin{aligned} z_{\text{molecule}} &= \sum_t \sum_r \sum_v \sum_e g_{t,i} g_{r,j} g_{v,k} g_{e,m} \exp\left(\frac{-\epsilon_{i,j,k,m}}{kT}\right) \\ &= \sum_t g_{t,i} exp\left(\frac{-\epsilon_{t,i}}{kT}\right) \sum_r g_{r,j} exp\left(\frac{-\epsilon_{r,j}}{kT}\right) \sum_v g_{v,k} exp\left(\frac{-\epsilon_{v,k}}{kT}\right) \sum_e g_{e,m} exp\left(\frac{-\epsilon_{e,m}}{kT}\right) \\ &= z_t z_r z_v z_e \end{aligned}$$

The magnitude of an individual partition function depends on the magnitudes of the energy levels associated with that kind of motion. Table 1 gives the contributions made to their partition functions by levels that have various energy values.

Table 1:

ϵ_i	$\frac{-\epsilon_i}{kT}$	$\exp\left(\frac{-\epsilon_i}{kT}\right)$	Type of Motion
$10^{-2}\ kT$	-10^{-2}	0.990	
$10^{-1}\ kT$	-10^{-1}	0.905	
kT	-1	0.365	translational
$5\ kT$	-5	0.0067	rotational
$10\ kT$	-10	4.5×10^{-5}	vibration
$100\ kT$	-100	3.7×10^{-44}	electronic

We see that only quantum states whose energy is less than kT can make substantial contributions to the magnitude of a partition function. Very approximately, we can say that the partition function is equal to the number of quantum states for which the energy is less than kT. Each such quantum state will contribute approximately one to the sum that comprises the partition function; the contribution of the corresponding energy level will be approximately equal to its degeneracy. If the energy of a quantum state is large compared to kT, the fraction of molecules occupying that quantum state will be small. This idea is often expressed by saying that such states are “unavailable” to the molecule. It is then said that the value of the partition function is approximately equal to the number of available quantum states. When most energy levels are non-degenerate, we can also say that the value of the partition function is approximately equal to the number of available energy levels.

4.2: Conditions under which Integrals Approximate Partition Functions

The Boltzmann equation gives the equilibrium fraction of particles in the i^{th} energy level, ϵ_i, as

$$\frac{N_i^{\bullet}}{N} = \frac{g_i}{z}\exp\left(\frac{-\epsilon_i}{kT}\right)$$

so the fraction of particles in energy levels less than ϵ_n is

$$f\left(\epsilon_n\right) = z^{-1}\sum_{i=1}^{n-1} g_i \exp\left(\frac{-\epsilon_i}{kT}\right)$$

where $z = \sum_{i=1}^{\infty} g_i \exp\left(\epsilon_i / kT\right)$. We can represent either of these sums as the area under a bar graph, where the height and width of each bar are $g_i \exp\left(\epsilon_i / kT\right)$ and unity, respectively. If g_i and ϵ_i can be approximated as continuous functions, this area can be approximated as the area under the continuous function $y\left(i\right) = g_i \exp\left(\epsilon_i / kT\right)$. That is,

$$\sum_{i=1}^{n-1} g_i \exp\left(\frac{-\epsilon_i}{kT}\right) \approx \int_{i=0}^{n} g_i \exp\left(\frac{-\epsilon_i}{kT}\right) di$$

To evaluate this integral, we must know how both g_i and ϵ_i depend on the quantum number, i.

Let us consider the case in which $g_i = 1$ and look at the constraints that the ϵ_i must satisfy in order to make the integral a good approximation to the sum. The graphical description of this case is sketched in Figure 1. Since $\epsilon_i > \epsilon_{i-1} > 0$, we have

$$e^{-\epsilon_{i-1}/kT} - e^{-\epsilon_i/kT} > 0$$

For the integral to be a good approximation, we must have

$$e^{-\epsilon_{i-1}/kT} \gg e^{-\epsilon_{i-1}/kT} - e^{-\epsilon_i/kT} > 0,$$

which means that

$$1 \gg 1 - e^{-\Delta\epsilon/kT} > 0$$

where $\Delta\epsilon = \epsilon_i - \epsilon_{i-1}$. Now,

$$e^x \approx 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \ldots$$

so that the approximation will be good if

$$1 \gg 1 - \left(1 - \frac{\Delta\epsilon}{kT} + \ldots\right)$$

or

$$1 \gg \frac{\Delta\epsilon}{kT}$$

or

$$kT \gg \Delta\epsilon$$

We can be confident that the integral is a good approximation to the exact sum whenever there are many pairs of energy levels, ϵ_i and ϵ_{i-1}, that satisfy the condition

$$\Delta\epsilon = \epsilon_i - \epsilon_{i-1} \ll kT.$$

If there are many energy levels that satisfy $\epsilon_i \ll kT$, there are necessarily many intervals, $\Delta\epsilon$, that satisfy $\Delta\epsilon \ll kT$. In short, if a large number of the energy levels of a system satisfy the criterion $\epsilon \ll kT$, we can use integration to approximate the sums that appear in the Boltzmann equation. In Section 24.3, we use this approach and the energy levels for a particle in a box to find the partition function for an ideal gas.

4.3: Probability Density Functions from the Energies of Classical-mechanical Models

Guided by our development of the Maxwell-Boltzmann probability density function for molecular velocities, we could postulate that similar probability density functions apply to other energies derived from classical-mechanical models for molecular motion. We will see that this can indeed be done. The results correspond to the results that we get from the Boltzmann equation, where we assume for both derivations that many energy levels satisfy $\epsilon \ll kT$. The essential point is that, at a sufficiently high temperature, the behavior predicted by the quantum mechanical model and that predicted from classical mechanics converge. This high-temperature approximation is a good one for translational motions but a very poor one for vibrational motions. These results further illuminate the differences between the classical-mechanical and the quantum-mechanical models for the behavior of molecules.

Let us look at how we can generate probability density functions based on the energies of classical-mechanical models for molecular motions. In the classical mechanical model, a particle moving in one dimension with velocity v has kinetic energy $mv^2/2$. From the discussion above, if many velocities satisfy $kT \gg mv^2/2$, we can postulate a probability density function of the form

$$\frac{df}{dv} = B_{\text{trans}} \exp\left(\frac{-mv^2}{2kT}\right)$$

where B_{trans} is fixed by the condition

$$\int_{-\infty}^{\infty} \left(\frac{df}{dv}\right) dv = B_{\text{trans}} \int_{-\infty}^{\infty} \exp\left(\frac{-mv^2}{2kT}\right) dv = 1$$

Evidently, this postulate assumes that each velocity constitutes a quantum state and that the degeneracy is the same for all velocities. This assumption is successful for one-dimensional translation, but not for translational motion in two or three dimensions. The definite integral is given in Appendix D. We find

$$B_{\text{trans}} = (m/2\pi kT)^{1/2}$$

and

$$\frac{df}{dv} = \left(\frac{m}{2\pi kT}\right)^{1/2} \exp\left(\frac{-mv^2}{2kT}\right)$$

With $m/kT = \lambda$, this is the same as the result that we obtain in Section 4.4. With B_{trans} in hand, we can calculate the average energy associated with the motion of a gas molecule in one dimension

$$\langle \epsilon \rangle = \int_{-\infty}^{\infty} \left(\frac{mv^2}{2}\right) \left(\frac{df}{dv}\right) dv = \left(\frac{m^3}{8\pi kT}\right)^{1/2} \int_{-\infty}^{\infty} v^2 \exp\left(\frac{-mv^2}{2kT}\right) dv$$

This definite integral is also given in Appendix D. We find

$$\langle \epsilon_{\text{trans}} \rangle = \frac{kT}{2}$$

We see that we can obtain the average kinetic energy for one degree of translational motion by a simple argument that uses classical-mechanical energies in the Boltzmann equation. We can make the same argument for each of the other two degrees of translational motion. We conclude that each degree of translational freedom contributes $kT/2$ to the average energy of a gas molecule. For three degrees of translational freedom, the total contribution is $3kT/2$, which is the result that we first obtained in Section 2.10.

Now let us consider a classical-mechanical model for a rigid molecule rotating in a plane. The classical kinetic energy is $\epsilon_{\text{rot}} = I\omega^2/2$, where I is the molecule's moment of inertia about the axis of rotation, and ω is the angular rotation rate. This has the same form as the translational kinetic energy, so if we assume $kT \gg I\omega^2/2$ and a probability density function of the form

$$\frac{df}{d\omega} = B_{\text{rot}} \exp\left(\frac{-I\omega^2}{2kT}\right)$$

finding B_{rot} and $\langle \epsilon_{\text{rot}} \rangle$ follows exactly as before, and the average rotational kinetic energy is

$$\langle \epsilon_{\text{rot}} \rangle = kT/2$$

for a molecule with one degree of rotational freedom.

For a classical harmonic oscillator, the vibrational energy has both kinetic and potential energy components. They are $mv^2/2$ and $kx^2/2$ where v is the oscillator's instantaneous velocity, x is its instantaneous location, and k is the force constant. Both of these have the same form as the translational kinetic energy equation. If we can assume that $kT \gg mv^2/2$, that $kT \gg kx^2/2$, and that the probability density functions are

$$\frac{df}{dv} = B_{\text{vib}}^{\text{kinetic}} \exp\left(\frac{-mv^2}{2kT}\right)$$

and

$$\frac{df}{dx} = B_{\text{vib}}^{\text{potential}} \exp\left(\frac{-kx^2}{2kT}\right)$$

the same arguments show that the average kinetic energy and the average potential energy are both $kT/2$:

$$\left\langle \epsilon_{\text{vib}}^{\text{kinetic}} \right\rangle = kT/2$$

and

$$\left\langle \epsilon_{\text{vib}}^{\text{potential}} \right\rangle = kT/2$$

so that the average total vibrational energy is

$$\left\langle \epsilon_{\text{vib}}^{\text{total}} \right\rangle = kT$$

In summary, because the energy for translational motion in one dimension, the energy for rotational motion about one axis, the energy for vibrational kinetic energy in one dimension, and the energy for vibrational potential energy in one dimension all have the same form ($\epsilon = Xu^2$) each of these modes can contribute $kT/2$ to the average energy of a molecule. For translation and rotation, the total is $kT/2$ for each degree of translational or rotational freedom. For vibration, because there is both a kinetic and a potential energy contribution, the total is kT per degree of vibrational freedom.

Let us illustrate this for the particular case of a non-linear, triatomic molecule. From our discussion in Section 18.4, we see that there are three degrees of translational freedom, three degrees of rotational freedom, and three degrees of vibrational freedom. The contributions to the average molecular energy are

- $3\,(kT/2)$ from translation
- $+3\,(kT/2)$ from rotation
- $+3kT$ from vibration
- $=6kT$ in total

Since the heat capacity is

$$C_V = \left(\frac{\partial \epsilon}{\partial T}\right)_v$$

each translational degree of freedom can contribute $k/2$ to the heat capacity. Each rotational degree of freedom can also contribute $k/2$ to the heat capacity. Each vibrational degree of freedom can contribute k to the heat capacity. It is important to remember that these results represent upper limits for real molecules. These limits are realized at high temperatures, or more precisely, at temperatures where many energy levels, ϵ_i, satisfy $\epsilon_i \ll kT$

4.4: Partition Functions and Average Energies at High Temperatures

It is enlightening to find the integral approximations to the partition functions and average energies for our simple quantum-mechanical models of translational, rotational, and vibrational motions. In doing so, however, it is important to remember that the use of integrals to approximate Boltzmann-equation sums assumes that there are a large number of energy levels, ϵ_i, for which $\epsilon_i \ll kT$. If we select a high enough temperature, the energy levels for any motion will always satisfy this condition. The energy levels for translational motion satisfy this condition even at sub-ambient temperatures. This is the reason that Maxwell's derivation of the probability density function for translational motion is successful.

Rotational motion is an intermediate case. At sub-ambient temperatures, the classical-mechanical derivation can be inadequate; at ordinary temperatures, it is a good approximation. This can be seen by comparing the classical-theory prediction to experimental values for diatomic molecules. For diatomic molecules, the classical model predicts a constant-volume heat capacity of $5k/2$ from 3 degrees of translational and 2 degrees of rotational freedom. Since this does not include the contributions from vibrational motions, constant-volume heat capacities for diatomic molecules must be greater than $5k/2$ if both the translational and rotational contributions are accounted for by the classical model. For diatomic molecules at 298 K, the experimental values are indeed somewhat larger than $5k/2$. (Hydrogen is an exception; its value is $2.47\ k$.)

Vibrational energies are usually so big that only a minor fraction of the molecules can be in higher vibrational levels at reasonable temperatures. If we try to increase the temperature enough to make the high-temperature approximation describe vibrational motions, most molecules decompose. Likewise, electronic partition functions must be evaluated from the defining equation.

The high-temperature limiting average energies can also be calculated from the Boltzmann equation and the appropriate quantum-mechanical energies. Recall that we find the following quantum-mechanical energies for simple models of translational, rotational, and vibrational motions:

Translation

$$\epsilon_{\text{trans}}^{(n)} = \frac{n^2h^2}{8m\ell^2}$$

(n = 1, 2, 3, ... Derived for a particle in a box)

Rotation

$$\epsilon_{\text{rot}}^{(m)} = \frac{m^2h^2}{8\pi^2 I}$$

(m = 1, 2, 3, ... Derived for rotation about one axis—each energy level is doubly degenerate)

Vibration

$$\epsilon_{\text{vibration}}^{(n)} = h\nu\left(n + \frac{1}{2}\right)$$

(n = 0, 1, 2, 3, ... Derived for simple harmonic motion in one dimension)

When we assume that the temperature is so high that many ϵ_i are small compared to kT, we find the following high-temperature limiting partition functions for these motions:

$$z_{\text{translation}} = \sum_{n=1}^{\infty} \exp\left(\frac{-n^2h^2}{8m\ell^2kT}\right) \approx \int_0^{\infty} \exp\left(\frac{-n^2h^2}{8m\ell^2kT}\right) dn = \left(\frac{2\pi mkT\ell^2}{h^2}\right)^{1/2}$$

$$z_{\text{rotation}} = \sum_{m=1}^{\infty} 2\exp\left(\frac{-m^2h^2}{8\pi^2IkT}\right) \approx 2\int_0^{\infty} \exp\left(\frac{-m^2h^2}{8\pi^2IkT}\right) dn = \left(\frac{8\pi^3IkT}{h^2}\right)^{1/2}$$

$$z_{\text{vibration}} = \sum_{n=0}^{\infty} \exp\left(\frac{-h\nu}{kT}\left(n + \frac{1}{2}\right)\right) \approx \int_0^{\infty} \exp\left(\frac{-h\nu}{kT}\left(n + \frac{1}{2}\right)\right) dn = \frac{kT}{h\nu}\exp\left(\frac{-h\nu}{2kT}\right)$$

We can then calculate the average energy for each mode as

$$\langle \epsilon \rangle = z^{-1} \int_0^\infty \epsilon_n \exp\left(\frac{-\epsilon_n}{kT}\right)\, dn$$

and find

$$\begin{aligned}
\langle \epsilon_{\text{translation}} \rangle &= z^{-1}_{\text{translation}} \int_0^\infty \left(\frac{n^2h^2}{8m\ell^2}\right)\, \exp\left(\frac{-n^2h^2}{8m\ell^2 kT}\right) dn \\
&= \frac{kT}{2} \\
\langle \epsilon_{\text{rotation}} \rangle &= z^{-1}_{\text{rotation}} \int_0^\infty 2\left(\frac{m^2h^2}{8\pi^2 I}\right)\, \exp\left(\frac{-m^2h^2}{8\pi^2 IkT}\right) dm \\
&= \frac{kT}{2} \\
\langle \epsilon_{\text{vibration}} \rangle &= z^{-1}_{\text{vibration}} \times \int_0^\infty h\nu\left(n+\frac{1}{2}\right)\, \exp\left(\frac{-h\nu}{kT}\left(n+\frac{1}{2}\right)\right) dn \\
&= kT + \frac{h\nu}{2} \\
&\approx kT
\end{aligned}$$

where the last approximation assumes that $h\nu/2 \ll kT$. In the limit as $T \to 0$, the average energy of the vibrational mode becomes just $h\nu/2$. This is just the energy of the lowest vibrational state, implying that all of the molecules are in the lowest vibrational energy level at absolute zero.

4.5: Energy Levels for a Three-dimensional Harmonic Oscillator

One of the earliest applications of quantum mechanics was Einstein's demonstration that the union of statistical mechanics and quantum mechanics explains the temperature variation of the heat capacities of solid materials. In Section 7.14, we note that the heat capacities of solid materials approach zero as the temperature approaches absolute zero. We also review the law of Dulong and Petit, which describes the limiting heat capacity of many solid elements at high (ambient) temperatures. The Einstein model accounts for both of these observations.

The physical model underlying Einstein's development is that a monatomic solid consists of atoms vibrating about fixed points in a lattice. The particles of this solid are distinguishable from one another, because the location of each lattice point is uniquely specified. We suppose that the vibration of any one atom is independent of the vibrations of the other atoms in the lattice. We assume that the vibration results from a Hooke's Law restoring force

$$\vec{F} = -\lambda \vec{r} = -\lambda \left(x \vec{i} + y \vec{j} + z \vec{k} \right)$$

that is zero when the atom is at its lattice point, for which $\vec{r} = (0, 0, 0)$. The potential energy change when the atom, of mass m, is driven from its lattice point to the point $\vec{r} = (x, y, x)$ is

$$V = \int_{\vec{r}=\vec{0}}^{\vec{r}} -\vec{F} \bullet d\vec{r} = \lambda \int_{x=0}^{x} x dx + \lambda \int_{y=0}^{y} y dy + \lambda \int_{z=0}^{z} z dz = \lambda \frac{x^2}{2} + \lambda \frac{y^2}{2} + \lambda \frac{z^2}{2}$$

The Schrödinger equation for this motion is

$$-\frac{h^2}{8\pi^2 m} \left[\frac{\partial^2 \psi}{\partial x^2} + \frac{\partial^2 \psi}{\partial y^2} + \frac{\partial^2 \psi}{\partial z^2} \right] + \lambda \left[\frac{x^2}{2} + \frac{y^2}{2} + \frac{z^2}{2} \right] \psi = \epsilon \psi$$

where ψ is a function of the three displacement coordinates; that is $\psi = \psi(x, y, z)$. We assume that motions in the x-, y-, and z-directions are completely independent of one another. When we do so, it turns out that we can express the three-dimensional Schrödinger equation as the sum of three one-dimensional Schrödinger equations

$$\begin{aligned} &\left[-\frac{h^2}{8\pi^2 m} \frac{\partial^2 \psi_x}{\partial x^2} + \lambda \frac{x^2 \psi_x}{2} \right] \\ &+ \left[-\frac{h^2}{8\pi^2 m} \frac{\partial^2 \psi_y}{\partial y^2} + \lambda \frac{y^2 \psi_y}{2} \right] \\ &+ \left[-\frac{h^2}{8\pi^2 m} \frac{\partial^2 \psi_z}{\partial z^2} + \lambda \frac{z^2 \psi_z}{2} \right] \\ &= \epsilon \psi_x + \epsilon \psi_y + \epsilon \psi_z \end{aligned}$$

where any wavefunction $\psi_x^{(n)}$ is the same function as $\psi_y^{(n)}$ and $\psi_z^{(n)}$, and the corresponding energies $\epsilon_x^{(n)}$, $\epsilon_y^{(n)}$, and $\epsilon_z^{(n)}$ have the same values. The energy of the three-dimensional atomic motion is simply the sum of the energies for the three one-dimensional motions. That is,

$$\epsilon_{n,m,p} = \epsilon_x^{(n)} + \epsilon_y^{(m)} + \epsilon_z^{(p)},$$

which, for simplicity, we also write as

$$\epsilon_{n,m,p} = \epsilon_n + \epsilon_m + \epsilon_p.$$

4.6: Energy and Heat Capacity of the "Einstein Crystal"

In Section 22.4, we find an approximate partition function for the harmonic oscillator at high temperatures. Because it is a geometric series, the partition function for the harmonic oscillator can also be obtained exactly at any temperature. By definition, the partition function for the harmonic oscillator is

$$z=\sum_{n=0}^{\infty}\exp\left(\frac{-h\nu}{kT}\left(n+\frac{1}{2}\right)\right)=\exp\left(\frac{-h\nu}{2kT}\right)\sum_{n=0}^{\infty}\exp\left(\frac{-nh\nu}{kT}\right)=\exp\left(\frac{-h\nu}{2kT}\right)\sum_{n=0}^{\infty}\left[\exp\left(\frac{-h\nu}{kT}\right)\right]^{n}$$

This is just the infinite sum

$$z=a\sum_{n=0}^{\infty}r^{n}=\frac{a}{1-r}$$

with

$$a=\exp\left(\frac{-h\nu}{2kT}\right)$$

and

$$r=\exp\left(\frac{-h\nu}{kT}\right)$$

Hence, the exact partition function for the one-dimensional harmonic oscillator is

$$z=\frac{\exp\left(-h\nu/2kT\right)}{1-\exp\left(-h\nu/kT\right)}$$

The partition function for vibration in each of the other two dimensions is the same. To get the partition function for oscillation in all three dimensions, we must sum over all possible combinations of the three energies. Distinguishing the energies associated with motion in the x-, y-, and z-directions by the subscripts n, m, and p, respectively, we have for the three-dimensional harmonic oscillator:

$$\begin{aligned}z_{3D}&=\sum_{p=0}^{\infty}\sum_{m=0}^{\infty}\sum_{n=0}^{\infty}\exp\left[\frac{-\left(\epsilon_n+\epsilon_m+\epsilon_p\right)}{kT}\right]\\&=\sum_{p=0}^{\infty}\exp\frac{-\epsilon_p}{kT}\sum_{m=0}^{\infty}\exp\frac{-\epsilon_m}{kT}\sum_{n=0}^{\infty}\exp\frac{-\epsilon_n}{kT}\\&=z^3\end{aligned}$$

Hence,

$$z_{3D}=\left[\frac{\exp\left(-h\nu/2kT\right)}{1-\exp\left(-h\nu/kT\right)}\right]^3$$

and the energy of a crystal of N, independent, distinguishable atoms is

$$\begin{aligned}E&=N\left\langle\epsilon\right\rangle\\&=NkT^2\left(\frac{\partial\ln z_{3D}}{\partial T}\right)_V\\&=\frac{3Nh\nu}{2}+\frac{3Nh\nu\exp\left(-h\nu/kT\right)}{1-\exp\left(-h\nu/kT\right)}\end{aligned}$$

Taking the partial derivative with respect to temperature gives the heat capacity of this crystal. The molar heat capacity can be expressed in two ways that are useful for our purposes:

$$
\begin{aligned}
C_V &= \left(\frac{\partial \overline{E}}{\partial T}\right)_V \\
&= 3\overline{N}k\left(\frac{h\nu}{kT}\right)^2\left[\frac{\exp\left(-h\nu/kT\right)}{\left(1-\exp\left(-h\nu/kT\right)\right)^2}\right] \\
&= 3\overline{N}k\left(\frac{h\nu}{kT}\right)^2\left[\frac{\exp\left(h\nu/kT\right)}{\left(\exp\left(h\nu/kT\right)-1\right)^2}\right]
\end{aligned}
$$

Consider the heat capacity at high temperatures. As the temperature becomes large, $h\nu/kT$ approaches zero. Then

$$
\exp\left(\frac{h\nu}{kT}\right) \approx 1+\frac{h\nu}{kT}
$$

Using this approximation in the second representation of C_V gives for the high temperature limit

$$
\begin{aligned}
C_V &\approx 3\overline{N}k\left(\frac{h\nu}{kT}\right)^2\left[\frac{1+h\nu/kT}{\left(1+h\nu/kT-1\right)^2}\right] \\
&\approx 3\overline{N}k\left(1+\frac{h\nu}{kT}\right) \\
&\approx 3\overline{N}k = 3R
\end{aligned}
$$

Since C_V and C_P are about the same for solids at ordinary temperatures, this result is essentially equivalent to the law stated by Dulong and Petit. Indeed, it suggests that the law would be more accurate if stated as a condition on C_V rather than C_P, and this proves to be the case.

At low temperatures, $h\nu/kT$ becomes arbitrarily large and $\exp\left(-h\nu/kT\right)$ approaches zero. From the first representation of C_V, we see that

$$
\lim_{T\to 0}\left(\frac{\partial \overline{E}}{\partial T}\right)_V = C_V = 0
$$

In Section 10.9, we see that $C_P - C_V \to 0$ as $T \to 0$. Hence, the theory also predicts that $C_P \to 0$ as $T \to 0$, in agreement with experimental results.

The Einstein model assumes that energy variations in a solid near absolute zero are entirely due to variations in the vibrational energy. From the assumption that all of these vibrational motions are characterized by a single frequency, it predicts the limiting values for the heat capacity of a solid at high and low temperatures. At intermediate temperatures, the quantitative predictions of the Einstein model leave room for improvement. An important refinement developed by Peter Debye assumes a spectrum of vibrational frequencies and results in excellent quantitative agreement with experimental values at all temperatures.

We can give a simple qualitative interpretation for the result that heat capacities decrease to zero as the temperature goes to absolute zero. The basic idea is that, at a sufficiently low temperature, essentially all of the molecules in the system are in the lowest available energy level. Once essentially all of the molecules are in the lowest energy level, the energy of the system can no longer decrease in response to a further temperature decrease. Therefore, in this temperature range, the heat capacity is essentially zero. Alternatively, we can say that as the temperature approaches zero, the fraction of the molecules that are in the lowest energy level approaches one, and the energy of the system of N molecules approaches the smallest value it can have.

The weakness in this qualitative view is that there is always a non-zero probability of finding molecules in a higher energy level, and this probability changes as the temperature changes. To firm up the simple picture, we need a way to show that the energy decreases more rapidly than the temperature near absolute zero. More precisely, we need a way to show that

$$
\lim_{T\to 0}\left(\frac{\partial \overline{E}}{\partial T}\right)_V = C_V = 0
$$

Since the Einstein model produces this result, it constitutes a quantitative validation of our qualitative model.

4.7: Applications of Other Entropy Relationships

In most cases, calculation of the entropy from information about the energy levels of a system is best accomplished using the partition function. Occasionally other entropy relationships are useful. We illustrate this by using the entropy relationship

$$S = -Nk\sum_{i=1}^{\infty} g_i \rho\left(\epsilon_i\right) \ln \rho\left(\epsilon_i\right) \ + Nk \ln g_1$$

to find the entropy of an N-molecule disordered crystal at absolute zero. To be specific, let us consider a crystal of carbon monoxide.

We can calculate the entropy of carbon monoxide at absolute zero from either of two perspectives. Let us first assume that the energy of a molecule is almost completely independent of the orientations of its neighbors in the crystal. Then the energy of any molecule in the crystal is essentially the same in either of the two orientations available to it. In this model for the system, we consider that there are two, non-degenerate, low-energy quantum states available to the molecule. We suppose that all other quantum states lie at energy levels whose probabilities are very small when the temperature is near absolute zero. We have $g_1 = g_2 = 1$, $\epsilon_2 \approx \epsilon_1$. Near absolute zero, we have $\rho\left(\epsilon_2\right) \approx \rho\left(\epsilon_1\right) \approx 1/2$; for $i > 2$, $\rho\left(\epsilon_i\right) \approx 0$. The entropy becomes

$$\begin{aligned} S &= -Nk\sum_{i=1}^{\infty} g_i \rho\left(\epsilon_i\right) \ln \rho\left(\epsilon_i\right) \ + Nk \ln g_1 \\ &= -Nk\left(\frac{1}{2}\right) \ln\left(\frac{1}{2}\right) \ - Nk\left(\frac{1}{2}\right) \ln\left(\frac{1}{2}\right) \\ &= -Nk \ln\left(\frac{1}{2}\right) \\ &= Nk \ln 2 \end{aligned}$$

Alternatively, we can consider that there is just one low-energy quantum state available to the molecule but that this quantum state is doubly degenerate. In this model, the energy of the molecule is the same in either of the two orientations available to it. We have $g_1 = 2$. Near absolute zero, we have $\rho\left(\epsilon_1\right) \approx 1$; for $i > 1$, $\rho\left(\epsilon_i\right) \approx 0$. The summation term vanishes, and the entropy becomes

$$S = Nk \ln g_1 \ = Nk \ln 2$$

Either perspective implies the same value for the zero-temperature entropy of the N-molecule crystal.

Either of these treatments involves a subtle oversimplification. In our first model, we recognize that the carbon monoxide molecule must have a different energy in each of its two possible orientations in an otherwise perfect crystal. The energy of the orientation that makes the crystal perfect is slightly less than the energy of the other orientation. We introduce an approximation when we say that $\rho\left(\epsilon_2\right) \approx \rho\left(\epsilon_1\right) \approx 1/2$. However, if ϵ_2 is not exactly equal to ϵ_1, this approximation cannot be valid at an arbitrarily low temperature. To see this, we let the energy difference between these orientations be $\epsilon_2 - \epsilon_1 = \Delta\epsilon > 0$. At relatively high temperatures, at which $\Delta\epsilon \ll kT$, we have

$$\frac{\rho\left(\epsilon_2\right)}{\rho\left(\epsilon_1\right)} = \exp\left(\frac{-\Delta\epsilon}{kT}\right) \ \approx 1$$

and $\rho\left(\epsilon_2\right) \approx \rho\left(\epsilon_1\right) \approx 1/2$. At such temperatures, the system behaves as if the lowest energy level were doubly degenerate, with $\epsilon_2 = \epsilon_1$. However, since T can be arbitrarily close to zero, this condition cannot always apply. No matter how small $\Delta\epsilon$ may be, there are always temperatures at which $\Delta\epsilon \gg kT$ and at which we have

$$\frac{\rho\left(\epsilon_2\right)}{\rho\left(\epsilon_1\right)} \approx 0$$

This implies that the molecule should always adopt the orientation that makes the crystal perfectly ordered when the temperature becomes sufficiently close to zero. This conclusion disagrees with the experimental observations.

Our second model assumes that the energy of a carbon monoxide molecule is the same in either of its two possible orientations. However, its interactions with the surrounding molecules cannot be exactly the same in each orientation; consequently, its energy cannot be exactly the same. From first principles, therefore, our second model cannot be strictly correct.

To resolve these apparent contradictions, we assume that the rate at which a carbon monoxide molecule can change its orientation within the lattice depends on temperature. For some temperature at which $\Delta\epsilon \ll kT$, the reorientation process occurs rapidly, and the two orientations are equally probable. As the temperature decreases, the rate of reorientation becomes very slow. If the reorientation process effectively ceases to occur while the condition $\Delta\epsilon \ll kT$ applies, the orientations of the component molecules remain those that occur at higher temperatures no matter how much the temperature decreases thereafter. This is often described by saying that molecular orientations become "frozen." The zero-temperature entropy of the system is determined by the energy-level probabilities that describe the system at the temperature at which reorientation effectively ceases to occur.

4.8: Problems

1. The gravitational potential energies available to a molecule near the surface of the earth are $\epsilon(h) = mgh$. Each height, h, corresponds to a unique energy, so we can infer that the degeneracy of $\epsilon(h)$ is unity. Derive the probability density function for the distribution of molecules in the earth's atmosphere. (See Problem 19 in Chapter 3.)

2. The value of the molecular partition function approximates the number of quantum states that are available to the molecule and whose energy is less than kT. How many such quantum states are available to a molecule of molecular weight 40 that is confined in a volume of 10^{-6} m^3 at 300 K?

CHAPTER OVERVIEW

5: The Ensemble Treatment

5.1: Ensembles of N-molecule Systems

When we begin our discussion of Boltzmann statistics in Chapter 20, we note that there exists, in principle, a Schrödinger equation for an N-molecule system. For any particular set of boundary conditions, the solutions of this equation are a set of infinitely many wavefunctions, $\Psi_{i,j}$, for the N-molecule system. For every such wavefunction, there is a corresponding system energy, E_i. The wavefunctions reflect all of the attractive and repulsive interactions among the molecules of the system. Likewise, the energy levels of the system reflect all of these interactions.

In Section 20.12, we introduce the symbol Ω_E to denote the degeneracy of the energy, E, of an N-molecule system. Because the constituent molecules are assumed to be distinguishable and non-interacting, we have

$$\Omega_E = \sum_{\{N_i\},E} W\left(N_i, g_i\right)$$

In the solution of the Schrödinger equation for a system of N interacting molecules, each system-energy level, E_i, can be degenerate. We again let Ω denote the degeneracy of an energy level of the system. We use Ω_i (rather than Ω_{E_i}) to represent the degeneracy of E_i. It is important to recognize that the symbol "Ω_i" now denotes an intrinsic quantum-mechanical property of the N-particle system.

In Chapters 21 and 22, we denote the parallel properties of an individual molecule by $\psi_{i,j}$ for the molecular wavefunctions, ϵ_i for the corresponding energy levels, and g_i for the degeneracy of the i^{th} energy level. We imagine creating an N-molecule system by collecting N non-interacting molecules in a fixed volume and at a fixed temperature.

In exactly the same way, we now imagine collecting $\hat{N}$ of these N-molecule, constant-volume, constant-temperature systems. An aggregate of many multi-molecule systems is called an ***ensemble***. Just as we assume that no forces act among the non-interacting molecules we consider earlier, we assume that no forces act among the systems of the ensemble. However, as we emphasize above, our model for the systems of an ensemble recognizes that intermolecular forces among the molecules of an individual system can be important. We can imagine specifying the properties of the individual systems in a variety of ways. A collection is called a ***canonical ensemble*** if each of the systems in the ensemble has the same values of N, V, and T. (The sense of this name is that by specifying constant N, V, and T, we create the ensemble that can be described most simply.)

The canonical ensemble is a collection of $\hat{N}$ identical systems, just as the N-molecule system is a collection of N identical molecules. We imagine piling the systems that comprise the ensemble into a gigantic three-dimensional stack. We then immerse the entire stack—the ensemble—in a constant temperature bath. The ensemble and its constituent systems are at the constant temperature T. The volume of the ensemble is $\hat{N}V$. Because we can specify the location of any system in the ensemble by specifying its x-, y-, and z-coordinates in the stack, the individual systems that comprise the ensemble are distinguishable from one another. Thus the ensemble is analogous to a crystalline N-molecule system, in which the individual molecules are distinguishable from one another because each occupies a particular location in the crystal lattice, the entire crystal is at the constant temperature, T, and the crystal volume is NV_{molecule}.

Since the ensemble is a conceptual construct, we can make the number of systems in the ensemble, $\hat{N}$, as large as we please. Each system in the ensemble will have one of the quantum-mechanically allowed energies, E_i. We let the number of systems that have energy E_1 be $\hat{N}_1$. Similarly, we let the number with energy E_2 be $\hat{N}_2$, and the number with energy E_i be $\hat{N}_i$. Thus at any given instant, the ensemble is characterized by a population set, $\{\hat{N}_1,\ \hat{N}_2,\ \ldots,\ \hat{N}_i, \ldots\}$, in exactly the same way that an N-molecule system is characterized by a population set, $\{N_1,\ N_2, \ldots,\ N_i, \ldots\}$. We have

$$\hat{N} = \sum_{i=1}^{\infty} \hat{N}_i$$

While all of the systems in the ensemble are immersed in the same constant-temperature bath, the energy of any one system in the ensemble is completely independent of the energy of any other system. This means that the total energy of the ensemble, $\hat{E}$, is given by

$$\hat{E} = \sum_{i=1}^{\infty} \hat{N}_i E_i$$

Property	System	Ensemble
Quantum entity	***Molecule*** at fixed volume and temperature	***System*** comprising a collection of N molecules at fixed volume and temperature
Aggregate of quantum entities	***System*** comprising a collection of N molecules at fixed volume and temperature	***Ensemble*** comprising $\hat{N}$ systems each of which contains N molecules
Number of quantum entities in aggregate	N	$\hat{N}$
Wave functions/quantum states	ψ_i	Ψ_i
Energy levels	ϵ_i	E_i
Energy level degeneracies	g_i	Ω_i
Probability that an energy level is occupied	P_i	$\hat{P}_i$
Number of quantum entities in the i^{th} energy level	N_i	$\hat{N}_i$
Probability that a quantum state is occupied	$\rho\left(\epsilon_i\right)$	$\hat{\rho}\left(E_i\right)$
Energy of the aggregate's k^{th} population set	$E_k = \sum N_{k,i}\epsilon_i$	$\hat{E}_k = \sum \hat{N}_{k,i}\epsilon_i$
Expected value of the energy of the aggregate	$\langle E \rangle = N \sum P_i \epsilon_i$	$\left\langle \hat{E} \right\rangle = \hat{N} \sum \hat{P}_i E_i$

The population set, $\{\hat{N}_1,\ \hat{N}_2,\ \ldots,\ \hat{N}_i, \ldots\}$, that characterizes the ensemble is not constant in time. However, by the same arguments that we apply to the N-molecule system, there is a population set

$$\{\hat{N}_1^{\bullet},\ \hat{N}_2^{\bullet}, \ldots,\ \hat{N}_i^{\bullet}, \ldots\}$$

which characterizes the ensemble when it is at equilibrium in the constant-temperature bath.

We define the probability, $\hat{P}_i$, that a system of the ensemble has energy E_i to be the fraction of the systems in the ensemble with this energy, when the ensemble is at equilibrium at the specified temperature. Thus, by definition,

$$\hat{P}_i = \frac{\hat{N}_i^{\bullet}}{\hat{N}}.$$

We define the probability that a system is in one of the states, $\Psi_{i,j}$, with energy E_i, as

$$\hat{\rho}\left(E_i\right) = \frac{\hat{P}_i}{\Omega_i}$$

The method we have used to construct the canonical ensemble insures that the entire ensemble is always at the specified temperature. If the component systems are at equilibrium, the ensemble is at equilibrium. The expected value of the ensemble energy is

$$\left\langle \hat{E} \right\rangle = \hat{N} \sum_{i=1}^{\infty} \hat{P}_i E_i = \sum_{i=1}^{\infty} \hat{N}_i^{\bullet} E_i$$

Because the number of systems in the ensemble, $\hat{N}$, is very large, we know from the central limit theorem that any observed value for the ensemble energy will be indistinguishable from the expected value. To an excellent approximation, we have at any time,

$$\hat{E} = \left\langle \hat{E} \right\rangle$$

and

$$\hat{N}_i^{\bullet} = \hat{N}_i$$

The table above summarizes the terminology that we have developed to characterize molecules, N-molecule systems, and $\hat{N}$-system ensembles of N-molecule systems.

We can now apply to an ensemble of $\hat{N}$, distinguishable, non-interacting systems the same logic that we applied to a system of N, distinguishable, non-interacting molecules. The probability that a system is in one of the energy levels is

$$1 = \hat{P}_1 + \hat{P}_2 + \cdots + \hat{P}_i + \ldots$$

The total probability sum for the constant-temperature ensemble is

$$1 = \left(\hat{P}_1 + \hat{P}_2 + \cdots + \hat{P}_i + \ldots\right)^{\hat{N}} = \sum_{\{\hat{N}_i\}} \hat{W}\left(\hat{N}_i, \Omega_i\right) \hat{\rho}\left(E_1\right)^{\hat{N}_1} \hat{\rho}\left(E_2\right)^{\hat{N}_2} \ldots \hat{\rho}\left(E_i\right)^{\hat{N}_i} \ldots$$

where

$$\hat{W}\left(\hat{N}_i, \Omega_i\right) = \hat{N}! \prod_{i=1}^{\infty} \frac{\Omega_i^{\hat{N}_i}}{\hat{N}_i!}$$

Moreover, we can imagine instantaneously isolating the ensemble from the temperature bath in which it is immersed. This is a wholly conceptual change, which we effect by replacing the fluid of the constant-temperature bath with a solid blanket of insulation. The ensemble is then an isolated system whose energy, $\hat{E}$, is constant. Every system of the isolated ensemble is immersed in a constant-temperature bath, where the constant-temperature bath consists of the $\hat{N} - 1$ systems that make up the rest of the ensemble. This is an important feature of the ensemble treatment. It means that any conclusion we reach about the systems of the constant-energy ensemble is also a conclusion about each of the $\hat{N}$ identical, constant-temperature systems that comprise the isolated, constant-energy ensemble.

Only certain population sets, $\{\hat{N}_1,\ \hat{N}_2,\ \ldots,\ \hat{N}_i, \ldots\}$, are consistent with the fixed value, $\hat{E}$, of the isolated ensemble. For each of these population sets, there are $\hat{W}\left(\hat{N}_i, \Omega_i\right)$ system states. The probability of each of these system states is proportional to $\hat{\rho}\left(E_1\right)^{\hat{N}_1} \hat{\rho}\left(E_2\right)^{\hat{N}_2} \ldots \hat{\rho}\left(E_i\right)^{\hat{N}_i} \ldots$. By the principle of equal *a priori* probability, every system state of the fixed-energy ensemble occurs with equal probability. We again conclude that the population set that characterizes the equilibrium state of the constant-energy ensemble, $\{\hat{N}_1^{\bullet},\ \hat{N}_2^{\bullet}, \ldots,\ \hat{N}_i^{\bullet}, \ldots\}$, is the one for which $\hat{W}$ or $\ln \hat{W}$ is a maximum, subject to the constraints

$$\hat{N} = \sum_{i=1}^{\infty} \hat{N}_i$$

and

$$\hat{E} = \sum_{i=1}^{\infty} \hat{N}_i E_i$$

The fact that we can make $\hat{N}$ arbitrarily large ensures that any term, $\hat{N}_i^{\bullet}$, in the equilibrium-characterizing population set can be very large, so that $\hat{N}_i^{\bullet}$ can be found using Stirling's approximation and Lagrange's method of undetermined multipliers. We have the mnemonic function

$$F_{mn} = \hat{N} \ln \hat{N} - \hat{N} + \sum_{i=1}^{\infty} \left(\hat{N}_i \ln \Omega_i\ - \hat{N}_i \ln \hat{N}_i\ + \hat{N}_i\right) + \alpha\left(\hat{N} - \sum_{i=1}^{\infty} \hat{N}_i\right) + \beta\left(\hat{E} - \sum_{i=1}^{\infty} \hat{N}_i E_i\right)$$

so that

$$\left(\frac{\partial F_{mn}}{\partial \hat{N}_i^{\bullet}}\right)_{j \neq i} = \ln \Omega_i\ - \frac{\hat{N}_i^{\bullet}}{\hat{N}_i^{\bullet}} - \ln \hat{N}_i^{\bullet}\ + 1 - \alpha - \beta E_i = 0$$

and

$$\ln \hat{N}_i^{\bullet} = \ln \Omega_i - \alpha - \beta E_i$$

or

$$\hat{N}_i^{\bullet} = \Omega_i \exp(-\alpha) \exp -\beta E_i$$

When we make use of the constraint on the total number of systems in the ensemble, we have

$$\hat{N} = \sum_{i=1}^{\infty} \hat{N}_i^{\bullet} = \exp(-\alpha) \sum_{i=1}^{\infty} \Omega_i \exp(-\beta E_i)$$

so that

$$\exp(-\alpha) = \hat{N} Z^{-1}$$

where the partition function for a system of N possibly-interacting molecules is

$$Z = \sum_{i=1}^{\infty} \Omega_i \exp(-\beta E_i)$$

The probability that a system has energy E_i is equal to the equilibrium fraction of systems in the ensemble that have energy E_i, so that

$$\hat{P}_i = \frac{\hat{N}_i^{\bullet}}{\hat{N}} = \frac{\Omega_i \exp(-\beta E_i)}{Z}$$

5.2: The Ensemble Entropy and the Value of ß

At equilibrium, the entropy of the $\hat{N}$-system ensemble, S_{ensemble}, must be a maximum. By arguments that parallel those in Chapter 20, $\hat{W}$ is a maximum for the ensemble population set that characterizes this equilibrium state. Applying the Boltzmann definition to the ensemble, the ensemble entropy is $S_{\text{ensemble}} = k\ln \hat{W}_{\max}$. Since all $\hat{N}$ systems in the ensemble have effectively the same entropy, S, we have $S_{\text{ensemble}} = \hat{N}S$. When we assume that $\hat{W}_{\max}$ occurs for the equilibrium population set, $\left\{\hat{N}_1^{\bullet},\ \hat{N}_2^{\bullet}, \ldots,\ \hat{N}_i^{\bullet}, \ldots\right\}$, we have

$$\hat{W}_{\max} = \hat{N}! \prod_{i=1}^{\infty} \frac{\Omega_i^{\hat{N}_i^{\bullet}}}{\hat{N}_i^{\bullet}!}$$

so that

$$S_{\text{ensemble}} = \hat{N}S = k\ln \hat{N}! + k\sum_{i=1}^{\infty} \hat{N}_i^{\bullet} \ln \Omega_i - k\sum_{i=1}^{\infty} \ln\left(\hat{N}_i^{\bullet}!\right)$$

From the Boltzmann distribution function, $\hat{N}_i^{\bullet}/\hat{N} = Z^{-1}\Omega_i \exp\left(-\beta E_i\right)$, we have

$$\ln \Omega_i \ = \ln Z \ + \ln \hat{N}_i^{\bullet} \ + \beta E_i - \ln \hat{N}$$

Substituting, and introducing Stirling's approximation, we find

$$\begin{aligned} \hat{N}S \ &= k\hat{N}\ln \hat{N} - k\hat{N} \ + k\sum_{i=1}^{\infty} \hat{N}_i^{\bullet} \left(\ln Z + \ln \hat{N}_i^{\bullet} \ + \beta E_i - \ln \hat{N} \right) - k\sum_{i=1}^{\infty} \left(\hat{N}_i^{\bullet} \ln \hat{N}_i^{\bullet} - \hat{N}_i^{\bullet} \right) \\ &= \hat{N}k\ln Z \ + k\beta \sum_{i=1}^{\infty} \hat{N}_i^{\bullet} E_i \end{aligned}$$

Since $\sum_{i=1}^{\infty} \hat{N}_i^{\bullet} E_i$ is the energy of the $\hat{N}$-system ensemble and the energy of each system is the same, we have

$$\sum_{i=1}^{\infty} \hat{N}_i^{\bullet} E_i = E_{\text{ensemble}} = \hat{N}E$$

Substituting, we find

$$S = k\beta E + k\ln Z$$

where S, E, and Z are the entropy, energy, and partition function for the N-molecule system. From the fundamental equation, we have

$$\left(\frac{\partial E}{\partial S}\right)_V = T$$

Differentiating $S = k\beta E + k\ln Z$ with respect to entropy at constant volume, we find

$$1 = k\beta \left(\frac{\partial E}{\partial S}\right)_V$$

and it follows that

$$\beta = \frac{1}{kT}$$

We have, for the N-molecule system

$$Z = \sum_{i=1}^{\infty} \Omega_i \exp\left(\frac{-E_i}{kT}\right)$$

(System partition function)

$$\hat{P}_i = Z^{-1}\Omega_i \exp\left(\frac{-E_i}{kT}\right)$$

(Boltzmann's equation)

$$S = \frac{E}{T} + k\ln Z$$

(Entropy of the N-molecule system)

5.3: The Thermodynamic Functions of the N-molecule System

With the results of Section 23.2 in hand, we can find the other thermodynamic functions for the N-molecule system from the equations for Z and $\hat{P}_i$ by the arguments we use in Chapters 20 and 21. Let us summarize these arguments. From

$$E = \sum_{i=1}^{\infty} \hat{P}_i E_i$$

we have

$$dE = \sum_{i=1}^{\infty} E_i d\hat{P}_i + \sum_{i=1}^{\infty} \hat{P}_i dE_i$$

We associate the first term with dq^{rev} and the second term with $dw = -PdV$; that is,

$$dq^{rev} = TdS = \sum_{i=1}^{\infty} E_i d\hat{P}_i = -kT \sum_{i=1}^{\infty} \ln\left(\frac{\hat{P}_i}{\Omega_i}\right) d\hat{P}_i - kT \ln Z \sum_{i=1}^{\infty} d\hat{P}_i$$

Where we substitute

$$E_i = -kT\ln\left(\frac{\hat{P}_i}{\Omega_i}\right) - kT\ln Z$$

which we obtain by taking the natural logarithm of the partition function. Since $\sum_{i=1}^{\infty} d\hat{P}_i = 0$, we have for each system,

$$dS = -k\sum_{i=1}^{\infty} \ln\left(\frac{\hat{P}_i}{\Omega_i}\right) d\hat{P}_i = -k\sum_{i=1}^{\infty} \left\{ \Omega_i d\left(\frac{\hat{P}_i}{\Omega_i} \ln \frac{\hat{P}_i}{\Omega_i}\right) - d\hat{P}_i \right\} = -k\sum_{i=1}^{\infty} d\left(\hat{P}_i \ln \frac{\hat{P}_i}{\Omega_i}\right)$$

The system entropy, S, and the system-energy-level probabilities, $\hat{P}_i$, are functions of temperature. Integrating from $T = 0$ to T and choosing the lower limits for the integrations on the right to be $\hat{P}_1(0) = 1$ and $\hat{P}_i(0) = 0$ for $i > 1$, we have

$$\int_{S_0}^{S} dS = -k\sum_{i=1}^{\infty} \int_{\hat{P}_i(0)}^{\hat{P}_i(T)} d\left(\hat{P}_i \ln \frac{\hat{P}_i}{\Omega_i}\right)$$

Letting $\hat{P}_i(T) = \hat{P}_i$, the result is

$$\begin{aligned} S - S_0 &= -k\hat{P}_1 \ln \frac{\hat{P}_1}{\Omega_1} + k\ln\frac{1}{\Omega_1} - k\sum_{i=2}^{\infty} \hat{P}_i \ln \frac{\hat{P}_i}{\Omega_i} \\ &= -k\sum_{i=1}^{\infty} \hat{P}_i \ln \frac{\hat{P}_i}{\Omega_i} - k\ln\Omega_1 \end{aligned}$$

From the partition function, we have

$$\ln\left(\frac{\hat{P}_i}{\Omega_i}\right) = -\frac{E_i}{kT} + \ln Z$$

so that

$$\begin{aligned} S - S_0 &= -k\sum_{i=1}^{\infty} \hat{P}_i \left(-\frac{E_i}{kT} + \ln Z\right) - k\ln\Omega_1 \\ &= \frac{1}{T}\sum_{i=1}^{\infty} \hat{P}_i E_i + k\ln Z \sum_{i=1}^{\infty} \hat{P}_i - k\ln\Omega_1 \\ &= \frac{E}{T} + k\ln Z - k\ln\Omega_1 \end{aligned}$$

We take the system entropy at absolute zero, S_0, to be

$$S_0 = k\ln\Omega_1$$

If the lowest energy state is non-degenerate, $\Omega_1 = 1$, and $S_0 = 0$, so that

$$S\left(T\right) = \frac{E}{T} + k\ln Z$$

As in Section 21.6, we observe that

$$E = \sum_{i=1}^{\infty}\hat{P}_i E_i = Z^{-1}\sum_{i=1}^{\infty}\Omega_i E_i \exp\left(\frac{-E_i}{kT}\right)$$

and that

$$\left(\frac{\partial \ln Z}{\partial T}\right)_V = Z^{-1}\sum_{i=1}^{\infty}\Omega_i\left(\frac{E_i}{kT^2}\right)\exp\left(\frac{-E_i}{kT}\right) = \frac{E}{kT^2}$$

so that

$$E = kT^2\left(\frac{\partial \ln Z}{\partial T}\right)_V$$

From $A = E - TS$ and the entropy equation, $S = E/T + k\ln Z$, the Helmholtz free energy of the system is

$$A = -kT\ln Z$$

For the system pressure, we find from

$$P = -\left(\frac{\partial A}{\partial V}\right)_T$$

that

$$P = kT\left(\frac{\partial \ln Z}{\partial V}\right)_T$$

From $H = E + PV$, we find

$$H = kT^2\left(\frac{\partial \ln Z}{\partial T}\right)_V + VkT\left(\frac{\partial \ln Z}{\partial V}\right)_T$$

and from $G = E + PV - TS$, we find

$$G = VkT\left(\frac{\partial \ln Z}{\partial V}\right)_T - kT\ln Z$$

For the chemical potential per molecule in the N-molecule system, we obtain

$$\mu = \left(\frac{\partial A}{\partial N}\right)_{VT} = -kT\left(\frac{\partial \ln Z}{\partial N}\right)_{VT}$$

Thus, we have found the principle thermodynamic functions for the N-molecule system expressed in terms of $\ln Z$ and its derivatives. The system partition function, Z, depends on the energy levels available to the N-molecule system. The thermodynamic functions we have obtained are valid for any system, including systems in which intermolecular forces make large contributions to the system energy. Of course, the system partition function, Z, must accurately reflect the effects of these forces.

In Chapter 24 we find that the partition function, Z, for a system of N, distinguishable, non-interacting molecules is related in a simple way to the molecular partition function, z. We find $Z = z^N$. When we substitute this result for Z into the system partition functions developed above, we recover the same results that we developed in Chapters 20 and 21 for the thermodynamic properties of a system of N, distinguishable, non-interacting molecules.

CHAPTER OVERVIEW

6: The Distribution of Gas Velocities

6.1: Distribution Functions for Gas-velocity Components

In Chapter 2, we assume that all of the molecules in a gas move with the same speed and use a simplified argument to conclude that this speed depends only on temperature. We now recognize that the individual molecules in a gas sample have a wide range of speeds; the velocities of gas molecules must be described by a distribution function. It is true, however, that the average speed depends only on temperature.

James Clerk Maxwell was the first to derive the distribution function for gas velocities. He did it about 1860. We follow Maxwell's argument. For a molecule moving in three dimensions, there are three velocity components. Maxwell's argument uses only one assumption: the speed of a gas molecule is independent of the direction in which it is moving. Equivalently, we can say that the components of the velocity of a gas molecule are independent of one another; knowing the value of one component of a molecule's velocity does not enable us to infer anything about the values of the other two components. When we use Cartesian coordinates, Maxwell's assumptionMaxwell's assumption means also that the same mathematical model must describe the distribution of each of the velocity components.

Since the velocity of a gas molecule has three components, we must treat the velocity distribution as a function of three random variables. To understand how this can be done, let us consider how we might find probability distribution functions for velocity components. We need to consider both spherical and Cartesian coordinate systems.

Let us suppose that we are able to measure the Cartesian-coordinate components v_x, v_y, and v_z of the velocities of a large number of randomly selected gas molecules in a particular constant-temperature sample. Then we can transform each set of Cartesian components to spherical-coordinate velocity covelocity componentsmponents v, θ, and φ. We imagine accumulating the results of these measurements in a table like Table 1. As a practical matter, of course, we cannot make the measurements to complete such a table. However, there is no doubt that, at every instant, every gas molecule can be characterized by a set of such velocity components; the values exist, even if we cannot measure them. We imagine that we have such data only as a way to clarify the properties of the distribution functions that we need.

Table 1. Molecular Velocity Components

Molecule Number	v_x	v_x	v_x	v	θ	φ
1	$v_x\,(1)$	$v_y\,(1)$	$v_z\,(1)$	$v\,(1)$	$\theta\,(1)$	$\varphi\,(1)$
2	$v_x\,(2)$	$v_y\,(2)$	$v_z\,(2)$	$v\,(2)$	$\theta\,(2)$	$\varphi\,(2)$
3	$v_x\,(3)$	$v_y\,(3)$	$v_z\,(3)$	$v\,(3)$	$\theta\,(3)$	$\varphi\,(3)$
4	$v_x\,(4)$	$v_y\,(4)$	$v_z\,(4)$	$v\,(4)$	$\theta\,(4)$	$\varphi\,(4)$
…	…	…	…	…	…	…
N	$v_x\,(N)$	$v_x\,(N)$	$v_z\,(N)$	$v\,(N)$	$\theta\,(N)$	$\varphi\,(N)$

These data have several important features. The scalar velocity, v, ranges from 0 to $+\infty$; v_x, v_y, and v_z range from $-\infty$ to $+\infty$. In §2, we see that θ varies from 0 to π; and φ ranges from 0 to 2π. Each column represents data sampled from the distribution of the corresponding random variable. In Chapter 3, we find that we can use such data to find mathematical models for such distributions. Here, we can find mathematical models for the cumulative distribution functions $f_x\,(v_x)$, $f_y\,(v_y)$, and $f_z\,(v_z)$. We can approximate the graph of $f_x\,(v_x)$ by plotting the rank probability of v_x versus v_x. We expect this plot to be sigmoid; at any v_x, the slope of this plot is the probability-density function, $df_x\,(v_x)/dv_x$. The probability density function for v_x depends only on v_x, because the value measured for v_x is independent of the values measured for v_y and v_z. However, by Maxwell's assumption, the functions describing the distribution of v_y and v_z are the same as those describing the distribution of v_x. While redundant, it is convenient to introduce additional symbols to represent these probability density functions. We define $\rho_x\,(v_x) = df_x\,(v_x)/dv_x$, $\rho_y\,(v_y) = df_y\,(v_y)/dv_y$, and $\rho_z\,(v_z) = df_z\,(v_z)/dv_z$.

When we find these one-dimensional distribution functions by modeling the experimental data in this way, each v_x datum that we use in our analysis comes from an observation on a molecule and is associated with particular v_y and v_z values. These values of v_y and v_z can be anything from $-\infty$ to $+\infty$. This is a significant point. The functions $f_x\,(v_x)$ and $df_x\,(v_x)/dv_x$ are independent of

v_y and v_z. We can also say that $df_x(v_x)/dv_x$ describes the distribution of v_x when v_y and v_z are averaged over all the values it is possible for them to have.

To clarify this, let us consider another cumulative probability distribution function, $f_{xyz}(v_x, v_y, v_z)$, which is just the fraction of all molecules whose respective Cartesian velocity components are less than v_x, v_y, v_z. Since $f_x(v_x)$, $f_y(v_y)$, and $f_z(v_z)$ are the fractions whose components are less than v_x, v_y, and v_z, respectively, their product is equal to $f_{xyz}(v_x, v_y, v_z)$ We have $f_{xyz}(v_x, v_y, v_z) = f_x(v_x) f_y(v_y) f_z(v_z)$. For the velocity of a randomly selected molecule, (v_x^*, v_y^*, v_z^*), to be included in the fraction represented by $f_{xyz}(v_x, v_y, v_z)$, the velocity must be in the particular range $+\infty < v_x^* < v_x$, $+\infty < v_y^* < v_y$, and $+\infty < v_z^* < v_z$.

However, for a velocity v_x^* to be included in $f_x(v_x)$, we must have $v_x^* < v_x$, $v_y^* < \infty$, and $v_z^* < \infty$; that is, the components v_y^* and v_z^* can have any values. Since the probability that v_x, v_y, and v_z satisfy $v_x^* < v_x$, $v_y^* < v_y$, and $v_z^* < v_z$ is

$$\begin{aligned} P(v_x^* < v_x, v_y^* < v_y, v_z^* < v_z) &= f_{xyz}(v_x, v_y, v_z) \\ &= f_x(v_x) f_y(v_y) f_z(v_z) \end{aligned}$$

the probability that v_x^* is included in $f_x(v_x)$ becomes

$$\begin{aligned} P(v_x^* < v_x, v_y^* < \infty, v_z^* < \infty) &= f_{xyz}(v_x, \infty, \infty) \\ &= f_x(v_x) f_y(\infty) f_z(\infty) \\ &= f_x(v_x) \end{aligned}$$

For our purposes, we need to be able to express the probability that the velocity lies within any range of velocities. Let us use υ to designate a particular "volume" region in velocity space and use $P(\upsilon)$ to designate the probability that the velocity of a randomly selected molecule is in this region. When we let υ be the region in velocity space in which x-components lie between v_x

and $v_x + dv_x$, y-components lie between v_y, and $v_y + dv_y$, and z-components lie between v_z and $v_z + dv_z$, $dP(\upsilon)$ denotes the probability that the velocity of a randomly chosen molecule, (v_x^*, v_y^*, v_z^*), satisfies the conditions $v_x < v_x^* < v_x + dv_x$, $v_y < v_y^* < v_y + dv_y$, and $v_z < v_z^* < v_z + dv_z$.

$dP(\upsilon)$ is an increment of probability. The dependence of $dP(\upsilon)$ on v_x, v_y, v_z, dv_x, dv_y, and dv_zcan be made explicit by introducing a new function, $\rho(v_x, v_y, v_z)$, defined by

$$dP(\upsilon) = \rho(v_x, v_y, v_z)\, dv_x dv_y dv_z$$

Since $dv_x dv_y dv_z$ is the volume available in velocity space for velocities whose x-components are between v_x and $v_x + dv_x$, whose y-components are between v_y, and $v_y + dv_y$, and whose z-components are between v_z and $v_z + dz$, we see that $\rho(v_x, v_y, v_z)$ is a probability density function in three dimensions. The value of $\rho(v_x, v_y, v_z)$ is the probability, per unit volume in velocity space, that a molecule has the velocity (v_x, v_y, v_z). For any velocity, (v_x, v_y, v_z), there is a value of $\rho(v_x, v_y, v_z)$; this value is just a number. If we want the probability of finding a velocity within some small volume of velocity space around (v_x, v_y, v_z), we can find it by multiplying $\rho(v_x, v_y, v_z)$ by this volume.

From the one-dimensional probability-density functions, the probability that the x-component of a molecular velocity lies between v_x and $v_x + dv_x$, is just $(df_x(v_x)/dv_x)\, dv_x$, whatever the values of v_y and v_z. The probability that the y-component lies between v_y and $v_y + dv_y$, is just $\left(df_y(v_y)/dv_y\right) dv_y$, whatever the values of v_x and v_z. The probability that the z-component lies between v_z and $v_z + dv_z$, is just $(df_z(v_z)/dv_z)\, dv_z$, whatever the values of v_x and v_y. When we interpret Maxwell's assumption to mean that these are independent probabilities, the probability that all three conditions are realized simultaneously is

$$dP(\upsilon) = \left(\frac{df_x(v_x)}{dv_x}\right)\left(\frac{df_y(v_y)}{dv_y}\right)\left(\frac{df_z(v_z)}{dv_z}\right) dv_x dv_y dv_z = \rho(v_x, v_y, v_z)\, dv_x dv_y dv_z$$

Evidently, the product of these three one-dimensional probability densities is the three-dimensional probability density. We have

$$\rho(v_x, v_y, v_z) = \left(\frac{df_x(v_x)}{dv_x}\right)\left(\frac{df_y(v_y)}{dv_y}\right)\left(\frac{df_z(v_z)}{dv_z}\right) = \rho_x(v_x)\, \rho_y(v_y)\, \rho_z(v_z)$$

From Maxwell's assumption, we have derived the conclusion that $\rho(v_x, v_y, v_z)$ can be expressed as a product of the one-dimensional probability densities $(df(v_x)/dv_x)\, dv_x$, $(df(v_y)/dv_y)\, dv_y$, and $(df(v_z)/dv_z)\, dv_z$. Since these are probability densities, we have

$$\int_{-\infty}^{\infty}\left(\frac{df_x(v_x)}{dv_x}\right)dv_x=\int_{-\infty}^{\infty}\left(\frac{df_y(v_y)}{dv_y}\right)dv_y=\int_{-\infty}^{\infty}\left(\frac{df_z(v_z)}{dv_z}\right)dv_z=1$$

and

$$\iiint_{-\infty}^{\infty}\rho(v_x,v_y,v_z)\,dv_xdv_ydv_z=1$$

Moreover, because the Cartesian coordinates differ from one another only in orientation, $(df(v_x)/dv_x)\,dv_x$, $(df(v_y)/dv_y)\,dv_y$, and $(df(v_z)/dv_z)\,dv_z$ must all be the same function.

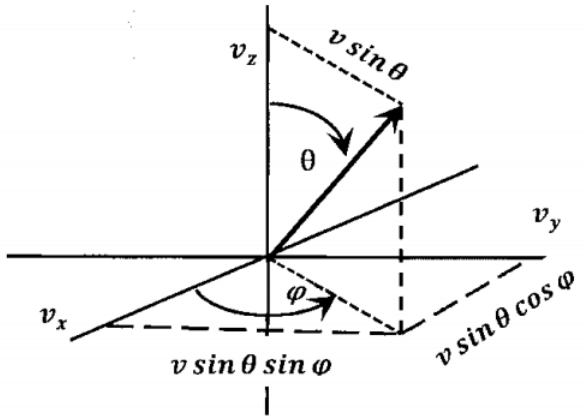


Figure 1. Transformation from Cartesian to spherical coordinates.

To summarize the development above, we define $\rho(v_x,v_y,v_z)$ independently of $df_x(v_x)/dv_x$, $df_y(v_y)/dv_y$, and $df_z(v_z)/dv_z$. Then, from Maxwell's assumption that the three one-dimensional probabilities are independent, we find

$$\begin{aligned}\rho(v_x,v_y,v_z)&=\left(\frac{df_x(v_x)}{dv_x}\right)\left(\frac{df_y(v_y)}{dv_y}\right)\left(\frac{df_z(v_z)}{dv_z}\right)\\&=\rho_x(v_x)\,\rho_y(v_y)\,\rho_z(v_z)\end{aligned}$$

Alternatively, we could take Maxwell's assumption to be that the three-dimensional probability density function is expressible as a product of three one-dimensional probability densities:

$$\rho(v_x,v_y,v_z)=\rho_x(v_x)\,\rho_y(v_y)\,\rho_z(v_z)$$

In this case, the relationships of $\rho_x(v_x)$, $\rho_y(v_y)$, and $\rho_z(v_z)$, to the one-dimensional cumulative probabilities ($f_x(v_x)$, *etc.*) must be deduced from the properties of $\rho(v_x,v_y,v_z)$. As emphasized above, our deduction of $f_x(v_x)$ from experimental data uses v_x values that are associated with all possible values of v_y and v_z. That is, what we determine in our (hypothetical) experiment is

$$\begin{aligned}f_x(v_x)&=\int_{v_x=-\infty}^{v_x}\iint_{v_{y,z}=-\infty}^{\infty}\rho(v_x,v_y,v_z)\,dv_xdv_ydv_z\\&=\int_{-\infty}^{v_x}\rho_x(v_x)\,dv_x\int_{-\infty}^{\infty}\rho_y(v_y)\,dv_y\int_{-\infty}^{\infty}\rho_z(v_z)\,dv_z\\&=\int_{-\infty}^{v_x}\rho_x(v_x)\,dv_x\end{aligned}$$

from which it follows that

$$\frac{df_x(v_x)}{dv_x}=\rho_x(v_x)$$

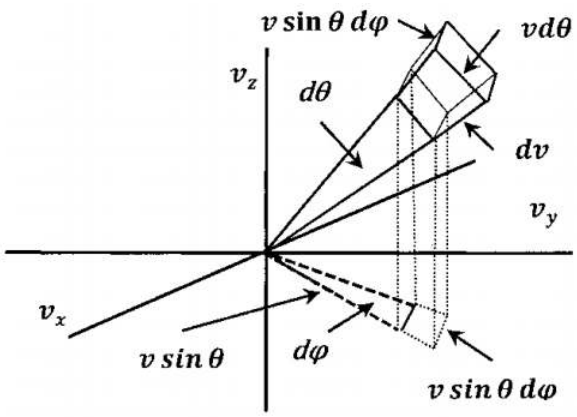


Figure 2. The differential volume element in spherical coordinates.

6.2: Probability Density Functions for Velocity Components in Spherical Coordinates

We introduce the idea of a three-dimensional probability-density function by showing how to find it from data referred to a Cartesian coordinates system. The probability density associated with a particular molecular velocity is just a number—a number that depends only on the velocity. Given a velocity, the probability density associated with that velocity must be independent of our choice of coordinate system. We can express the three-dimensional probability density using any coordinate system. We turn now to expressing velocities and probability density functions using spherical coordinates.

Just as we did for the Cartesian velocity components, we deduce the cumulative probability functions $f_v(v)$, $f_\theta(\theta)$, and $f_\varphi(\varphi)$ for the spherical-coordinate components. Our deduction of $f_v(v)$ from the experimental data uses v-values that are associated with all possible values of θ and φ. Corresponding statements apply to our deductions of $f_\theta(\theta)$, and $f_\varphi(\varphi)$. We also obtain their derivatives, the probability-density functions $df_v(v)/dv$, $df_\theta(\theta)/d\theta$, and $df_\varphi(\varphi)/d\varphi$. From the properties of probability-density functions, we have

$$\int_0^\infty \left(\frac{df_v(v)}{dv}\right)dv = \int_0^\pi \left(\frac{df_\theta(\theta)}{d\theta}\right)d\theta = \int_0^{2\pi} \left(\frac{df_\varphi(\varphi)}{d\varphi}\right)d\varphi = 1$$

Let $\upsilon\prime$ be the arbitrarily small increment of volume in velocity space in which the v-, θ-, and φ-components of velocity lie between v and $v+dv$, θ and $\theta+d\theta$, and φ and $\varphi+d\varphi$. Then the probability that the velocity of a randomly selected molecule lies within $\upsilon\prime$ is

$$dP(\upsilon\prime) = \left(\frac{df_v(v)}{dv}\right)\left(\frac{df_\theta(\theta)}{d\theta}\right)\left(\frac{df_\varphi(\varphi)}{d\varphi}\right)dvd\theta d\varphi$$

Note that the product

$$\left(\frac{df_v(v)}{dv}\right)\left(\frac{df_\theta(\theta)}{d\theta}\right)\left(\frac{df_\varphi(\varphi)}{d\varphi}\right)$$

is not a three-dimensional probability density function. This is most immediately appreciated by recognizing that $dvd\theta d\varphi$ is not an incremental "volume" in velocity space. That is, $\upsilon\prime \neq dvd\theta d\varphi$

We let $\rho(v,\ \theta,\varphi)$ be the probability-density function for the velocity vector in spherical coordinates. When v, θ, and φ specify the velocity, $\rho(v,\ \theta,\varphi)$ is the probability per unit volume at that velocity. We want to use $\rho(v,\ \theta,\varphi)$ to express the probability that an arbitrarily selected molecule has a velocity vector whose magnitude lies between v and $v+dv$, while its θ-component lies between θ and $\theta+d\theta$, and its φ-component lies between φ and $\varphi+d\varphi$. This is just $\rho(v,\ \theta,\varphi)$ times the velocity-space "volume" included by these ranges of v, θ, and φ.

When we change from Cartesian coordinates, $\vec{v} = (v_x, v_y, v_z)$, to spherical coordinates, $\vec{v} = (v,\theta,\varphi)$, the transformation is $v_x = v\sin\theta\cos\varphi$, $v_y = v\sin\theta\sin\varphi$, $v_z = v\cos\theta$. (See Figure 1.) As sketched in Figure 2, an incremental increase in each of the coordinates of the point specified by the vector $(v,\ \theta,\varphi)$ advances the vector to the point $(v+dv, \theta+d\theta, \varphi+d\varphi)$. When dv, $d\theta$, and $d\varphi$ are arbitrarily small, these two points specify the diagonally opposite corners of a rectangular parallelepiped, whose edges have the lengths dv, $vd\theta$, and $v\sin\theta\, d\varphi$. The volume of this parallelepiped is $v^2\sin\theta\, dvd\theta d\varphi$. Hence, the differential volume elementdifferential volume element in Cartesian coordinates, $dv_x dv_y dv_z$, becomes $v^2\sin\theta\, dvd\theta d\varphi$ in spherical coordinates.

Mathematically, this conversion is obtained using the absolute value of the ***Jacobian***, $J\left(\frac{v_x,v_y,v_z}{v,\theta,\varphi}\right)$, of the transformation. That is,

$$dv_x dv_y dv_z = \left|J\left(\frac{v_x, v_y, v_z}{v, \theta, \varphi}\right)\right| dvd\theta d\varphi$$

where the Jacobian is a determinate of partial derivatives

$$J\left(\frac{v_x, v_y, v_z}{v, \theta, \varphi}\right) = \begin{vmatrix} \partial v_x/\partial v & \partial v_x/\partial\theta & \partial v_x/\partial\varphi \\ \partial v_y/\partial v & \partial v_y/\partial\theta & \partial v_y/\partial\varphi \\ \partial v_z/\partial v & \partial v_z/\partial\theta & \partial v_z/\partial\varphi \end{vmatrix}$$

$$= v^2\sin\theta$$

Since the differential unit of volume in spherical coordinates is $v^2 \sin\theta \; dv d\theta d\varphi$, ***the probability that the velocity components lie within the indicated ranges*** is

$$dP\left(\upsilon\prime\right) = \rho\left(v,\theta,\varphi\right) v^2 \sin\theta \, dv d\theta d\varphi$$

We can develop the next step in Maxwell's argument by taking his assumption to mean that the three-dimensional probability density function is expressible as a product of three one-dimensional functions. That is, we take Maxwell's assumption to assert the existence of independent functions $\rho_v\left(v\right)$, $\rho_\theta\left(\theta\right)$, and $\rho_\varphi\left(\varphi\right)$ such that $\rho\left(v,\ \theta,\varphi\right) = \rho_v\left(v\right)\rho_\theta\left(\theta\right)\rho_\varphi\left(\varphi\right)$. The probability that the v-, θ-, and φ-components of velocity lie between v and $v+dv$, θ and $\theta+d\theta$, and φ and $\varphi+d\varphi$ becomes

$$\begin{aligned} dP\left(\upsilon\prime\right) &= \left(\frac{df_v\left(v\right)}{dv}\right)\left(\frac{df_\theta\left(\theta\right)}{d\theta}\right)\left(\frac{df_\varphi\left(\varphi\right)}{d\varphi}\right) dv d\theta d\varphi \\ &= \rho\left(v,\theta,\varphi\right) v^2 \sin\theta dv d\theta d\varphi \\ &= \rho_v\left(v\right)\rho_\theta\left(\theta\right)\rho_\varphi\left(\varphi\right) v^2 \sin\theta dv d\theta d\varphi \end{aligned}$$

Since v, θ, and φ are independent, it follows that

$$\frac{df_v\left(v\right)}{dv} = v^2 \rho_v\left(v\right)$$

$$\frac{df_\theta\left(\theta\right)}{d\theta} = \rho_\theta\left(\theta\right)\sin\theta$$

$$\frac{df_\varphi\left(\varphi\right)}{d\varphi} = \rho_\varphi\left(\varphi\right)$$

Moreover, the assumption that velocity is independent of direction means that $\rho_\theta\left(\theta\right)$ must actually be independent of θ; that is, $\rho_\theta\left(\theta\right)$ must be a constant. We let this constant be α_θ; so $\rho_\theta\left(\theta\right) = \alpha_\theta$. By the same argument, we set $\rho_\varphi\left(\varphi\right) = \alpha_\varphi$. Each of these probability-density functions must be normalized. This means that

$$1 = \int_0^\infty v^2 \rho_v\left(v\right) dv$$

$$1 = \int_0^\pi \alpha_\theta \sin\theta \, d\theta = 2\alpha_\theta$$

$$1 = \int_0^{2\pi} \alpha_\varphi d\varphi = 2\pi\alpha_\varphi$$

from which we see that $\rho_\theta\left(\theta\right) = \alpha_\theta = 1/2$ and $\rho_\varphi\left(\varphi\right) = \alpha_\varphi = 1/2\pi$. It is important to recognize that, while $\rho_x\left(v_x\right)$, $\rho_y\left(v_y\right)$, and $\rho_z\left(v_z\right)$ are probability density functions, $\rho_\theta\left(\theta\right)$ and $\rho_v\left(v\right)$ are not. (However, $\rho_\varphi\left(\varphi\right)$ is a probability density function.) We can see this by noting that, if $\rho_\theta\left(\theta\right)$ were a probability density, its integral over all possible values of θ $\left(0 < \theta < \pi\right)$ would be one. Instead, we find

$$\int_0^\pi \rho_\theta\left(\theta\right) d\theta = \int_0^\pi d\theta/2 = \pi/2$$

Similarly, when we find $\rho_v\left(v\right)$, we can show explicitly that

$$\int_0^\infty \rho_v\left(v\right) dv \neq 1$$

Our notation now allows us to express the probability that an arbitrarily selected molecule has a velocity vector whose magnitude lies between v and $v+dv$, while its θ-component lies between θ and $\theta+d\theta$, and its φ-component lies between φ and $\varphi+d\varphi$ using three equivalent representations of the probability density function:

$$dP\left(\upsilon\prime\right) = \rho\left(v,\theta,\varphi\right) v^2 \sin\theta dv d\theta d\varphi \; - \rho_v\left(v\right)\rho_\theta\left(\theta\right)\rho_\varphi\left(\varphi\right) v^2 \sin\theta dv d\theta d\varphi \; = \left(\frac{1}{4\pi}\right)\rho_v\left(v\right) v^2 \sin\theta \, dv d\theta d\varphi$$

The three-dimensional probability-density function in spherical coordinates is

$$\rho\left(v,\ \theta,\varphi\right)=\rho_{v}\left(v\right)\rho_{\theta}\left(\theta\right)\rho_{\varphi}\left(\varphi\right)=\frac{\rho_{v}\left(v\right)}{4\pi}$$

This shows explicitly that $\rho\left(v,\ \theta,\varphi\right)$ is independent of θ and φ; if the speed is independent of direction, the probability density function that describes velocity must be independent of the coordinates, θ and φ, that specify its direction.

6.3: Maxwell's Derivation of the Gas-velocity Probability-density Function

To this point, we have been developing our ability to characterize the gas-velocity distribution functions. We now want to use Maxwell's argument to find them. We have already introduced the first step, which is the recognition that three-dimensional probability-density functions can be expressed as products of independent one-dimensional functions, and that $\rho_\theta(\theta)$, and $\rho_\varphi(\varphi)$ are the constants $1/2$ and $1/2\pi$. Now, because the probability density associated with any given velocity is just a number that is independent of the coordinate system, we can equate the three-dimensional probability-density functions for Cartesian and spherical coordinates: $\rho(v_x, v_y, v_z) = \rho(v,\ \theta, \varphi)$ so that

$$\rho_x(v_x)\,\rho_y(v_y)\,\rho_z(v_z) = \frac{\rho_v(v)}{4\pi}$$

We take the partial derivative of this last equation with respect to v_x. The probability densities $\rho_y(v_y)$ and $\rho_z(v_z)$ are independent of v_x. However, v is a function of v_x, because $v^2 = v_x^2 + v_y^2 + v_z^2$. We find

$$\frac{d\rho_x(v_x)}{dv_x}\rho_y(v_y)\,\rho_z(v_z) = \frac{1}{4\pi}\left(\frac{\partial\rho_v(v)}{\partial v_x}\right)_{v_y v_v} = \frac{1}{4\pi}\left(\frac{d\rho_v(v)}{dv}\right)\left(\frac{\partial v}{\partial v_x}\right)_{v_y v_z}$$

Since $v^2 = v_x^2 + v_y^2 + v_z^2$, $2v(\partial v/\partial v_x)_{v_y v_z} = 2v_x$ and

$$\left(\frac{\partial v}{\partial v_x}\right)_{v_y v_z} = \frac{v_x}{v}$$

Making this substitution and dividing by the original equation gives

$$\frac{d\rho_x(v_x)}{dv_x}\frac{\rho_y(v_y)\,\rho_z(v_z)}{\rho_x(v_x)\,\rho_y(v_y)\,\rho_z(v_z)} = \frac{v_x}{v}\frac{1}{\rho_v(v)}\frac{d\rho_v(v)}{dv}$$

Cancellation and rearrangement of the result leads to an equation in which the independent variables v_x and v are separated. This means that each term must be equal to a constant, which we take to be $-\lambda$. We find

$$\left(\frac{1}{v_x\rho_x(v_x)}\right)\frac{d\rho_x(v_x)}{dv_x} = \left(\frac{1}{v\rho_v(v)}\right)\frac{d\rho_v(v)}{dv} = -\lambda$$

so that

$$\frac{d\rho_x(v_x)}{\rho_x(v_x)} = -\lambda v_x dv_x$$

and

$$\frac{d\rho_v(v)}{\rho_v(v)} = -\lambda v dv$$

From the first of these equations, we obtain the probability density function for the distributions of one-dimensional velocities. (See Section 4.4.) The three-dimensional probability density function can be deduced from the one-dimensional function. (See Section 4.5.)

From the second equation, we obtain the three-dimensional probability-density function directly. Integrating from $v = 0$, where $\rho_v(0)$ has a fixed value, to an arbitrary scalar velocity, v, where the scalar-velocity function is $\rho_v(v)$, we have

$$\int_{\rho_v(0)}^{\rho_v(v)}\frac{d\rho_v(v)}{\rho_v(v)} = -\lambda\int_0^v v dv$$

or

$$\rho_v(v) = \rho_v(0)\,exp\left(\frac{-\lambda v^2}{2}\right)$$

The probability-density function for the scalar velocity becomes

$$\frac{df_v(v)}{dv} = v^2\rho_v(v) = \rho_v(0)\,v^2 exp\left(\frac{-\lambda v^2}{2}\right)$$

This is the result we want, except that it contains the unknown parameters $\rho_v(0)$ and λ. The value of $\rho_v(0)$ must be such as to make the integral over all velocities equal to unity. We require

$$\begin{aligned} 1 &= \int_0^\infty \left(\frac{df_v(v)}{dv}\right) dv \\ &= \rho_v(0)\int_0^\infty v^2 \exp\left(\frac{-\lambda v^2}{2}\right) dv \\ &= \frac{\rho_v(0)}{4\pi}\left(\frac{2\pi}{\lambda}\right)^{3/2} \end{aligned}$$

so that

$$\rho_v(0) = 4\pi\left(\frac{\lambda}{2\pi}\right)^{3/2}$$

where we use the definite integral $\int_0^\infty x^2 \exp\left(-ax^2\right) dx = (1/4)\sqrt{\pi/a^3}$. (See Appendix D.) The scalar-velocity function in the three-dimensional probability-density function becomes

$$\rho_v(v) = 4\pi\left(\frac{\lambda}{2\pi}\right)^{3/2} \exp\left(\frac{-\lambda v^2}{2}\right)$$

The probability-density function for the scalar velocity becomes

$$\begin{aligned} \frac{df_v(v)}{dv} &= v^2\rho_v(v) \\ &= 4\pi\left(\frac{\lambda}{2\pi}\right)^{3/2} v^2 \exp\left(\frac{-\lambda v^2}{2}\right) \end{aligned}$$

The three-dimensional probability density in spherical coordinates becomes

$$\begin{aligned} \rho(v,\ \theta, \varphi) &= \rho_v(v)\,\rho_\theta(\theta)\,\rho_\varphi(\varphi) \\ &= \left(\frac{\lambda}{2\pi}\right)^{3/2} \exp\left(\frac{-\lambda v^2}{2}\right) \end{aligned}$$

The probability that an arbitrarily selected molecule has a velocity vector whose magnitude lies between v and $v + dv$, while its θ-component lies between θ and $\theta + d\theta$, and its φ-component lies between φ and $\varphi + d\varphi$ becomes

$$\begin{aligned} dP(\upsilon\prime) &= \left(\frac{df_v(v)}{dv}\right)\left(\frac{df_\theta(\theta)}{d\theta}\right)\left(\frac{df_\varphi(\varphi)}{d\varphi}\right) dv d\theta d\varphi \\ &= \rho(v,\theta,\varphi)\,v^2 \sin\theta dv d\theta d\varphi \\ &= \left(\frac{1}{4\pi}\right)\rho_v(v)\,v^2 \sin\theta dv d\theta d\varphi \\ &= \left(\frac{\lambda}{2\pi}\right)^{3/2} v^2 exp\left(\frac{-\lambda v^2}{2}\right) \sin\theta dv d\theta d\varphi \end{aligned}$$

In Section 4.6, we again derive Boyle's law and use the ideal gas equation to show that $\lambda = m/kT$.

6.4: The Probability-density Function for Gas Velocities in One Dimension

In Section 4.3, we find a differential equation in the function $\rho_x(v_x)$. Unlike the velocity, which takes values from zero to infinity, the x-component, v_x, takes values from minus infinity to plus infinity. The probability density at an infinite velocity, in either direction, is necessarily zero. Therefore, we cannot evaluate the integral of $d\rho_x(v_x)/\rho_x(v_x)$ from $v_x = -\infty$ to an arbitrary velocity, v_x. However, we know from Maxwell's assumption that the probability density for v_x must be independent of whether the molecule is traveling in the direction of the positive x-axis or the negative x-axis. That is, $\rho_x(v_x)$ must be an even function; the probability density function must be symmetric around $v_x = 0$; $\rho_x(v_x) = \rho_x(-v_x)$. Hence, we can express $\rho_x(v_x)$ relative to its fixed value,$\rho_x(0)$, at $v_x = 0$. We integrate $d\rho_x(v_x)/\rho_x(v_x)$ from $\rho_x(0)$ to $\rho_x(v_x)$ as v_x goes from zero to an arbitrary velocity, v_x, to find

$$\int_{\rho_x(0)}^{\rho_x(v_x)} \frac{d\rho_x(v_x)}{\rho_x(v_x)} = -\lambda \int_0^{v_x} v_x dv_x$$

or

$$\rho_x(v_x) = \frac{df_x(v_x)}{dv_x} = \rho_x(0)\exp\left(\frac{-\lambda v_x^2}{2}\right)$$

The value of $\rho_x(0)$ must be such as to make the integral of $\rho_x(v_x)$ over all possible values of v_x, \(-\infty < v_x <\infty \), equal to unity. That is, we must have

$$\begin{aligned} 1 &= \int_{-\infty}^{\infty} \rho_x(v_x)\, dv_x \\ &= \int_{-\infty}^{\infty} \frac{df_x(v_x)}{dv_x} dv_x \\ &= \rho_x(0) \int_{-\infty}^{\infty} \exp\left(\frac{-\lambda v_x^2}{2}\right) dv_x \\ &= \rho_x(0)\sqrt{\frac{2\pi}{\lambda}} \end{aligned}$$

where we use the definite integral $\int_{-\infty}^{\infty} \exp\left(-ax^2\right) dx = \sqrt{\pi/a}$. (See Appendix D.) It follows that $\rho_x(0) = (\lambda/2\pi)^{1/2}$. The one-dimensional probability-density function becomes

$$\begin{aligned} \rho_x(v_x) &= \frac{df_x(v_x)}{dv_x} \\ &= \left(\frac{\lambda}{2\pi}\right)^{1/2} \exp\left(\frac{-\lambda v_x^2}{2}\right) \end{aligned}$$

Note that this is the normal distribution with $\mu = 0$ and $\sigma^2 = \lambda^{-1}$. So λ^{-1} is the variance of the normal one-dimensional probability-density function. As noted above, in Section 4.6 we find that $\lambda = m/kT$.

6.5: Combining the One-dimensional Probability Density Functions

In Section 4.4, we derive the probability density function for one Cartesian component of the velocity of a gas molecule. The probability density functions for the other two Cartesian components are the same function. For $\vec{v} = (v_x, v_y, v_z)$, we have $v^2 = v_x^2 + v_y^2 + v_z^2$, and

$$\frac{df_x(v_x)}{dv_x} = \left(\frac{\lambda}{2\pi}\right)^{1/2} \exp\left(\frac{-\lambda v_x^2}{2}\right)$$
$$\frac{df_y(v_y)}{dv_y} = \left(\frac{\lambda}{2\pi}\right)^{1/2} \exp\left(\frac{-\lambda v_y^2}{2}\right)$$
$$\frac{df_z(v_z)}{dv_z} = \left(\frac{\lambda}{2\pi}\right)^{1/2} \exp\left(\frac{-\lambda v_z^2}{2}\right)$$

We now want to derive the three-dimensional probability density function from these relationships. Given these probability density functions for the Cartesian components of $\vec{v}$, we can find the probability density function in spherical coordinates

$$\left(\tfrac{df_x(v_x)}{dv_x}\right)\left(\tfrac{df_y(v_y)}{dv_y}\right)\left(\tfrac{df_z(v_z)}{dv_z}\right)$$
$$= \left(\tfrac{\lambda}{2\pi}\right)^{3/2} \exp\left(\tfrac{-\lambda v_x^2}{2}\right) exp\left(\tfrac{-\lambda v_y^2}{2}\right) exp\left(\tfrac{-\lambda v_z^2}{2}\right)$$
$$= \left(\tfrac{\lambda}{2\pi}\right)^{3/2} \exp\left(\tfrac{-\lambda v^2}{2}\right)$$
$$= \rho(v, \theta, \varphi)$$

Since the differential volume element in spherical coordinates is $v^2 \sin\theta \, dv d\theta d\varphi$, the probability that a molecule has a a velocity vector whose magnitude lies between v and $v + dv$, while its θ-component lies between θ and $\theta + d\theta$, and its φ-component lies between φ and $\varphi + d\varphi$ becomes

$$\left(\tfrac{df_v(v)}{dv}\right)\left(\tfrac{df_\theta(\theta)}{d\theta}\right)\left(\tfrac{df_\varphi(\varphi)}{d\varphi}\right) dv d\theta d\varphi$$
$$= \rho(v, \theta, \varphi) \, v^2 \sin\theta dv d\theta d\varphi$$
$$= \left(\tfrac{\lambda}{2\pi}\right)^{3/2} v^2 \exp\left(\tfrac{-\lambda v^2}{2}\right) \sin\theta dv d\theta d\varphi$$

(We found the same result in Section 4.3, of course.) We can find the probability-density function for the scalar velocity by eliminating the dependence on the angular components. To do this, we need only sum up, at a given value of v, the contributions from all possible values of θ and φ, recalling that $0 \le \theta < \pi$ and $0 \le \varphi < 2\pi$. This sum is just

$$\frac{df_v(v)}{dv} \int_{\theta=0}^{\pi} \left(\frac{df_\theta(\theta)}{d\theta}\right) d\theta \int_{\varphi=0}^{2\pi} \left(\frac{df_\varphi(\varphi)}{d\varphi}\right) d\varphi =$$
$$= \left(\frac{\lambda}{2\pi}\right)^{3/2} v^2 exp\left(\frac{-\lambda v^2}{2}\right) \int_{\theta=0}^{\pi} \sin\theta d\theta \int_{\varphi=0}^{2\pi} d\varphi$$

Since $\int_{\theta=0}^{\pi} \left(\tfrac{df_\theta(\theta)}{d\theta}\right) d\theta = \int_{\varphi=0}^{2\pi} \left(\tfrac{df_\varphi(\varphi)}{d\varphi}\right) d\varphi = 1$, $\int_0^{\pi} \sin\theta d\theta = 2$, and $\int_0^{2\pi} d\varphi = 2\pi$, we again obtain the Maxwell-Boltzmann probability-density function for the scalar velocity:

$$\frac{df_v(v)}{dv} = 4\pi \left(\frac{\lambda}{2\pi}\right)^{3/2} v^2 exp\left(\frac{-\lambda v^2}{2}\right)$$

Unlike the distribution function for the Cartesian components of velocity, the Maxwell-Boltzmann distribution for scalar velocities is not a normal distribution. Possible speeds lie in the interval $0 \le v < \infty$. Because of the v^2 term, the Maxwell-Boltzmann equation is asymmetric; it has a pronounced tail at high velocities.

6.6: Boyle's Law from the Maxwell-Boltzmann Probability Density

In Chapter 2, we derive Boyle's lawBoyle's law using simplifying assumptions. We are now able to do this derivation much more rigorously. We consider the collisions of gas molecules with a small portion of the wall of their container. We suppose that the wall is smooth, so that we can select a small and compact segment of it that is arbitrarily close to being planar. We denote both the segment of the wall and its area as A. A can have any shape so long as it is a smooth, flat surface enclosed by a smooth curve.

Let the volume of the container be V and the number of gas molecules in the container be N. We imagine that we follow the trajectory of one particular molecule as it moves to hit the wall somewhere within A. We begin our observations at time $t = 0$ and suppose that the collision occurs at time t.

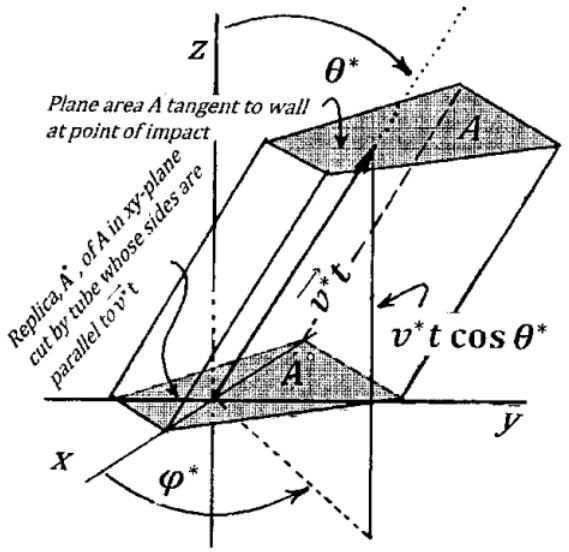


Figure 3. Trajectory of a molecule colliding with a wall of its container.

As sketched in Figure 3, we erect a Cartesian coordinate system with its origin at the location in space of the molecule at time $t = 0$. We orient the axes of this coordinate system so that the xy-plane is parallel to the plane of A, and the z-axis is pointed toward the wall. Then the unit vector along the z-axis and a vector perpendicular to A are parallel to one another. It is convenient to express the velocity of the selected molecule in spherical coordinates. We suppose that, referred to the Cartesian coordinate system we have erected, the velocity vector of the selected molecule is $(v^*, \theta^*, \varphi^*)$. The vector $\overrightarrow{v^*}\, t$, drawn from the origin of our Cartesian system to the point of impact on the wall, follows the trajectory of the molecule from time zero to time t. The z-component of the molecular velocity vector is normal to the plane of A at the point of impact; the magnitude of the z-component $v^*\cos\theta^*$. The perpendicular distance from the plane of A to the xy-plane of the Cartesian system is $v^*t\cos\theta^*$.

We assume that the collision is perfectly elastic. Before collision, the velocity component perpendicular to the wall is $v_z = v^*\cos\theta^*$. Afterward, it is $v_z = -v^*\cos\theta^*$. Only this change in the v_z component contributes to the force on the wall within A. (The v_x and v_y components are not changed by the collision.) During the collision, the molecule's momentum change is $-2mv^*\cos\theta^*$. During our period of observation, the average force on the molecule is thus $(-2mv^*\cos\theta^*)/t$. The force that the molecule exerts on the wall is $(2mv^*\cos\theta^*)/t$, and hence the contribution that this particular collision—by one molecule traveling at velocity v^*— makes to the pressure on the wall is

$$P_1(v^*) = \frac{2mv^*\cos\theta^*}{At}$$

We want to find the pressure on segment A of the wall that results from all possible impacts. To do so, we recognize that any other molecule whose velocity components are v^*, θ^*, and φ^*, and whose location at time $t = 0$ enables it to reach A within time t, makes the same contribution to the pressure as the selected molecule does. Let us begin by assuming that the velocities of all N of the molecules in the volume, V, are the same as that of the selected molecule. In this case, we can find the number of the molecules in the container that can reach A within time t by considering a tubular segment of the interior of the container. The long axis of this tube is parallel to the velocity vector of the selected molecule. The sides of this tube cut the container wall along the perimeter

of A. This tube also cuts the xy-plane (the $z = 0$ plane) of our coordinate system in such a way as to make an exact replica of A in this plane. Call this replica A^o.

The area of A^o is A; the plane of A^o is parallel to the plane of A; and the perpendicular distance between the plane of A and the plane of A^o is $v^* t \cos\theta^*$. The volume of this tube is therefore $Av^* t \cos\theta^*$. Since there are N/V molecules per unit volume, the total number of molecules in the tube is $(ANv^* t \cos\theta^*)/V$. When we assume that every molecule has velocity components v^*, θ^*, and φ^*, all of the molecules in the tube reach A within time t, because each of them travels parallel to the selected molecule, and each of them is initially at least as close to A as is the selected molecule. Therefore, each molecule in the tube contributes $P_1(v^*) = 2mv^* \cos\theta^* / At$ to the pressure at A. The total pressure is the pressure per molecule multiplied by the number of molecules:

$$\left(\frac{2mv^* \cos\theta^*}{At}\right)\left(\frac{ANv^* t \cos\theta^*}{V}\right) = \frac{2mN(v^* \cos\theta^*)^2}{V}$$

However, the molecular velocities are not all the same, and the pressure contribution $2mN(v^* \cos\theta^*)^2/V$ is made only by that fraction of the molecules whose velocity components lie in the intervals $\theta^* < \theta < \theta^* + d\theta$ and $\varphi^* < \varphi < \varphi^* + d\varphi$. This fraction is

$$\rho(v^*, \theta^*, \varphi^*)(v^*)^2 \sin\theta^* dv d\theta d\varphi = \left(\frac{\lambda}{2\pi}\right)^{3/2} (v^*)^2 \exp\left(\frac{-\lambda(v^*)^2}{2}\right) \sin\theta^* dv d\theta d\varphi$$

so that the pressure contribution from molecules whose velocity components lie in these ranges is

$$dP = \frac{2mN(v^* \cos\theta^*)^2}{V} \times \left(\frac{\lambda}{2\pi}\right)^{3/2} (v^*)^2 \exp\left(\frac{-\lambda(v^*)^2}{2}\right) \sin\theta^* dv d\theta d\varphi$$

The total pressure at A is just the sum of the contributions from molecules with all possible combinations of velocities v^*, θ^*, and φ^*. To find this sum, we integrate over all possible velocity vectors. The allowed values of v are $0 \le v < \infty$. There are no constraints on the values of φ; we have $0 \le \varphi < 2\pi$. However, since all of the impacting molecules must have a velocity component in the positive z-direction, the possible values of θ lie in the interval $0 \le \theta < \pi/2$. We designate the velocity of the original molecule as $(v^*, \theta^*, \varphi^*)$ and retain this notation to be as specific as possible in describing the tube bounded by A and A^o. However, the velocity components of an arbitrary molecule can have any of the allowed values. To integrate (See Appendix D) over the allowed values, we drop the superscripts. The pressurepressure:on wall at A becomes

$$P = \frac{2mN}{V}\left(\frac{\lambda}{2\pi}\right)^{3/2} \times$$

$$\int_0^\infty v^4 exp\left(\frac{-\lambda v^2}{2}\right) dv \int_0^{\pi/2} \cos^2\theta \sin\theta \, d\theta \int_0^{2\pi} d\varphi$$

$$= \frac{2mN}{V}\left(\frac{\lambda}{2\pi}\right)^{3/2}\left[\frac{3}{8}\left(\frac{2}{\lambda}\right)^2\left(\frac{2\pi}{\lambda}\right)^{1/2}\right]\left[\frac{1}{3}\right][2\pi] = mN/V\lambda$$

and the pressure–volume product becomes

$$PV = \frac{mN}{\lambda}$$

Since m, N, and λ are constants, this is Boyle's law. Equating this pressure–volume product to that given by the ideal gas equation, we have $mN/\lambda = NkT$ so that

$$\lambda = \frac{m}{kT}$$

Finally, the ***Maxwell-Boltzmann equation*** becomes

$$\frac{df_v(v)}{dv} = 4\pi\left(\frac{m}{2\pi kT}\right)^{3/2} v^2 \exp\left(\frac{-mv^2}{2kT}\right)$$

and the probability density becomes

$$\rho\left(v,\theta,\varphi\right)=\left(\frac{m}{2\pi kT}\right)^{3/2}v^2\exp\left(\frac{-mv^2}{2kT}\right)$$

This derivation can be recast as a computation of the expected value of the pressurepressure:expected value. To do so, we rephrase our description of the system: A molecule whose velocity components are (v^*,θ^*,φ^*) creates a pressure $2mv^*\cos\theta^*\ /At$ on the area A with a probability of $Av^*t\cos\theta^*/V$. (The latter term is the probability that a molecule, whose velocity is (v^*,θ^*,φ^*), is, at time $t=0$, in a location from which it can reach A within time t. If the molecule is to hit the wall within time t, at time $t=0$ the molecule must be within the tubular segment of volume is $Av^*t\cos\theta^*$. The probability that the molecule is within this tubular segment is equal to the fraction of the total volume that this segment occupies.) Therefore, the product

$$\left(\frac{2mv^*\cos\theta^*}{At}\right)\left(\frac{Av^*t\cos\theta^*}{V}\right)=\frac{2m}{V}\left(v^*\cos\theta^*\right)^2$$

is the pressure contribution of a molecule with velocity (v^*,θ^*,φ^*), when θ^* is in the interval $0\le\theta^*<\pi/2$. The total pressure per molecule is the expected value of this pressure contribution; the expected value is the integral, over the entire volume of velocity space, of the pressure contribution times the probability density function for velocities.

It is useful to view the Maxwell-Boltzmann equation as the product of a term

$$\exp\left(-mv^2/2kT\right)$$

—called the ***Boltzmann factor***Boltzmann factor—and a pre-exponential term that is proportional to the ***number of ways*** that a molecule can have a given velocity, v. If there were no constraints on a molecule's speed, we would expect that the number of molecules with speeds between v and $v+dv$ would increase as v increases, because the probability that a molecule has a speed between v and $v+dv$ is proportional to the volume in velocity space of a spherical shell of thickness dv. The volume of a spherical shell of thickness dv is $4\pi v^2dv$, which increases as the square of v. However, the number of molecules with large values of v is constrained by the conservation of energy. Since the total energy of a collection of molecules is limited, only a small proportion of the molecules can have very large velocities. The Boltzmann factor introduces this constraint. A molecule whose mass is m and whose scalar velocity is v has kinetic energy $\epsilon=mv^2/2$. The Boltzmann factor is often written as $\exp\left(-\epsilon/kT\right)$.

6.7: Experimental Test of the Maxwell-Boltzmann Probability Density

There are numerous applications of the Maxwell-Boltzmann equation. These include predictions of collision frequencies, mean-free paths, effusion and diffusion rates, the thermal conductivity of gases, and gas viscosities. These applications are important, but none of them is a direct test of the validity of the Maxwell-Boltzmann equation.

The validity of the equation has been demonstrated directly in experiments in which a gas of metal atoms is produced in an oven at a very high temperature. As sketched in Figure 4, the gas is allowed to escape into a vacuum chamber through a very small hole in the side of the oven. The escaping atoms impinge on one or more metal plates. Narrow slits cut in these plates stop any metal atoms whose flight paths do not pass though the slits. This produces a beambeam:of metal atoms of metal atoms whose velocity distribution is the same as that of the metal-atom gas inside the oven. The rate at which metal atoms arrive at a detector is measured. Various methods are used to translate the atom-arrival rate into a measurement of their speed.

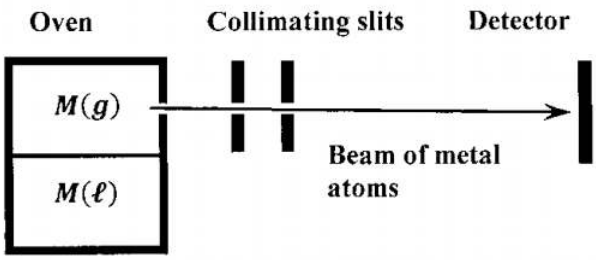


Figure 4. Producing a beam of metal atoms.

One device uses a solid cylindrical drum, which rotates on its cylindrical axis. As sketched in Figure 5, a spiral groove is cut into the cylindrical face of this drum. This groove is cut with a constant pitch. When the drum rotates at a constant rate, an atom traveling at a constant

velocity parallel to the cylindrical axis can traverse the length of the drum while remaining within the groove. That is, for a given rotation rate, there is one critical velocity at which an atom can travel in a straight line while remaining in the middle of the groove all the way from one end of the drum to the other. If the atom moves significantly faster or slower than this critical velocity, it collides with—and sticks to—one side or the other of the groove.

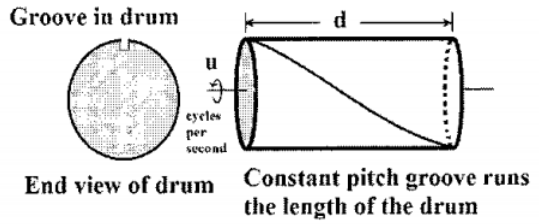


Figure 5. Device to select metal atoms having a specified velocity.

Since the groove has a finite width, atoms whose velocities lie in a narrow range about the critical velocity can traverse the groove without hitting one of the sides.

Let us assume that the groove is cut so that the spiral travels half way around the cylinder. That is, if we project the spiral onto one of the circular faces of the drum, the projection traverses an angle of 180^o on the face. In order to remain in the middle of this groove all the way from one end of the drum to the other, the atom must travel the length of the cylindrical drum in exactly the same time that it takes the drum to make a half-rotation. Let the critical velocity be $v_{critical}$. Then the time required for the atom to traverse the length, d, of the drum is $d/v_{critial}$. If the drum rotates at u cycles/sec, the time required for the drum to make one-half rotation is $1/2u$. Thus, the atom will remain in the middle of the groove all the way through the drum if

$$v_{critial} = 2ud$$

By varying the rotation rate, we can vary the critical velocity.

Because the groove has a finite width, atoms whose velocities are in a range \(v_{min}<v_{max}\) can successfully traverse the groove. Whether or not a particular atom can do so depends on its velocity, where it enters the groove, and the width of the groove. Let the width of the groove be w and the radius of the drum be r, where the drum is constructed with $r \gg w$. A slower atom that

enters the groove at the earliest possible time—when the leading edge of the groove first encounters the beam of atoms—can traverse the length of the groove in a longer time, t_{max}. A point on the circumference of the drum travels with speed $2\pi ru$. The slowest atom traverses the length of the drum while a point on the circumference of the drum travels a distance $\pi r + w$. (To intercept the slowest atom, the trailing edge of the groove must travel a distance equal to half the circumference of the drum, πr, plus the width of the groove, w.) The time required for this rotation is the maximum time a particle can take to traverse the length, so

$$t_{max} = (\pi r + w) / (2\pi ru)$$

and

$$v_{min} = d/t_{max} = 2\pi rud/ (\pi r + w)$$

A fast atom that enters the groove at the last possible moment—when the trailing edge of the grove just leaves the beam of atoms—can still traverse the groove if it does so in the time, t_{min} that it takes the trailing edge of the groove to travel a distance $\pi r - w$. So,

$$t_{min} = (\pi r - w) / (2\pi ru)$$

and

$$v_{max} = d/t_{min} = 2\pi rud/ (\pi r - w)$$

At a given rotation rate, the drum will pass atoms whose speeds are in the range

$$\Delta v = v_{max} - v_{min} = 2ud \left(\frac{\pi r}{\pi r - w} - \frac{\pi r}{\pi r + w} \right) = 2ud \left(\frac{2\pi rw}{(\pi r)^2 - w^2} \right) \approx v_{critical} \left(\frac{2w}{\pi r} \right)$$

So that

$$\frac{\Delta v}{v_{critial}} \approx \frac{2w}{\pi r}$$

The fraction of the incident atoms that successfully traverse the groove is equal to the fraction that have velocities in the interval Δv centered on the critical velocity, $v_{critical} = 2ud$.

6.8: Statistics for Molecular Speeds

Expected values for several quantities can be calculated from the Maxwell-Boltzmann probability density function. The required definite integrals are tabulated in Appendix D.

The ***most probable speed***, v_{mp}, is the speed at which the Maxwell-Boltzmann equation takes on its maximum value. At this speed, we have

$$0 = \frac{d}{dv}\left(\frac{df(v)}{dv}\right) = \frac{d}{dv}\left[4\pi\left(\frac{m}{2\pi kT}\right)^{3/2} v^2 \exp\left(\frac{-mv^2}{2kT}\right)\right]$$
$$= \left[4\pi\left(\frac{m}{2\pi kT}\right)^{3/2} \exp\left(\frac{-mv^2}{2kT}\right)\right]\left[2v - \frac{mv^3}{kT}\right]$$

from which

$$v_{mp} = \sqrt{\frac{2kT}{m}} \approx 1.414\sqrt{\frac{kT}{m}}$$

The ***average speed***, $\bar{v}$ or $\langle v\rangle$, is the expected value of the scalar velocity ($g(v) = v$). We find

$$\bar{v} = \langle v\rangle = \int_0^\infty 4\pi\left(\frac{m}{2\pi kT}\right)^{3/2} v^3 exp\left(\frac{-mv^2}{2kT}\right) dv = \sqrt{\frac{8kT}{\pi m}} \approx 1.596\sqrt{\frac{kT}{m}}$$

The ***mean-square speed***, $\overline{v^2}$ or $\langle v^2\rangle$, is the expected value of the velocity squared ($g(v) = v^2$):

$$\overline{v^2} = \langle v^2\rangle = \int_0^\infty 4\pi\left(\frac{m}{2\pi kT}\right)^{3/2} v^4 exp\left(\frac{-mv^2}{2kT}\right) dv = \frac{3kT}{m}$$

and the ***root mean-square speed***, v_{rms}, is

$$v_{rms} = \sqrt{\langle v^2\rangle} = \sqrt{\frac{3kT}{m}} \approx 1.732\sqrt{\frac{kT}{m}}$$

Example 6.8.1

Figure 6 shows the velocity distribution 300 K for nitrogen molecules at 300 K.

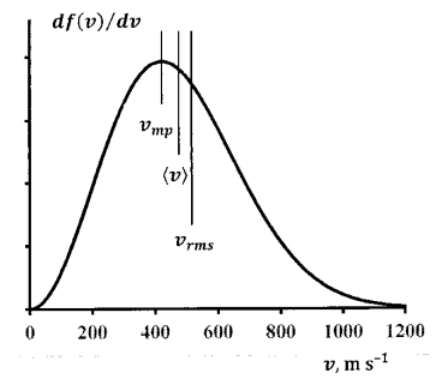


Figure 6. The Maxwell-Boltzmann distribution function for N_2 at 300K.

Solution

Finally, let us find the variance of the velocity; that is, the expected value of $(v - \langle v\rangle)^2$:

$$\text{variance}\,(v) = \sigma_v^2$$

$$
\begin{aligned}
&= \int_0^\infty (v - \langle v \rangle)^2 \left(\frac{df(v)}{dv} \right) dv \\
&= \int_0^\infty v^2 \left(\frac{df}{dv} \right) dv - 2\langle v \rangle \int_0^\infty v \left(\frac{df}{dv} \right) dv + \langle v \rangle^2 \int_0^\infty \left(\frac{df}{dv} \right) \\
&= \left\langle v^2 \right\rangle - 2\langle v \rangle \langle v \rangle + \langle v \rangle^2 \\
&= \left\langle v^2 \right\rangle - \langle v \rangle^2
\end{aligned}
$$

For N_2 at 300 K, we calculate:

$$v_{mp} = 422 \text{ m s}^{-1}$$

$$\langle v \rangle = \bar{v} = 476 \text{ m s}^{-1}$$

$$v_{rms} = 517 \text{ m s}^{-1}$$

$$\text{Variance}\,(v) = \sigma_v^2 = 40.23 \times 10^{-3} \text{ m s}^{-1}$$

$$\sigma_v = 201 \text{ m s}^{-1}$$

6.9: Pressure Variations for Macroscopic Samples

At 300 K, the standard deviation of N_2 speeds is about 40% of the average speed. Clearly the relative variation among molecular speeds in a sample of ordinary gas is very large. Why do we not observe macroscopic effects from this variation? In particular, if we measure the pressure at a small area of the container wall, why do we not observe pressure variations that reflect the wide variety of speeds with which molecules strike the wall?

Qualitatively, the answer is obvious. A single molecule whose scalar velocity is v contributes $P_1(v) = mv^2/3V$ to the pressure on the walls of its container. (See problem 20.) When we measure pressure, we measure an average squared velocity. Even if we measure the pressure over a very small area and a very short time, the

number of molecules striking the wall during the time of the measurement is very large. Consequently, the average speed of the molecules hitting the wall during any one such measurement is very close to the average speed in any other such measurement.

We are now able to treat this question quantitatively. For N_2 gas at 300 K and 1 bar, roughly 3×10^{15} molecules collide with a square millimeter of wall every microsecond. (See problem 12.) The standard deviation of the velocity of an N_2 molecule is $201\ \mathrm{m\ s^{-1}}$. Using the central limit theorem, the standard deviation of the average of 3×10^{15} molecular speeds is

$$\frac{201\ m\ s^{-1}}{\sqrt{3 \times 10^{15}}} \approx 4 \times 10^{-6} \mathrm{ms}^{-1}$$

The distribution of the average of 3×10^{15} molecular speeds is very narrow indeed.

Similarly, when molecular velocities follow the Maxwell-Boltzmann distribution function, we can show that the expected value of the pressure for a single-molecule collision is $\langle P_1(v)\rangle = kT/V$. (See problem 21.) The variance of the distribution of these individual pressure measurements is $\sigma^2_{P_1(v)} = 2k^2T^2/3V^2$, so that the magnitude of the standard deviation is comparable to that of the average:

$$\sigma_{P_1(v)} / \langle P_1(v)\rangle = \sqrt{2/3}$$

For the distribution of averages of 3×10^{15} pressure contributions, we find

$$\begin{aligned} P_{avg} &= \langle P_1(v)\rangle \\ &= \sqrt{3/2}\sigma_{P_1(v)} \end{aligned}$$

$$\sigma_{avg} = \frac{\sigma_{P_1(v)}}{\sqrt{3 \times 10^{15}}}$$

and

$$\frac{\sigma_{avg}}{P_{avg}} \approx 1.5 \times 10^{-8}$$

6.10: Collisions between Gas Molecules Relative Velocity Coordinates

The pressure of a gas depends on the frequency with which molecules collide with the wall of their container. The rate at which gas molecules escape through a very small opening in their container is called the ***effusion rate***. The effusion rate rate also depends on the frequency of collisions with the wall. (See problem 4.12.) Other gas properties depend not on the rate of collision with the wall, but on the rate with which gas molecules collide with one another. We turn now to some of these properties. For these considerations, we need to describe the motion of one molecule relative to another. We need the probability density function for the ***relative velocity*** of two particles.

To describe the relative velocity of two particles, we introduce ***relative velocity coordinates***. Let us begin by considering a Cartesian coordinate frame, with x-, y-, and z-axes, whose origin is at a point O; we will use $Oxyz$ to designate this set of axes. We specify the location of particle 1 by the vector $\vec{r}_1 = (x_1, y_1, z_1) = x_1\vec{i} + y_1\vec{j} + z_1\vec{k}$ and that of particle 2 by $\vec{r}_2 = (x_2, y_2, z_2)$. We let the location of the center of mass of this two-particle

system be specified by $\vec{r}_0 = (x_0, y_0, z_0)$. The vector from particle 1 to particle 2, $\vec{r}_{12} = (x_{12}, y_{12}, z_{12})$, is the vector difference

$$\vec{r}_{12} = \vec{r}_2 - \vec{r}_1 = (x_2 - x_1, y_2 - y_1, z_2 - z_1)$$

When the particles are moving, these vectors and their components are functions of time. Using the notation $\dot{x}_1 = dx_1/dt$, we can specify the velocity of particle 1, for example, as $\vec{v}_1 = d\vec{r}_1/dt = (\dot{x}_1, \dot{y}_1, \dot{z}_1)$. Our goal is to find the relative velocity vector, $\vec{v}_{12} = d\vec{r}_{12}/dt$. We call the components of $\vec{v}_{12}$ the relative velocity coordinates.

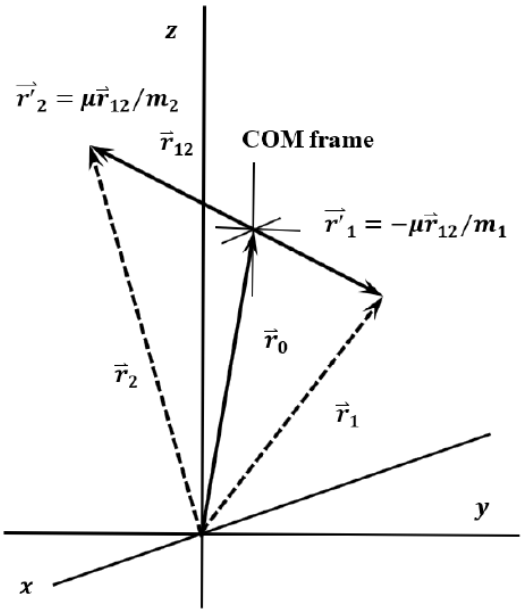


Figure 7. The center of mass frame

The essential idea underlying relative velocity coordinates is that the vectors $\vec{r}_0$ and $\vec{r}_{12}$ contain the same information as the vectors $\vec{r}_1$ and $\vec{r}_2$. This is equivalent to saying that we can transform the locations as specified by (x_1, y_1, z_1) and (x_2, y_2, z_2) to the same locations as specified by (x_0, y_0, z_0) and (x_{12}, y_{12}, z_{12}), and *vice versa*. To accomplish this, we write the equation defining the x-component of the center of mass, x_0:

$$m_1(x_1 - x_0) + m_2(x_2 - x_0) = 0$$

which we rearrange to

$$\frac{x_1}{m_2} + \frac{x_2}{m_1} = \left(\frac{1}{m_1} + \frac{1}{m_2}\right) x_0$$

Corresponding relationships can be written for the y- and z-components. It proves to be useful to introduce the ***reduced mass***, μ, defined by

$$\frac{1}{\mu} = \frac{1}{m_1} + \frac{1}{m_2}$$

Using the reduced mass, we can express the coordinates of the center of mass in terms of the coordinates of the individual particles. That is,

$$x_0 = \left(\frac{\mu}{m_2}\right) x_1 + \left(\frac{\mu}{m_1}\right) x_2$$

$$y_0 = \left(\frac{\mu}{m_2}\right) y_1 + \left(\frac{\mu}{m_1}\right) y_2$$

$$z_0 = \left(\frac{\mu}{m_2}\right) z_1 + \left(\frac{\mu}{m_1}\right) z_2$$

Since, by definition, we also have

$$x_{12} = x_2 - x_1$$

$$y_{12} = y_2 - y_1$$

$$z_{12} = z_2 - z_1$$

we have developed the transformation from (x_0, y_0, z_0) and (x_{12}, y_{12}, z_{12}) to (x_1, y_1, z_1) and (x_2, y_2, z_2). The inverse transformation is readily found to be

$$x_1 = x_0 - (\mu/m_1)\, x_{12}$$

$$y_1 = y_0 - (\mu/m_1)\, y_{12}$$

$$z_1 = z_0 - (\mu/m_1)\, z_{12}$$

$$x_2 = x_0 + (\mu/m_2)\, x_{12}$$

$$y_2 = y_0 + (\mu/m_2)\, y_{12}$$

$$z_2 = z_0 + (\mu/m_2)\, z_{12}$$

Now we can create two new Cartesian coordinate frames. Which of these is more useful depends on the objective of the particular analysis we have at hand. We call the first one the ***center of mass frame***, $O_O x' y' z'$. It is sketched in Figure 7. The x'-, y'-, and z'-axes of $O_O x' y' z'$ are parallel to the corresponding axes of $Oxyz$, but their origin, O_O, is always at the point occupied by the center of mass of the two-particle system. In this reference frame, the coordinates of particles 1 and 2 are their displacements from the center of mass:

$$x_1' = x_1 - x_0 = -(\mu/m_1)\, x_{12}$$

$$y_1' = y_1 - y_0 = -(\mu/m_1)\, y_{12}$$

$$z_1' = z_1 - z_0 = -(\mu/m_1)\, z_{12}$$

$$x_2' = x_2 - x_0 = (\mu/m_2)\, x_{12}$$

$$y_2' = y_2 - y_0 = (\mu/m_2)\, y_{12}$$

$$z_2' = z_2 - z_0 = (\mu/m_2)\, z_{12}$$

The center of mass frame is particularly useful for analyzing interactions between colliding particles.

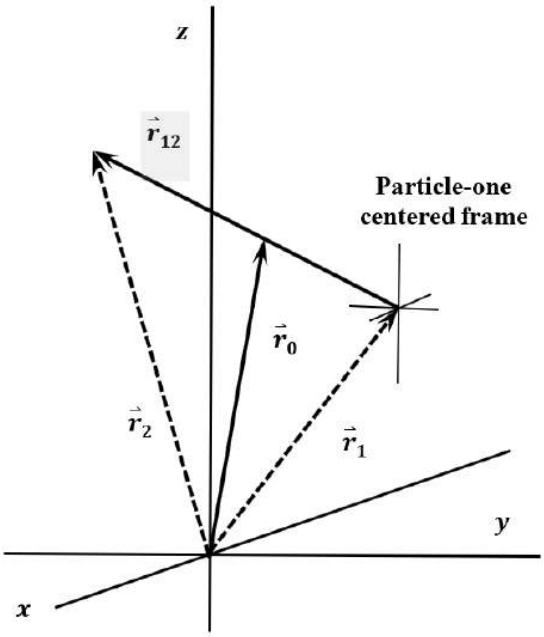


Figure 8. The particle-one centered frame.

For our purposes, a third Cartesian coordinate frame, which we will denote the $O_1x''y''z''$ frame, is more useful. It is sketched in Figure 8. The x''-, y''-, and z''-axes of $O_1x''y''z''$ are parallel to the corresponding axes of $Oxyz$, but their origin, O_1, is always at the point occupied by particle 1. In this reference frame, the coordinates of particles 1 and 2 are

$$x_1'' = 0$$

$$y_1'' = 0$$

$$z_1'' = 0$$

$$x_2'' = x_2 - x_1 = x_{12}$$

$$y_2'' = y_2 - y_1 = y_{12}$$

$$z_2'' = z_2 - z_1 = z_{12}$$

and the coordinates of the center of mass are

$$x_0'' = x_0 - x_1 = \mu x_{12}/m_1$$

$$y_0'' = y_0 - y_1 = \mu y_{12}/m_1$$

$$z_0'' = z_0 - z_1 = \mu z_{12}/m_1$$

The $O_1x''y''z''$ frame is sometimes called the center of mass frame also. To avoid confusion, we call $O_1x''y''z''$ the ***particle-one centered frame***. In the particle-one centered frame, particle 1 is stationary at the origin. With its tail at the origin, the vector $\vec{r}_{12} = (x_{12}, y_{12}, z_{12})$ specifies the position of particle 2.

We are interested in the relative velocity of particles 1 and 2. The velocity components for particles 1 and 2, and for their relative velocity, are obtained by finding the time-derivatives of the corresponding displacement components. Since the transformations of the displacement coordinates are linear, the velocity components transform from one reference frame to another in exactly the same way that the displacement components do. We have

$$\vec{v}_0 = d\vec{r}_0/dt = (\dot{x}_0, \dot{y}_0, \dot{z}_0)$$

and

$$\vec{v}_{12} = d\vec{r}_{12}/dt = (\dot{x}_{12}, \dot{y}_{12}, \dot{z}_{12})$$

The vector $\vec{v}_{12}$ specifies the velocity of particle 2, relative to a stationary particle 1. Just as $\vec{r}_0$ and $\vec{r}_{12}$ contain the same information as the vectors $\vec{r}_1$ and $\vec{r}_2$, the vectors $\vec{v}_0$ and $\vec{v}_{12}$ contain the same information as $\vec{v}_1$ and $\vec{v}_2$. Since a parallel displacement leaves a vector unchanged, each of these vectors is the same in any of the three reference frames. In §11, we find the probability density function for the magnitude of the scalar relative velocity, $v_{12} = \left|\vec{v}_{12}\right|$. Since the probability is independent of direction, the probability that two molecules have relative velocity $\vec{v}_{12}$ is the same as that they have relative velocity $-\vec{v}_{12}$. (In spherical coordinates, if $\vec{v}_{12} = (v_{12}, \theta, \varphi)$, then $-\vec{v}_{12} = (v_{12}, \theta + \pi, \varphi + \pi)$.) The probability and magnitude of the relative velocity are independent of which particle—if either—we choose to view as being stationary; they are independent of whether the particles are approaching or receding from one another.

6.11: The Probability Density Function for the Relative Velocity

From our development of the Maxwell-Boltzmann probability density functions, we can express the probability that the velocity components of particle 1 lie in the intervals v_{1x} to $v_{1x}+dv_{1x}$; v_{1y} to $v_{1y}+dv_{1y}$; v_{1z} to $v_{1z}+dv_{1z}$; while those of particle 2 simultaneously lie in the intervals v_{2x} to $v_{2x}+dv_{2x}$; v_{2y} to $v_{2y}+dv_{2y}$; v_{2z} to $v_{2z}+dv_{2z}$ as

$$\begin{aligned}\left(\frac{df\left(v_{1x}\right)}{dv_{1x}}\right)\left(\frac{df\left(v_{1y}\right)}{dv_{1y}}\right)\left(\frac{df\left(v_{1z}\right)}{dv_{1z}}\right)\left(\frac{df\left(v_{2x}\right)}{dv_{2x}}\right)\left(\frac{df\left(v_{2y}\right)}{dv_{2y}}\right)\left(\frac{df\left(v_{2z}\right)}{dv_{2z}}\right)&\\ \times dv_{1x}dv_{1y}dv_{1z}dv_{2x}dv_{2y}dv_{2z}&\\ =\left(\frac{df\left(v_1\right)}{dv_1}\right)\left(\frac{df\left(v_2\right)}{dv_2}\right)dv_1dv_2&\end{aligned}$$

We want to express this probability using the relative velocity coordinates. Since the velocity of the center of mass and the relative velocity are independent, we might expect that the Jacobian of this transformation is just the product of the two individual Jacobians. This turns out to be the case. The Jacobian of the transformation

$$\left(\dot{x}_1,\dot{y}_1,\dot{z}_1,\dot{x}_2,\dot{y}_2,\dot{z}_2\right)\rightarrow\left(\dot{x}_0,,\dot{y}_0,\dot{z}_0,\dot{x}_{12},\dot{y}_{12},\dot{z}_{12}\right)$$

is a six-by-six determinate. It is messy, but straightforward, to show that it is equal to the product of two three-by-three determinants and that the absolute value of this product is one. Therefore, we have

$$\begin{aligned}dv_{1x}dv_{1y}dv_{1z}dv_{2x}dv_{2y}dv_{2z}=&d\dot{x}_1d\dot{y}_1d\dot{z}_1d\dot{x}_2d\dot{y}_2d\dot{z}_2\\ =&d\dot{x}_0d\dot{y}_0d\dot{z}_0d\dot{x}_{12}d\dot{y}_{12}d\dot{z}_{12}\end{aligned}$$

We transform the probability density by substituting into the one-dimensional probability density functions. That is,

$$\begin{aligned}\left(\frac{df\left(v_1\right)}{dv_1}\right)\left(\frac{df\left(v_2\right)}{dv_2}\right)&=\left(\frac{m_1}{2\pi kT}\right)^{3/2}\exp\left(\frac{-m_1\left(v_{1x}^2+v_{1y}^2+v_{1z}^2\right)}{2kT}\right)\times\left(\frac{m_2}{2\pi kT}\right)^{3/2}\exp\left(\frac{-m_2\left(v_{2x}^2+v_{2y}^2+v_{2z}^2\right)}{2kT}\right)\\ &=\left(\frac{m_1m_2}{4\pi^2k^2T^2}\right)^{3/2}\times\exp\left(\frac{-m_1\left(v_{1x}^2+v_{1y}^2+v_{1z}^2\right)-m_2\left(v_{2x}^2+v_{2y}^2+v_{2z}^2\right)}{2kT}\right)\\ &=\left(\frac{m_1m_2}{4\pi^2k^2T^2}\right)^{3/2}\times\exp\left(\frac{-\frac{m_1m_2}{\mu}\left(\dot{x}_0^2+\dot{y}_0^2+z_0^2\right)-\mu\left(\dot{x}_{12}^2+\dot{y}_{12}^2+\dot{z}_{12}^2\right)}{2kT}\right)\end{aligned}$$

where the last expression specifies the probability density as a function of the relative velocity coordinates.

Next, we make a further transformation of variables. We convert the velocity of the center of mass, $\left(\dot{x}_0,\dot{y}_0,\dot{z}_0\right)$, and the relative velocity, $\left(\dot{x}_{12},\dot{y}_{12},\dot{z}_{12}\right)$, from Cartesian coordinates to spherical coordinates, referred to the $Oxyz$ axis system. (The motion of the center of mass is most readily visualized in the original frame $Oxyz$. The relative motion, $\overrightarrow{v}_{12}$, is most readily visualized in the Particle-One Centered Frame, $O_1x''y''z''$. In $O_1x''y''z''$, the motion of particle 2 is specified by $\dot{x}_2''=\dot{x}_{12}$, $\dot{y}_2''=\dot{y}_{12}$, and $\dot{z}_2''=\dot{z}_{12}$. The motion of the center of mass is specified by $\dot{x}_0''=\mu\dot{x}_{12}/m_1$, $\dot{y}_0''=\mu\dot{y}_{12}/m_1$, and $\dot{z}_0''=\mu\dot{z}_{12}/m_1$. Since it is the relative motion that is actually of interest, it might seem that we should refer the spherical coordinates to the $O_1x''y''z''$ frame. This is an unnecessary distinction because all three coordinate frames are parallel to one another, and $\overrightarrow{r}_0$ and $\overrightarrow{r}_{12}$ are the same vectors in all three frames.) Letting

$$v_0^2=\dot{x}_0^2+\dot{y}_0^2+z_0^2$$

$$v_{12}^2=\dot{x}_{12}^2+\dot{y}_{12}^2+z_{12}^2$$

the Cartesian velocity components are expressed in spherical coordinates by

$$\dot{x}_0=v_0\sin\theta_0\cos\varphi_0$$

$$\dot{y}_0=v_0\sin\theta_0\sin\varphi_0$$

$$\dot{z}_0 = v_0 \cos\theta_0$$

$$\dot{x}_{12} = v_{12} \sin\theta_{12}\ \cos\varphi_{12}$$

$$\dot{y}_{12} = v_{12} \sin\theta_{12}\ \sin\varphi_{12}$$

$$\dot{z}_{12} = v_{12} \cos\theta_{12}$$

The angles θ_0, θ_{12}, φ_0, and φ_{12} are defined in the usual manner relative to the $Oxyz$ axis system. The Jacobian of this transformation is a six-by-six determinate; which can again be converted to the product of two three-by-three determinates. We find

$$d\dot{x}_0 d\dot{y}_0 d\dot{z}_0 d\dot{x}_{12} d\dot{y}_{12} d\dot{z}_{12} =$$

$$= v_0^2 \sin\theta_0\ dv_0 d\theta_0 d\varphi_0 v_{12}^2 \sin\theta_{12}\ dv_{12} d\theta_{12} d\varphi_{12}$$

The probability that the components of the velocity of the center of mass lie in the intervals v_0 to $v_0 + dv_0$; θ_0 to $\theta_0 + d\theta_0$; φ_0 to $\varphi_0 + d\varphi_0$; while the components of the relative velocity lie in the intervals v_{12} to $v_{12} + dv_{12}$; θ_{12} to $\theta_{12} + d\theta_{12}$; φ_{12} to $\varphi_{12} + d\varphi_{12}$; becomes

$$\left(\frac{m_1 m_2}{4\pi^2 k^2 T^2}\right)^{3/2} exp\left(\frac{-m_1 m_2 v_0^2}{2\mu k T}\right) exp\left(\frac{-\mu v_{12}^2}{2kT}\right) \times$$

$$v_0^2 \sin\theta_0\ dv_0 d\theta_0 d\varphi_0 v_{12}^2 \sin\theta_{12}\ dv_{12} d\theta_{12} d\varphi_{12}$$

We are interested in the probability increment for the relative velocityrelative velocity:probability density function irrespective of the velocity of the center of mass. To sum the contributions for all possible motions of the center of mass, we integrate this expression over the possible ranges of v_0, θ_0, and φ_0. We have

$$\left(\frac{df\left(v_{12}\right)}{dv_{12}}\right)\left(\frac{df\left(\theta_{12}\right)}{d\theta_{12}}\right)\left(\frac{df\left(\varphi_{12}\right)}{d\varphi_{12}}\right) dv_{12} d\theta_{12} d\varphi_{12} =$$

$$= \left(\frac{m_1 m_2}{4\pi^2 k^2 T^2}\right)^{3/2} \int_0^\infty v_0^2 exp\left(\frac{-m_1 m_2 v_0^2}{2\mu k T}\right) dv_0 \times \int_0^\pi \sin\theta_0\ d\theta_0 \int_0^{2\pi} d\varphi_0$$

$$\times \left[v_{12}^2 exp\left(\frac{-\mu v_{12}^2}{2kT}\right) \sin\theta_{12}\ dv_{12} d\theta_{12} d\varphi_{12}\right]$$

$$= \left(\frac{\mu}{2\pi k T}\right)^{3/2} v_{12}^2 exp\left(\frac{-\mu v_{12}^2}{2kT}\right) \sin\theta_{12}\ dv_{12} d\theta_{12} d\varphi_{12}$$

This is the same as the probability increment for a single-particle velocity—albeit with μ replacing m; v_{12} replacing v; θ_{12} replacing θ; and φ_{12} replacing φ. As in the single-particle case, we can obtain the probability increment for the scalar component of the relative velocity by integrating over all possible values of θ_{12} and φ_{12}. We find

$$\frac{df\left(v_{12}\right)}{dv_{12}} = 4\pi\left(\frac{\mu}{2\pi k T}\right)^{3/2} v_{12}^2 exp\left(\frac{-\mu v_{12}^2}{2kT}\right) dv_{12}$$

In §8, we find the most probable velocity, the mean velocity, and the root-mean-square velocity for a gas whose particles have mass m. By identical arguments, we obtain the most probable relative velocity, the mean relative velocity, and the root-mean-square relative velocity. To do so, we can simply substitute μ for m in the earlier results. In particular, the mean relative velocity is

$$\bar{v}_{12} = \langle v_{12} \rangle = \left(\frac{8kT}{\pi\mu}\right)^{1/2} \approx 1.596 \left(\frac{kT}{\pi\mu}\right)^{1/21}$$

If particles 1 and 2 have the same mass, m, the reduced mass becomes $\mu = m/2$. In this case, we have

$$\langle v_{12} \rangle = \left(\frac{2\left(8kT\right)}{\pi m}\right)^{1/2} = \sqrt{2}\,\langle v \rangle$$

We can arrive at this same conclusion by considering the relative motion of two particles that represents the average case. As illustrated in Figure 9, this occurs when the two particles have the same speed, $\langle v \rangle$, but are moving at 90-degree angles to one another. In this situation, the length of the resultant vector—the relative speed— is just

$$|\bar{v}_{12}| = \langle v_{12} \rangle = \sqrt{2}\,\langle v \rangle\,.$$

6.12: The Frequency of Collisions between Unlike Gas Molecules

Thus far in our theoretical development of the properties of gases, we have assumed that ideal gas molecules are point masses. While they can collide with the walls of their container, point masses cannot collide with one another. As we saw in our discussion of van der Waals equation, the deviation of real gases from ideal gas behavior is one indication that an individual gas molecule occupies a finite volume.

To develop a model for molecular collisions, we need to know the size and shape of the colliding molecules. For a general model, we want to use the simplest possible size and shape. Accordingly, we consider a model in which gas molecules are spheres with well-defined radii. We let the radii of molecules 1 and 2 be σ_1 and σ_2, respectively. See Figure 10. When such molecules collide, their surfaces must come into contact, and the distance between their centers must be $\sigma_{12} = \sigma_1 + \sigma_2$. We call σ_{12} the ***collision radius***.

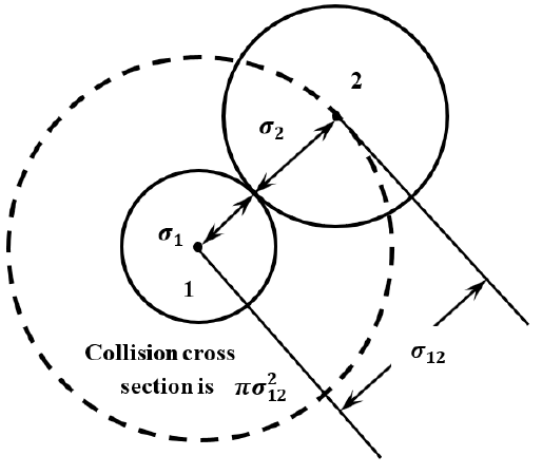


Figure 10. The molecular collision radius.

Let us consider a molecule of type 1 in a container with a large number of molecules of type 2. We suppose that there are N_2 molecules of type 2 per unit volume. Every molecule of type 2 has some velocity, v_{12}, relative to the molecule of type 1. From our development above, we know both the probability density function for v_{12} and the expected value $\langle v_{12} \rangle$. Both molecule 1 and all of the molecules of type 2 are moving with continuously varying speeds. However, it is reasonable to suppose that—on average—the encounters between molecule 1 and molecules of type 2 are the same as they would be if all of the type 2 molecules were fixed at random locations in the volume, and molecule 1 moved among them with a speed equal to the average relative velocity, $\langle v_{12} \rangle$.

Under this assumption, a molecule 1 travels a distance equal to $\langle v_{12} \rangle$ in unit time. As it does so, it collides with any type 2 molecule whose center is within a distance σ_{12} of its own center. For the moment, let us suppose that the trajectory of molecule 1 is unaffected by the collisions it experiences. Then, in unit time, molecule 1 sweeps out a cylinder whose length is $\langle v_{12} \rangle$ and whose cross-sectional area is $\pi\sigma_{12}^2$. The volume of this cylinder is $\pi\sigma_{12}^2 \langle v_{12} \rangle$. (See Figure 11.)

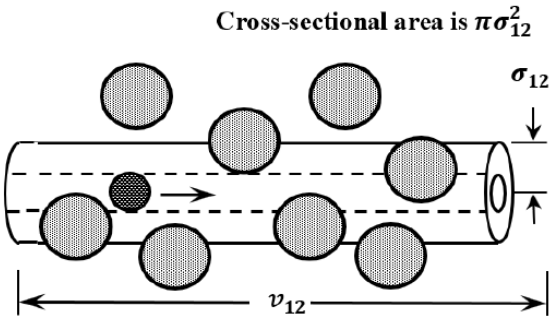


Figure 11. The collision volume of a gas molecule in unit time.

Since there are N_2 molecules of type 2 per unit volume, the number of type 2 molecules in the cylinder is $N_2\pi\sigma_{12}^2 \langle v_{12} \rangle$. Each of these molecules is a molecule of type 2 that experiences a collision with molecule 1 in unit time. Letting $\tilde{\nu}_{12}$ be the frequency (number of collision per unit time) with which molecule 1 collides with molecules of type 2, we have

$$\tilde{\nu}_{12} = N_2 \pi \sigma_{12}^2 \langle v_{12} \rangle = N_2 \pi \sigma_{12}^2 \left(\frac{8kT}{\pi \mu} \right)^{1/2}$$

$$= N_2 \sigma_{12}^2 \left(\frac{8\pi kT}{\mu} \right)^{1/2}$$

Two additional parameters that are useful for characterizing molecular collisions are τ_{12}, the ***mean time between collisions***, and λ_{12}, the ***mean distance*** that molecule 1 travels between collisions with successive molecules of type 2. λ_{12} is called the ***mean free path***. The mean time between collisions is simply the reciprocal of the collision frequency,

$$\tau_{12} = 1/\tilde{\nu}_{12}$$

and the mean free path for molecule 1 is the distance that molecule 1 actually travels in this time, which is $\langle v_1 \rangle$, not $\langle v_{12} \rangle$, so that

$$\lambda_{12} = \langle v_1 \rangle \tau_{12} = \frac{(\mu/m)^{1/2}}{N_2 \pi \sigma_{12}^2}$$

Now, we need to reevaluate the assumption that the trajectory of a molecule of a molecule 1 is unaffected by its collisions with molecules of type 2. Clearly, this is not the case. The path of molecule 1 changes abruptly at each collision. The actual cylinder that molecule 1 sweeps out will have numerous kinks, as indicated in Figure 12. The kinked cylinder can be produced from a straight one by making a series oblique cuts (one for each kink) across the straight cylinder and then rotating the ends of each cut into convergence. If we think of the cylinder as a solid rod, its volume is unchanged by these cuttings and rotations. The volume of the kinked cylinder is the same as that of the straight cylinder. Thus, our conclusions about the collision frequency, the mean time between collisions, and the mean free path are not affected by the fact that the trajectory of molecule 1 changes at each collision.

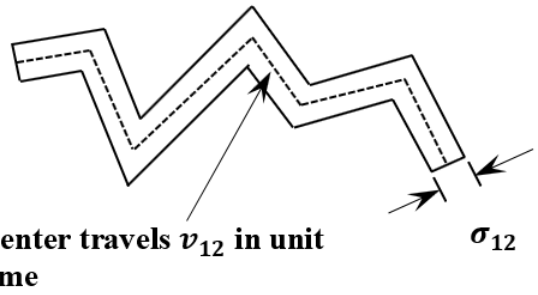


Figure 12. The collision volume is unaffected by collisions.

6.13: The Rate of Collisions between Unlike Gas Molecules

We define the collision frequencycollision frequency, $\tilde{\nu}_{12}$, as the number of collision per unit time between a single molecule of type 1 and any of the molecules of type 2 present in the same container. We find $\tilde{\nu}_{12} = N_2 \pi \sigma_{12}^2 \langle v_{12} \rangle$. If there are N_1 molecules of type 1 present in a unit volume of the gas, the total number of collisions between type 1 molecules and type 2 molecules is N_1 times greater. For clarity, let us refer to the total number of such collisions, per unit volume and per unit time, as the ***collision rate***, ρ_{12}. We have

$$\rho_{12} = N_1 \tilde{\nu}_{12} = N_1 N_2 \pi \sigma_{12}^2 \langle v_{12} \rangle = N_1 N_2 \sigma_{12}^2 \left(\frac{8 \pi k T}{\mu} \right)^{1/2}$$

6.14: Collisions between like Gas Molecules

When we consider collisions between different gas molecules of the same substance, we can denote the relative velocity and the expected value of the relative velocity as v_{11} and $\langle v_{11} \rangle$, respectively. By the argument we make above, we can find the number of collisions between any one of these molecules and all of the others. Letting this collision frequency be $\tilde{\nu}_{11}$, we find

$$\tilde{\nu}_{11} = N_1 \pi \sigma_{11}^2 \langle v_{11} \rangle ,$$

where $\sigma_{11} = 2\sigma_1$. Since we have

$$\langle v_{11} \rangle = \sqrt{2} \langle v_1 \rangle ,$$

while

$$\langle v_1 \rangle = \sqrt{8kT/\pi m_1},$$

we have $\langle v_{11} \rangle = 4\sqrt{kT/\pi m_1}$. The frequency of collisions between molecules of the same substance becomes

$$\tilde{\nu}_{11} = N_1 \pi \sigma_{11}^2 \langle v_{11} \rangle = 4 N_1 \sigma_{11}^2 \left(\frac{\pi kT}{m_1} \right)^{1/2}$$

The mean time between collisions, τ_{11}, is

$$\tau_{11} = 1/\tilde{\nu}_{11}$$

and the mean free path, λ_{11},

$$\lambda_{11} = \langle v_1 \rangle \tau_{11} = \frac{1}{\sqrt{2} N_1} \pi \sigma_{11}^2$$

When we consider the rate of collisions between all of the molecules of type 1 in a container, ρ_{11}, there is a minor complication. If we multiply the collision frequency per molecule, $\tilde{\nu}_{11}$, by the number of molecules available to undergo such collisions, N_1, we count each collision twice, because each such collision involves two type 1 molecules. To find the collision rate among like molecules, we must divide this product by 2. That is,

$$\rho_{11} = \frac{N_1 \tilde{\nu}_{11}}{2} = 2 N_1^2 \sigma_{11}^2 \left(\frac{\pi kT}{m_1} \right)^{1/2}$$

6.15: The Geometry of A Collision between Spherical Molecules

Thus far we have not concerned ourselves with the relative orientation of a pair of colliding molecules. We want to develop a more detailed model for the collision process[1] itself, and the first step is to specify what we mean by relative orientation.

As before, we consider a molecule of type 1 moving with the relative velocity v_{12} through a gas of stationary type 2 molecules. In unit time, molecule 1 travels a distance v_{12} and collides with many molecules of type 2. We can characterize each such collision by the angle, θ, between the velocity vector and the line of centers of the colliding pair. For glancing collisions, we have $\theta = \pi/2$. For head-on collisions, we have $\theta = 0$. All else being equal, the collision will be more violent the smaller the angle θ. Evidently, we can describe the average effect of collisions more completely if we can specify the frequency of collisions as a function of θ. More precisely, we want to find the frequency of collisions in which this angle lies between θ and $\theta + d\theta$.

When a collision occurs, the distance between the molecular centers is σ_{12}. We can say that the center of molecule 2 is at a particular point on the surface of a sphere, of radius σ_{12}, circumscribed about molecule 1. As sketched in Figure 13, we can rotate the line of centers around the velocity vector, while keeping the angle between them constant at θ. As we do so, the line of centers traces out a circle on the surface of the sphere; collisions that put the center of molecule 2 at any two points on this circle are completely equivalent. Letting the radius of this circle be r, we see that $r = \sigma_{12}\sin\theta$. Evidently, for spherical molecules, specifying θ specifies the relative orientation at the time of collision.

If we now allow θ to vary by $d\theta$, the locus of equivalent points on the circumscribed sphere expands to a band. Measured along the surface of the sphere, the width of this band is $\sigma_{12}d\theta$. As molecule 1 moves through the gas of stationary type 2 molecules, this band sweeps out a cylindrical shell. Molecule 1 collides, at an angle between θ and $\theta + d\theta$, with every type 2 molecule in this cylindrical shell. Conversely, every type 2 molecule in this cylindrical shell collides with molecule 1 at an angle between θ and $\theta + d\theta$. (Molecule 1 also collides with many other type 2 molecules, but those collisions are at other angles; they have different orientations.) In unit time, the length of the cylindrical shell is v_{12}. The volume of the cylindrical shell is its length times its cross-sectional area.

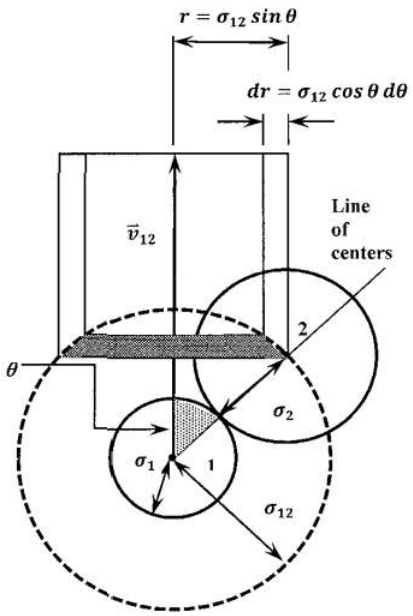


Figure 13. The geometry of collisions between spheres.

The cross-section of the cylindrical shell is a circular annulus. Viewing the annulus as a rectangular strip whose length is the circumference of the shell and whose width is the radial thickness of the annulus, the area of the annulus is the circumference times the radial thickness. Since the radius of the shell is $r = \sigma_{12}\sin\theta$, its circumference is $2\pi\sigma_{12}\sin\theta$. The radial thickness of the annulus is just the change in the distance, $r = \sigma_{12}\sin\theta$, between the velocity vector and the wall of the cylinder when θ changes by a small amount $d\theta$. This is

$$dr = \left(\frac{dr}{d\theta}\right) d\theta = \sigma_{12}\cos\theta d\theta$$

Therefore, the area of the annulus is

$$2\pi\sigma_{12}^2 \sin\theta\cos\theta d\theta$$

and the volume of the cylindrical shell swept out by a type 1 molecule (traveling at exactly the speed v_{12}) in unit time is

$$2\pi\sigma_{12}^2 v_{12} \sin\theta\cos\theta d\theta$$

We again let N_2 be the number of molecules of type 2 per unit volume. The number of collisions, per unit time, between a molecule of type 1, traveling at exactly v_{12}, and molecules of type 2, in which the collision angle lies between θ and $\theta + d\theta$ is

$$2\pi N_2 \sigma_{12}^2 v_{12} \sin\theta\cos\theta d\theta$$

We need to find the number of such collisions in which the relative velocity lies between v_{12} and $v_{12} + dv_{12}$. The probability of finding v_{12} in this interval is $\left(df\left(v_{12}\right)/dv_{12}\right) dv_{12}$. Let $d\tilde{\nu}_{12}\left(\nu_{12}, \theta\right)$ be the number of collisions made in unit time, by a type 1 molecule, with molecules of type 2, in which the collision angle is between θ and $\theta + d\theta$, and the scalar relative velocity is between v_{12} and $v_{12} + dv_{12}$. This is just the number of collisions when the relative velocity is v_{12} multiplied by the probability that the relative velocity is between v_{12} and $v_{12} + dv_{12}$. We have the result we need:

$$\begin{aligned} d\tilde{\nu}_{12}\left(v_{12}, \theta\right) &= 2\pi N_2 \sigma_{12}^2 v_{12} \left(\frac{df\left(v_{12}\right)}{dv_{12}}\right) \sin\theta\, \cos\theta\, d\theta dv_{12} \\ &= 8\pi^2 N_2 \sigma_{12}^2 \left(\frac{\mu}{2\pi kT}\right)^{3/2} v_{12}^3 exp\left(\frac{-\mu v_{12}^2}{2kT}\right) \times \sin\theta\, \cos\theta\, d\theta dv_{12} \end{aligned}$$

Recognizing that possible values of θ lie in the range $0 \leq \theta < \pi/2$ and that possible values of v_{12} lie in the range $0 \leq v_{12} < \infty$, we can find the frequency of all possible collisions, $\tilde{\nu}_{12}$, by summing over all possible values of θ and v_{12}. That is,

$$\begin{aligned} \tilde{\nu}_{12} &= 8\pi^2 N_2 \sigma_{12}^2 \left(\frac{\mu}{2\pi kT}\right)^{3/2} \int_0^\infty v_{12}^3 exp\left(\frac{-\mu v_{12}^2}{2kT}\right) dv_{12} \times \int_0^{\pi/2} \sin\theta\, \cos\theta\, d\theta \\ &= 8\pi^2 N_2 \sigma_{12}^2 \left(\frac{\mu}{2\pi kT}\right)^{3/2} \left[2\left(\frac{kT}{\mu}\right)^2\right] \left[\frac{1}{2}\right] \\ &= N_2 \sigma_{12}^2 \left(\frac{8\pi kT}{\mu}\right)^{1/2} \end{aligned}$$

In Section 4.12, we obtained this result by a slightly different argument, in which we did not explicitly consider the collision angle, θ.

6.16: The Energy of A Collision between Gas Molecules

It is useful to extend our model of molecular collisions to suppose that one or both of the molecules can undergo chemical change as a result of the collision. In doing so, we are introducing some ideas that we develop further in Chapter 5.

When we ask about the factors that determine whether such a reaction can occur, there can be several possibilities. We want to focus on one such factor—the violence of the collision. We expect that a collision is more likely to result in a reaction the harder the two molecules hit one another. When we try to formulate our basis for this expectation, we see that the underlying idea is that a collision deforms the colliding molecules. The more violent the collision, the greater the deformation, and the greater the likelihood of reaction becomes.

To proceed, we need to be more precise about what we mean by the violence of the collision. Evidently, what we have in mind has two components: the relative velocity and the collision angle. If the collision is a glancing one, $\theta = \pi/2$, we expect the effect on the molecules to be minimal, even if the relative velocity is high. On the other hand, a direct collision, $\theta \approx 0$, might lead to reaction even if the relative velocity is comparatively low. With these ideas in mind, we see that a reasonable model is to suppose that forces acting along the line of centers can lead to reaction, whereas forces acting perpendicular to the line of centers cannot. If the colliding molecules have complex shapes, this may be a poor assumption.

We also need a way to specify how much deformation occurs in a collision. If we want to specify the deformation by describing specific changes in the molecular structures, this is a complex problem. For a general model, however, we can avoid this level of detail. To do so, we recognize that any deformation can proceed only until the work done in deforming the molecules equals the energy that can be expended to do this work. As the molecules are deformed, their potential energies change. The maximum change in this potential energy is just the amount of kinetic energy that the colliding molecules can use to effect this deformation. We can identify this amount of kinetic energy with the component of the molecules' kinetic energy that is associated with their relative motion along the line of centers.

If we now associate a threshold level of deformation with the occurrence of a chemical change, the kinetic energy required to effect this deformation determines whether the change can occur. If the available kinetic energy is less than that required to achieve the threshold level of deformation, reaction cannot occur. If the available kinetic energy exceeds this minimum, reaction takes place. We call the minimum kinetic energy the ***activation energy*** and usually represent it by the symbol ϵ_a. (In discussing reaction rates, we usually express the activation energy per mole and represent it as E_a, where

$$E_a = \overline{N}\epsilon_a.)$$

We can apply these ideas to our model for collision between spherical molecules. In Section 4.10, we develop relative velocity coordinates. It follows that we can partition kinetic energy of the two-particle system into a component that depends on the velocity of the center of mass and a component that depends on the relative velocity. That is, we have

$$\begin{aligned} KE &= \frac{m_1 v_1^2}{2} + \frac{m_2 v_2^2}{2} \\ &= \frac{m_1 m_2}{2\mu}\left(\dot{x}_0^2 + \dot{y}_0^2 + z_0^2\right) + \mu\left(\dot{x}_{12}^2 + \dot{y}_{12}^2 + \dot{z}_{12}^2\right) \\ &= \frac{m_1 m_2 v_0^2}{2\mu} + \frac{\mu v_{12}^2}{2} \end{aligned}$$

Only the component that depends on the relative velocity can contribute to the deformation of the colliding molecules. The relative velocity can be resolved into components parallel and perpendicular to the line of centers. The parallel component is the projection of the velocity vector onto the line of centers. This is $v_{12}\cos\theta$, and the perpendicular component is $v_{12}\sin\theta$. We see that the kinetic energy associated with the relative motion of particles 1 and 2 has a component

$$\frac{\mu v_{12}^2 \cos^2\theta}{2}$$

parallel to the line of centers and a component

$$\frac{\mu v_{12}^2 \sin^2\theta}{2}$$

perpendicular to it.

The idea that the kinetic energy parallel to the line of centers must exceed ε_a for reaction to occur can now be expressed as the requirement that

$$\epsilon_a < \frac{\mu v_{12}^2 \cos^2\theta}{2}$$

When we consider all possible collisions between molecules 1 and 2, the collision angle varies from 0 to $\pi/2$. However, only those collisions for which v_{12} satisfies the inequality above will have sufficient kinetic energy along the line of centers for reaction to occur. The smallest value of v_{12} that can satisfy this inequality occurs when $\theta = 0$. This minimum relative velocity is

$$v_{12}^{minimun} = (2\epsilon_a/\mu)^{1/2}$$

For relative velocities in excess of this minimum, collisions are effective only when

$$\cos\theta > \left(2\epsilon_a/\mu v_{12}^2\right)^{1/2}$$

so that

$$\theta < \cos^{-1}\left(2\epsilon_a/\mu v_{12}^2\right)^{1/2}$$

Let us designate the frequency of collisions satisfying these constraints as $\tilde{\nu}_{12}(\epsilon_a)$. Recalling that

$$d\tilde{\nu}_{12}(\nu_{12},\theta) = 8\pi^2 N_2 \sigma_{12}^2 \left(\frac{\mu}{2\pi kT}\right)^{3/2} v_{12}^3 \exp\left(\frac{-\mu v_{12}^2}{2kT}\right) \times \sin\theta\cos\theta\, d\theta dv_{12}$$

we see that

$$\tilde{\nu}_{12}(\varepsilon_a) = 8\pi^2 N_2 \sigma_{12}^2 \left(\frac{\mu}{2\pi kT}\right)^{3/2} \times \int_{v_{12}=(2\epsilon_a/\mu)^{1/2}}^{\infty} \int_{\theta=0}^{\cos^{-1}\left(2\epsilon_a/\mu v_{12}^2\right)^{1/2}} v_{12}^3 \exp\left(\frac{-\mu v_{12}^2}{2kT}\right) \times \sin\theta\cos\theta\, d\theta dv_{12}$$

The integral involving θ is

$$\int_{\theta=0}^{\cos^{-1}\left(2\epsilon_a/\mu v_{12}^2\right)^{1/2}} \sin\theta\cos\theta\, d\theta = \left[\frac{\sin^2\theta}{2}\right]_0^{\cos^{-1}\left(2\epsilon_a/\mu v_{12}^2\right)^{1/2}} = \frac{1}{2}\left[1 - \frac{2\epsilon_a}{\mu v_{12}^2}\right]$$

where, to evaluate the integral at its upper limit, we note that the angle $\theta = \cos^{-1}\left(2\epsilon_a/\mu v_{12}^2\right)^{1/2}$ lies in a triangle whose sides have lengths as indicated in Figure 14.

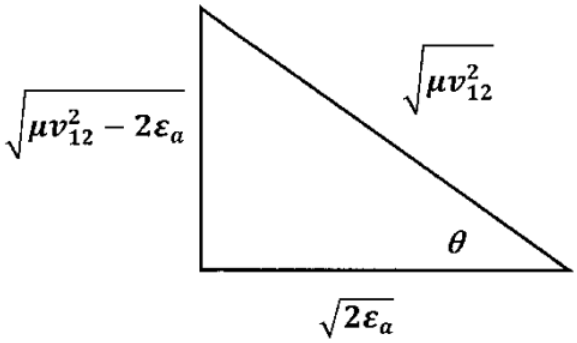


Figure 14. Maximum angle for an effective collision.

The collision frequency becomes

$$\tilde{\nu}_{12}(\epsilon_a) = 4\pi^2 N_2 \sigma_{12}^2 \left(\frac{\mu}{2\pi kT}\right)^3/2 \times \int_{v_{12}=(2\epsilon_a/\mu)^{1/2}}^{\infty} \left[1 - \frac{2\epsilon_a}{\mu v_{12}^2}\right] v_{12}^3 \exp\left(\frac{-\mu v_{12}^2}{2kT}\right) dv_{12}$$

This integral can be evaluated by making the substitution $v_{12} = (2\epsilon/\mu)^{1/2}$. The lower limit of integration becomes ϵ_a; we have

$$\begin{aligned}\int_{v_{12}=(2\epsilon_a/\mu)^{1/2}}^{\infty}\left[1-\frac{2\epsilon_a}{\mu v_{12}^2}\right]v_{12}^3\exp\left(\frac{-\mu v_{12}^2}{2kT}\right)dv_{12}&\\=\frac{2}{\mu^2}\int_{\epsilon_a}^{\infty}(\epsilon-\epsilon_a)\exp\left(\frac{-\epsilon}{kT}\right)d\epsilon&\\=2\left(\frac{kT}{\mu}\right)^2\exp\left(\frac{-\epsilon_a}{kT}\right)&\end{aligned}$$

Then

$$\begin{aligned}\tilde{\nu}_{12}\left(\varepsilon_a\right)&=4\pi^2N_2\sigma_{12}^2\left(\frac{\mu}{2\pi kT}\right)^{3/2}\times 2\left(\frac{kT}{\mu}\right)^2\exp\left(\frac{-\epsilon_a}{kT}\right)\\&=N_2\sigma_{12}^2\left(\frac{8\pi kT}{\mu}\right)^{1/2}\exp\left(\frac{-\epsilon_a}{kT}\right)\end{aligned}$$

Note that when $\epsilon_a=0$, this reduces to the same expression for $\tilde{\nu}_{12}$ that we have obtained twice previously. The frequency of collisions having kinetic energy along the line of centers in excess of ϵ_a depends exponentially on $-\epsilon_a/kT$. All else being equal, this frequency increases as the temperature increases; it decreases as the activation energy increases.

6.17: Problems

1. For an oxygen molecule at 25 C, calculate (a) the most probable velocity, (b) the average velocity, (c) the root-mean-square velocity.

2. For a gas of oxygen molecules at 25 C and 1.00 bar, calculate (a) the collision frequency, (b) the mean time between collisions, (c) the mean free path. The diameter of an oxygen molecule, as estimated from gas-viscosity measurements, is 3.55 x 10^{-10} m.

3. For oxygen molecules at 25 C, calculate (a) the fraction with speeds between 150 and 151 m s^{-1}, (b) the fraction with speeds between 400 and 401 m s^{-1}, (c) the fraction with speeds between 550 and 551 m s^{-1}.

4. For a hydrogen molecule at 100 C, calculate (a) the most probable velocity, (b) the average velocity, (c) the root-mean-square velocity.

5. For a gas of hydrogen molecules at 100 C and 1.00 bar, calculate (a) the collision frequency, (b) the mean time between collisions, (c) the mean free path. The diameter of a hydrogen molecule, as estimated from gas-viscosity measurements, is 2.71 x 10^{-10} m.

6. For a uranium hexafluoride (UF_6) molecule at 100 C, calculate (a) the most probable velocity, (b) the average velocity, (c) the root-mean-square velocity.

7. For a gas of uranium hexafluoride molecules at 100 C and 1.00 bar, calculate (a) the collision frequency, (b) the mean time between collisions, (c) the mean free path. Assume that the diameter of a uranium hexafluoride molecule is 7.0×10^{-10} m.

8. What is the average kinetic energy of hydrogen molecules at 100 C? What is the average kinetic energy of uranium hexafluoride (UF_6) molecules at 100 C?

9. Assuming the temperature in interstellar space is 2.73 K, calculate, for a hydrogen atom, (a) the most probable velocity, (b) the average velocity, (c) the root-mean-square velocity.

10. Assuming that interstellar space is occupied entirely by hydrogen atoms at a particle density of 10^2 molecules m^{-3}, calculate (a) the collision frequency, (b) the mean number of years between collisions, (c) the mean free path. Assume that the diameter of a hydrogen atom is 2.40×10^{-10} m.

11. Ignoring any effects attributable to its charge and assuming that the temperature is 2.73 K, calculate, for an electron in interstellar space, (a) the most probable velocity, (b) the average velocity, (c) the root-mean-square velocity.

12. If a wall of a gas-filled container contains a hole, gas molecules escape through the hole. If all of the molecules that hit the hole escape, but the hole is so small that the number escaping has no effect on the velocity distribution of the remaining gas molecules, we call the escaping process **effusion**. That is, we call the process effusion only if it satisfies three rather stringent criteria. First, the hole must be large enough (and the wall must be thin enough) so that most molecules passing through the hole do not hit the sides of the hole. Second, a molecule that passes through the hole must not collide with anything on the other side that can send it back through the hole into the original container. Third, the hole must be small enough so that the escaping molecules do not create a pressure gradient; the rate at which gas molecules hit the hole and escape must be determined entirely by the equilibrium distribution of gas velocities and, of course, the area of the hole. Show that the number of molecules effusing through a hole of area A in time t is

$$At\left(\frac{N}{V}\right)\left(\frac{kT}{2\pi m}\right)^{1/2}$$

where (N/V) is the number density of molecules in the container, and m is their molecular mass.

13. A vessel contains hydrogen and oxygen at 350 K and partial pressures of 0.50 bar and 1.50 bar, respectively. These gases effuse into a vacuum. What is the ratio of hydrogen to oxygen in the escaping gas?

14. How could we use effusion to estimate the molecular weight of an unknown substance?

15. An equimolar mixture of $^{235}UF_6$ and $^{238}UF_6$ is subjected to effusion. What is the ratio of ^{235}U to ^{238}U in the escaping gas?

16. Calculate the number of nitrogen molecules that collide with 10^{-6} m^2 of wall in 10^{-6} s, if the pressure is 1.00 bar and the temperature is 300 K.

17. Air is approximately 20% oxygen and 80% nitrogen by volume. Assume that oxygen and nitrogen molecules both have a radius of 1.8×10^{-8} m. For air at 1.0 bar and 298 K, calculate:

(a) The number of collisions that one oxygen molecule makes with nitrogen molecules every second.

(b) The number of collisions that occur between oxygen and nitrogen molecules in one cubic meter of air every second.

(c) The number of collisions that one oxygen molecule makes with other oxygen molecules every second.

(d) The number of collisions that occur between oxygen molecules in one cubic meter of air every second.

(e) The number of collisions that occur between oxygen and nitrogen molecules in one cubic meter each second in which the kinetic energy along the line of centers exceeds 100 kJ mol^{-1} or 1.66×10^{-19} J per collision.

(f) The number of oxygen-nitrogen collisions that occur in which the kinetic energy along the line of centers exceeds 50 kJ mol^{-1}.

18. Show that $\int_0^\infty \rho_v\left(v\right) dv \neq 1$.

19. For what volume element, υ, is

$$P\left(\upsilon\right) = f_{xyz}\left(v_x, v_y, v_z\right)?$$

20. Using the model we develop in Section 2.10:

(a) Show that the pressure, $P_1\left(v\right)$, attributable to a single molecule of mass m and velocity v in a container of volume V is

$$P_1\left(v\right) = \frac{mv^2}{3V}$$

(b) In Section 4.6, we find that this pressure is

$$\delta P_1\left(v\right) = \frac{2mv^2 \cos^2\theta}{V}$$

for a molecule whose velocity vector lies between θ and $\theta + d\theta$ and between φ and $\varphi + d\varphi$. This angular region comprises a solid angle whose magnitude is $d\Omega = \sin\theta d\theta d\varphi$. Since the solid angle surrounding a given point is 4π, the probability that a randomly oriented velocity vector lies between θ and $\theta + d\theta$ and between φ and $\varphi + d\varphi$ is

$$\frac{d\Omega}{4\pi} = \frac{\sin\theta d\theta d\varphi}{4\pi}$$

Therefore, given that the scalar component of a molecule's velocity is v, its contribution to the pressure at A is

$$dP_1\left(v\right) = \left(\frac{mv^2}{2\pi V}\right) \cos^2\theta \sin\theta \, d\theta d\varphi$$

To find the pressure contribution made by this molecule irrespective of the values of θ and φ, we must integrate $dP_1\left(v\right)$ over all values of θ and φ that allow the molecule to impact the wall at A. Recalling that these ranges are $0 \leq \theta < \pi/2$ and $0 \leq \varphi < 2\pi$, show that

$$P_1\left(v\right) = \frac{mv^2}{3V}$$

21. Taking $P_1\left(v\right) = mv^2/3V$ as the contribution made to the pressure by one molecule whose velocity is v:

(a) Show that the expected value for the contribution made to the pressure by one molecule when the Maxwell–Boltzmann distribution function describes the distribution of molecular velocities is

$$\left\langle P_1\left(v\right)\right\rangle = \frac{kT}{V}$$

(b) Show that the variance of the contribution made to the pressure by one molecule is

$$\sigma^2_{P_1(v)} = \frac{2k^2T^2}{3V^2}$$

What is the standard deviation, $\sigma_{P_1(v)}$?

(c) What is the value of the ratio

$$\frac{\sigma_{P_1(v)}}{\langle P_1\,(v)\rangle}$$

(d) Taking 3×10^{15} as the number of collisions of N_2 molecules at 1 bar and 300 K with one one square millimeter per microsecond, what pressure, P_{avg}, would we find if we could measure the individual contribution made by each collision and compute their average? What would be the variance, σ^2_{avg}, of this average? The standard deviation, σ_{avg}? The ratio σ_{avg}/P_{avg}?

22. Let $\epsilon = mv^2/2$ be the translational kinetic energy of a gas molecule whose mass is m. Show that the probability density function for ϵ is

$$\frac{df}{d\epsilon} = 2\pi\left(\frac{1}{\pi kT}\right)^{3/2}\epsilon^{1/2}exp\left(\frac{-\epsilon}{kT}\right)$$

Letting the translational kinetic energy per mole be $E = \overline{N}\epsilon$, show that

$$\frac{df}{dE} = 2\pi\left(\frac{1}{\pi RT}\right)^{3/2}E^{1/2}exp\left(\frac{-E}{RT}\right)$$

Notes

[1] Our collision model and quantitative treatment of the role of activation energy in chemical reaction rates follow those given by Arthur A. Frost and Ralph G. Pearson, ***Kinetics and Mechanism***, 2^{nd} Ed., John Wiley and Sons, New York, 1961, pp 65-68. See also R. H. Fowler, ***Statistical Mechanics***, Cambridge University Press, New York, 1936.

CHAPTER OVERVIEW

7: Diffusion

CHAPTER OVERVIEW

8: Chemical Kinetics

Chemical kinetics is the study of the speed with which a *chemical* reaction occurs and the factors that affect this speed. This information is especially useful for determining how a reaction occurs.

8.1: Reaction Rates

In this Module, the quantitative determination of a reaction rate is demonstrated. Reaction rates can be determined over particular time intervals or at a given point in time. A rate law describes the relationship between reactant rates and reactant concentrations. Reaction rates are usually expressed as the concentration of reactant consumed or the concentration of product formed per unit time. The units are thus moles per liter per unit time, written as M/s, M/min, or M/h. To measure reaction rates, chemists initiate the reaction, measure the concentration of the reactant or product at different times as the reaction progresses, perhaps plot the concentration as a function of time on a graph, and then calculate the change in the concentration per unit time.

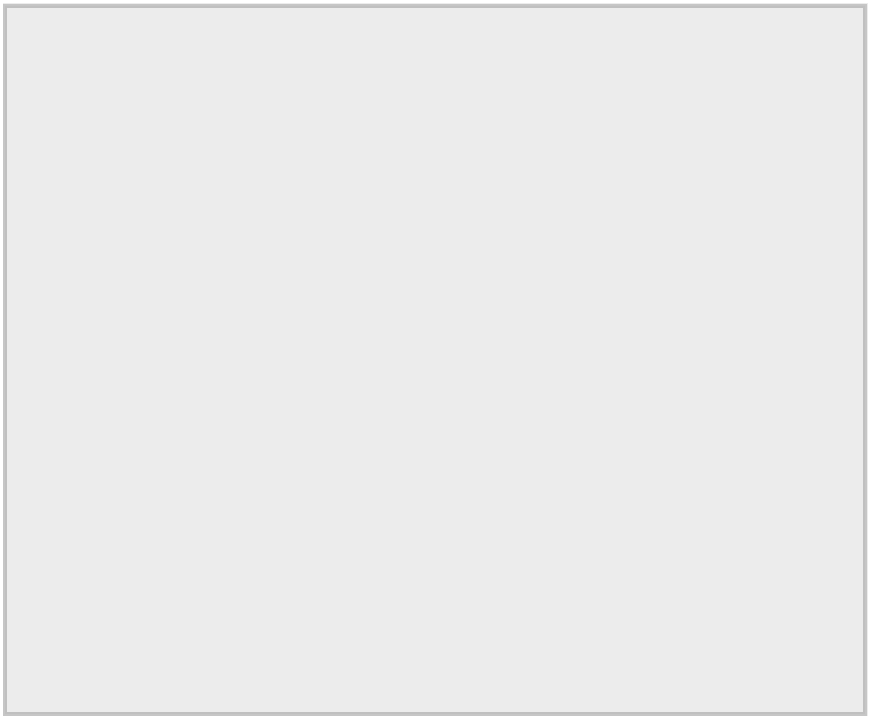


Figure 8.1.1 The Progress of a Simple Reaction (A → B). The mixture initially contains only A molecules (purple). Over time, the number of A molecules decreases and more B molecules (green) are formed (top). The graph shows the change in the number of A and B molecules in the reaction as a function of time over a 1 min period (bottom).

The progress of a simple reaction (A → B) is shown in Figure 8.1.1; the beakers are snapshots of the composition of the solution at 10 s intervals. The number of molecules of reactant (A) and product (B) are plotted as a function of time in the graph. Each point in the graph corresponds to one beaker in Figure 8.1.1. The reaction rate is the change in the concentration of either the reactant or the product over a period of time. The concentration of A decreases with time, while the concentration of B increases with time.

$$\text{rate} = \frac{\Delta[\text{B}]}{\Delta t} = -\frac{\Delta[\text{A}]}{\Delta t} \tag{8.1.1}$$

Square brackets indicate molar concentrations, and the capital Greek delta (Δ) means “change in.” Because chemists follow the convention of expressing all reaction rates as positive numbers, however, a negative sign is inserted in front of Δ[A]/Δt to convert that expression to a positive number. The reaction rate calculated for the reaction A → B using Equation 8.1.1 is different for each interval (this is not true for every reaction, as shown below). A greater change occurs in [A] and [B] during the first 10 s interval, for example, than during the last, meaning that the reaction rate is greatest at first.

> *Reaction rates generally decrease with time as reactant concentrations decrease.*

Study 1: The Hydrolysis of Aspirin

We can use Equation 8.1.1 to determine the reaction rate of hydrolysis of aspirin, probably the most commonly used drug in the world (more than 25,000,000 kg are produced annually worldwide). Aspirin (acetylsalicylic acid) reacts with water (such as water in body fluids) to give salicylic acid and acetic acid, as shown in Figure 8.1.2.

Figure 8.1.2

Because salicylic acid is the actual substance that relieves pain and reduces fever and inflammation, a great deal of research has focused on understanding this reaction and the factors that affect its rate. Data for the hydrolysis of a sample of aspirin are in Table 8.1.1 and are shown in the graph in Figure 8.1.3. These data were obtained by removing samples of the reaction mixture at the indicated times and analyzing them for the concentrations of the reactant (aspirin) and one of the products (salicylic acid).

Table 8.1.1: Data for Aspirin Hydrolysis in Aqueous Solution at pH 7.0 and 37°C*

Time (h)	[Aspirin] (M)	[Salicylic Acid] (M)
0	5.55×10^{-3}	0
2.0	5.51×10^{-3}	0.040×10^{-3}
5.0	5.45×10^{-3}	0.10×10^{-3}
10	5.35×10^{-3}	0.20×10^{-3}
20	5.15×10^{-3}	0.40×10^{-3}
30	4.96×10^{-3}	0.59×10^{-3}
40	4.78×10^{-3}	0.77×10^{-3}
50	4.61×10^{-3}	0.94×10^{-3}
100	3.83×10^{-3}	1.72×10^{-3}
200	2.64×10^{-3}	2.91×10^{-3}
300	1.82×10^{-3}	3.73×10^{-3}

*The reaction at pH 7.0 is very slow. It is much faster under acidic conditions, such as those found in the stomach.

The **average reaction rate** for a given time interval can be calculated from the concentrations of either the reactant or one of the products at the beginning of the interval (time = t_0) and at the end of the interval (t_1). Using salicylic acid, the reaction rate for the interval between t = 0 h and t = 2.0 h (recall that change is always calculated as final minus initial) is calculated as follows:

$$\text{rate}_{(t=0-2.0\text{ h})} = \frac{[\text{salicyclic acid}]_2 - [\text{salicyclic acid}]_0}{2.0\text{ h} - 0\text{ h}} \tag{8.1.2}$$

$$= \frac{0.040 \times 10^{-3}\text{ M} - 0\text{ M}}{2.0\text{ h}} = 2.0 \times 10^{-5}\text{ M/h} \tag{8.1.3}$$

The reaction rate can also be calculated from the concentrations of aspirin at the beginning and the end of the same interval, remembering to insert a negative sign, because its concentration decreases:

$$\text{rate}_{(t=0-2.0\text{ h})} = -\frac{[\text{aspirin}]_2 - [\text{aspirin}]_0}{2.0\text{ h} - 0\text{ h}} \tag{8.1.4}$$

$$= -\frac{(5.51 \times 10^{-3}\text{ M}) - (5.55 \times 10^{-3}\text{ M})}{2.0\text{ h}} \tag{8.1.5}$$

$$= 2 \times 10^{-5}\text{ M/h} \tag{8.1.6}$$

If the reaction rate is calculated during the last interval given in Table 8.1.1(the interval between 200 h and 300 h after the start of the reaction), the reaction rate is significantly slower than it was during the first interval (t = 0–2.0 h):

$$\text{rate}_{(t=200-300\text{h})} = \frac{[\text{salicyclic acid}]_{300} - [\text{salicyclic acid}]_{200}}{300\text{ h} - 200\text{ h}} \tag{8.1.7}$$

$$= -\frac{(3.73 \times 10^{-3}\text{ M}) - (2.91 \times 10^{-3}\text{ M})}{100\text{ h}} \tag{8.1.8}$$

$$= 8.2 \times 10^{-6}\text{ M/h} \tag{8.1.9}$$

Study 2: The Fermentation of Sucrose

In the preceding example, the stoichiometric coefficients in the balanced chemical equation are the same for all reactants and products; that is, the reactants and products all have the coefficient 1. Consider a reaction in which the coefficients are not all the same, the fermentation of sucrose to ethanol and carbon dioxide:

$$\underset{\text{sucrose}}{C_{12}H_{22}O_{11}(aq)} + H_2O(l) \rightarrow 4C_2H_5OH(aq) + 4CO_2(g) \tag{8.1.10}$$

The coefficients indicate that the reaction produces four molecules of ethanol and four molecules of carbon dioxide for every one molecule of sucrose consumed. As before, the reaction rate can be found from the change in the concentration of any reactant or product. In this particular case, however, a chemist would probably use the concentration of either sucrose or ethanol because gases are usually measured as volumes and the volume of CO_2 gas formed depends on the total volume of the solution being studied and the solubility of the gas in the solution, not just the concentration of sucrose. The coefficients in the balanced chemical equation tell us that the reaction rate at which ethanol is formed is always four times faster than the reaction rate at which sucrose is consumed:

$$\frac{\Delta[C_2H_5OH]}{\Delta t} = -\frac{4\Delta[\text{sucrose}]}{\Delta t} \tag{8.1.11}$$

The concentration of the reactant—in this case sucrose—*decreases* with time, so the value of Δ[sucrose] is negative. Consequently, a minus sign is inserted in front of Δ[sucrose] in Equation 8.1.11 so the rate of change of the sucrose concentration is expressed as a positive value. Conversely, the ethanol concentration *increases* with time, so its rate of change is automatically expressed as a positive value.

Often the reaction rate is expressed in terms of the reactant or product with the smallest coefficient in the balanced chemical equation. The smallest coefficient in the sucrose fermentation reaction (Equation 8.1.10) corresponds to sucrose, so the reaction rate is generally defined as follows:

$$\text{rate} = -\frac{\Delta[\text{sucrose}]}{\Delta t} = \frac{1}{4}\left(\frac{\Delta[C_2H_5OH]}{\Delta t}\right) \tag{8.1.12}$$

Example 8.1.1

Consider the thermal decomposition of gaseous N_2O_5 to NO_2 and O_2 via the following equation:

$$2N_2O_5(g) \xrightarrow{\Delta} 4NO_2(g) + O_2(g)$$

Write expressions for the reaction rate in terms of the rates of change in the concentrations of the reactant and each product with time.

Given: balanced chemical equation

Asked for: reaction rate expressions

Strategy:

A. Choose the species in the equation that has the smallest coefficient. Then write an expression for the rate of change of that species with time.
B. For the remaining species in the equation, use molar ratios to obtain equivalent expressions for the reaction rate.

Solution

A Because O_2 has the smallest coefficient in the balanced chemical equation for the reaction, define the reaction rate as the rate of change in the concentration of O_2 and write that expression.

B The balanced chemical equation shows that 2 mol of N_2O_5 must decompose for each 1 mol of O_2 produced and that 4 mol of NO_2 are produced for every 1 mol of O_2 produced. The molar ratios of O_2 to N_2O_5 and to NO_2 are thus 1:2 and 1:4,

respectively. This means that the rate of change of $[N_2O_5]$ and $[NO_2]$ must be divided by its stoichiometric coefficient to obtain equivalent expressions for the reaction rate. For example, because NO_2 is produced at four times the rate of O_2, the rate of production of NO_2 is divided by 4. The reaction rate expressions are as follows:

$$\text{rate} = \frac{\Delta[O_2]}{\Delta t} = \frac{\Delta[NO_2]}{4\Delta t} = -\frac{\Delta[N_2O_5]}{2\Delta t}$$

? Exercise 8.1.1

The key step in the industrial production of sulfuric acid is the reaction of SO_2 with O_2 to produce SO_3.

$$2SO_{2(g)} + O_{2(g)} \rightarrow 2SO_{3(g)} \tag{8.1.13}$$

Write expressions for the reaction rate in terms of the rate of change of the concentration of each species.

Answer

$$\text{rate} = -\frac{\Delta[O_2]}{\Delta t} = -\frac{\Delta[SO_2]}{2\Delta t} = \frac{\Delta[SO_3]}{2\Delta t}$$

✔ Example 8.1.2

Using the reaction shown in Example 8.1.1, calculate the reaction rate from the following data taken at 56°C:

$$2N_2O_{5(g)} \rightarrow 4NO_{2(g)} + O_{2(g)} \tag{8.1.14}$$

Time (s)	$[N_2O_5]$ (M)	$[NO_2]$ (M)	$[O_2]$ (M)
240	0.0388	0.0314	0.00792
600	0.0197	0.0699	0.0175

Given: balanced chemical equation and concentrations at specific times

Asked for: reaction rate

Strategy:

A. Using the equations in Example 8.1.1, subtract the initial concentration of a species from its final concentration and substitute that value into the equation for that species.
B. Substitute the value for the time interval into the equation. Make sure your units are consistent.

Solution

A Calculate the reaction rate in the interval between t_1 = 240 s and t_2 = 600 s. From Example 8.1.1, the reaction rate can be evaluated using any of three expressions:

$$\text{rate} = \frac{\Delta[O_2]}{\Delta t} = \frac{\Delta[NO_2]}{4\Delta t} = -\frac{\Delta[N_2O_5]}{2\Delta t}$$

Subtracting the initial concentration from the final concentration of N_2O_5 and inserting the corresponding time interval into the rate expression for N_2O_5,

$$\text{rate} = -\frac{\Delta[N_2O_5]}{2\Delta t} = -\frac{[N_2O_5]_{600} - [N_2O_5]_{240}}{2(600\text{ s} - 240\text{ s})}$$

B Substituting actual values into the expression,

$$\text{rate} = -\frac{0.0197\text{ M} - 0.0388\text{ M}}{2(360\text{ s})} = 2.65 \times 10^{-5}\text{ M/s}$$

Similarly, NO_2 can be used to calculate the reaction rate:

$$\text{rate} = \frac{\Delta[NO_2]}{4\Delta t} = \frac{[NO_2]_{600} - [NO_2]_{240}}{4(600\text{ s} - 240\text{ s})} = \frac{0.0699\text{ M} - 0.0314\text{ M}}{4(360\text{ s})} = 2.67 \times 10^{-5}\text{ M/s} \tag{8.1.15}$$

Allowing for experimental error, this is the same rate obtained using the data for N_2O_5. The data for O_2 can also be used:

$$\text{rate} = \frac{\Delta[O_2]}{\Delta t} = \frac{[O_2]_{600} - [O_2]_{240}}{600\ \text{s} - 240\ \text{s}} = \frac{0.0175\ \text{M} - 0.00792\ \text{M}}{360\ \text{s}} = 2.66 \times 10^{-5}\ \text{M/s} \tag{8.1.16}$$

Again, this is the same value obtained from the N_2O_5 and NO_2 data. Thus, the reaction rate does not depend on which reactant or product is used to measure it.

? Exercise 8.1.2

Using the data in the following table, calculate the reaction rate of SO_2(g) with O_2(g) to give SO_3(g).

$$2SO_{2(g)} + O_{2(g)} \rightarrow 2SO_{3(g)} \tag{8.1.17}$$

Time (s)	[SO_2] (M)	[O_2] (M)	[SO_3] (M)
300	0.0270	0.0500	0.0072
720	0.0194	0.0462	0.0148

Answer 9.0×10^{-6} M/s

Instantaneous Rates of Reaction

The **instantaneous rate** of a reaction is the reaction rate at any given point in time. As the period of time used to calculate an average rate of a reaction becomes shorter and shorter, the average rate approaches the instantaneous rate. Comparing this to calculus, the instantaneous rate of a reaction at a given time corresponds to the slope of a line tangent to the concentration-versus-time curve at that point—that is, the derivative of concentration with respect to time.

$$rate = \lim_{\Delta t \to 0} \frac{-\Delta[R]}{\Delta t} = -\frac{d[R]}{dt} \tag{8.1.18}$$

The distinction between the instantaneous and average rates of a reaction is similar to the distinction between the actual speed of a car at any given time on a trip and the average speed of the car for the entire trip. Although the car may travel for an extended period at 65 mph on an interstate highway during a long trip, there may be times when it travels only 25 mph in construction zones or 0 mph if you stop for meals or gas. The average speed on the trip may be only only 50 mph, whereas the instantaneous speed on the interstate at a given moment may be 65 mph. Whether the car can be stopped in time to avoid an accident depends on its instantaneous speed, not its average speed. There are important differences between the speed of a car during a trip and the speed of a chemical reaction, however. The speed of a car may vary unpredictably over the length of a trip, and the initial part of a trip is often one of the slowest. In a chemical reaction, the initial interval typically has the fastest rate (though this is not always the case), and the reaction rate generally changes smoothly over time.

> *Chemical kinetics generally focuses on one particular instantaneous rate, which is the initial reaction rate, t = 0. Initial rates are determined by measuring the reaction rate at various times and then extrapolating a plot of rate versus time to t = 0.*

8.2: Reaction Order

The kinetic theory of gases can be used to model the frequency of collisions between hard-sphere molecules, which is proportional to the reaction rate. Most systems undergoing a chemical reaction, however, are much more complex. The reaction rates may be dependent on specific interactions between reactant molecules, the phase(s) in which the reaction takes place, etc. The field of chemical kinetics is thus by-and-large based on empirical observations. From experimental observations, scientists have established that reaction rates almost always have a power-law dependence on the concentrations of one or more of the reactants. In the following sections, we will discuss different power laws that are commonly observed in chemical reactions.

0^{th} Order Reaction Kinetics

Consider a closed container initially filled with chemical species A. At $t = 0$, a stimulus, such as a change in temperature, the addition of a catalyst, or irradiation, causes an irreversible chemical reaction to occur in which A transforms into product B:

$$a\mathrm{A} \longrightarrow b\mathrm{B} \tag{8.2.1}$$

The rate that the reaction proceeds, r, can be described as the change in the concentrations of the chemical species with respect to time:

$$r = -\frac{1}{a}\frac{d\left[\mathrm{A}\right]}{dt} = \frac{1}{b}\frac{d\left[\mathrm{B}\right]}{dt} \tag{8.2.2}$$

where $\left[\mathrm{X}\right]$ denotes the molar concentration of chemical species X with units of $\frac{\mathrm{mol}}{\mathrm{L}^3}$.

Let us first examine a reaction $\mathrm{A} \longrightarrow \mathrm{B}$ in which the reaction rate, r, is constant with time:

$$r = -\frac{d\left[\mathrm{A}\right]}{dt} = k \tag{8.2.3}$$

where k is a constant, also known as the rate constant with units of $\frac{\mathrm{mol}}{\mathrm{m}^3\mathrm{s}}$. Such reactions are called zeroth order reactions because the reaction rate depends on the concentrations of species A and B to the 0^{th} power. Integrating $\left[\mathrm{A}\right]$ with respect to T, we find that

$$\left[\mathrm{A}\right] = -kt + c_1 \tag{8.2.4}$$

At $t = 0$, $\left[\mathrm{A}\right](0) = \left[\mathrm{A}\right]_0$. Plugging these values into the equation, we find that $c_1 = \left[\mathrm{A}\right]_0$. The final form of the equation is:

$$\left[\mathrm{A}\right] = \left[\mathrm{A}\right]_0 - kt \tag{8.2.5}$$

A plot of the concentration of species A with time for a 0^{th} order reaction is shown in Figure 8.2.1, where the slope of the line is $-k$ and the y-intercept is $\left[\mathrm{A}\right]_0$. Such reactions in which the reaction rates are independent of the concentrations of products and reactants are rare in nature. An example of a system displaying 0^{th} order kinetics would be one in which a reaction is mediated by a catalyst present in small amounts.

1^{st} Order Reaction Kinetics

Experimentally, it is observed than when a chemical reaction is of the form

$$\sum_i \nu_i \mathrm{A}_i = 0 \tag{8.2.6}$$

the reaction rate can be expressed as

$$r = k \prod_{\text{reactants}} \left[\mathrm{A}_i\right]^{\nu_i} \tag{8.2.7}$$

where it is assumed that the stoichiometric coefficients ν_i of the reactants are all positive. Thus, or a reaction $\mathrm{A} \longrightarrow \mathrm{B}$, the reaction rate depends on $\left[\mathrm{A}\right]$ raised to the first power:

$$r = \frac{d\left[\mathrm{A}\right]}{dt} = -k\left[\mathrm{A}\right] \tag{8.2.8}$$

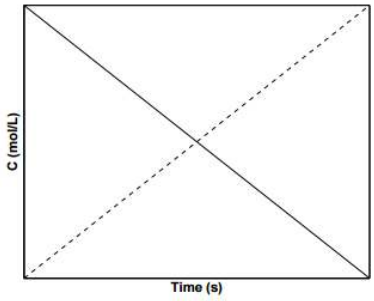


Figure 8.2.1: Plots of $[\mathrm{A}]$ (solid line) and $[\mathrm{B}]$ (dashed line) over time for a 0^{th} order reaction.

For first order reactions, k has the units of $\frac{1}{\mathrm{s}}$. Integrating and applying the condition that at $t = 0\ \mathrm{s}$, $[\mathrm{A}] = [\mathrm{A}]_0$, we arrive at the following equation:

$$[\mathrm{A}] = [\mathrm{A}]\, e^{-kt} \tag{8.2.9}$$

Figure 8.2.2 displays the concentration profiles for species A and B for a first order reaction. To determine the value of k from

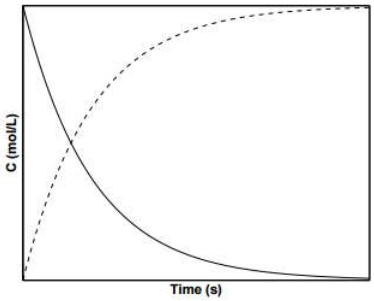


Figure 8.2.2: Plots of $[\mathrm{A}]$ (solid line) and $[\mathrm{B}]$ (dashed line) over time for a 1^{st} order reaction.

experimental data, it is convenient to take the natural log of Equation 8.2.9:

$$\ln\left([\mathrm{A}]\right) = \ln\left([\mathrm{A}]_0\right) - kt \tag{8.2.10}$$

For a first order irreversible reaction, a plot of $\ln\left([\mathrm{A}]\right)$ vs. t is straight line with a slope of $-k$ and a y-intercept of $\ln\left([\mathrm{A}]_0\right)$.

2^{nd} Order Reaction Kinetics

Another type of reaction depends on the square of the concentration of species A - these are known as second order reactions. For a second order reaction in which $2\mathrm{A} \longrightarrow \mathrm{B}$, we can write the reaction rate to be

$$r = -\frac{1}{2}\frac{d\,[\mathrm{A}]}{dt} = k[\mathrm{A}]^2 \tag{8.2.11}$$

For second order reactions, k has the units of $\frac{\mathrm{m}^3}{\mathrm{mol}\cdot\mathrm{s}}$. Integrating and applying the condition that at $t = 0\ \mathrm{s}$, $[\mathrm{A}] = [\mathrm{A}]_0$, we arrive at the following equation for the concentration of A over time:

$$[\mathrm{A}] = \frac{1}{2kt + \dfrac{1}{[\mathrm{A}]_0}} \tag{8.2.12}$$

Figure 8.2.3 shows concentration profiles of A and B for a second order reaction. To determine k from experimental data for

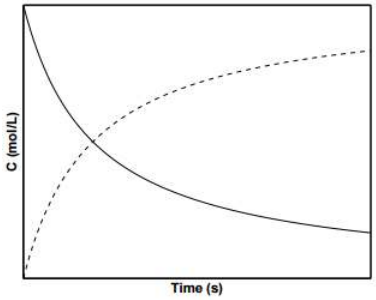


Figure 8.2.3: Plots of $[\mathrm{A}]$ (solid line) and $[\mathrm{B}]$ (dashed line) over time for a 2^{nd} order reaction.

second-order reactions, it is convenient to invert Equation 8.2.12:

$$\frac{1}{[\mathrm{A}]} = \frac{1}{[\mathrm{A}]_0} + 2kt \tag{8.2.13}$$

A plot of $1/[\mathrm{A}]$ vs. t will give rise to a straight line with slope k and intercept $1/[\mathrm{A}]_0$.

Second order reaction rates can also apply to reactions in which two species react with each other to form a product:

$$\mathrm{A} + \mathrm{B} \xrightarrow{k} \mathrm{C} \tag{8.2.14}$$

In this scenario, the reaction rate will depend on the concentrations of both A and B to the first order:

$$r = -\frac{d[\mathrm{A}]}{dt} = -\frac{d[\mathrm{B}]}{dt} = k[\mathrm{A}][\mathrm{B}] \tag{8.2.15}$$

To integrate the above equation, we need to write it in terms of one variable. Since the concentrations of A and B are related to each other via the chemical reaction equation, we can write:

$$[\mathrm{B}] = [\mathrm{B}]_0 - ([\mathrm{A}]_0 - [\mathrm{A}]) = [\mathrm{A}] + [\mathrm{B}]_0 - [\mathrm{A}]_0 \tag{8.2.16}$$

$$\frac{d[\mathrm{A}]}{dt} = -k[\mathrm{A}]([\mathrm{A}] + [\mathrm{B}]_0 - [\mathrm{A}]_0) \tag{8.2.17}$$

We can then use partial fractions to integrate:

$$kdt = \frac{d[\mathrm{A}]}{[\mathrm{A}]([\mathrm{A}] + [\mathrm{B}]_0 - [\mathrm{A}]_0)} = \frac{1}{[\mathrm{B}]_0 - [\mathrm{A}]_0}\left(\frac{d[\mathrm{A}]}{[\mathrm{A}]} - \frac{d[\mathrm{A}]}{[\mathrm{B}]_0 - [\mathrm{A}]_0 + [\mathrm{A}]}\right) \tag{8.2.18}$$

$$kt = \frac{1}{[\mathrm{A}]_0 - [\mathrm{B}]_0}\ln\frac{[\mathrm{A}][\mathrm{B}]_0}{[\mathrm{B}][\mathrm{A}]_0} \tag{8.2.19}$$

Figure 8.2.4 displays the concentration profiles of species A, B, and C for a second order reaction in which the initial concentrations of A and B are not equal.

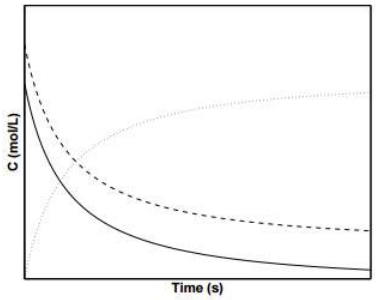


Figure 8.2.4: Plots of $[\mathrm{A}]$ (solid line), $[\mathrm{B}]$ (dashed line) and $[\mathrm{C}]$ (dotted line) over time for a 2^{nd} order reaction in which the initial concentrations of the reactants, $[\mathrm{A}]_0$ and $[\mathrm{B}]_0$, are not equal.

Contributors

- Mark Tuckerman (New York University)

8.3: Molecularity of a Reaction

In general, it is necessary to experimentally measure the concentrations of species over time in order to determine the apparent rate law governing the reaction. If the reactions are elementary reactions, (i.e. they cannot be expressed as a series of simpler reactions), then we can directly define the rate law based on the chemical equation. For example, an elementary reaction in which a single reactant transforms into a single product, is unimolecular reaction. These reactions follow 1^{st} order rate kinetics. An example of this type of reaction would be the isomerization of butane:

$$nC_4H_{10} \longrightarrow iC_4H_{10} \tag{8.3.1}$$

From the chemical reaction equation, we can directly write the rate law as

$$\frac{d\left[nC_4H_{10}\right]}{dt} = -k\left[nC_4H_{10}\right] \tag{8.3.2}$$

without the need to carry out experiments.

Elementary bimolecular reactions that involve two molecules interacting to form one or more products follow second order rate kinetics. An example would be the following reaction between a nitrate molecule and carbon monoxide to form nitrogen dioxide and carbon dioxide:

$$NO_3 + CO \longrightarrow NO_2 + CO_2 \tag{8.3.3}$$

For the above elementary reaction, we can directly write the rate law as:

$$\frac{d\left[NO_3\right]}{dt} = -k\left[NO_3\right]\left[CO\right] \tag{8.3.4}$$

Trimolecular elementary reactions involving three reactant molecules to form one or more products are rare due to the low probability of three molecules simultaneously colliding with one another.

8.4: More Complex Reactions

A major goal in chemical kinetics is to determine the sequence of elementary reactions, or the reaction mechanism, that comprise complex reactions. For example, Sherwood Rowland and Mario Molina won the Nobel Prize in Chemistry in 1995 for proposing the elementary reactions involving chlorine radicals that contribute to the overall reaction of $O_3 \rightarrow O_2$ in the troposphere. In the following sections, we will derive rate laws for complex reaction mechanisms, including reversible, parallel and consecutive reactions.

Parallel Reactions

Consider the reaction in which chemical species A undergoes one of two irreversible first order reactions to form either species B or species C :

$$\mathrm{A} \xrightarrow{k_1} \mathrm{B} \tag{8.4.1}$$

$$\mathrm{A} \xrightarrow{k_2} \mathrm{C} \tag{8.4.2}$$

The overall reaction rate for the consumption of A can be written as:

$$\frac{d[\mathrm{A}]}{dt} = -k_1[\mathrm{A}] - k_2[\mathrm{A}] = -(k_1 + k_2)[\mathrm{A}] \tag{8.4.3}$$

Integrating $[\mathrm{A}]$ with respect to t, we obtain the following equation:

$$[\mathrm{A}] = [\mathrm{A}]_0 e^{-(k_1+k_2)t} \tag{8.4.4}$$

Plugging this expression into the equation for $\frac{d[\mathrm{B}]}{dt}$, we obtain:

$$\frac{d[\mathrm{B}]}{dt} = k_1[\mathrm{A}] = k_1[\mathrm{A}]_0 e^{-(k_1+k_2)t} \tag{8.4.5}$$

Integrating $[\mathrm{B}]$ with respect to t, we obtain:

$$[\mathrm{B}] = -\frac{k_1[\mathrm{A}]_0}{k_1+k_2}\left(e^{-(k_1+k_2)t}\right) + c_1 \tag{8.4.6}$$

At $t = 0$, $[\mathrm{B}] = 0$. Therefore,

$$c_1 = \frac{k_1[\mathrm{A}]_0}{k_1+k_2} \tag{8.4.7}$$

$$[\mathrm{B}] = \frac{k_1[\mathrm{A}]_0}{k_1+k_2}\left(1 - e^{-(k_1+k_2)t}\right) \tag{8.4.8}$$

Likewise,

$$[\mathrm{C}] = \frac{k_2[\mathrm{A}]_0}{k_1+k_2}\left(1 - e^{-(k_1+k_2)t}\right) \tag{8.4.9}$$

The ratio of $[\mathrm{B}]$ to $[\mathrm{C}]$ is simply:

$$\frac{[\mathrm{B}]}{[\mathrm{C}]} = \frac{k_1}{k_2} \tag{8.4.10}$$

An important parallel reaction in industry occurs in the production of ethylene oxide, a reagent in many chemical processes and also a major component in explosives. Ethylene oxide is formed through the partial oxidation of ethylene:

$$2\ C_2H_4 + O_2 \xrightarrow{k_1} 2\ C_2H_4O \tag{8.4.11}$$

However, ethylene can also undergo a combustion reaction:

$$C_2H_4 + 3\ O_2 \xrightarrow{k_2} 2\ CO_2 + 2\ H_2O \tag{8.4.12}$$

To select for the first reaction, the oxidation of ethylene takes place in the presence of a silver catalyst, which significantly increases k_1 compared to k_2. Figure 8.4.1 displays the concentration profiles for species A, B, and C in a parallel reaction in which $k_1 > k_2$.

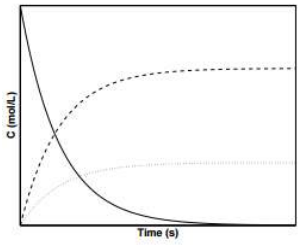


Figure 8.4.1: Plots of $[\mathrm{A}]$ (solid line), $[\mathrm{B}]$ (dashed line) and $[\mathrm{C}]$ (dotted line) over time for a parallel reaction.

Consecutive Reactions

Consider the following series of first-order irreversible reactions, where species A reacts to form an intermediate species, I, which then reacts to form the product, P:

$$\mathrm{A} \xrightarrow{k_1} \mathrm{I} \xrightarrow{k_2} \mathrm{P} \tag{8.4.13}$$

We can write the reaction rates of species A, I and P as follows:

$$\frac{d[\mathrm{A}]}{dt} = -k_1[\mathrm{A}] \tag{8.4.14}$$

$$\frac{d[\mathrm{I}]}{dt} = k_1[\mathrm{A}] - k_2[\mathrm{I}] \tag{8.4.15}$$

$$\frac{d[\mathrm{P}]}{dt} = k_2[\mathrm{I}] \tag{8.4.16}$$

As before, integrating $[\mathrm{A}]$ with respect to t leads to:

$$[\mathrm{A}] = [\mathrm{A}]_0 e^{-k_1 t} \tag{8.4.17}$$

The concentration of species I can be written as

$$[\mathrm{I}] = \frac{k_1[\mathrm{A}]_0}{k_2 - k_1}\left(e^{-k_1 t} - e^{-k_2 t}\right) \tag{8.4.18}$$

Then, solving for $[\mathrm{P}]$, we find that:

$$[\mathrm{P}] = [\mathrm{A}]_0\left[1 + \frac{1}{k_1 - k_2}\left(k_2 e^{-k_1 t} - k_1 e^{-k_2 t}\right)\right] \tag{8.4.19}$$

Figure 8.4.2 displays the concentration profiles for species A, I, and P in a consecutive reaction in which $k_1 = k_2$. As can be seen from the figure, the concentration of species I reaches a maximum at some time, $t_{\max}$. Oftentimes, species I is the desired product. Returning to the oxidation of ethylene into ethylene oxide, it is important to note another reaction in which ethylene oxide can decompose into carbon dioxide and water through the following reaction

$$C_2H_4O + \frac{5}{2}\,O_2 \xrightarrow{k_3} 2\,CO_2 + 2\,H_2O \tag{8.4.20}$$

Thus, to maximize the concentration of ethylene oxide, the oxidation of ethylene is only allowed proceed to partial completion before the reaction is stopped.

Finally, in the limiting case when $k_2 \gg k_1$, we can write the concentration of P as

$$[\mathrm{P}] \approx [\mathrm{A}]_0\left\{1 + \frac{1}{-k_2}k_2 e^{-k_1 t}\right\} = [\mathrm{A}]_0\left(1 - e^{-k_1 t}\right) \tag{8.4.21}$$

Thus, when $k_2 \gg k_1$, the reaction can be approximated as $\mathrm{A} \rightarrow \mathrm{P}$ and the apparent rate law follows 1^{st} order kinetics.

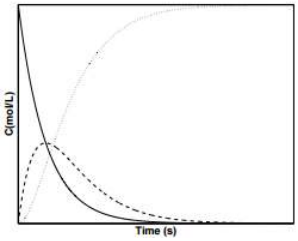


Figure 8.4.2: Plots of $[\mathrm{A}]$ (solid line), $[\mathrm{I}]$ (dashed line) and $[\mathrm{P}]$ (dotted line) over time for consecutive first order reactions.

Consecutive Reactions With an Equilibrium

Consider the reactions

$$\mathrm{A} \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} \mathrm{I} \overset{k_2}{\rightarrow} \mathrm{P} \tag{8.4.22}$$

We can write the reaction rates as:

$$\frac{d[\mathrm{A}]}{dt} = -k_1[\mathrm{A}] + k_{-1}[\mathrm{I}] \tag{8.4.23}$$

$$\frac{d[\mathrm{I}]}{dt} = k_1[\mathrm{A}] - k_{-1}[\mathrm{I}] - k_2[\mathrm{I}] \tag{8.4.24}$$

$$\frac{d[\mathrm{P}]}{dt} = k_2[\mathrm{I}] \tag{8.4.25}$$

The exact solutions of these is straightforward, in principle, but rather involved, so we will just state the exact solutions, which are

$$[\mathrm{A}](t) = \frac{[\mathrm{A}]_0}{2\lambda}\left[(\lambda - k_1 + K)\, e^{-(k_1+K-\lambda)t/2} + (\lambda + k_1 - K)\, e^{-(k_1+K+\lambda)t/2}\right] \tag{8.4.26}$$

$$[\mathrm{I}](t) = \frac{k_1[\mathrm{A}]_0}{\lambda}\left[e^{-(k_1+K-\lambda)t/2} - e^{-(k_1+K+\lambda)t/2}\right] \tag{8.4.27}$$

$$[\mathrm{P}](t) = 2k_1k_2[\mathrm{A}]_0\left[\frac{2}{(k_1+K)^2 - \lambda^2} - \frac{1}{\lambda}\left(\frac{e^{-(k_1+K-\lambda)t/2}}{k_1+K-\lambda} - \frac{e^{-(k_1+K+\lambda)t/2}}{k_1+K+\lambda}\right)\right] \tag{8.4.28}$$

where

$$K = k_2 + k_{-1} \tag{8.4.29}$$

$$\lambda = \sqrt{(k_1 - K)^2 - 4k_1k_{-1}} \tag{8.4.30}$$

Steady-State Approximations

Consider the following consecutive reaction in which the first step is reversible:

$$\mathrm{A} \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} \mathrm{I} \overset{k_2}{\rightarrow} \mathrm{P} \tag{8.4.31}$$

We can write the reaction rates as:

$$\frac{d[\mathrm{A}]}{dt} = -k_1[\mathrm{A}] + k_{-1}[\mathrm{I}] \tag{8.4.32}$$

$$\frac{d[\mathrm{I}]}{dt} = k_1[\mathrm{A}] - k_{-1}[\mathrm{I}] - k_2[\mathrm{I}] \tag{8.4.33}$$

$$\frac{d[\mathrm{P}]}{dt} = k_2[\mathrm{I}] \tag{8.4.34}$$

These equations can be solved explicitly in terms of $[\mathrm{A}]$, $[\mathrm{I}]$, and $[\mathrm{P}]$, but the math becomes very complicated quickly. If, however, $k_2 + k_{-1} \gg k_1$ (in other words, the rate of consumption of I is much faster than the rate of production of I), we can make the approximation that the concentration of the intermediate species, I, is small and constant with time:

$$\frac{d[\mathrm{I}]}{dt} \approx 0 \tag{8.4.35}$$

Equation 21.22 can now be written as

$$\frac{d[\mathrm{I}]}{dt} = k_1[\mathrm{A}] - k_{-1}[\mathrm{I}]_{ss} - k_2[\mathrm{I}]_{ss} \approx 0 \tag{8.4.36}$$

where $[\mathrm{I}]_{ss}$ is a constant represents the steady state concentration of intermediate species, $[\mathrm{I}]$. Solving for $[\mathrm{I}]_{ss}$,

$$[\mathrm{I}]_{ss} = \frac{k_1}{k_{-1}+k_2}[\mathrm{A}] \tag{8.4.37}$$

We can then write the rate equation for species A as

$$\frac{d[\mathrm{A}]}{dt} = -k_1[\mathrm{A}] + k_{-1}[\mathrm{I}]_{ss} = -k_1[\mathrm{A}] + k_{-1}\frac{k_1}{k_{-1}+k_2}[\mathrm{A}] = -\frac{k_1 k_2}{k_{-1}+k_2}[\mathrm{A}] \tag{8.4.38}$$

Integrating,

$$[\mathrm{A}] = [\mathrm{A}]_0 e^{-\frac{k_1 k_2}{k_{-1}+k_2}t} \tag{8.4.39}$$

Equation 21.28 is the same equation we would obtain for apparent 1^{st} order kinetics of the following reaction:

$$\mathrm{A} \xrightarrow{k'} \mathrm{P} \tag{8.4.40}$$

where

$$k' = \frac{k_1 k_2}{k_{-1}+k_2} \tag{8.4.41}$$

Figure 8.4.3 displays the concentration profiles for species, A, I, and P with the condition that $k_2 + k_{-1} \gg k_1$. These types of reaction kinetics appear when the intermediate species, I, is highly reactive.

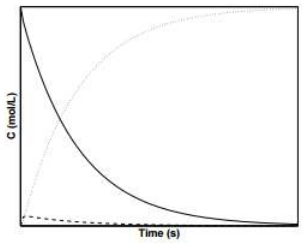


Figure 8.4.3: Plots of $[\mathrm{A}]$ (solid line), $[\mathrm{I}]$ (dashed line) and $[\mathrm{P}]$ (dotted line) over time for $k_2 + k_{-1} \gg k_1$.

Lindemann Mechanism

Consider the isomerization of methylisonitrile gas, CH_3NC, to acetonitrile gas, CH_3CN:

$$CH_3NC \xrightarrow{k} CH_3CN \tag{8.4.42}$$

If the isomerization is a unimolecular elementary reaction, we should expect to see 1^{st} order rate kinetics. Experimentally, however, 1^{st} order rate kinetics are only observed at high pressures. At low pressures, the reaction kinetics follow a 2^{nd} order rate law:

$$\frac{d[CH_3NC]}{dt} = -k[CH_3NC]^2 \tag{8.4.43}$$

To explain this observation, J.A. Christiansen and F.A. Lindemann proposed that gas molecules first need to be energized via intermolecular collisions before undergoing an isomerization reaction. The reaction mechanism can be expressed as the following two elementary reactions

$$\mathrm{A} + \mathrm{M} \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} \mathrm{A}^* + \mathrm{M} \tag{8.4.44}$$

$$\mathrm{A}^* \overset{k_2}{\rightarrow} \mathrm{B} \tag{8.4.45}$$

where M can be a reactant molecule, a product molecule or another inert molecule present in the reactor. Assuming that the concentration of A^* is small, or $k_1 \ll k_2 + k_{-1}$, we can use a steady-state approximation to solve for the concentration profile of species B with time:

$$\frac{d\left[\mathrm{A}^*\right]}{dt} = k_1\left[\mathrm{A}\right]\left[\mathrm{M}\right] - k_{-1}\left[\mathrm{A}^*\right]_{ss}\left[\mathrm{M}\right] - k_2\left[\mathrm{A}^*\right]_{ss} \approx 0 \tag{8.4.46}$$

Solving for $[\mathrm{A}^*]$,

$$\left[\mathrm{A}^*\right] = \frac{k_1\left[\mathrm{M}\right]\left[\mathrm{A}\right]}{k_2 + k_{-1}\left[\mathrm{M}\right]} \tag{8.4.47}$$

The reaction rates of species A and B can be written as

$$-\frac{d\left[\mathrm{A}\right]}{dt} = \frac{d\left[\mathrm{B}\right]}{dt} = k_2\left[\mathrm{A}^*\right] = \frac{k_1 k_2\left[\mathrm{M}\right]\left[\mathrm{A}\right]}{k_2 + k_{-1}\left[\mathrm{M}\right]} = k_{\mathrm{obs}}\left[\mathrm{A}\right] \tag{8.4.48}$$

where

$$k_{\mathrm{obs}} = \frac{k_1 k_2\left[\mathrm{M}\right]}{k_2 + k_{-1}\left[\mathrm{M}\right]} \tag{8.4.49}$$

At high pressures, we can expect collisions to occur frequently, such that $k_{-1}\left[\mathrm{M}\right] \gg k_2$. Equation 21.33 then becomes

$$-\frac{d\left[\mathrm{A}\right]}{dt} = \frac{k_1 k_2}{k_{-1}}\left[\mathrm{A}\right] \tag{8.4.50}$$

which follows 1^{st} order rate kinetics.

At low pressures, we can expect collisions to occurs infrequently, such that $k_{-1}\left[\mathrm{M}\right] \ll k_2$. In this scenario, equation 21.33 becomes

$$-\frac{d\left[\mathrm{A}\right]}{dt} = k_1\left[\mathrm{A}\right]\left[\mathrm{M}\right] \tag{8.4.51}$$

which follows second order rate kinetics, consistent with experimental observations.

Equilibrium Approximations

Consider again the following consecutive reaction in which the first step is reversible:

$$\mathrm{A} \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} \mathrm{I} \overset{k_2}{\rightarrow} \mathrm{P} \tag{8.4.52}$$

Now let us consider the situation in which $k_2 \ll k_1$ and k_{-1}. In other words, the conversion of I to P is slow and is the *rate-limiting step*. In this situation, we can assume that $[\mathrm{A}]$ and $[\mathrm{I}]$ are in equilibrium with each other. As we derived before for a reversible reaction in equilibrium,

$$K_{\mathrm{eq}} = \frac{k_1}{k_{-1}} \approx \frac{[\mathrm{I}]}{[\mathrm{A}]} \tag{8.4.53}$$

or, in terms of $[\mathrm{I}]$,

$$[\mathrm{I}] = K_{\mathrm{eq}}\left[\mathrm{A}\right] \tag{8.4.54}$$

These conditions also result from the exact solution when we set $k_2 \approx 0$. When this is done, we have the approximate expressions from the exact solution:

$$K \approx k_{-1} \tag{8.4.55}$$

$$\lambda \approx \sqrt{\left(k_1 - k_{-1}\right)^2 + 4k_1k_{-1}} = \sqrt{k_1^2 + 2k_1k_{-1} + k_{-1}^2} = k_1 + k_{-1} \tag{8.4.56}$$

$$\lambda - k_1 + K \approx k_1 + k_{-1} + k_1 - k_{-1} = 2k_1 \tag{8.4.57}$$

$$\lambda + k_1 - K \approx k_1 + k_{-1} + k_1 - k_{-1} = 2k_1 \tag{8.4.58}$$

$$k_1 + K - \lambda \approx k_1 + k_{-1} - k_1 - k_{-1} = 0 \tag{8.4.59}$$

$$k_1 + K + \lambda \approx k_1 + k_{-1} + k_1 + k_{-1} = 2\left(k_1 + k_{-1}\right) \tag{8.4.60}$$

and the approximate solutions become

$$[\mathrm{A}]\left(t\right) = \frac{[\mathrm{A}]_0}{2\left(k_1 + k_{-1}\right)}\left[2k_{-1} + 2k_1 e^{-\left(k_1 + k_{-1}\right)t}\right] \tag{8.4.61}$$

$$[\mathrm{I}]\left(t\right) = \frac{k_1[\mathrm{A}]_0}{\left(k_1 + k_{-1}\right)}\left[1 - e^{-\left(k_1 + k_{-1}\right)t}\right] \tag{8.4.62}$$

In the long-time limit, when equilibrium is reached and transient behavior has decayed away, we find

$$\frac{[\mathrm{I}]}{[\mathrm{A}]} \equiv K_{\mathrm{eq}} \rightarrow \frac{k_1}{k_{-1}} \tag{8.4.63}$$

Plugging the above equation into the expression for $d\left[\mathrm{P}\right]/dt$,

$$\frac{d\left[\mathrm{P}\right]}{dt} = k_2\left[\mathrm{I}\right] = k_2 K_{\mathrm{eq}}\left[\mathrm{A}\right] = \frac{k_1 k_2}{k_{-1}}[\mathrm{A}] \tag{8.4.64}$$

The reaction can thus be approximated as a 1^{st} order reaction

$$\mathrm{A} \xrightarrow{k'} \mathrm{P} \tag{8.4.65}$$

with

$$k' = \frac{k_1 k_2}{k_{-1}} \tag{8.4.66}$$

Figure 8.4.4 displays the concentration profiles for species, A, I, and P with the condition that $k_2 \ll k_1 = k_{-1}$. When $k_1 = k_{-1}$, we expect $[\mathrm{A}] = [\mathrm{I}]$. As can be seen from the figure, after a short initial startup time, the concentrations of species A and I are approximately equal during the reaction.

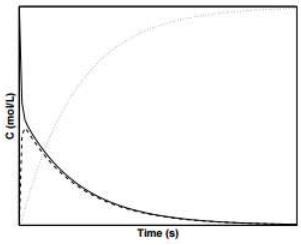


Figure 8.4.4: Plots of $[\mathrm{A}]$ (solid line), $[\mathrm{I}]$ (dashed line) and $[\mathrm{P}]$ (dotted line) over time for $k_2 \ll k_1 = k_{-1}$.

8.5: The Effect of Temperature on Reaction Rates

When molecules collide, the kinetic energy of the molecules can be used to stretch, bend, and ultimately break bonds, leading to chemical reactions. If molecules move too slowly with little kinetic energy, or collide with improper orientation, they do not react and simply bounce off each other. However, if the molecules are moving fast enough with a proper collision orientation, such that the kinetic energy upon collision is greater than the minimum energy barrier, then a reaction occurs. The minimum energy requirement that must be met for a chemical reaction to occur is called the activation energy, E_a.

Figure 8.5.1: In Greek mythology Sisyphus was punished by being forced roll an immense boulder up a hill, only to watch it roll back down, and to repeat this action forever. If this were a chemical reaction, then it would never be observed, since the reactants must overcome the energy barrier to get to the other side (products).

The reaction pathway is similar to what happens in Figure 8.5.1. To get to the other end of the road, an object must roll with enough speed to completely roll over the hill of a certain height. The faster the object moves, the more kinetic energy it has. If the object moves too slowly, it does not have enough kinetic energy necessary to overcome the barrier; as a result, it eventually rolls back down. In the same way, there is a minimum amount of energy needed in order for molecules to break existing bonds during a chemical reaction. If the kinetic energy of the molecules upon collision is greater than this minimum energy, then bond breaking and forming occur, forming a new product (provided that the molecules collide with the proper orientation).

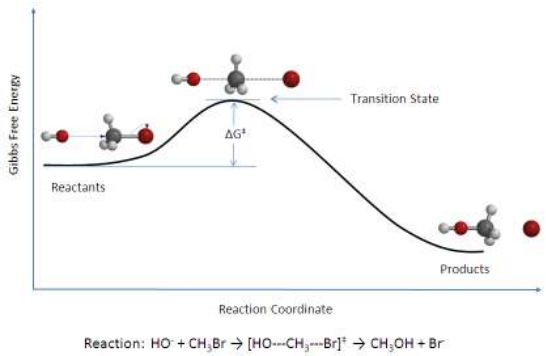


Figure 8.5.2: Reaction coordinate diagram for the bimolecular nucleophilic substitution (S_N2) reaction between bromomethane and the hydroxide anion. from Wikipedia.

The activation energy (E_a), labeled $\Delta G^{\ddagger}$ in Figure 8.5.2, is the energy difference between the reactants and the activated complex, also known as transition state. In a chemical reaction, the transition state is defined as the highest-energy state of the system. If the molecules in the reactants collide with enough kinetic energy and this energy is higher than the transition state energy, then the reaction occurs and products form. In other words, the higher the activation energy, the harder it is for a reaction to occur and vice versa.

Overcoming the energy barrier from thermal energy involves addressing the fraction of the molecules that possess enough kinetic energy to react at a given temperature. According to kinetic molecular theory, a population of molecules at a given temperature is distributed over a variety of kinetic energies that is described by the Maxwell-Boltzman distribution law.

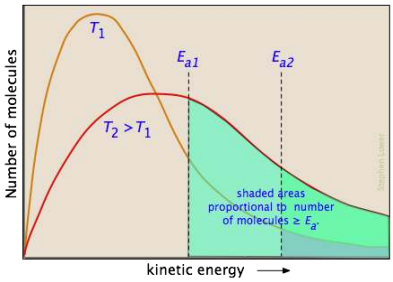
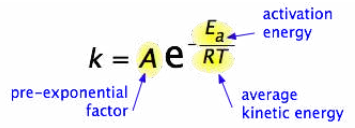


Figure 8.5.3*: Kinetic energy distributions (similar to Maxwell-Boltzman distributions for velocity) for a gas at two temperatures and critical energies for overcoming an activation barrier.*

The two distribution plots shown here are for a lower temperature T_1 and a higher temperature T_2. The area under each curve represents the total number of molecules whose energies fall within particular range. The shaded regions indicate the number of molecules which are sufficiently energetic to meet the requirements dictated by the two values of E_a that are shown. It is clear from these plots that the fraction of molecules whose kinetic energy exceeds the activation energy increases quite rapidly as the temperature is raised. This the reason that virtually all chemical reactions (and all elementary reactions) proceed more rapidly at higher temperatures.

Arrhenius Equation

By 1890 it was common knowledge that higher temperatures speed up reactions, often doubling the rate for a 10-degree rise, but the reasons for this were not clear. Finally, in 1899, the Swedish chemist Svante Arrhenius (1859-1927) combined the concepts of activation energy and the Boltzmann distribution law into one of the most important relationships in physical chemistry:



Take a moment to focus on the meaning of this equation, neglecting the *A* factor for the time being. First, note that this is another form of the exponential decay law discussed in the previous section of this series. What is "decaying" here is not the concentration of a reactant as a function of time, but the magnitude of the rate constant as a function of the exponent $-E_a/RT$. And what is the significance of this quantity? Recalling that *RT* is the *average kinetic energy*, it becomes apparent that the exponent is just the ratio of the activation energy E_a to the average kinetic energy. The larger this ratio, the smaller the rate (hence the negative sign). This means that high temperature and low activation energy favor larger rate constants, and thus speed up the reaction. Because these terms occur in an exponent, their effects on the rate are quite substantial.

Svante August Arrhenius

Svante August Arrhenius (19 February 1859 – 2 October 1927) was a Swedish scientist, originally a physicist, but often referred to as a chemist, and one of the founders of the science of physical chemistry. He received the Nobel Prize for Chemistry in 1903, becoming the first Swedish Nobel laureate, and in 1905 became director of the Nobel Institute where he remained until his death. The Arrhenius equation, Arrhenius definition of an acid, lunar crater Arrhenius, the mountain of Arrheniusfjellet and the Arrhenius Labs at Stockholm University are named after him. Today, Arrhenius is best known for his study published in 1896, on the greenhouse effect.

The two plots in Figure 8.5.4 show the effects of the activation energy (denoted here by $E^{\ddagger}$) on the rate constant. Even a modest activation energy of 50 kJ/mol reduces the rate by a factor of 10^8.

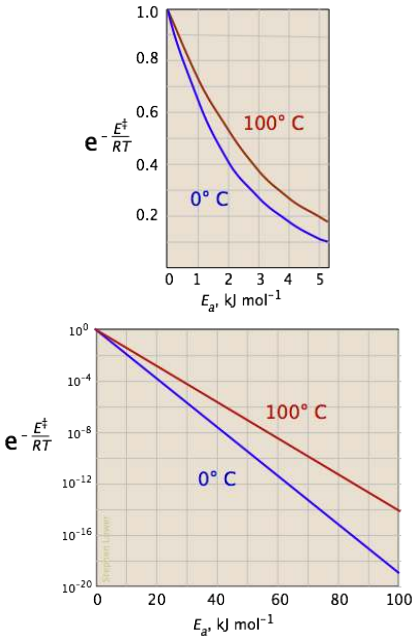


Figure 8.5.4: Arrhenius plots. The logarithmic scale in the right-hand plot leads to nice straight lines.

Looking at the role of temperature, a similar effect is observed. (If the *x*-axis were in "kilodegrees" the slopes would be more comparable in magnitude with those of the kilojoule plot at the above right.)

Determining the activation energy

The Arrhenius equation

$$k = Ae^{-E_a/RT} \tag{8.5.1}$$

can be written in a non-exponential form that is often more convenient to use and to interpret graphically (Figure 8.5.4). Taking the logarithms of both sides and separating the exponential and pre-exponential terms yields

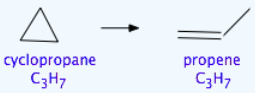
$$\ln k = \ln\left(Ae^{-E_a/RT}\right) = \ln A + \ln\left(e^{-E_a/RT}\right) \tag{8.5.2}$$

$$\ln k = \ln A + \frac{-E_a}{RT} = \left(\frac{-E_a}{R}\right)\left(\frac{1}{T}\right) + \ln A \tag{8.5.3}$$

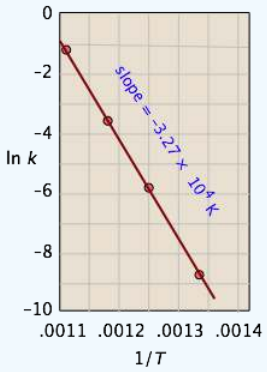
which is the equation of a straight line whose slope is $-E_a\ /R$. This affords a simple way of determining the activation energy from values of *k* observed at different temperatures, by plotting $\ln k$ as a function of $1/T$.

Example 8.5.1: Isomerization of Cyclopropane

For the isomerization of cyclopropane to propene,



the following data were obtained (calculated values shaded in pink):



T, °C	477	523	577	623
$1/T$, $K^{-1} \times 10^3$	1.33	1.25	1.18	1.11
k, s^{-1}	0.00018	0.0027	0.030	0.26
ln k	–8.62	–5.92	–3.51	–1.35

From the calculated slope, we have

$-(E_a/R) = -3.27 \times 10^4$ K

$E_a = -(8.314 \text{ J mol}^{-1} \text{ K}^{-1})(-3.27 \times 10^4 \text{ K}) = 273 \text{ kJ mol}^{-1}$

This activation energy is high, which is not surprising because a carbon-carbon bond must be broken in order to open the cyclopropane ring. (C–C bond energies are typically around 350 kJ/mol.) This is why the reaction must be carried out at high temperature.

Calculating E_a without a plot

Because the ln k vs.-1/T plot yields a straight line, it is often convenient to estimate the activation energy from experiments at only two temperatures. The ln A term is eliminated by subtracting the expressions for the two ln-k terms via the following steps:

$$\ln k_1 = \ln A - \frac{E_a}{k_B T_1} \tag{8.5.4}$$

at T_1 and

$$\ln k_2 = \ln A - \frac{E_a}{k_B T_2} \tag{8.5.5}$$

at T_2. By rewriting the second equation:

$$\ln A = \ln k_2 + \frac{E_a}{k_B T_2} \tag{8.5.6}$$

and substitute for $\ln A$ into the first equation:

$$\ln k_1 = \ln k_2 + \frac{E_a}{k_B T_2} - \frac{E_a}{k_B T_1} \tag{8.5.7}$$

This simplifies to:

$$\ln k_1 - \ln k_2 = -\frac{E_a}{k_B T_1} + \frac{E_a}{k_B T_2} \tag{8.5.8}$$

$$\ln \frac{k_1}{k_2} = -\frac{E_a}{k_B}\left(\frac{1}{T_1} - \frac{1}{T_2}\right) \tag{8.5.9}$$

Example 8.5.2

A widely used rule-of-thumb for the temperature dependence of a reaction rate is that a **ten degree** rise in the temperature approximately doubles the rate. This is not generally true, especially when a strong covalent bond must be broken. For a reaction that does show this behavior, what would the activation energy be?

Solution

Center the ten degree interval at 300 K. Substituting into the above expression yields

$$E_a = \frac{(8.314)(\ln 2/1)}{\frac{1}{295} - \frac{1}{305}} = \frac{(8.314)(0.693)}{0.00339 K^{-1} - 0.00328\ K^{-1}} \tag{8.5.10}$$

= (5.76 J mol^{-1} K^{-1}) / (0.00011 K^{-1}) = 52400 J mol^{-1} = **52.4 kJ mol^{-1}**

Example 8.5.3

It takes about 3.0 minutes to cook a hard-boiled egg in Los Angeles, but at the higher altitude of Denver, where water boils at 92°C, the cooking time is 4.5 minutes. Use this information to estimate the activation energy for the coagulation of egg albumin protein.

Solution

The ratio of the rate constants at the elevations of Los Angeles and Denver is 4.5/3.0 = 1.5, and the respective temperatures are 373 K and 365 K. With the subscripts 2 and 1 referring to Los Angeles and Denver respectively:

$$E_a = \frac{(8.314)(\ln 1.5)}{\frac{1}{365\ \text{K}} - \frac{1}{373\ \text{K}}} = \frac{(8.314)(0.405)}{0.00274\ \text{K}^{-1} - 0.00268\ \text{K}^{-1}} \tag{8.5.11}$$

$$= \frac{(3.37\ \text{J mol}^{-1}\text{K}^{-1})}{5.87 \times 10^{-5}\ \text{K}^{-1}} = 5740\ \text{J mol}^{-1} = 5.73\ \text{kJ mol}^{-1} \tag{8.5.12}$$

Comment: This low value seems reasonable because thermal denaturation of proteins primarily involves the disruption of relatively weak hydrogen bonds; no covalent bonds are broken (although disulfide bonds can interfere with this interpretation).

The pre-exponential factor

Up to this point, the pre-exponential term, A in the Arrhenius equation, has been ignored because it is not directly involved in relating temperature and activation energy, which is the main practical use of the equation.

$$k = Ae^{-E_a/RT}$$

However, because A multiplies the exponential term, its value clearly contributes to the value of the rate constant and thus of the rate. Recall that the exponential part of the Arrhenius equation expresses the fraction of reactant molecules that possess enough kinetic energy to react, as governed by the Maxwell-Boltzmann law. This fraction can run from zero to nearly unity, depending on the magnitudes of E_a and of the temperature.

If this fraction were 0, the Arrhenius law would reduce to

$$k = A \tag{8.5.13}$$

In other words, A is the fraction of molecules that would react if either the activation energy were zero, or if the kinetic energy of all molecules exceeded E_a — admittedly, an uncommon scenario (although barrierless reactions have been characterized).

The Role of Collisions

What would limit the rate constant if there were no activation energy requirements? The most obvious factor would be the rate at which reactant molecules come into contact. This can be calculated from kinetic molecular theory and is known as the *frequency-* or *collision factor*, Z.

$$k = A e^{-\frac{E_a}{RT}}$$

$A = Z \times \rho$ (frequency factor, steric factor)

In some reactions, the relative orientation of the molecules at the point of collision is important, so a geometrical or *steric factor* (commonly denoted by ρ (Greek lower case *rho*) can be defined. In general, we can express A as the product of these two factors:

$$A = Z\rho \tag{8.5.14}$$

Values of ρ are generally very difficult to assess; they are sometime estimated by comparing the observed rate constant with the one in which A is assumed to be the same as Z. Usually, the more complex the reactant molecules, the lower the steric factors. The deviation from unity can have different causes: the molecules are not spherical, so different geometries are possible; not all the kinetic energy is delivered into the right spot; the presence of a solvent (when applied to solutions) and other factors (Figure 8.5.4).

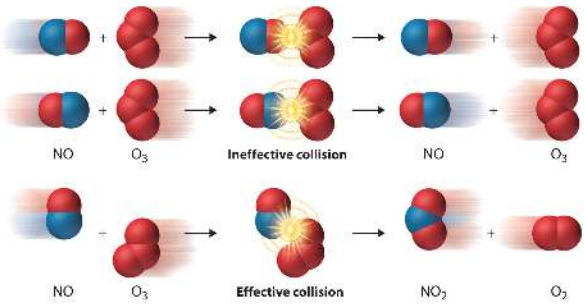


Figure 8.5.5: The Effect of Molecular Orientation on the Reaction of NO and O_3. *Most collisions of NO and O_3 molecules occur with an incorrect orientation for a reaction to occur. Only those collisions in which the N atom of NO collides with one of the terminal O atoms of O_3 are likely to produce NO_2 and O_2, even if the molecules collide with $E > E_a$.*

Contributors

- Stephen Lower, Professor Emeritus (Simon Fraser U.) Chem1 Virtual Textbook
- Wikipedia

8.6: Potential Energy Surfaces

A potential energy surface (PES) describes the potential energy of a system, especially a collection of atoms, in terms of certain parameters, normally the positions of the atoms. The surface might define the energy as a function of one or more coordinates; if there is only one coordinate, the surface is called a potential energy curve or energy profile. It is helpful to use the analogy of a landscape: for a system with two degrees of freedom (e.g. two bond lengths), the value of the energy (analogy: the height of the land) is a function of two bond lengths (analogy: the coordinates of the position on the ground). The Potential Energy Surface represents the concepts that each geometry (both external and internal) of the atoms of the molecules in a chemical reaction is associated with it a unique potential energy. This creates a smooth energy “landscape” and chemistry can be viewed from a topology perspective (of particles evolving over "valleys""and passes").

Potential Energy Curves (1-D Potential Energy Surfaces)

The PES is the energy of a molecule as a function of the positions of its nuclei r. This energy of a system of two atoms depends on the distance between them. At large distances the energy is zero, meaning “no interaction”. At distances of several atomic diameters attractive forces dominate, whereas at very close approaches the force is repulsive, causing the energy to rise. The attractive and repulsive effects are balanced at the minimum point in the curve. Plots that illustrate this relationship are quite useful in defining certain properties of a chemical bond.

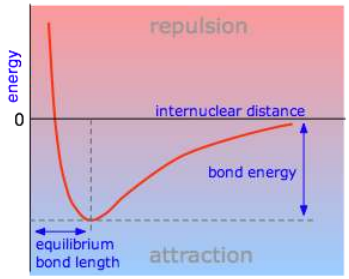


Figure 8.6.1: A potential Energy Curve for a covalent bond.

The internuclear distance at which the potential energy minimum occurs defines the **bond length**. This is more correctly known as the *equilibrium* bond length, because thermal motion causes the two atoms to vibrate about this distance. In general, the stronger the bond, the smaller will be the bond length.

Attractive forces operate between all atoms, but unless the potential energy minimum is at least of the order of *RT*, the two atoms will not be able to withstand the disruptive influence of thermal energy long enough to result in an identifiable molecule. Thus we can say that a chemical bond exists between the two atoms in H_2. The weak attraction between argon atoms does not allow Ar_2 to exist as a molecule, but it does give rise to the *van Der Waals force* that holds argon atoms together in its liquid and solid forms.

Potential, Kinetic, and Total Energy for a System

Potential energy and kinetic energy Quantum theory tells us that an electron in an atom possesses kinetic energy K as well as potential energy V, so the total energy E is always the sum of the two: $E = V + K$. The relation between them is surprisingly simple: $K = -0.5V$. This means that when a chemical bond forms (an exothermic process with $\Delta E < 0$), the decrease in potential energy is accompanied by an increase in the kinetic energy (embodied in the momentum of the bonding electrons), but the magnitude of the latter change is only half as much, so the change in potential energy always dominates. The bond energy $-\Delta E$ has half the magnitude of the fall in potential energy.

Mathematical definition and computation

The geometry of a set of atoms can be described by a vector, r, whose elements represent the atom positions. The vector r could be the set of the Cartesian coordinates of the atoms, or could also be a set of inter-atomic distances and angles. Given r, the energy as a function of the positions, $V(r)$, is the value of $V(r)$ for all values of r of interest. Using the landscape analogy from the introduction, $V(r)$ gives the height on the "energy landscape" so that the concept of a potential energy surface arises. An example is the PES for water molecule (Figure 8.6.1) that show the energy minimum corresponding to optimized molecular structure for water- O-H bond length of 0.0958 nm and H-O-H bond angle of 104.5°

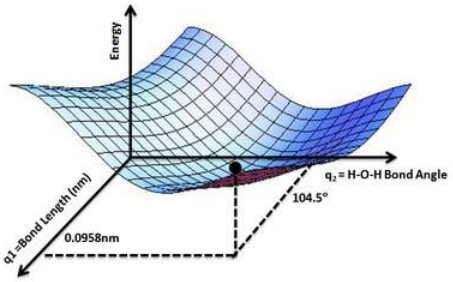


***Figure* 8.6.2**: *PES for water molecule: Shows the energy minimum corresponding to optimized molecular structure for water- O-H bond length of 0.0958nm and H-O-H bond angle of 104.5°. of Wikipedia (Credit: Aimnature).*

The Dimensionality of a Potential Energy Surface

To define an atom's location in 3-dimensional space requires three coordinates (e.g., x, y,and z or r, θ and phi in Cartesian and Spherical coordinates) or *degrees of freedom*. However, a reaction and hence the corresponding PESs do not depend of the absolute position of the reaction, only the relative positions (internal degrees). Hence both translation and rotation of the entire system can be removed (each with 3 degree of freedom, assuming non-linear geometries). So the dimensionality of a PES is

$$3N-6 \tag{8.6.1}$$

where N is the number of atoms involves in the reaction, i.e., the number of atoms in each reactants). The PES is a hypersurface with many degrees of freedom and typically only a few are plotted at any one time for understanding. See Calculate Number of Vibrational Modes to get a more details picture of how this applies to calculating the number of vibrations in a molecule

To study a chemical reaction using the PES as a function of atomic positions, it is necessary to calculate the energy for **every atomic** arrangement of interest. Methods of calculating the energy of a particular atomic arrangement of atoms are well described in the computational chemistry article, and the emphasis here will be on finding approximations of $(V(r)$ to yield fine-grained energy-position information.

For very simple chemical systems or when simplifying approximations are made about inter-atomic interactions, it is sometimes possible to use an analytically derived expression for the energy as a function of the atomic positions. An example is

$$H+H_2 \rightarrow H_2+H \tag{8.6.2}$$

system as a function of the **three** H-H distances. For more complicated systems, calculation of the energy of a particular arrangement of atoms is often too computationally expensive for large scale representations of the surface to be feasible.

Application of Potential Energy Surfaces

A PES is a conceptual tool for aiding the analysis of molecular geometry and chemical reaction dynamics. Once the necessary points are evaluated on a PES, the points can be classified according to the first and second derivatives of the energy with respect to position, which respectively are the gradient and the curvature. Stationary points (or points with a zero gradient) have physical meaning: energy minima correspond to physically stable chemical species and saddle points correspond to transition states, the highest energy point on the reaction coordinate (which is the lowest energy pathway connecting a chemical reactant to a chemical product). Three

- PES do not show kinetic energy, only potential energy.
- At T = 0 K (no KE), species will want to be at the lowest possible potential energy, (i.e., at a minimum on the PES).
- Between any two minima (valley bottoms) the lowest energy path will pass through a maximum at a **saddle point**, which we call that saddle point a transition-state structure.

The PES concept finds application in fields such as chemistry and physics, especially in the theoretical sub-branches of these subjects. It can be used to theoretically explore properties of structures composed of atoms, for example, finding the minimum energy shape of a molecule or computing the rates of a chemical reaction.

Contributors

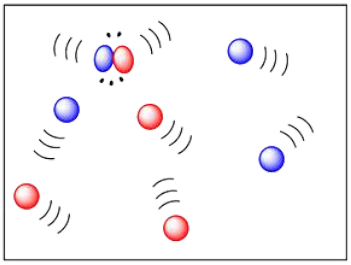
- Wikipedia
- Stephen Lower, Professor Emeritus (Simon Fraser U.) Chem1 Virtual Textbook

8.7: Theories of Reaction Rates

The macroscopic discussion of kinetics discussed in previous sections can be now expanded into a more microscopic picture in terms of molecular level properties (e..g, mass and velocities) involving two important theories: (1) collision theory and (2) transition-state theory.

Collision Theory

If two molecules need to collide in order for a reaction to take place, then factors that influence the ease of collisions will be important. The more energy there is available to the molecules, the faster they will move around, and the more likely they are to bump into each other. Higher temperatures ought to lead to more collisions and a greater frequency of reactions between molecules. In the drawing below, the cold, sluggish molecules on the left are not likely to collide, but the energetic molecules on the right are due to collide at any time.



The rate at which molecules collide which is the frequency of collisions is called the collision frequency, Z, which has units of collisions per unit of time. Given a container of molecules A and B, the collision frequency between A and B is defined by:

$$Z = N_A N_B \sigma_{AB} \sqrt{\frac{8k_B T}{\pi \mu_{AB}}} \tag{8.7.1}$$

where:

- N_A and N_B are the numbers of molecules A and B, and is directly related to the concentrations of A and B.
- The mean speed of molecules obtained from the Maxwell-Boltzmann distribution for thermalized gases

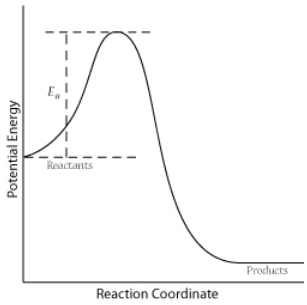
$$\sqrt{\frac{8k_B T}{\pi \mu_{AB}}} \tag{8.7.2}$$

- σ_{AB} is the averaged sum of the collision cross-sections of molecules A and B. The collision cross section represents the collision region presented by one molecule to another.
- μ is the reduced mass and is given by

$$\mu = \frac{m_A m_B}{m_A + m_B} \tag{8.7.3}$$

The concepts of collision frequency can be applied in the laboratory: (1) The temperature of the environment affects the average speed of molecules. Thus, reactions are heated to increase the reaction rate. (2) The initial concentration of reactants is directly proportional to the collision frequency; increasing the initial concentration will speed up the reaction.

For a successful collision to occur, the reactant molecules must collide with enough kinetic energy to break original bonds and form new bonds to become the product molecules. This energy is called the activation energy for the reaction; it is also often referred to as the energy barrier.



The fraction of collisions with enough energy to overcome the activation barrier is given by:

$$f = e^{\frac{-E_a}{RT}} \tag{8.7.4}$$

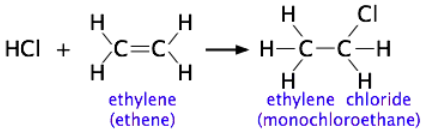
where:

- f is the fraction of collisions with enough energy to react
- E_a is the activation energy

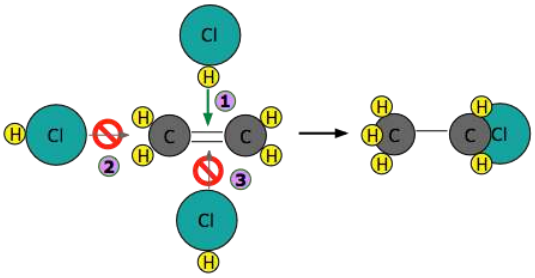
The fraction of successful collisions is directly proportional to the temperature and inversely proportional to the activation energy.

The fraction of successful collisions is directly proportional to the temperature and inversely proportional to the activation energy.

The more complicated the structures of the reactants, the more likely that the value of the rate constant will depend on the trajectories at which the reactants approach each other. This kind of *electrophilic addition reaction* is well-known to all students of organic chemistry. Consider the addition of a hydrogen halide such as HCl to the double bond of an alkene, converting it to a chloroalkane.



Experiments have shown that the reaction only takes place when the HCl molecule approaches the alkene with its hydrogen-end, and in a direction that is approximately perpendicular to the double bond, as shown at (1) below.



The reason for this becomes apparent when we recall that HCl is highly polar owing to the high electronegativity of chlorine, so that the hydrogen end of the molecule is slightly positive. The steric factor, ρ is then introduced to represent is the probability of the reactant molecules colliding with the right orientation and positioning to achieve a product with the desirable geometry and stereospecificity. Values of ρ are generally very difficult to assess and range from 0 to 1, but are sometime estimated by comparing the observed rate constant with the one in which the preexponential constant A is assumed to be the same as Z.

The lesson you should take from this example is that once you start combining a variety of chemical principles, you gradually develop what might be called "chemical intuition" which you can apply to a wide variety of problems. This is far more important than memorizing specific examples.

All Three Factors Combined

The rate constant of the gas-phase reaction is proportional to the product of the *collision frequency* and the *fraction of successful reactions*. As stated above, sufficient kinetic energy is required for a successful reaction; however, they must also collide properly. Compare the following equation to the Arrhenius equation:

$$k = Z\rho e^{\frac{-E_a}{RT}} \tag{8.7.5}$$

where

- k is the rate constant for the reaction
- ρ is the steric factor.
- Zρ is the pre-exponential factor, A, of the Arrhenius equation. It is the frequency of total collisions that collide with the right orientation. In practice, it is the pre-exponential factor that is directly determined by experiment and then used to calculate the steric factor.
- E_a is activation energy
- T is absolute temperature
- R is gas constant.

Although the collision theory deals with gas-phase reactions, its concepts can also be applied to reactions that take place in solvents; however, the properties of the solvents (for example: solvent cage) will affect the rate of reactions. Ultimately, collision theory illustrates how reactions occur; it can be used to approximate the rate constants of reactions, and its concepts can be directly applied in the laboratory. Read this for a more detailed discussion of Collision Theory.

Transition-State Theory

Transition state theory (TST) provides a more accurate alternative to the previously used Arrhenius equation and the collision theory. The transition state theory attempts to provide a greater understanding of activation energy, E_a, and the thermodynamic properties involving the transition state. Collision theory of reaction rate, although intuitive, lacks an accurate method to predict the probability factor for the reaction. The theory assumes that reactants are hard spheres rather than molecules with specific structures. In 1935, Henry Eyring helped develop a new theory called the transition state theory to provide a more accurate alternative to the previously used Arrhenius equation and the collision theory. The Eyring equation involves the statistical frequency factory, v, which is fundamental to the theory.

According to TST, between the state where molecules are reactants and the state where molecules are products, there is a state known as the transition state. In the transition state, the reactants are combined in a species called the activated complex. The theory suggests that there are three major factors that determine whether a reaction will occur:

1. The concentration of the activated complex
2. The rate at which the activated complex breaks apart
3. The way in which the activated complex breaks apart: whether it breaks apart to reform the reactants or whether it breaks apart to form a new complex, the products.

Collision theory proposes that not all reactants that combine undergo a reaction. However, assuming the stipulations of the collision theory are met and a successful collision occurs between the molecules, transition state theory allows one of two outcomes: a return to the reactants, or a rearranging of bonds to form the products.

Consider a bimolecular reaction:

$$A + B \rightarrow C \tag{8.7.6}$$

$$K = \frac{[C]}{[A][B]} \tag{8.7.7}$$

where K is the equilibrium constant. In the transition state model, the activated complex AB is formed:

$$A + B \rightleftharpoons AB^{\ddagger} \rightarrow C \tag{8.7.8}$$

$$K^{\ddagger} = \frac{[AB]^{\ddagger}}{[A][B]} \tag{8.7.9}$$

There is an energy barrier, called activation energy, in the reaction pathway. A certain amount of energy is required for the reaction to occur. The transition state, $AB^{\ddagger}$, is formed at maximum energy. This high-energy complex represents an unstable intermediate. Once the energy barrier is overcome, the reaction is able to proceed and product formation occurs.

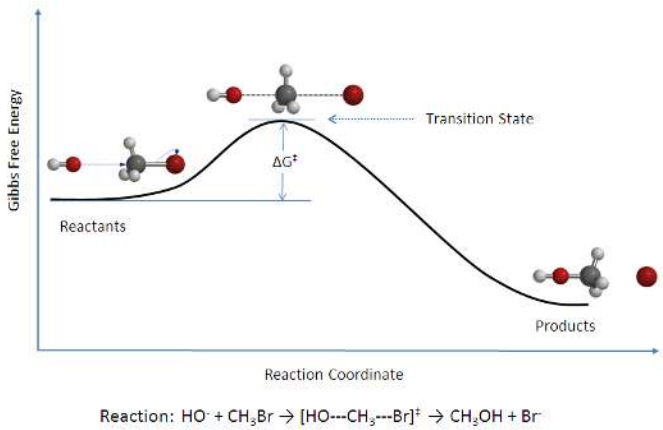


Figure 8.7.1: Reaction coordinate diagram for the bimolecular nucleophilic substitution (S_N2) reaction between bromomethane and the hydroxide anion. form Wikipedia.

The rate of a reaction is equal to the number of activated complexes decomposing to form products. Hence, it is the concentration of the high-energy complex multiplied by the frequency of it surmounting the barrier.

$$rate = v[AB^{\ddagger}] \tag{8.7.10}$$
$$= v[A][B]K^{\ddagger} \tag{8.7.11}$$

The rate can be rewritten:

$$rate = k[A][B] \tag{8.7.12}$$

Combining Equations 8.7.12 and 8.7.11 gives:

$$k[A][B] = v[A][B]K^{\ddagger} \tag{8.7.13}$$
$$k = vK^{\ddagger} \tag{8.7.14}$$

where

- v is the frequency of vibration,
- k is the rate constant and
- $K^{\ddagger}$ is the thermodynamic equilibrium constant.

Statistical mechanics (not shown) provides that the frequency, v, is equivalent to the thermal energy, k_BT, divided by Planck's constant, h.

$$v = \frac{k_BT}{h} \tag{8.7.15}$$

where

- k_B is the Boltzmann's constant (1.381 x 10^{-23} J/K),
- T is the absolute temperature in Kelvin (K) and
- h is Planck's constant (6.626 x 10^{-34} Js).

Substituting Equation 8.7.15 into Equation 8.7.14:

$$k = \frac{k_BT}{h}K^{\ddagger} \tag{8.7.16}$$

Equation ref is often tagged with another term (M^{1-m}) that makes the units equal with M is the molarity and m is the molecularly of the reaction.

$$k = \frac{k_BT}{h}K^{\ddagger}(M^{1-m}) \tag{8.7.17}$$

It is important to note here that the equilibrium constant $K^{\ddagger}$ can be calculated by absolute, fundamental properties such as bond length, atomic mass, and vibration frequency. This gives the transition rate theory the alternative name absolute rate theory, because the rate constant, k, can be calculated from fundamental properties.

Thermodynamics of Transition State Theory

To reveal the thermodynamics of the theory, $K^{\ddagger}$ must be expressed in terms of $\Delta G^{\ddagger}$. $\Delta G^{\ddagger}$ is simply,

$$\Delta G^{o\ddagger} = G^{o}(transitionstate) - G^{o}(reactants) \tag{8.7.18}$$

By definition, at equilibrium, $\Delta G^{\ddagger}$ can be expressed as:

$$\Delta G^{\ddagger} = -RT \ln K^{\ddagger} \tag{8.7.19}$$

Rearrangement gives:

$$[K]^{\ddagger} = e^{-\frac{\Delta G^{\ddagger}}{RT}} \tag{8.7.20}$$

From Equation 8.7.17

$$k = ve^{-\frac{\Delta G^{\ddagger}}{RT}}(M^{1-m}) \tag{8.7.21}$$

It is also possible to obtain terms for the change in enthalpy and entropy for the transition state. Because

$$\Delta G^{\ddagger} = \Delta H^{\ddagger} - T\Delta S^{\ddagger} \tag{8.7.22}$$

it follows that the derived equation becomes,

$$k = \frac{k_B T}{h} e^{\Delta S^{\ddagger}/R} e^{-\Delta H^{\ddagger}/RT} M^{1-m} \tag{8.7.23}$$

Equation 8.7.23 is known as the Eyring Equation and was developed by Henry Eyring in 1935, is based on transition state theory and is used to describe the relationship between reaction rate and temperature. It is similar to the Arrhenius Equation, which also describes the temperature dependence of reaction rates.

The linear form of the Eyring Equation is given below:

$$\ln \frac{k}{T} = \frac{-\Delta H^{\dagger}}{R}\frac{1}{T} + \ln \frac{k_B}{h} + \frac{\Delta S^{\ddagger}}{R} \tag{8.7.24}$$

The values for $\Delta H^{\ddagger}$ and $\Delta S^{\ddagger}$ can be determined from kinetic data obtained from a $\ln \frac{k}{T}$ vs. $\frac{1}{T}$ plot. The Equation is a straight line with negative slope, $\frac{-\Delta H^{\ddagger}}{R}$, and a y-intercept, $\frac{\Delta S^{\ddagger}}{R} + \ln \frac{k_B}{h}$.

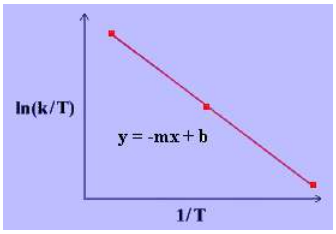


Figure 8.7.2: Linearized TST theory

Conclusion

In this article, the complete thermodynamic formulation of the transition state theory was derived. This equation is more reliable than either the Arrhenius equation and the equation for the Collision Theory. However, it has its limitations, especially when considering the concepts of quantum mechanics. Quantum mechanics implies that tunneling can occur, such that particles can bypass the energy barrier created by the transition state. This can especially occur with low activation energies, because the probability of tunneling increases when the barrier height is lowered.

In addition, transition state theory assumes that an equilibrium exists between the reactants and the transition state phase. However, in solution non-equilibrium situations can arise, upsetting the theory. Several more complex theories have been presented to correct for these and other discrepancies. This theory still remains largely useful in calculating the thermodynamic properties of the transition state from the overall reaction rate. This presents immense usefulness in medicinal chemistry, in which the study of transition state analogs is widely implemented.

References

1. Chang, Raymond. Physical Chemistry for the Biosciences. University Science Books, California 2005
2. Truhlar, D. G.; Garrett, B. C.; Klippenstein, S. J., Current Status of Transition-State Theory. The Journal of physical chemistry 1996, 100, (31), 12771-12800
3. Laidler, K.; King, C., Development of transition-state theory. The Journal of physical chemistry 1983, 87, (15), 2657
4. Lienhard, Gustav, Enzyme Catalysis and Transition -State Theory: Transition State Analogs. Cold Spring Harb Symp Quant Biol 1972. 36:45-51

Contributors

- Phong Dao, Bilal Latif, Lu Zhao (UC Davis)

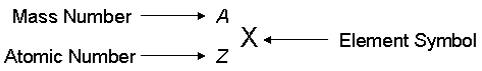
8.8: Isotope Effects in Chemical Reactions

The **kinetic isotope effect** (KIE) is a phenomenon associated with isotopically substituted molecules exhibiting different reaction rates. Isotope effects such as KIEs are invaluable tools in both physical and biological sciences and are used to aid in the understanding of reaction kinetics, mechanisms, and solvent effects.

Introduction

Research was first introduced on this topic over 50 years ago and has grown into an enormous field. The scientists behind much of the understanding and development of kinetic isotope effects were Jacob Bigeleisen and Maria Goeppert Mayer who published the first paper on isotope effects [J. Chem. Phys., 15, 261 (1947)]. Kinetic isotope effects specifically explore the change in rate of a reaction due to isotopic substitution.

An element is identified by its symbol, mass number, and atomic number. The atomic number is the number of protons in the nucleus while the mass number is the total number of protons and neutrons in the nucleus. Isotopes are two atoms of the same element that have the same number of protons but different numbers of neutrons. Isotopes are specified by the mass number.



As an example consider the two isotopes of chlorine, you can see that their mass numbers vary, with ^{35}Cl being the most abundant isotope, while their atomic numbers remain the same at 17.

$$^{35}Cl \text{ and } ^{37}Cl \tag{8.8.1}$$

The most common isotope used in light atom isotope effects is hydrogen (^{1}H) commonly replaced by its isotope deuterium (^{2}H). Note: Hydrogen also has a third isotope, tritium (^{2}H). Isotopes commonly used in heavy atom isotope effects include carbon (^{12}C, ^{13}C, nitrogen (^{14}N, ^{15}N), oxygen, sulfur, and bromine. Not all elements exhibit reasonably stable isotopes (i.e. Fluorine, ^{19}F), but those that due serve as powerful tools in isotope effects.

Potential Energy Surfaces

Understanding potential energy surfaces is important in order to be able to understand why and how isotope effects occur as they do. The harmonic oscillator approximation is used to explain the vibrations of a diatomic molecule. The energies resulting from the quantum mechanic solution for the harmonic oscillator help to define the internuclear potential energy of a diatomic molecule and are

$$E_n = \left(n + \frac{1}{2}\right) h\nu \tag{8.8.2}$$

where

- n is a positive integer (n=1,2,3...),
- h is Planck's constant and
- ν is the frequency of vibration.

The Morse potential is an analytic expression that is used as an approximation to the intermolecular potential energy curves:

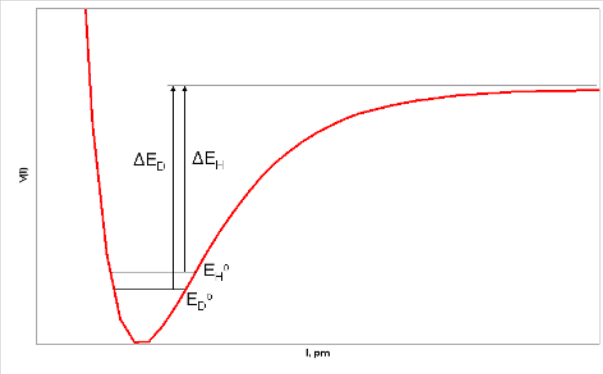
$$V(l) = D_e \left(1 - e^{-\beta(l-l_o)}\right)^2 \tag{8.8.3}$$

where

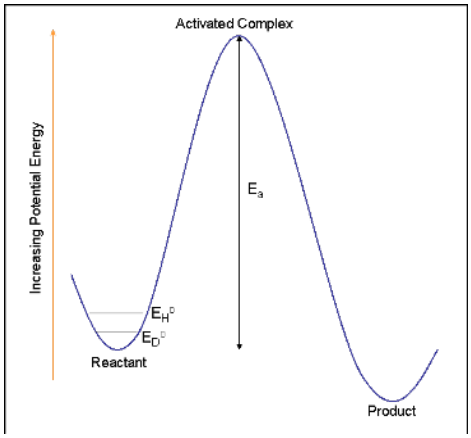
- $V(l)$ is the potential energy,
- D_e is the dissociation energy of the molecule,
- β is the measure of the curvature of the potential at its minimum,
- l is displacement, and
- l_o is the equilibrium bond length.

The D_e, β, and l_o variables can be looked up in a textbook or CRC handbook.

Below is an example of a Morse potential curve with the zero point vibrational energies of two isotopic molecules (for example R-H and R-D where R is a group/atom that is much heavier than H or D). The y-axis is potential energy and the x axis is internuclear distance. In this figure E_D^o and E_H^o correspond to the zero point energies of deuterium and hydrogen. The zero point energy is the lowest possible energy of a system and equates to the ground state energy. Zero point energy is dependent upon the reduced mass of the molecule as will be shown in the next section. The heavier the molecule or atom, the lower the frequency of vibration and the smaller the zero point energy. Lighter molecules or atoms have a greater frequency of vibration and a higher zero point energy. We see this is the figure below where deuterium is heavier than hydrogen and therefore has the lower zero point energy.



This results in different bond dissociation energies for R-D and R-H. The bond dissociation energy for R-D (E_D) is greater than the bond dissociation energy of R-H (E_H). This difference in energy due to isotopic replacement results in differing rates of reaction, the effect that is measured in kinetic isotope effects. The reaction rate for the conversion of R-D is slower than the reaction rate for the conversion of R-H.



It is important to note that isotope replacement does not change the electronic structure of the molecule or the potential energy surfaces of the reactions the molecule may undergo. Only the rate of the reaction is affected.

Activation Energies

The energy of the vibrational levels of a vibration (i.e., a bond) in a molecule is given by

$$E_n = \left(n + \frac{1}{2}\right) h\nu \tag{8.8.4}$$

where we assume that the molecule is in its ground state and we can compare zero-point vibrational energies,

$$E_o = \left(\frac{1}{2}\right) h\nu \tag{8.8.5}$$

Using the harmonic oscillator approximation the fundamental vibrational frequency is

$$\nu = \frac{1}{2\pi}\sqrt{\frac{k}{\mu}} \tag{8.8.6}$$

where

- k is the force constant of the bond and
- μ is the reduced mass

$$\mu = \frac{m_1 m_2}{m_1 + m_2} \tag{8.8.7}$$

The Arrhenius equation is used to determine reaction rates and activation energies and since we are interested in the change in rate of reactions with different isotopes, this equation is very important,

$$k = Ae^{-\frac{E_a}{kT}} \tag{8.8.8}$$

where

- k is the reaction rate,
- E_a is the activation energy, and
- A is the Arrhenius constant.

The Arrhenius equation can be used to compare the rates of a reaction with R-H and R-D,

$$k_H = A_H e^{-\frac{E_a^H}{kT}} \tag{8.8.9}$$

$$k_D = A_D e^{-\frac{E_a^D}{kT}} \tag{8.8.10}$$

where k_H and k_D are the rates of reaction associated with R-H and the isotope substituted R-D. We will then assume the Arrhenius constants are equal ($A_H = A_D$). The ratio of the rates of reaction gives an approximation for the isotope effect resulting in:

$$\frac{k_H}{k_D} = e^{-\frac{E_a^H - E_a^D}{kT}} \tag{8.8.11}$$

By using the relationship that for both R-H and R-D

$$E_o = \left(\frac{1}{2}\right) h\nu \tag{8.8.12}$$

a substitution can be made resulting in

$$\frac{k_H}{k_D} = e^{\frac{h(\nu_H - \nu_D)}{2kT}} \tag{8.8.13}$$

The vibrational frequency (Equation 5) can then be substituted for R-H and R-D and the value of the expected isotope effect can be calculated.

$$\frac{k_H}{k_D} = e^{\frac{h\left(\frac{k_{RH}}{\mu_{RH}} - \frac{k_{RD}}{\mu_{RD}}\right)}{4\pi kT}} \tag{8.8.14}$$

The same general procedure can be followed for any isotope substitution.

In summary, the greater the mass the more energy is needed to break bonds. A heavier isotope forms a stronger bond. The resulting molecule has less of a tendency to dissociate. The increase in energy needed to break the bond results in a slower reaction rate and the observed isotope effect.

Kinetic Isotope Effects

Kinetic Isotope Effects (KIEs) are used to determine reaction mechanisms by determining rate limiting steps and transition states and are commonly measured using NMR to detect isotope location or GC/MS to detect mass changes. In a KIE experiment an atom is replaced by its isotope and the change in rate of the reaction is observed. A very common isotope substitution is when hydrogen

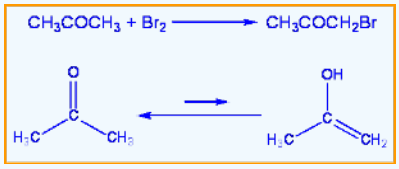
is replaced by deuterium. This is known as a deuterium effect and is expressed by the ratio k_H/k_D (as explained above). Normal KIEs for the deuterium effect are around 1 to 7 or 8. Large effects are seen because the percentage mass change between hydrogen and deuterium is great. Heavy atom isotope effects involve the substitution of carbon, oxygen, nitrogen, sulfur, and bromine, with effects that are much smaller and are usually between 1.02 and 1.10. The difference in KIE magnitude is directly related to the percentage change in mass. Large effects are seen when hydrogen is replaced with deuterium because the percentage mass change is very large (mass is being doubled) while smaller percent mass changes are present when an atom like sulfur is replaced with its isotope (increased by two mass units).

Primary KIEs

Primary kinetic isotope effects are rate changes due to isotopic substitution at a site of bond breaking in the rate determining step of a reaction.

Example

Consider the bromination of acetone: kinetic studies have been performed that show the rate of this reaction is independent of the concentration of bromine. To determine the rate determining step and mechanism of this reaction the substitution of a deuterium for a hydrogen can be made.



When hydrogen was replaced with deuterium in this reaction a $\frac{k_H}{k_D}$ of 7 was found. Therefore the rate determining step is the tautomerization of acetone and involves the breaking of a C-H bond. Since the breaking of a C-H bond is involved, a substantial isotope effect is expected.

Heavy Atom Isotope Effects

A rule of thumb for heavy atom isotope effects is that the maximum isotopic rate ratio is proportional to the square root of the inverse ratio of isotopic masses.

- Expected:

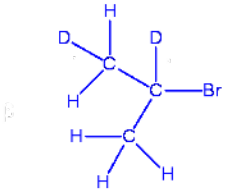
$$\frac{k_{32}}{k_{34}} = \sqrt{\frac{34}{32}} = 1.031 \tag{8.8.15}$$

- Experimental:

$$\frac{k_{32}}{k_{34}} = 1.072 \tag{8.8.16}$$

Secondary KIEs

Secondary kinetic isotope effects are rate changes due to isotopic substitutions at a site other than the bond breaking site in the rate determining step of the reaction. These come in three forms: α, β, and γ effects.



β secondary isotope effects occur when the isotope is substituted at a position next to the bond being broken.

$$(CH_3)_2CHBr + H_2O \xrightarrow{k_H} (CH_3)_2CHOH \quad (8.8.17)$$

$$(CD_3)_2CHBr + H_2O \xrightarrow{k_D} (CD_3)_2CHOH \quad (8.8.18)$$

This is thought to be due to hyperconjugation in the transition state. Hyperconjugation involves a transfer of electron density from a sigma bond to an empty p orbital (for more on hyperconjugation see outside links).

Solvent Effects in Reactions

Reactions may be affected by the type of solvent used (for example H_2O to D_2O or ROH to ROD). There are three main ways solvents effect reactions:

1. The solvent can act as a reactant resulting in a primary isotope effect.
2. Rapid hydrogen exchange can occur between substrate molecules labeled with deuterium and hydrogen atoms in the solvent. Deuterium may change positions in the molecule resulting in a new molecule that is then reacted in the rate determining step of the reaction.
3. The nature of solvent and solute interactions may also change with differing solvents. This could change the energy of the transition state and result in a secondary isotope effects.

References

1. Baldwin, J.E., Gallagher, S.S., Leber, P.A., Raghavan, A.S., Shukla, R.; J. Org. Chem. **2004**, 69, 7212-7219 (This is a great paper using kinetic isotope effects to determine a reaction mechanism. It will interest the organic chemistry oriented reader.)
2. Bigeleisen, J., Goeppert, M., J. Chem. Phys. **1947**, 15, 261.
3. Chang, R.; Physical Chemistry for the Chemical and Biological Sciences; University Science Books: Sausalito, CA, 2000, pp 480-483.
4. Isaacs, N.; Physical Organic Chemistry; John Wiley & Sons Inc.: New York, NY; **1995**, 2nd ed, pp 287-313.
5. March, J., Smith, M.B.; March's Advanced Organic Chemistry; John Wiley & Sons, Inc.: Hoboken, NJ, **2007**; 6th ed.
6. McMurry, J.; Organic Chemistry; Brooks & Cole: Belmont, CA; **2004**, 6th ed.
7. McQuarrie, D.; Quantum Chemistry; University Science Books: Sausalito, CA, **2008**, 2nd ed.
8. Rouhi, A.; C&EN. 1997, 38-42.

Problems

1. Describe the difference between primary and secondary kinetic isotope effects.
2. Estimate the k_{N-H}/k_{N-D} for a deuterium substitution on nitrogen given that v_H=9.3x1013 Hz and the activation energy is equal to 5.31 kJ/mol.
3. Using the 'rule of thumb' for heavy isotope effects, calculate the expected effect for a bromine isotope substitution, ^{79}Br and ^{81}Br.
4. Explain some of the main ways kinetic isotope effects are used.
5. As discussed, the rate-limiting step in the bromination of acetone is the breaking of a carbon-hydrogen bond. Estimate kC-H/KC-D for this reaction at 285 K. (Given: vtilde$_{C-H}$=3000 cm^{-1} and vtilde$_{C-D}$=2100 cm^{-1})

Solutions

1. Primary isotope effects involve isotopic substitution at the bond being broken in a reaction, while secondary isotope effects involve isotopic Substitution on bonds adjacent to the bond being broken.
2. 8.5
3. 1.0126
4. To determine reaction mechanisms, to determine rate limiting steps in reactions, to determine transition states in reactions.
5. 9.685

8.9: Reactions in Solution

Learning Objectives

Make sure you thoroughly understand the following essential ideas:

- Describe some of the major differences between the kinetics of reactions in the gas phase, compared with those in liquid solutions.
- What role do solvent cages play in solution kinetics?
- Explain the distinction between diffusion-control and activation-control of reaction rates in solutions.
- How can the polarity of a solvent affect the energetics of a reaction mechanism?

The kinetics fundamentals we covered in the earlier sections of this lesson group relate to processes that take place in the gas phase. But chemists and biochemists are generally much more concerned with solutions. This lesson will take you through some of the extensions of basic kinetics that you need in order to understand the major changes that occur when reactions take place in liquid solutions.

What's different about kinetics in liquid solutions?

Most of the added complications of kinetics and rate processes in liquid solutions arise from the *much higher density* of the liquid phase. In a typical gas at atmospheric pressure, the molecules occupy only about 0.2 per cent of the volume; the other 99.8 percent is empty space. In a liquid, molecules may take up more than half the volume, and the "empty" spaces are irregular and ever-changing as the solvent molecules undergo thermal motions of their own.

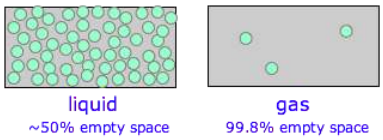


Figure 8.9.1: Liquid vs. Gases.

In a typical liquid solution, the solvent molecules massively outnumber the reactant solute molecules, which tend to find themselves momentarily (~10^{-11} sec) confined to a "hole" within the liquid. This trapping becomes especially important when the solvent is strongly hydrogen-bonded as is the case with water or alcohol.

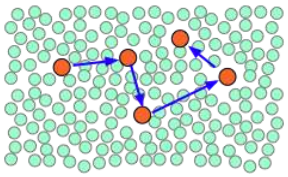
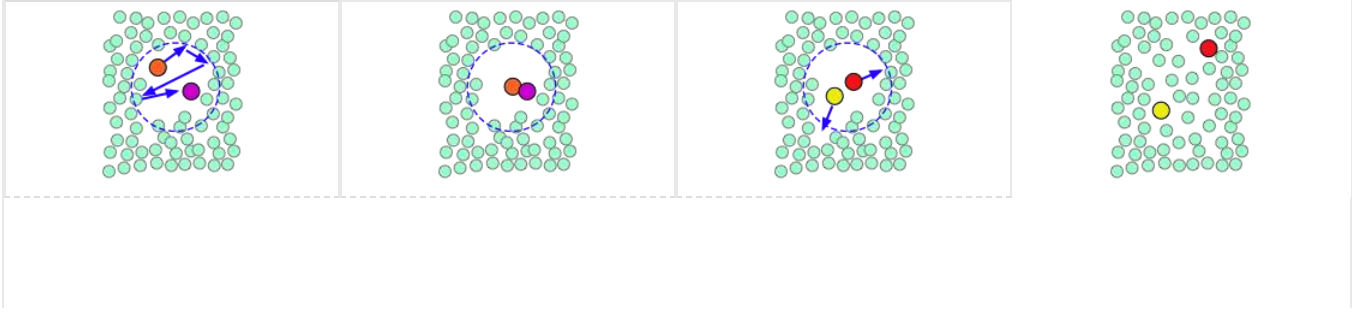


Figure 8.9.2: Brownian motion of a particle in solution

When thermal motions occasionally release a solute molecule from this trap, it will jump to a new location. The jumps are very fast (10^{-12} - 10^{-13} sec) and short (usually a few solvent-molecule diameters), and follow an entirely random pattern, very much as in **Brownian motion**. Consider a simple bimolecular process A + B → products. The reactant molecules will generally be jumping from hole to hole in the solvent matrix, only occasionally finding themselves in the same *solvent cage* where thermal motions are likely to bring them into contact.

Table 8.9.1: Solvent cages and encounter pairs



A pair of reactants end up in the same solvent cage, where they bounce around randomly and exchange kinetic energy with the solvent molecules.	Eventually the two reactants form an *encounter pair*. If they fail to react the first time, they have many more opportunities during the lifetime of the cage.	The products form and begin to move away from each other.	Finally, after about 10^{-11} sec, the solvent cage breaks up and the products diffuse away.

The process can be represented as

$$A + B \rightarrow \{AB\} \rightarrow \text{products} \tag{8.9.1}$$

in which the $\{AB\}$ term represents the caged reactants including the *encounter pair* and the activated complex.

Contrast this scenario with a similar reaction taking place in the gas phase; the molecules involved in the reaction will often be the only ones present, so a significant proportion of the collisions will be A-B encounters. However, if the collision should fail to be energetically or geometrically viable, the reactant molecules fly apart and are unlikely to meet again anytime soon. In a liquid, however, the solute molecules are effectively in a constant state of collision — if not with other reactants, then with solvent molecules which can exchange kinetic energy with the reactants. So once an A-B encounter pair forms, the two reactants get multiple whacks at each other, greatly increasing the probability that they will obtain the kinetic energy needed to kick them over the activation hump before the encounter pair disintegrates.

Limiting Cases: Diffusion-Controlled and Activation-Controlled Reactions

The encounter pair model introduces some new rate parameters:

$$\mathrm{A + B} \underset{k_{-1}}{\overset{k1}{\rightleftharpoons}} \mathrm{AB} \longrightarrow \text{products} \tag{8.9.2}$$

The first step is an equilibrium between the reactants outside and inside the solvent cage. The rate constants k_1 and k_2 reflect those relating to diffusion of molecules through the solvent; their values are strongly dependent on the viscosity (and thus the temperature) of the solvent. (Note that k_1 is a second-order rate constant, while k_2 is first-order.)

Diffusion is the transport of a substance through a concentration gradient; that is, from a region of higher concentration to one of lower concentration. Think of the way the color of tea spreads out when a tea bag is immersed in hot water. Diffusion occurs because random thermal motions are statistically more likely to move molecules out of a region of higher concentration than in the reverse direction, simply because in the latter case fewer molecules are available to make the reverse trip. Eventually the concentrations become uniform and equilibrium is attained.

As molecules diffuse through a liquid, they must nudge neighboring molecules out of the way. The work required to do this amounts to an activation energy, so diffusion can be thought of as a kinetic process with its own rate constant k_d and activation energy. These parameters depend on the sizes of the solute and solvent molecules and on how strongly the latter interact with each other. This suggests two important *limiting cases* for reactions in solution.

For water at room temperature, k_1 is typically 10^9-10^{10} $\mathrm{dm^{-3}\ mol^{-1}\ s^{-1}}$ and k_2 is around 10^{-9}-10^{-10} $\mathrm{dm^{-3}\ mol^{-1}\ s^{-1}}$. Given these values, $k_3 > 10^{12}$ $\mathrm{s^{-1}}$ implies diffusion control, while values $< 10^9$ $\mathrm{s^{-1}}$ are indicative of activation control.

- **Diffusion Controlled** ($k_3 \gg k_2$): If the activation energy of the A+B reaction is very small or if escape of molecules from the {AB} cage is difficult, the kinetics will be dominated by k_1, and thus by the activation energy of diffusion. Such a process is said to be *diffusion controlled*. Reactions in aqueous solution in which $E_a > 20$ kJ/mol are likely to fall into this category.
- **Activation Controlled** *(*$k_3 \ll k_2$): Alternatively, if the activation energy of the A+B reaction dominates the kinetics, and the reaction is *activation-controlled*.

Several general kinds of reactions are consistently very "fast" and thus are commonly found to be diffusion-controlled in most solvents:

Gas-phase rate constants are normally expressed in units of mol $\mathrm{s^{-1}}$, but rate constants of reactions in solution are conventionally given in mol/L units, or $\mathrm{dm^3\ mol^{-1}\ s^{-1}}$. Conversion between them depends on a number of assumptions and is non-trivial.

- **Recombination of atoms and radicals**

For example the formation of I_2 from I atoms in hexane at 298 K has $k_3 = 1.3\times10^{12}$ dm^3 mol^{-1} s^{-1}.

- **Acid-base reactions** which involve the transport of H^+ and OH^- ions tend to be very fast.

  The most famous of these is one of the fastest reactions known:

$$H^+ + OH^- \rightarrow H_2O$$

  for which $k_3 = 1.4\times10^{11}$ dm^3 mol^{-1} s^{-1} at 298 K.

Polar solvents such as water and alcohols interact with ions and polar molecules through attractive dipole-dipole and ion-dipole interactions, leading to lower-energy solvated forms which stabilize these species. In this way, a polar solvent can alter both the thermodynamics and kinetics (rate) of a reaction.

Solvent Thermodynamic Effect

If the products of the reaction are markedly more or less polar than the reactants, solvent polarity can change the overal thermodynamics (equilibrium constant) of the reaction. Nowhere is this more apparent than when an ionic solid such as salt dissolves in water. The Na^+ and Cl^- ions are bound together in the solid through strong coulombic forces; pulling the solid apart in a vacuum or in a nonpolar solvent is a highly endothermic process. In contrast, dissolution of NaCl in water is slightly exothermic and proceeds spontaneously.

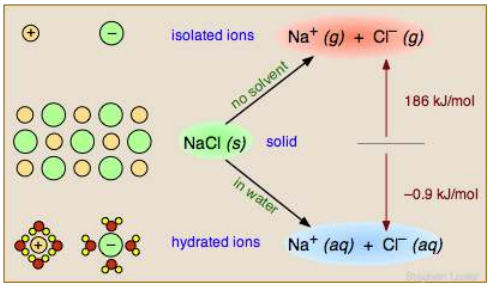
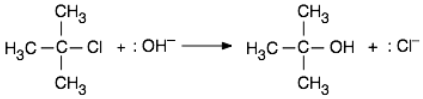


Figure 8.9.3: The solvent thermodynamics effect for table salt in water.

The water facilitates this process in two important ways. First, its high dielectric constant of 80 reduces the force between the separated ions to 1/80 of its normal value. Secondly, the water molecules form a *solvation shell* around the ions (lower left), rendering them energetically (thermodynamically) more stable than they were in the NaCl solid.

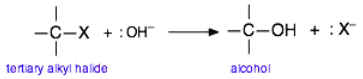
Solvent Kinetic Effect

In the same way, a reaction whose mechanism involves the formation of an intermediate or activated complex having a polar or ionic character will have its activation energy, and thus its rate, subject to change as the solvent polarity is altered. As an example we will consider an important class of reactions that you will hear much about if you take a course in organic chemistry. When an aqueous solution of a strong base such as KOH is added to a solution of *tertiary*-butyl chloride in ethanol, the chlorine is replaced by a hydroxyl group, leaving *t*-butyl alcohol as a product:

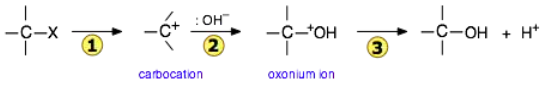


This reaction is one of a large and important class known as *S_N1 nucleophilic substitution processes* that are discussed in most organic chemistry courses. In these reactions, a species that possesses a pair of non-bonding electrons (also called a *nucleophile* or *Lewis base*) uses them to form a new bond with an *electrophile* — a compound in which a carbon atom has a partial positive charge owing to its bonds to electron-withdrawing groups. In the example here, other nucleophiles such as NH_3 or even H_2O would serve as well.

In order to reflect the generality of this process and to focus on the major changes that take place, we will represent this reaction as



Extensive studies of this class of reactions in the 1930's revealed that it proceeds in two activation energy-controlled steps, followed by a simple dissociation into the products:



In step ①, which is rate-determining, the chlorine leaves the alkyl chloride which becomes an intermediate known as a *carbocation* ("cat-ion"). These ions, in which the central carbon atom lacks a complete octet, are highly reactive, and in step ② the carbocation is attacked by the hydroxide ion which supplies the missing electron. The immediate product is another cation in which the positive charge is on the oxygen atom. This *oxonium ion* is unstable and rapidly dissociates (③)into the alcohol and a hydrogen ion.

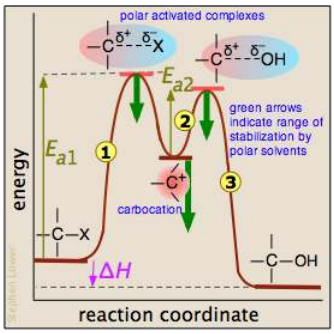


Figure 8.9.4: reaction Coordinate

The reaction coordinate diagram helps us understand the effect of solvent polarity on this reaction. Polar solvent molecules interact most strongly with species in which the electric charge is concentrated in one spot. Thus the carbocation is stabilized to a greater degree than are the activated complexes in which the charge is spread out between the positive and negative ends. As the heavy green arrows indicate, a more polar solvent will stabilize the carbocation more than it will either of the activated complexes; the effect is to materially reduce the activation energy of the rate-determining step, and thus speed up the reaction. Because neither the alkyl chloride nor the alcohol is charged, the change in solvent polarity has no effect on the equilibrium constant of the reaction. This is dramatically illustrated by observing the rate of the reaction in solvents composed of ethanol and water in varying amounts:

Table 8.9.2: Data

% water	10	20	30	40	50	60
$k_1 \times 10^6$	1.7	9.1	40.3	126	367	1294

8.10: Fast Reactions in Solution

The traditional experimental methods described above all assume the possibility of following the reaction after its components have combined into a homogeneous mixture of known concentrations. But what can be done if the time required to complete the mixing process is comparable to or greater than the time needed for the reaction to run to completion?

Flow methods

Flow instruments are a rapid mixing devices used to study the chemical kinetics of fast reactions in solution. There are different flavors that can be implement depending on the nature of the reaction as discussed below.

Continuous Flow Approach

For reactions that take place in milliseconds, the standard approach since the 1950s has been to employ a flow technique of some kind. An early example was used to study fast gas-phase reactions in which one of the reactants is a free radical such as OH that can be produced by an intense microwave discharge acting on a suitable source gas mixture. This gas, along with the other reactant being investigated, is made to flow through a narrow tube at a known velocity.

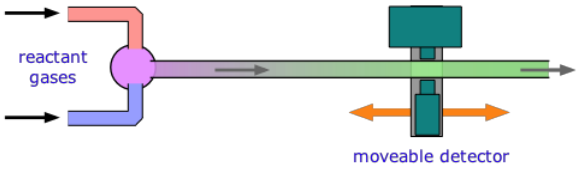


Figure 8.10.1: A continuous flow fast kinetic system.

If the distance between the point at which the reaction is initiated and the product detector is known, then the time interval can be found from the flow rate. By varying this distance, the time required to obtain the maximum yield can then be determined. Although this method is very simple in principle, it can be complicated in practice.

Stopped Flow Approach

Owing to the rather large volumes required, continuous flow method is more practical for the study of gas-phase reactions than for solutions, for which the stopped-flow method described below is generally preferred. These are by far the most common means of studying fast solution-phase reactions over time intervals of down to a fraction of a millisecond. The use of reasonably simple devices is now practical even in student laboratory experiments. These techniques make it possible to follow not only changes in the concentrations of reactants and products, but also the buildup and decay of reaction intermediates.

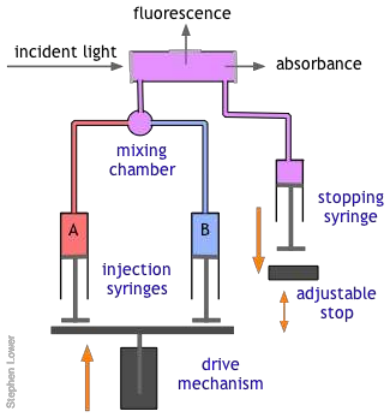


Figure 8.10.2: A stop flow fast kinetic system.

The basic stopped-flow apparatus consists of two or more coupled syringes that rapidly inject the reactants into a small mixing chamber and then through an observation cell that can be coupled to instruments that measure absorption, fluorescence, light scattering, or other optical or electrical properties of the solution. As the solution flows through the cell, it empties into a stopping syringe that, when filled, strikes a backstop that abruptly stops the flow. The volume that the stopping syringe can accept is adjusted so that the mixture in the cell has just become uniform and has reached a steady state; at this point, recording of the cell measurement begins and its change is followed.

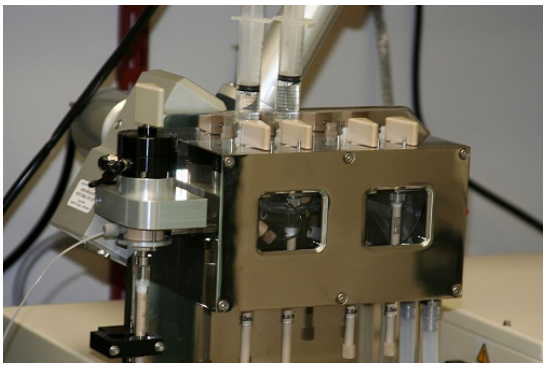


Figure 8.10.3: Stop-flow equipment at a biochemistry research laboratory for measuring rapid reactions and properties such as enzyme kinetics. from Wladimir Labeikovsky.

Quenched Flow Approach

In a quenched-flow instrument, the reaction is stopped after a certain amount of time has passed after mixing. The stopping of the reaction is called quenching and it can be achieved by various means, for example by mixing with another solution, which stops the reaction (chemical quenching), quickly lowering the temperature (freeze quenching) or even by exposing the sample to light of a certain wavelength (optical quenching).

Of course, there are many reactions that cannot be followed by changes in light absorption or other physical properties that are conveniently monitored. In such cases, it is often practical to *quench* (stop) the reaction after a desired interval by adding an appropriate quenching agent. For example, an enzyme-catalyzed reaction can be stopped by adding an acid, base, or salt solution that denatures (destroys the activity of) the protein enzyme. Once the reaction has been stopped, the mixture is withdrawn and analyzed in an appropriate manner.

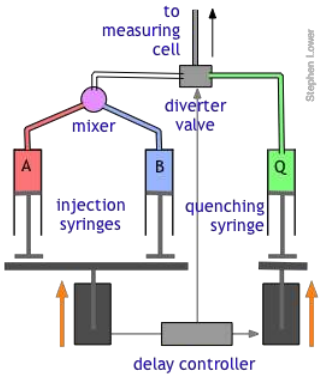


Figure 8.10.4: A quench flow fast kinetic system.

The quenched-flow technique works something like the stopped-flow method described above, with a slightly altered plumbing arrangement. The reactants A and B are mixed and fed directly through the diverter valve to the measuring cell, which is not shown in this diagram. After a set interval that can vary from a few milliseconds to 200 sec or more, the controller activates the quenching syringe and diverter valve, flooding the cell with the quenching solution.

Relaxation Methods

To investigate reactions that are complete in less than a millisecond, one can start with a pre-mixed sample in which one of active reactants is generated *in situ*. Alternatively, a rapid change in pressure or temperature can alter the composition of a reaction that has already achieved equilibrium.

Flash Photolysis

Many reactions are known which do not take place without light of wavelength sufficiently short to supply the activation energy needed to break a bond, often leading to the creation of a highly reactive radical. A good example is the combination of gaseous Cl_2 with H_2, which proceeds explosively when the system is illuminated with visible light. In *flash photolysis*, a short pulse of light is used to initiate a reaction whose progress can be observed by optical or other means.

Photolysis refers to the use of light to decompose a molecule into simpler units, often ions or free radicals. In contrast to *thermolysis* (decomposition induced by high temperature), photolysis is able to inject energy into a molecule almost instantaneously and can be much "cleaner," meaning that there are fewer side reactions that often lead to complex mixtures of products. Photolysis can also be highly *specific*; the wavelength of the light that triggers the reaction can often be adjusted to activate one particular kind of molecule without affecting others that might be present.

Norrish and Porter

All this had been known for a very long time, but until the mid-1940's there was no practical way of studying the kinetics of the reactions involving the highly reactive species produced by photolysis. In 1945, Ronald Norrish of Cambridge University and his graduate student George Porter conceived the idea of using a short-duration flash lamp to generate gas-phase CH_2 radicals, and then following the progress of the reaction of these radicals with other species by means of absorption spectroscopy.

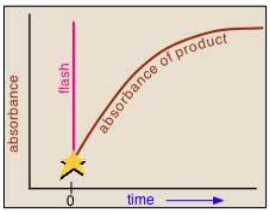


Figure 8.10.5: Basic principle of a flash-photolysis relaxation experiment where an excitation pulse purturbed a system at equilibrium and the subsequent dynamics are resolved in time.

In a flash photolysis experiment, recording of the absorbance of the sample cell contents is timed to follow the flash by an interval that can be varied in order to capture the effects produced by the product or intermediate as it is formed or decays. Norrish and Porter shared the 1967 Nobel Prize in Chemistry for this work.

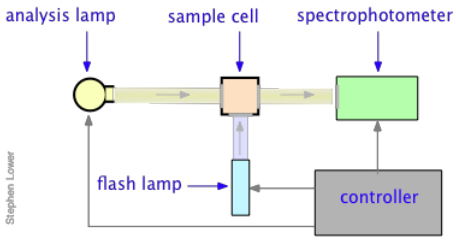


Figure 8.10.6: A flash-photolysis relaxation experiment

Many reactions, especially those that take place in solution, occur too rapidly to follow by flow techniques, and can therefore only be observed when they are already at equilibrium. The classical examples of such reactions are two of the fastest ones ever observed, the dissociation of water

$$2H_2O \rightarrow H_3O^+ + OH^- \tag{8.10.1}$$

and the formation of the triiodide ion in aqueous solution

$$I^- + I_2 \rightarrow I_3^- \tag{8.10.2}$$

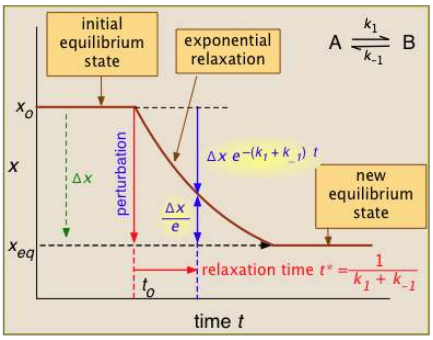
Reactions of these kinds could not be studied until the mid-1950s when techniques were developed to shift the equilibrium by imposing an abrupt physical change on the system.

Temperature Jumps

The rate constants of reversible reactions can be measured using a relaxation method. In this method, the concentrations of reactants and products are allowed to achieve equilibrium at a specific temperature. Once equilibrium has been achieved, the temperature is rapidly changed, and then the time needed to achieve the new equilibrium concentrations of reactants and products is measured. For example, if the reaction

$$\mathrm{A} \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} \mathrm{B} \tag{8.10.3}$$

is endothermic, then according to the Le Chatelier principle, subjecting the system to a rapid jump in temperature will shift the equilibrium state to one in which the product B has a higher concentration. The composition of the system will than begin to shift toward the new equilibrium composition at a rate determined by the kinetics of the process.



For the general case illustrated here, the quantity "*x*" being plotted is a measurable quantity such as light absorption or electrical conductivity that varies linearly with the composition of the system. **In a first-order process**, *x* will vary with time according to

$$x_t = x_o e^{-kt} \tag{8.10.4}$$

After the abrupt perturbation at time t_o, the relaxation time *t** is defined as the half-time for the return to equilibrium — that is, as the time required for x_o to decrease by \(Δ*x*/e = Δ*x*/2.718\). The derivation of *t** and the relations highlighted in yellow can be found in most standard kinetics textbooks. Temperature jumps are likely most commonly used.

The rate law for the reversible reaction in Equation 8.10.3 can be written as

$$\frac{d\,[\mathrm{B}]}{dt} = k_1\,[\mathrm{A}] - k_{-1}\,[\mathrm{B}] \tag{8.10.5}$$

Consider a system comprising A and B that is allowed to achieve equilibrium concentrations at a temperature, T_1. After equilibrium is achieved, the temperature of the system is instantaneously lowered to T_2 and the system is allowed to achieve new equilibrium concentrations of A and B, $[\mathrm{A}]_{\mathrm{eq},2}$ and $[\mathrm{B}]_{\mathrm{eq},2}$. During the transition time from the first equilibrium state to the second equilibrium state, we can write the instantaneous concentration of A as

$$[\mathrm{A}] = [\mathrm{B}]_{\mathrm{eq},1} - [\mathrm{B}] \tag{8.10.6}$$

The rate of change of species B can then be written as

$$\frac{d\,[\mathrm{B}]}{dt} = k_1\left([\mathrm{B}]_{\mathrm{eq},1} - [\mathrm{B}]\right) - k_{-1}\,[\mathrm{B}] = k_1[\mathrm{B}]_{\mathrm{eq},1} - (k_1 + k_{-1})\,[\mathrm{B}] \tag{8.10.7}$$

At equilibrium, $d\,[\mathrm{B}]\,/dt = 0$ and $[\mathrm{B}] = [\mathrm{B}]_{\mathrm{eq},2}$, allowing us to write

$$k_1[\mathrm{B}]_{\mathrm{eq},1} = (k_1 + k_{-1})\,[\mathrm{B}]_{\mathrm{eq},2} \tag{8.10.8}$$

Using the above equation, we can rewrite the rate equation as

$$\frac{d\mathrm{B}}{\left([\mathrm{B}]_{\mathrm{eq},2} - [\mathrm{B}]\right)} = (k_1 + k_{-1})\,dt \tag{8.10.9}$$

Integrating yields

$$-\ln\left([\mathrm{B}] - [\mathrm{B}]_{\mathrm{eq},2}\right) = -(k_1 + k_{-1})\,t + C \tag{8.10.10}$$

We can rearrange the above equation in terms of B

$$[\mathrm{B}] = Ce^{-(k_1 + k_{-1})t} + [\mathrm{B}]_{\mathrm{eq},2} \tag{8.10.11}$$

At $t = 0$, $[\mathrm{B}] = [\mathrm{B}]_{\mathrm{eq},1}$, so $C = [\mathrm{B}]_{\mathrm{eq},1} - [\mathrm{B}]_{\mathrm{eq},2}$. Plugging the the value of C, we arrive at

$$[\mathrm{B}] - [\mathrm{B}]_{\mathrm{eq},2} = \left([\mathrm{B}]_{\mathrm{eq},1} - [\mathrm{B}]_{\mathrm{eq},2} \right) e^{-(k_1 + k_{-1})t} \tag{8.10.12}$$

which can also be expressed as

$$\Delta [\mathrm{B}] = \Delta [\mathrm{B}]_0 e^{-(k_1 + k_{-1})t} = \Delta [\mathrm{B}]_0 e^{-t/\tau} \tag{8.10.13}$$

where $\Delta [\mathrm{B}]$ is the difference in the concentration of B from the final equilibrium concentration after the perturbation, and τ is the *relaxation time*. A plot of $\ln (\Delta [\mathrm{B}] / \Delta [\mathrm{B}]_0)$ versus t will be linear with a slope of $-(k_1 + k_{-1})$, where k_1 and k_{-1} are the rate constants at temperature, T_2.

Contributors

- Mark Tuckerman (New York University)
- Stephen Lower, Professor Emeritus (Simon Fraser U.) Chem1 Virtual Textbook

8.11: Oscillating Reactions

It should be clear by now that chemical kinetics is governed by the mathematics of systems of differential equations. Thus far, we have only looked at reaction systems that give rise to purely *linear* differential equations, however, in many instances the rate equations are nonlinear. When the differential equations are nonlinear, the behavior is considerably more complex. In particular, nonlinear equations can lead to oscillatory solutions and can also exhibit the phenomenon of *chaos*. Chaotic systems are systems that are highly sensitive to small changes in the parameters of the equations or in the initial conditions. Basically, this means that the behavior of a chaotic system can be unpredictable, since such small changes can occur in the form of small errors in determining the parameters (rounding to the nearest tenth or hundredth) or in specifying the initial conditions, and these small changes can cause the system to evolve in time in a very different way.

The Iodine Clock Reaction

The iodine clock reaction is a popular chemistry experiment in which one can visualize how different rate constants in consecutive reactions affect the concentration of species during the reaction. Iodine anions (I^-) are colorless. When I^- is reacted with hydrogen peroxide and protons, triiodide is formed, which has a dark blue color. Consider the following series of irreversible reactions:

$$H_2O_2 + 3\ I^- + 2\ H^+ \xrightarrow{k_1} I_3^- + 2\ H_2O \tag{8.11.1}$$

$$I_3^- + 2\ S_2O_3^{2-} \xrightarrow{k_2} 3\ I^- + S_4O_6^{2-} \tag{8.11.2}$$

The rate laws for this system are

$$\frac{d\left[I_3^-\right]}{dt} = k_1\left[I^-\right]^3 - k_2\left[I_3^-\right]\left[S_2O_3^{2-}\right]^2 \tag{8.11.3}$$

$$\frac{d\left[I^-\right]}{dt} = -k_1\left[I^-\right]^3 + 3k_2\left[I_3^-\right]\left[S_2O_3^{2-}\right]^2 \tag{8.11.4}$$

$$\frac{d\left[S_2O_3^{2-}\right]}{dt} = -k_2\left[I_3^-\right]\left[S_2O_3^{2-}\right]^2 \tag{8.11.5}$$

In order to make the equations look a little simpler, let us introduce the variables:

$$x = \left[I^-\right], \qquad y = \left[I_3^-\right], \qquad z = \left[S_2O_3^{2-}\right] \tag{8.11.6}$$

In terms of these, the rate equations are

$$\frac{dx}{dt} = -k_1x^3 + 3k_2yz^2 \tag{8.11.7}$$

$$\frac{dy}{dt} = k_1x^3 - k_2yz^2 \tag{8.11.8}$$

$$\frac{dz}{dt} = -k_2yz^2 \tag{8.11.9}$$

If we solve these numerically, we find the following time dependence of the three concentrations: This is a clear example of nonlinearity. Note how the concentration of I_3^- remains close to 0 for a period of time and then suddenly starts to increase. In a sense, think of the “straw that broke the camel’s back”. As we pile straws on the back of the camel, the camel remains upright until that last straw, which suddenly breaks the back of the camel, and the camel suddenly falls to the ground. This is also an illustration of nonlinearity.

***Video* 8.11.1 :** *Famous iodine clock reaction: oxidation of potassium iodide by hydrogen peroxide (https://www.youtube.com/watch?v=_qhYDuJt8fI).*

Despite the complexity of the rate equations, we can still analyze the approximately and predict the behavior seen in Figure 8.11.1. In this reaction mechanism, $k_2 \gg k_1$. Given the rate law for I_3^- ,

$$\frac{d\left[I_3^-\right]}{dt} = k_1\left[I^-\right]^3 - k_2\left[I_3^-\right]\left[S_2O_3^{2-}\right]^2 \tag{8.11.10}$$

if we use the steady-state approximation, we can set the $\frac{d\left[I_3^-\right]}{dt}$ equal to 0, yielding

$$\left[I_3^-\right] = \frac{k_1}{k_2}\frac{\left[I^-\right]^3}{\left[S_2O_3^{2-}\right]^2} \tag{8.11.11}$$

Since $k_2 \gg k_1$, the concentration of $\left[I_3^-\right]$ is approximately 0 as long as there are $S_2O_3^{2-}$ ions present. As soon as all of the $S_2O_3^{2-}$ is consumed, the concentration of I_3^- can build up in the solution, changing the solution to a dark blue color. Figure 8.11.1 displays the concentration profiles for I^-, I_3^- , and $S_3O_3^{2-}$. As can be seen from the figure, the concentration of I_3^- (red line) remains at approximately 0 $\mathrm{mol/L}$ until all of the $S_2O_3^{2-}$ (blue line) has been depleted.

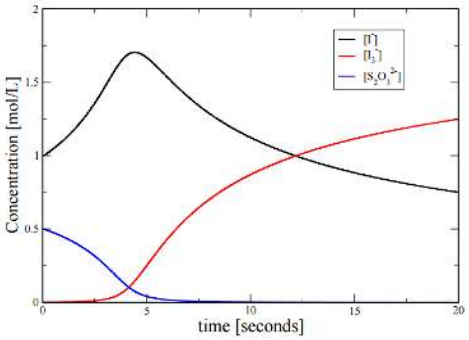


Figure 8.11.1: Concentrations as functions of time of the three species in the iodine clock reaction.

Oscillating Reactions

In all of the examples we have seen thus far, the concentration of intermediate species displays a single maximum during the course of the reaction. There is another class of reactions called *oscillating reactions* in which the concentration of intermediate species oscillates with time. Consider the following series of reactions

$$\mathrm{A} + \mathrm{Y} \xrightarrow{k_1} \mathrm{X} \tag{8.11.12}$$

$$\mathrm{X} + \mathrm{Y} \xrightarrow{k_2} \mathrm{P} \tag{8.11.13}$$

$$\mathrm{B} + \mathrm{X} \xrightarrow{k_3} 2\,\mathrm{X} + \mathrm{Z} \tag{8.11.14}$$

$$2\,\mathrm{X} \xrightarrow{k_4} \mathrm{Q} \tag{8.11.15}$$

$$\mathrm{Z} \xrightarrow{k_5} \mathrm{Y} \tag{8.11.16}$$

In the above reaction mechanism, A and B are reactants; X, Y, and Z are intermediates; and P and Q are products. The third reaction in which B and X react to form X and Z is known as an "autocatalytic reaction" in which at least one of the reactants is also a product. Such reactions are a key feature of oscillating reactions, as will be discussed below.

***Video* 8.11.2:** *The famous Belousov Zhabotinsky chemical reaction in a petri dish. The action is speeded up 8 x from real life. Pacemaker nucleation sites emit circular waves. Breaking the wavefront with a wire triggers pairs of spiral defects which emit more closely spaced waves which eventually fill the container.*

Let us assume that the concentrations of A and B are large, such that we can approximate them to be constant with time. The rate equation for species X can be written as

$$\frac{d\left[\mathrm{X}\right]}{dt} = k_1\left[\mathrm{A}\right]\left[\mathrm{Y}\right] - k_2\left[\mathrm{X}\right]\left[\mathrm{Y}\right] + k_3\left[\mathrm{B}\right]\left[\mathrm{X}\right] - 2k_4[\mathrm{X}]^2 \tag{8.11.17}$$

Using the steady-state approximation, we can set $d\mathrm{X}/dt = 0$ and rewrite Equation 8.11.17 as

$$\left(-2k_4\right)\left[\mathrm{X}\right]^2 + \left(k_2\left[\mathrm{Y}\right] - k_3\left[\mathrm{B}\right]\right)\left[\mathrm{X}\right] + k_1\left[\mathrm{A}\right]\left[\mathrm{Y}\right] = 0 \tag{8.11.18}$$

We can then use the quadratic formula to solve for X:

$$\left[\mathrm{X}\right] = -\frac{\left(k_2\left[\mathrm{Y}\right] - k_3\left[\mathrm{B}\right]\right) \pm \sqrt{\left(k_2\left[\mathrm{Y}\right] - k_3\left[\mathrm{B}\right]\right)^2 - 4\left(-2k_4\right)\left(k_1\left[\mathrm{A}\right]\left[\mathrm{Y}\right]\right)}}{2\left(-2k_4\right)} \tag{8.11.19}$$

Thus, there are two solutions for the concentration of X accessible to the reaction system. To examine solutions for $[\mathrm{X}]$, let us first assume that $[\mathrm{Y}]$ is large. Under these conditions, the first two reactions in the reaction mechanism largely determine the concentration of $[\mathrm{X}]$. We can thus approximate Equation 8.11.17 as

$$0 \approx k_1\left[\mathrm{A}\right]\left[\mathrm{Y}\right] - k_2\left[\mathrm{X}\right]\left[\mathrm{Y}\right] \tag{8.11.20}$$

Solving for $[\mathrm{X}]$ yields

$$\left[\mathrm{X}\right] \approx \frac{k_1\left[\mathrm{A}\right]}{k_2} \tag{8.11.21}$$

As the reaction continues, species Y is depleted and the assumption that $[\mathrm{Y}]$ is large becomes invalid. Instead the 3^{rd} and 4^{th} steps of the reaction mechanism determine the concentration of X. In this limit, we can approximate Equation 8.11.17 as

$$0 \approx k_3\left[\mathrm{B}\right]\left[\mathrm{X}\right] - 2k_4[\mathrm{X}]^2 \tag{8.11.22}$$

Solving for $[\mathrm{X}]$ yields

$$\left[\mathrm{X}\right] \approx \frac{k_3\left[\mathrm{B}\right]}{2k_4} \tag{8.11.23}$$

In the second mechanism, the autocatalytic reaction step leads to an increase in the concentration of X and Z, which in turn leads to an increase in the concentration of Y. The feedback loop between the production of species X and Y leads to oscillatory behavior in the system. This reaction mechanism is known as the Belousov-Zhabotinksii reaction first discovered in the 1950s.

- The autocatalytic reaction of X in which X reacts with B to form more X in reaction 3
- The regeneration of species Y in reaction 5
- The competition between reaction 2 and 3 for the consumption of X and the involvement of Y in reaction 2

The actual Belousov-Zhabotinskii reaction is complex, involving many individual steps, and involves the oscillation between the concentration of $HBrO_2$ and Br^-. The reaction equations are

$$BrO_3^- + Br^- + 2\ H^+ \longrightarrow HBrO_2 + HOBr \tag{8.11.24}$$
$$HBrO_2 + Br^- + H^+ \longrightarrow 2\ HOBr \tag{8.11.25}$$
$$HOBr + Br^- + H^+ \longrightarrow Br_2 + H_2O \tag{8.11.26}$$
$$2\ HBrO_2 \longrightarrow BrO_3^- + HOBr + H^+ \tag{8.11.27}$$
$$BrO_3^- + HBrO_2 + H^+ \longrightarrow 2\ BrO_2 + H_2O \tag{8.11.28}$$
$$BrO_2 + Ce^{3+} + H^+ \longrightarrow HBrO_2 + Ce^{4+} \tag{8.11.29}$$
$$BrO_2 + Ce^{4+} + H_2O \longrightarrow BrO_3^- + Ce^{3+} + 2\ H^+ \tag{8.11.30}$$

The essential steps in this mechanism can be reduced to the following set of reactions. Note that we leave this unbalanced and only include the species whose concentrations as functions of time we seek.

$$BrO_3^- + Br^- \xrightarrow{k_1} HBrO_2 + HOBr \tag{8.11.31}$$
$$HBrO_2 + Br^- \xrightarrow{k_2} 2\ HOBr \tag{8.11.32}$$
$$BrO_3^- + HBrO_2 \xrightarrow{k_3} 2\ HBrO_2 + 2\ Ce^{4+} \tag{8.11.33}$$
$$2\ HBrO_2 \xrightarrow{k_4} BrO_3^- + HOBr \tag{8.11.34}$$
$$Ce^{4+} \xrightarrow{k_5} f Br^- \tag{8.11.35}$$

Setting the variables as follows:

$$x = [HBrO_2], \qquad y = \left[Br^-\right], \qquad z = \left[Ce^{4+}\right] \tag{8.11.36}$$

We make the approximation that $\left[BrO_3^-\right]$ to be a constant a. In this case, the rate equations become

$$\frac{dx}{dt} = k_1 ay - k_2 xy + k_3 ax - k_4 x^2 \tag{8.11.37}$$
$$\frac{dy}{dt} = -k_1 ay - k_2 xy + f k_5 z \tag{8.11.38}$$
$$\frac{dz}{dt} = 2k_3 ax - k_5 z \tag{8.11.39}$$

Solving these equations numerically, we obtain the trajectory of two of the species show in the Figure 8.11.2

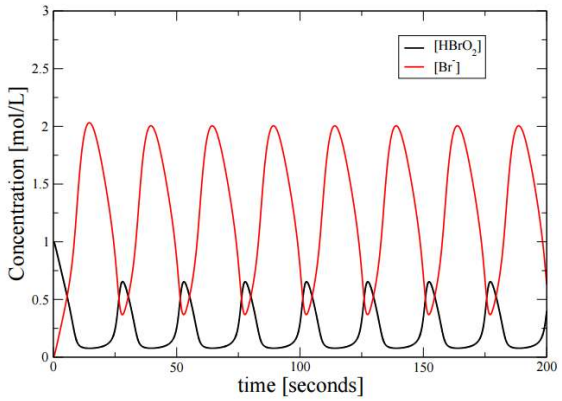


Figure 8.11.2: Oscillating pattern of the concentrations in the Belousov-Zhabotinskii reaction.

On the other hand, we can drive this system to become chaotic by changing the parameters a little. When this is done, we find the follow plot of the concentration of x:

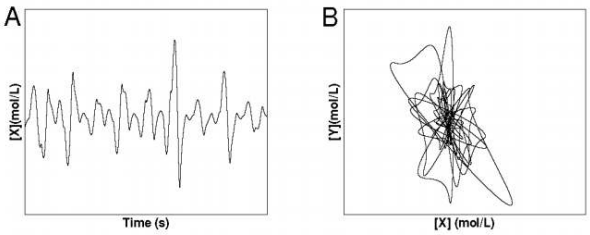


Figure 8.11.3: Chaotic behavior in the Belousov-Zhabotinskii reaction.

Contributors

- Mark Tuckerman (New York University)

8.E: Chemical Kinetics (Exercises)

CHAPTER OVERVIEW

9: Enzyme Kinetics

Enzyme kinetics is the study of the chemical reactions that are catalyzed by enzymes. In enzyme kinetics, the reaction rate is measured and the effects of varying the conditions of the reaction are investigated.

9.1: General Principles of Catalysis

As can be seen from the Arrhenius equation, the magnitude of the activation energy, E_a, determines the value of the rate constant, k, at a given temperature and thus the overall reaction rate. Catalysts provide a means of reducing E_a and increasing the reaction rate. Catalysts are defined as substances that participate in a chemical reaction but are not changed or consumed. Instead they provide a new mechanism for a reaction to occur which has a lower activation energy than that of the reaction without the catalyst. *Homogeneous catalysis* refers to reactions in which the catalyst is in solution with at least one of the reactants whereas *heterogeneous catalysis* refers to reactions in which the catalyst is present in a different phase, usually as a solid, than the reactants. Figure $9.1.1$ shows a comparison of energy profiles of a reaction in the absence and presence of a catalyst.

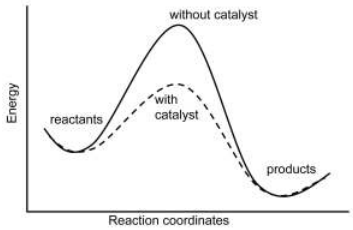


Figure 9.1.1: Comparison of energy profiles with and without catalyst present.

Consider a non-catalyzed elementary reaction

$$\mathrm{A} \xrightarrow{k} \mathrm{P} \tag{9.1.1}$$

which proceeds at rate k at a certain temperature. The reaction rate can be expressed as

$$\frac{d\left[\mathrm{A}\right]}{dt} = -k\left[\mathrm{A}\right] \tag{9.1.2}$$

In the presence of a catalyst C, we can write the reaction as

$$\mathrm{A} + \mathrm{C} \xrightarrow{k_{\mathrm{cat}}} \mathrm{P} + \mathrm{C} \tag{9.1.3}$$

and the reaction rate as

$$\frac{d\left[\mathrm{A}\right]}{dt} = -k\left[\mathrm{A}\right] - k_{\mathrm{cat}}\left[\mathrm{A}\right]\left[\mathrm{C}\right] \tag{9.1.4}$$

where the first term represents the uncatalyzed reaction and the second term represents the catalyzed reaction. Because the reaction rate of the catalyzed reaction is often magnitudes larger than that of the uncatalyzed reaction (i.e. $k_{\mathrm{cat}} \gg k$), the first term can often be ignored.

Acid Catalysis

A common example of homogeneous catalysts are acids and bases. For example, given an overall reaction is $\mathrm{S} \rightarrow \mathrm{P}$. If k is the rate, then

$$\frac{d\left[\mathrm{P}\right]}{dt} = k\left[\mathrm{S}\right] \tag{9.1.5}$$

The purpose of an enzyme is to enhance the rate of production of the product P. The equations of the acid-catalyzed reaction are

$$\mathrm{S} + \mathrm{A}H \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} \mathrm{S}H^+ + \mathrm{A}^- \tag{9.1.6}$$

$$\mathrm{S}H^+ + H_2O \xrightarrow{k_2} \mathrm{P} + H_3O^+ \tag{9.1.7}$$

$$H_3O^+ + \mathrm{A}^- \underset{k_{-3}}{\overset{k_3}{\rightleftharpoons}} \mathrm{A}H + H_2O \tag{9.1.8}$$

The full set of kinetic equations is

$$\frac{d[\mathrm{S}]}{dt} = -k_1[\mathrm{S}][\mathrm{A}H] + k_{-1}\left[\mathrm{S}H^+\right]\left[\mathrm{A}^-\right] \tag{9.1.9}$$

$$\frac{d[\mathrm{A}H]}{dt} = -k_1[\mathrm{S}][\mathrm{A}H] + k_{-1}\left[\mathrm{S}H^+\right]\left[\mathrm{A}^-\right] - k_{-3}[\mathrm{A}H] + k_3\left[H_3O^+\right]\left[\mathrm{A}^-\right] \tag{9.1.10}$$

$$\frac{d[\mathrm{S}H^+]}{dt} = k_1[\mathrm{S}][\mathrm{A}H] - k_{-1}\left[\mathrm{S}H^+\right]\left[\mathrm{A}^-\right] - k_2\left[\mathrm{S}H^+\right] \tag{9.1.11}$$

$$\frac{d\left[\mathrm{A}^-\right]}{dt} = k_1[\mathrm{S}][\mathrm{A}H] - k_{-1}\left[\mathrm{S}H^+\right]\left[\mathrm{A}^-\right] - k_2\left[\mathrm{A}^-\right]\left[H_3O^+\right] + k_{-3}[\mathrm{A}H] \tag{9.1.12}$$

$$\frac{d[\mathrm{P}]}{dt} = k_2\left[\mathrm{S}H^+\right] \tag{9.1.13}$$

$$\frac{d[H_3O^+]}{dt} = -k_2\left[\mathrm{S}H^+\right] - k_3\left[H_3O^+\right]\left[\mathrm{A}^-\right] + k_{-3}[\mathrm{A}H] \tag{9.1.14}$$

We cannot easily solve these, as they are nonlinear. However, let us consider two cases $k_2 \gg k_{-1}\left[\mathrm{A}^-\right]$ and $k_2 \ll k_{-1}\left[\mathrm{A}^-\right]$. In both cases, $\mathrm{S}H^+$ is consumed quickly, and we can apply a steady-state approximation:

$$\frac{d[\mathrm{S}H^+]}{dt} = k_1[\mathrm{S}][\mathrm{A}H] - k_{-1}\left[\mathrm{A}^-\right]\left[\mathrm{S}H^+\right] - k_2\left[\mathrm{S}H^+\right] = 0 \tag{9.1.15}$$

Rearranging in terms of $\mathrm{S}H^+$ yields

$$\left[\mathrm{S}H^+\right] = \frac{k_1[\mathrm{S}][\mathrm{A}H]}{k_{-1}\left[\mathrm{A}^-\right] + k_2} \tag{9.1.16}$$

and the rate of production of P can be written as

$$\frac{d[\mathrm{P}]}{dt} = k_2\left[\mathrm{S}H^+\right] = \frac{k_1k_2[\mathrm{S}][\mathrm{A}H]}{k_{-1}\left[\mathrm{A}^-\right] + k_2} \tag{9.1.17}$$

In the case where $k_2 \gg k_{-1}\left[\mathrm{A}^-\right]$, Equation 9.1.17 can be written as

$$\frac{d[\mathrm{P}]}{dt} = k_1[\mathrm{S}][\mathrm{A}H] \tag{9.1.18}$$

which is known as a general acid-catalyzed reaction. On the other hand, if $k_2 \ll k_{-1}\left[\mathrm{A}^-\right]$, we can use an equilibrium approximation to write the rate of production of P as

$$\frac{d[\mathrm{P}]}{dt} = \frac{k_1k_2[\mathrm{S}][\mathrm{A}H]}{k_{-1}\left[\mathrm{A}^-\right]} = \frac{k_1k_2}{k_{-1}K}[\mathrm{S}]\left[H^+\right] \tag{9.1.19}$$

where K is the acid dissociation constant:

$$K = \frac{\left[\mathrm{A}^-\right][H^+]}{[\mathrm{A}H]} \tag{9.1.20}$$

In this case, the reaction is hydrogen ion-catalyzed.

Enzyme Catalysis

To live, grow, and reproduce, microorganisms undergo a variety of chemical changes. They alter nutrients so they can enter the cell and they change them once they enter in order to synthesize cell parts and obtain energy.

Metabolism refers to all of the organized chemical reactions in a cell. Reactions in which chemical compounds are broken down are called **catabolic reactions** while reactions in which chemical compounds are synthesized are termed **anabolic reactions**. All of these reactions are under the control of enzymes.

Enzymes are substances present in the cell in small amounts that function to **speed up or catalyze chemical reactions**. On the surface of the enzyme is usually a small crevice that functions as an **active site** or catalytic site to which one or two specific substrates are able to bind. (Anything that an enzyme normally combines with is called a

substrate.) The binding of the substrate to the enzyme causes the flexible enzyme to change its shape slightly through a process called **induced fit** to form a tempore intermediate called an **enzyme-substrate complex** (Figure 1).

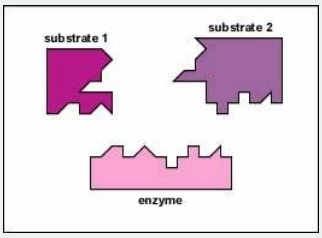
Figure 9.1.2 *: Enzymes are substances present in the cell in small amounts which speed up or catalyze chemical reactions. Enzymes speed up the rate of chemical reactions because they lower the energy of activation, the energy that must be supplied in order for molecules to react with one another. Enzymes lower the energy of activation by forming an enzyme-substrate complex.*

Enzymes speed up the rate of chemical reactions because they lower the energy of activation, the energy that must be supplied in order for molecules to react with one another (Figure 9.1.3). Like homogeneous catalysts discussed above, enzymes lower the energy of activation by forming an enzyme-substrate complex allowing products of the enzyme reaction to be formed and released (Figure 9.1.2).

Figure 9.1.3 *: An enzyme speeds up a chemical reaction by lowering its energy of activation, the energy that must be supplied in order for molecules to react with one another.*

Enzyme-Substrate Reactions

Enzymes are substances present in the cell in small amounts which speed up or catalyze chemical reactions. Enzymes speed up the rate of chemical reactions because they lower the energy of activation, the energy that must be supplied in order for molecules to react with one another. Enzymes lower the energy of activation by forming an enzyme-substrate complex.



Many enzymes require a nonprotein **cofactor** to assist them in their reaction. In this case, the protein portion of the enzyme, called an **apoenzyme**, combines with the cofactor to form the whole enzyme or **haloenzyme** (Figure 9.1.4). Some cofactors are ions such as Ca^{++}, Mg^{++}, and K^{+}; other cofactors are organic molecules called **coenzymes** which serve as carriers for chemical groups or electrons. NAD^{+}, $NADP^{+}$, FAD, and coenzyme A (CoA) are examples of coenzymes.

Figure 9.1.4 : An apoenzyme and cofactor combine to form a haloenzyme. If the cofactor is an organic molecule, it is called a coenzyme.

Enzymes are generally globular proteins. (Some RNA molecules called ribozymes can also be enzymes. These are usually found in the nuclear region of cells and catalyze the splitting of RNA molecules.). Enzymes are catalysts that breakdown or synthesize more complex chemical compounds. They allow chemical reactions to occur fast enough to support life. Enzymes speed up the rate of chemical reactions because they lower the energy of activation, the energy that must be supplied in order for molecules to react with one another. Anything that an enzyme normally combines with is called a substrate. Enzymes are very efficient with a typically enzyme generally able to catalyze between 1 and 10,000 molecules of substrate per second. The means that enzymes are only have to be present in small amounts in the cell since. They are not altered during their reaction and are highly specific for their substrate, with generally one specific enzyme dedicated for each specific chemical reaction.

Factors that affect the rate of enzyme reactions

Enzyme activity is affected by a number of factors including:

a. **The concentration of enzyme**: Assuming a sufficient concentration of substrate is available, increasing enzyme concentration will increase the enzyme reaction rate.
b. **The concentration of substrate**: At a constant enzyme concentration and at lower concentrations of substrates, the substrate concentration is the limiting factor. As the substrate concentration increases, the enzyme reaction rate increases. However, at very high substrate concentrations, the enzymes become saturated with substrate and a higher concentration of substrate does not increase the reaction rate.

c. **Inhibitors:** inhibitors will inhibit the activity of enzyme and decrease the rate of reaction. Enzyme inhibitors will bind to enzyme active sites and could modify the chemistry of an active site which can stop a substrate from entering.
d. **The temperature**: Each enzyme has an optimum temperature at which it works best. A higher temperature generally results in an increase in enzyme activity (Arrhenius kinetics). As the temperature increases, molecular motion increases resulting in more molecular collisions. If, however, the temperature rises above a certain point, the heat will denature the enzyme, causing it to lose its three-dimensional functional shape by denaturing its hydrogen bonds. Cold temperature, on the other hand, slows down enzyme activity by decreasing molecular motion.
e. **The pH**: Each enzyme has an optimal pH that helps maintain its three-dimensional shape. Changes in pH may denature enzymes by altering the enzyme's charge. This alters the ionic bonds of the enzyme that contribute to its functional shape.
f. **The salt concentration**: Each enzyme has an optimal salt concentration. Changes in the salt concentration may also denature enzymes.

Applications of Enzymes

Enzymes are essential to maintain homeostasis because any malfunction of an enzyme could lead to diseases. Therefore, pharmaceutical companies study enzyme to manipulate and synthesis new medicine. Besides their medicinal applications, enzymes in industry are important because enzymes help breaking down cellulose, wastes, etc. Enzymes are essential in the process of making new products in many industries such as pharmaceutical, food, paper, wine, etc.

References

1. "Chapter 4 Enzyme."Copyrighted IHW, Oct. 2005. Web. 5 Mar. 2011. http://biologymad.com/resources/Ch%2...0-20enzyme.pdf
2. Chang, Raymond. Physical Chemistry for the Biosciences. Sausalito, CA: University Science Books, 2005. Pages 363-384.

Contributors

- Mark Tuckerman (New York University)
- Dr. Gary Kaiser (COMMUNITY COLLEGE OF BALTIMORE COUNTY, CATONSVILLE CAMPUS)

9.2: The Equations of Enzyme Kinetics

In biological systems, enzymes act as catalysts and play a critical role in accelerating reactions, anywhere from 10^3 to 10^{17} times faster than the reaction would normally proceed. Enzymes are high-molecular weight proteins that act on a substrate, or reactant molecule, to form one or more products.

Michaelis-Menten Enzyme Kinetics

Enzymes are highly specific catalysts for biochemical reactions, with each enzyme showing a selectivity for a single reactant, or **substrate**. For example, the enzyme acetylcholinesterase catalyzes the decomposition of the neurotransmitter acetylcholine to choline and acetic acid. Many enzyme–substrate reactions follow a simple mechanism that consists of the initial formation of an enzyme–substrate complex, ES, which subsequently decomposes to form product, releasing the enzyme to react again.

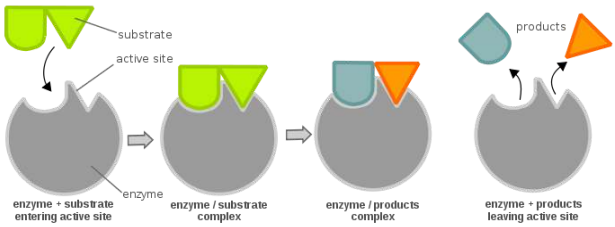


Figure 9.2.1: *An enzyme catalyzes the reaction of two substrates and to form one product. from Wikipedia.*

This is described within the following multi-step mechanism

$$E + S \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} ES \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} E + P \tag{9.2.1}$$

where k_1, k_{-1}, k_2, and k_{-2} are rate constants. The reaction's rate law for generating the product $[P]$ is

$$rate = \frac{d[P]}{dt} = k_2[ES] - k_{-2}[E][P] \tag{9.2.2}$$

However, if we make measurement early in the reaction, the concentration of products is negligible, i.e.,

$$[P] \approx 0 \tag{9.2.3}$$

and we can ignore the back reaction (second term in right side of Equation 9.2.2). Then under these conditions, the reaction's rate is

$$rate = \frac{d[P]}{dt} = k_2[ES] \tag{9.2.4}$$

To be analytically useful we need to write Equation 9.2.4 in terms of the reactants (e.g., the concentrations of enzyme and substrate). To do this we use the **steady-state approximation**, in which we assume that the concentration of ES remains essentially constant. Following an initial period, during which the enzyme–substrate complex first forms, the rate at which ES forms

$$\frac{d[ES]}{dt} = k_1[E][S] = k_1([E]_0 - [ES])[S] \tag{9.2.5}$$

is equal to the rate at which it disappears

$$-\frac{d[ES]}{dt} = k_{-1}[ES] + k_2[ES] \tag{9.2.6}$$

where $[E]_0$ is the enzyme's original concentration. Combining Equations 9.2.5 and 9.2.6 gives

$$k_1([E]_0 - [ES])[S] = k_{-1}[ES] + k_2[ES] \tag{9.2.7}$$

which we solve for the concentration of the enzyme–substrate complex

$$[ES] = \frac{[E]_0[S]}{\frac{k_{-1}+k_2}{k_1}+[S]} = \frac{[E]_0[S]}{K_m+[S]} \tag{9.2.8}$$

where K_m is the **Michaelis constant**. Substituting Equation 9.2.8 into Equation 9.2.4 leaves us with our final rate equation.

$$\frac{d[P]}{dt} = \frac{k_2[E]_0[S]}{K_m+[S]} \tag{9.2.9}$$

A plot of Equation 9.2.9, as shown in Figure 9.2.1, is instructive for defining conditions where we can use the rate of an enzymatic reaction for the quantitative analysis of an enzyme or substrate.

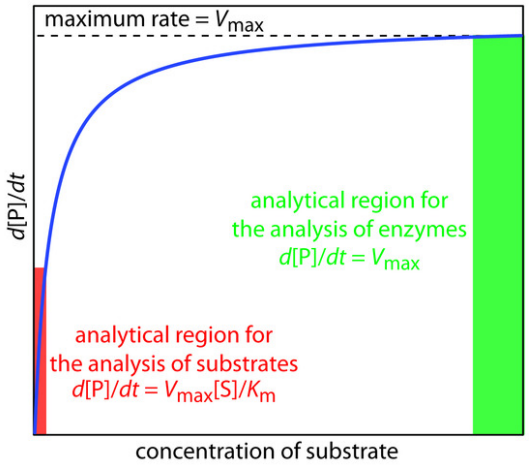


Figure 9.2.1: *Plot of Equation 9.2.9 showing limits for the analysis of substrates and enzymes in an enzyme-catalyzed chemical kinetic method of analysis. The curve in the region highlighted in red obeys equation 9.2.11 and the curve in the area highlighted in green follows Equation 9.2.10.*

For high substrate concentrations, where $[S] \gg K_m$, Equation 9.2.9 simplifies to

$$\frac{d[P]}{dt} = \frac{k_2[E]_0[S]}{K_m+[S]} \approx \frac{k_2[E]_0[S]}{[S]} = k_2[E]_0 = V_{max} \tag{9.2.10}$$

where V_{max} is the maximum rate for the catalyzed reaction. Under these conditions the reaction is zero-order in substrate and we can use V_{max} to calculate the enzyme's concentration, typically using a variable-time method. At lower substrate concentrations, where $[S] \ll K_m$, Equation 9.2.9 becomes

$$\frac{d[P]}{dt} = \frac{k_2[E]_0[S]}{K_m+[S]} \approx \frac{k_2[E]_0[S]}{K_m} = \frac{V_{max}[S]}{K_m} \tag{9.2.11}$$

The reaction is now first-order in substrate, and we can use the rate of the reaction to determine the substrate's concentration by a fixed-time method.

The Michaelis constant K_m is the substrate concentration at which the reaction rate is at half-maximum, and is an inverse measure of the substrate's affinity for the enzyme—as a small K_m indicates high affinity, meaning that the rate will approach V_{max} more quickly. The value of K_m is dependent on both the enzyme and the substrate, as well as conditions such as temperature and pH.

> *The Michaelis constant K_m is the substrate concentration at which the reaction rate is at half-maximum.*

From the last two terms in Equation 9.2.11, we can express V_{max} in terms of a **turnover number** (k_{cat}):

$$V_{max} = k_{cat}[E]_o \tag{9.2.12}$$

where $[E]_0$ is the enzyme concentration and k_{cat} is the turnover number, defined as the maximum number of substrate molecules converted to product per enzyme molecule per second. Hence, the turnover number is defined as the maximum number of chemical conversions of substrate molecules per second that a single catalytic site will execute for a given enzyme concentration $[E]_o$.

Example 9.2.1: Turnover number of acetylcholinesterase

Acetylcholinesterase (AChE) may be one of the fastest enzymes. It hydrolyzes acetylcholine to choline and an acetate group. One of the earliest values of the turnover number was 3×10^7 (molecules of acetylcholine) per minute per molecule of enzyme. A more recent value at 25°C, pH = 7.0, acetylcholine concentration of $2.5 \times 10^{-3}\ M$, was found to be $7.4 \times 10^5\ min^{-1}$ (*J Biol Chem. **236** (8): 2292–5.*).

There may be some 30 active centers per molecule. AChE is a serine hydrolase that reacts with acetylcholine at close to **the diffusion-controlled rate**. A diffusion-controlled reaction occurs so quickly that the reaction rate is the rate of transport of the reactants through the solution; as quickly as the reactants encounter each other, they react.

The Significance of K_M and V_{max}

The Michaelis-Menten model is used in a variety of biochemical situations other than enzyme-substrate interaction, including antigen-antibody binding, DNA-DNA hybridization, and protein-protein interaction. It can be used to characterize a generic biochemical reaction, in the same way that the **Langmuir equation** can be used to model generic adsorption of biomolecular species. When an empirical equation of this form is applied to microbial growth. The experimentally determined parameters values vary wildly between enzymes (Table 9.2.1):

Table 9.2.1: *Enzyme Kinetic parameters*

Enzyme	K_m (M)	k_{cat} (1/s)	k_{cat}/K_m (1/M.s)
Chymotrypsin	1.5×10^{-2}	0.14	9.3
Pepsin	3.0×10^{-4}	0.50	1.7×10^3
Tyrosyl-tRNA synthetase	9.0×10^{-4}	7.6	8.4×10^3
Ribonuclease	7.9×10^{-3}	7.9×10^2	1.0×10^5
Carbonic anhydrase	2.6×10^{-2}	4.0×10^5	1.5×10^7
Fumarase	5.0×10^{-6}	8.0×10^2	1.6×10^8

While K_m is equal to the substrate concentration at which the enzyme converts substrates into products at half its maximal rate and hence is related to the affinity of the substrate for the enzyme. The catalytic rate k_{cat} is the rate of product formation when the enzyme is saturated with substrate and therefore reflects the enzyme's maximum rate. The rate of product formation is dependent on both how well the enzyme binds substrate and how fast the enzyme converts substrate into product once substrate is bound. For a kinetically perfect enzyme, every encounter between enzyme and substrate leads to product and hence the reaction velocity is only limited by the rate the enzyme encounters substrate in solution. From Equation 9.2.8, the catalytic efficiency of a protein can be evaluated.

$$\frac{k_{cat}}{K_m} = \frac{k_2}{K_m} = \frac{k_1 k_2}{k_{-1} + k_2} \tag{9.2.13}$$

This k_{cat}/K_m ratio is called the specificity constant measure of how efficiently an enzyme converts a substrate into product. It has a theoretical upper limit of 10^8 – 10^{10} /M.s; enzymes working close to this, such as fumarase, are termed superefficient (Table 9.2.1).

Determining V_m and K_m from experimental data can be difficult and the most common way is to determine initial rates, v_0, from experimental values of $[P]$ or $[S]$ as a function of time. Hyperbolic graphs of v_0 vs. $[S]$ can be fit or transformed as we explored with the different mathematical transformations of the hyperbolic binding equation to determine K_d. These included:

- nonlinear hyperbolic fit (e.g., Figure 9.2.1)
- double reciprocal plot (e.g., Lineweaver–Burk plot discussed below
- Eadie-Hofstee plot

Lineweaver–Burk plot

Another commonly-used plot in examining enzyme kinetics is the **Lineweaver-Burk plot**, in with the inverse of the reaction rate, $1/r$, is plotted against the inverse of the substrate concentration $1/\left[\mathrm{S}\right]$. Rearranging Equation 9.2.10,

$$\frac{1}{r} = \frac{K_M + [\mathrm{S}]}{k_2[\mathrm{E}]_0\,[\mathrm{S}]} = \frac{K_M}{k_2[\mathrm{E}]_0}\frac{1}{[\mathrm{S}]} + \frac{1}{k_2[\mathrm{E}]_0} \tag{9.2.14}$$

Tthe Lineweaver–Burk plot (or double reciprocal plot) is a graphical representation of the Lineweaver–Burk equation of enzyme kinetics, described by Hans Lineweaver and Dean Burk in 1934 (Figure 9.2.2). The Lineweaver-Burk plot results in a straight line with the slope equal to $K_M/k_2[\mathrm{E}]_0$ and y-intercept equal to $1/k_2[\mathrm{E}]_0$ which is $1/V_{max}$ via Equation 9.2.10.

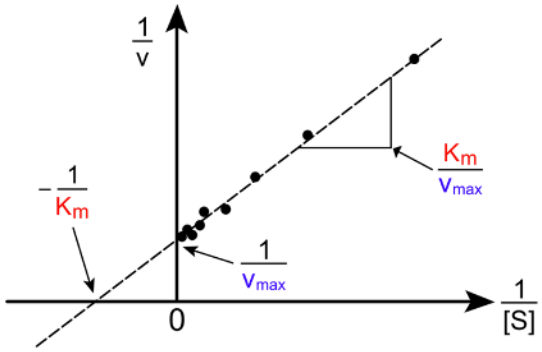


Figure 9.2.2*: Lineweaver–Burk plot of Michaelis–Menten kineitcs.*

The plot provides a useful graphical method for analysis of the Michaelis–Menten equation:

$$V = \frac{V_{\max}[S]}{K_m + [S]} \tag{9.2.15}$$

Taking the reciprocal gives

$$\frac{1}{V} = \frac{K_m + [S]}{V_{max}[S]} = \frac{K_m}{V_{max}}\frac{1}{[S]} + \frac{1}{V_{max}} \tag{9.2.16}$$

where

- V is the reaction velocity (the reaction rate),
- K_m is the Michaelis–Menten constant,
- V_{max} is the maximum reaction velocity, and
- $[S]$ is the substrate concentration.

The Lineweaver–Burk plot was widely used to determine important terms in enzyme kinetics, such as K_m and V_{max}, before the wide availability of powerful computers and non-linear regression software. The y-intercept of such a graph is equivalent to the inverse of V_{max}; the x-intercept of the graph represents $-1/K_m$. It also gives a quick, visual impression of the different forms of enzyme inhibition.

Example 9.2.2

The reaction between nicotineamide mononucleotide and ATP to form nicotineamide–adenine dinucleotide and pyrophosphate is catalyzed by the enzyme nicotinamide mononucleotide adenylyltransferase. The following table provides typical data obtained at a pH of 4.95. The substrate, S, is nicotinamide mononucleotide and the initial rate, *v*, is the μmol of nicotinamide–adenine dinucleotide formed in a 3-min reaction period.

[S] (mM)	v (μmol)	[S] (mM)	v (μmol)
0.138	0.148	0.560	0.324
0.220	0.171	0.766	0.390
0.291	0.234	1.460	0.493

Determine values for V_{max} and K_m.

Solution

Figure 13.12 shows the Lineweaver–Burk plot for this data and the resulting regression equation. Using the *y*-intercept, we calculate V_{max} as

$$V_{max} = 1 / y\text{−intercept} = 1 / 1.708 \text{ mol} = 0.585 \text{ mol}$$

and using the slope we find that K_m is

$$K_m = \text{slope} \times V_{max} = 0.7528 \text{ molimM} \times 0.585 \text{ mol} = 0.440 \text{ mM}$$

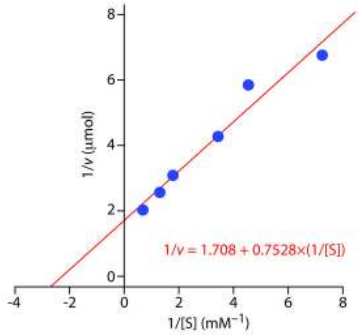
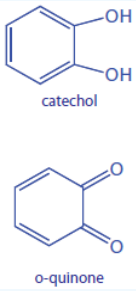


Figure 13.12: Linweaver–Burk plot and regression equation for the data in Example 13.6.

Exercise 9.2.2: o-diphenyl oxidase

The following data are for the oxidation of catechol (the substrate) to *o*-quinone by the enzyme *o*-diphenyl oxidase. The reaction is followed by monitoring the change in absorbance at 540 nm. The data in this exercise are adapted from jkimball.



[catechol] (mM)	0.3	0.6	1.2	4.8
rate (ΔAU/min)	0.020	0.035	0.048	0.081

Determine values for V_{max} and K_m.

Click here to review your answer to this exercise.

The double reciprocal plot distorts the error structure of the data, and it is therefore unreliable for the determination of enzyme kinetic parameters. Although it is still used for representation of kinetic data, non-linear regression or alternative linear forms of the Michaelis–Menten equation such as the **Hanes-Woolf plot** or **Eadie–Hofstee plot** are generally used for the calculation of parameters.

Problems with the Method

The Lineweaver–Burk plot is classically used in older texts, but is prone to error, as the *y*-axis takes the reciprocal of the rate of reaction – in turn increasing any small errors in measurement. Also, most points on the plot are found far to the right of the *y*-axis (due to limiting solubility not allowing for large values of $[S]$ and hence no small values for $1/[S]$), calling for a large extrapolation back to obtain *x*- and *y*-intercepts.

When used for determining the type of enzyme inhibition, the Lineweaver–Burk plot can distinguish competitive, non-competitive and uncompetitive inhibitors. Competitive inhibitors have the same y-intercept as uninhibited enzyme (since V_{max} is unaffected by competitive inhibitors the inverse of V_{max} also doesn't change) but there are different slopes and x-intercepts between the two data sets. Non-competitive inhibition produces plots with the same x-intercept as uninhibited enzyme (K_m is unaffected) but different slopes and y-intercepts. Uncompetitive inhibition causes different intercepts on both the y- and x-axes but the same slope.

Eadie–Hofstee Plot

The Eadie–Hofstee plot is a graphical representation of enzyme kinetics in which reaction rate is plotted as a function of the ratio between rate and substrate concentration and can be derived from the Michaelis–Menten equation (9.2.9) by inverting and multiplying with V_{max}:

$$\frac{V_{max}}{v} = \frac{V_{max}(K_m + [S])}{V_{max}[S]} = \frac{K_m + [S]}{[S]} \tag{9.2.17}$$

Rearrange:

$$V_{max} = \frac{vK_m}{[S]} + \frac{v[S]}{[S]} = \frac{vK_m}{[S]} + v \tag{9.2.18}$$

Isolate v:

$$v = -K_m \frac{v}{[S]} + V_{max} \tag{9.2.19}$$

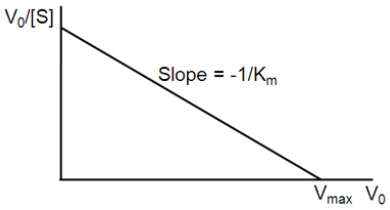


Figure 9.2.3*: The Eadie-Hofstee plot is a more accurate linear plotting method with v is plotted against v/[S].*

A plot of v against $v/[S]$ will hence yield V_{max} as the y-intercept, V_{max}/K_m as the x-intercept, and K_m as the negative slope (Figure 9.2.3). Like other techniques that linearize the Michaelis–Menten equation, the Eadie-Hofstee plot was used historically for rapid identification of important kinetic terms like K_m and V_{max}, but has been superseded by nonlinear regression methods that are significantly more accurate and no longer computationally inaccessible. It is also more robust against error-prone data than the Lineweaver–Burk plot, particularly because it gives equal weight to data points in any range of substrate concentration or reaction rate (the Lineweaver–Burk plot unevenly weights such points). Both Eadie-Hofstee and Lineweaver–Burk plots remain useful as a means to present data graphically.

Problems with the Method

One drawback from the Eadie–Hofstee approach is that neither ordinate nor abscissa represent independent variables: both are dependent on reaction rate. Thus any experimental error will be present in both axes. Also, experimental error or uncertainty will propagate unevenly and become larger over the abscissa thereby giving more weight to smaller values of v/[S]. Therefore, the typical measure of goodness of fit for linear regression, the correlation coefficient R, is not applicable.

Contributors

David Harvey (DePauw University)

- World Public Library
- Wikipedia
- Dr. S.K. Khare (IIT Delhi) via NPTEL

9.3: Chymotrypsin- A Case Study

Chymotrypsin is a digestive enzyme belonging to a super family of enzymes called serine proteases. It uses an active serine residue to perform hydrolysis on the C-terminus of the aromatic amino acids of other proteins. Chymotrypsin is a protease enzyme that cleaves on the C-terminal phenylalanine (F), tryptophan (W), and tyrosine (Y) on peptide chains. It shows specificity for aromatic amino acids because of its hydrophobic pocket.

Introduction

Chymotrypsin is one of the most studied enzymes due to its two phase kinetics: pre-steady-state and steady state. The study of these two kinetic states gives evidence of the "Ping-Pong" mechanism, the formation of covalent complexes leading to covalent hydrolysis reactions, and the rate of the catalyzed reactions. Synthesis of chymotrypsin occurs primarily in the pancreas. Instead of the active form, however, it is produced as an inactive zymogen called chymotrypsinogen to prevent its protease activity from digesting the pancreas. Upon secretion into the lumen of the small intestine, it is converted to its active form by another enzyme called trypsin. This dependence of a different enzyme for the activation of a protease is a common way for the body to prevent the digestion of organs and other harmful enzymatic side-effects.

Chymotrypsin operates through a general mechanism known as the ping-pong mechanism (Figure 9.3.1) whereby the enzyme reacts with a substrate to form an enzyme intermediate. This intermediate has different properties than the initial enzyme, so to regenerate the initial enzymatic activity, it must react with a secondary substrate. This process is illustrated below:

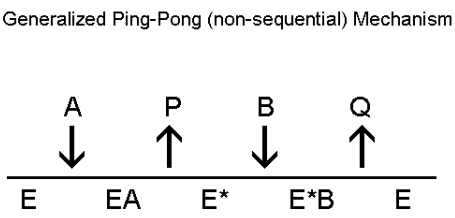
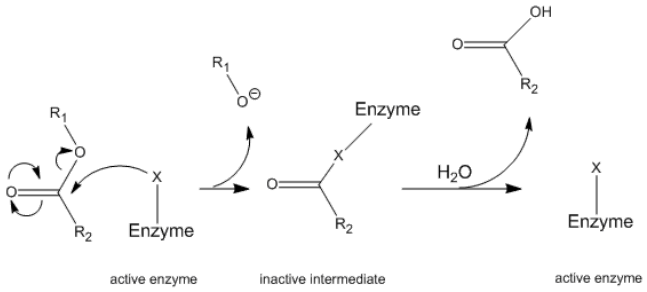


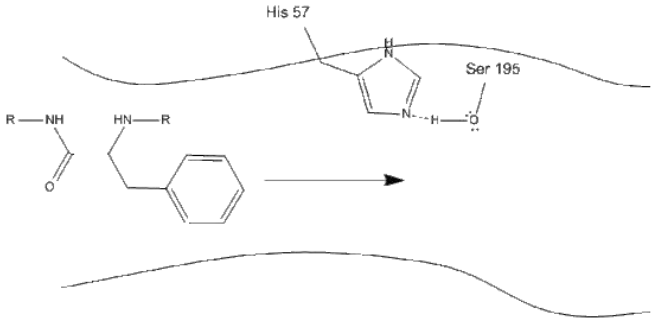
Figure 9.3.1: Generic Ping-Pong Mechanism

More specifically, chymotrypsin operates through a particular type of ping-pong mechanism called covalent hydrolysis. This means that the enzyme first forms a covalent bond with the target substrate, displacing the more stable moiety into solution. This enzyme-substrate complex is called the enzyme intermediate. The intermediate then reacts with water, which displaces the remaining part of the initial substrate and reforms the initial enzyme.

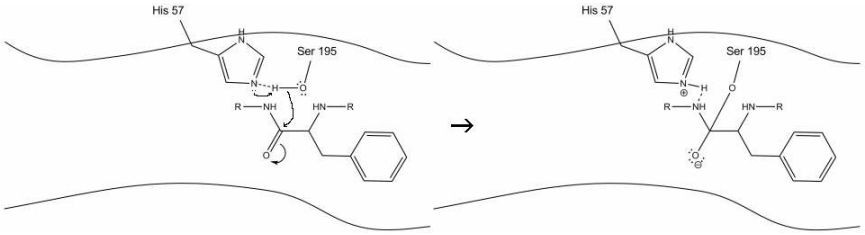


Chymotrypsin, like most enzymes, is specific in the types of substrates with which it reacts. As a protease, it cleaves polypeptides, and its inherent specificity allows it to act only on the carboxy-terminal of aromatic residues. It is a somewhat complicated mechanism, and is best explained in a series of steps.

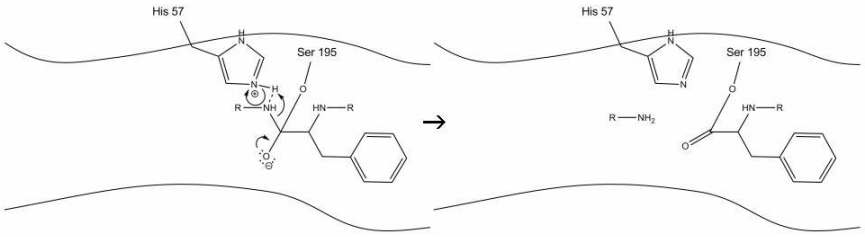
Step 1: The target enters the active site of chymotrypsin, and it is held there by hydrophobic interactions between exposed non-polar groups of enzyme residues and the non-polar aromatic side-chain of the substrate. It is important to note the hydrogen bond between the Schiff nitrogen on histidine-57 and the oxygen side-chain of serine-195.



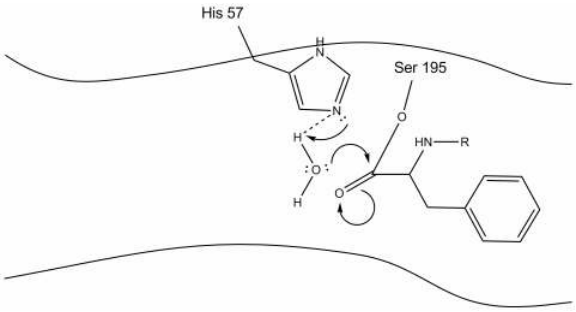
Step 2: Aided by the histidine-serine hydrogen bonding, the hydroxyl group on serine-195 performs a nucleophilic attack on the carbonyl carbon of an aromatic amino acid while simultaneously transferring the hydroxyl hydrogen to the histidine Schiff nitrogen. This attack pushes the pi carbonyl electrons onto the carbonyl oxygen, forming a short-lived intermediate consisting of a c-terminal carbon with four single bonds: an oxygen anion, the beta-carbon of the aromatic amino acid, the n-terminus of the subsequent amino acid of the substrate protein, and the serine-195 side-chain oxygen.



Step 3: This intermediate is short-lived, as the oxyanion electrons reform the pi bond with the c-terminus of the aromatic amino acid. The bond between the carboxy-terminus of the aromatic amino acid and the n-terminus of the subsequent residue is cleaved, and its electrons are used to extract the hydrogen of the protonated Schiff nitrogen on histidine-57. The bonds between the carbonyl carbon and the serine-195 oxygen remain in an ester configuration. This is called the acyl-enzyme intermediate. The c-terminal side of the polypeptide is now free to dissociate from the active site of the enzyme.

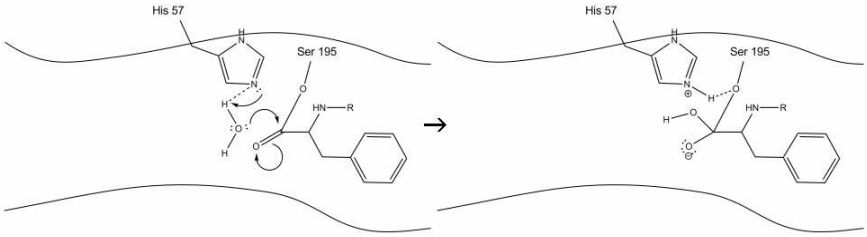


Step 4: Water molecules are now able to enter and bind to active site through hydrogen bonding between the hydrogen atoms of water and the histidine-57 Schiff nitrogen.

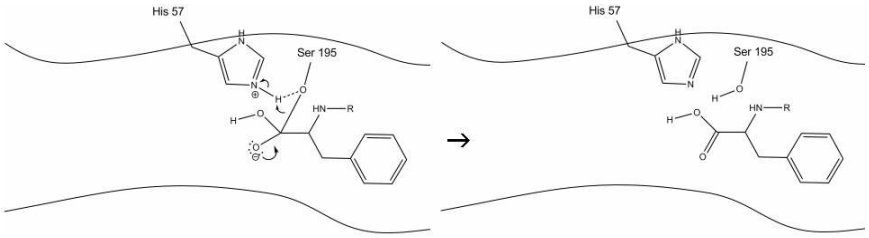


Step 5: The water oxygen now makes a nucleophilic attack on the carbonyl carbon of the acyl-enzyme intermediate, pushing the carbonyl's pi electrons onto the carbonyl carbon as histidine-57 extracts one proton from water. This forms another quaternary

carbon covalently bonded with serine, a hydroxyl, an oxyanion, and the aromatic amino acid. The proton on the recently protonated histidine-57 is now able to make a hydrogen bond with the serine oxygen.



Step 6: The oxyanion electrons reform the carbonyl pi bond, cleaving the bond between the carbonyl carbon and the serine hydroxyl. The electrons in this bond are used by the serine oxygen to deprotonate the histidine Schiff nitrogen and reform the original enzyme. The substrate no longer has affinity for the active site, and it soon dissociates from the complex.



Kinetics

Experiments were conducted in 1953 by B.S. Hartley and B.A. Kilby to investigate the kinetics of chymotrypsin-catalyzed hydrolysis. Instead of using a poly-peptide chain as a substrate, they used a nitro-phenyl ester, p-nitrophenyl acetate, that resembles an aromatic amino acid. Hydrolysis of this compound by chymotrypsin at the carbonyl group yields acetate and nitrophenolate, the latter of which absorbs near 400 nm light and its concentration can thus be measured by spectrophotometry (Figure 9.3.2).

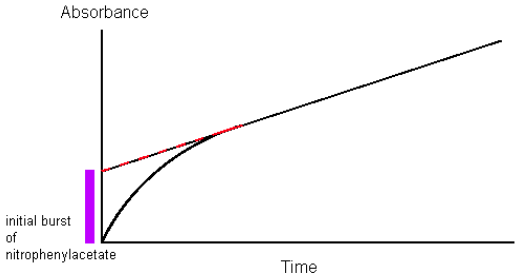
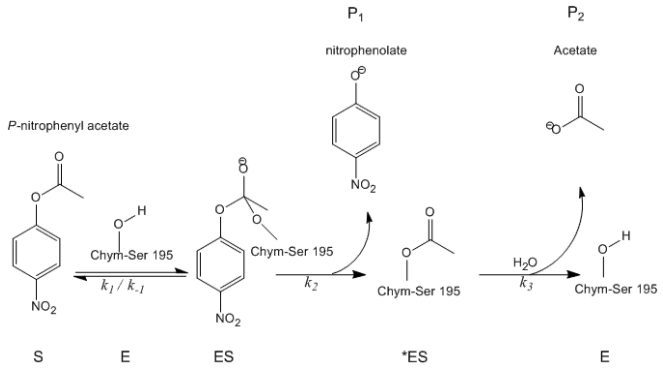


Figure 9.3.2*: Catalytic activity of he hydrolysis of p-nitropenolate under*

Spectrophotometric analysis of chymotrypsin acting on nitrophenylacetate showed that nitrophenolate was produced at a rate independent of substrate concentration, proving that the only factor contributing to the rate of product formation is the concentration of enzyme; this is typical for enzyme-substrate kinetics. However, when the slope of the 0-order absorbance plot was traced back to the starting point (time = 0), it was found that the initial concentration of nitrophenolate was not 0. In fact, it showed a 1:1 stoichiometric ratio with the amount of chymotrypsin used in the assay. This can only be explained by the fact that hydrolysis by chymotrypsin is biphasic in nature, meaning that it proceeds in two distinct steps.

1. The first step, which describes the **initial burst** of nitrophenolate seen in Hartley and Kilby's absorbance plot, is the fastest. The attack of the nitrophenyl acetate substrate by chymotrypsin immediately cleaves the nitrophenolate moiety and leaves the acetate group attached to chymotrypsin, rendering the enzyme inactive.
2. The second step has been deduced to involve the hydrolysis of the acetate group from the inactivated chymotrypsin to regenerate the original enzyme.

To analytically determine the rate of catalysis, all substrates, products, and intermediates must be defined. Refer to the figure below:



Using these abbreviations, kinetic analysis becomes less cumbersome.

1. The initial amount of enzyme can be represented as the sum of the free enzyme, the bound enzyme, and the inactive intermediate.

$$[E]_o = [ES] + [^*ES] + [E]$$

2. Assuming the initial step is the faster than the subsequent steps, the rate of nitrophenolate production can be described as:

$$\frac{d[P_1]}{dt} = k_2[ES]$$

3. Likewise, the rate of acetate formation can be represented by the equation:

$$\frac{d[P_2]}{dt} = k_3[^*ES]$$

4. Therefore, the net change in concentration of the inactive intermediate can be deduced:

$$\frac{d[*ES]}{dt} = k_2[ES] - k_3[^*ES]$$

5. The last inference that can be made from analysis of the measured kinetics data (Figure $9.3.2$) is that the first step of the reaction equilibrates rapidly, and thus the change in bound substrate can be described in the following equation. This is a principal tenet in analyzing **the kinetics of chymotrypsin and is a ubiquitous mechanism in biological enzyme** catalysis.

$$\frac{d[ES]}{dt} = k_1[E][S] - k_{-1}[ES] = 0$$

6. Where:

$$\frac{k_{-1}}{k_1} = K_s = \frac{[E][S]}{[ES]}$$

7. Combining all of these quantities, we can deduce the catalytic rate constant as:

$$k_{cat} = \frac{k_2 k_3}{k_2 + k_3}$$

8. In ester hydrolysis, $k_3 >> k_2$, so the resultant catalytic rate constant simplifies to:

$$k_{cat} = k_2$$

which is in agreement with the observed zeroth-order kinetics of Figure $9.3.2$.

Table 9.3.1: Kinetic Constants of the Chymotrypsin-Catalyzed Hydrolysis of p-Nitrophenyl Trimethylacetate at pH 8.2. 0.01 M tris-HCL buffer, ionic strength 0.06, 25.6 ± 0.1 °C, 1.8% (v/v) acetonitrile-water. From M.L. Bender, F.J. Kezdy and F.C. Wedler, J. Chem. Educ. 44, 84 (1967

k_2	$0.37 \pm 0.11\ s^{-1}$
k_3	$(1.3 \pm 0.02) \times 10^{-4}\ s^{-1}$
K_s	$(1.6 \pm 0.5) \times 10^{-3}$
k_{cat}	$1.3 \times 10\text{-}4\ s^{-1}$
K_M	5.6×10^{-7}

Exercise 9.3.1

Speculate on how the catalytic rate constant can be determined from the spectrophotogram.

Answer

The catalytic rate constant can be deduced from the graph by simply determining the slope of the line where the reaction demonstrates 0-order kinetics (the linear part)

Exercise 9.3.2

How can product be consistently produced if the rate of change of the ES complex is 0?

Answer

This is pre-equilibrium kinetics in action. The ES complex is formed from E and S at a faster rate than any other step in the reaction. As soon as ES is converted to *ES, another mole of ES is produced from an infinite supply of E + S. This means that the amount of ES and E + S is constantly at equilibrium, and thus the change of either with respect to time is 0.

Exercise 9.3.3

How would the rate of product formation change if:

a. the substrate concentration was doubled.
b. the enzyme concentration was doubled.
c. The reaction was carried out in mono-deuterated water instead of H2O (comment qualitatively).

Answer

a. No change.
b. Two fold increase.
c. Since water is involved in the final, slowest step of the mechanism, deuterating the water would decrease the rate of the overall reaction from 5 to 30-fold.

Exercise 9.3.4

Explain the role of hydrogen bonding in protein hydrolysis catalyzed by chymotrypsin.

Answer

Initially, hydrogen bonding between the enzymes histidine and serine side chains weakens the bond of serine's O-H. This allows for a facilitated nucleophilic attack of the hydroxyl Oxygen on the substrates carbonyl group. Conversely, in the final step of the reaction, the bound serine oxygen forms a hydrogen bond with a protonated histidine, which allows for easier cleavage from the substrate.

Exercise 9.3.5

What would the spectrophotogram look like if the reaction proceeded via a steady-state mechanism instead of pre-equilibrium.

Answer

The graph would show similar 0-order kinetics, but the line would intercept the Y-axis at an absorbance of 0 instead of the 1:1 mole ratio of nitrophenolate to enzyme.

References

1. Chang, Raymond. Physical Chemistry for the Biosciences. Sansalito, CA: University Science, 2005.
2. Segel, Irwin H., Enzyme Kinetics: Behavior and Analysis of Rapid Equilibrium and Steady-State Enzyme Systems (Wiley Classics Library)
3. Garret, Reginald. Grishman, Charles. Biochemistry. University of Virginia, 2007
4. Whitaker, John., Princles of Enzymology for the Food Sciences. University of California, Davis. 1994
5. “Chymotrypsin Mechanism” viewable at: www.sumanasinc.com/webcontent...motrypsin.html
6. Sumanas Inc, viewed on February 25th, 2009.
7. M.L. Bender, F.J. Kezdy and F.C. Wedler, J. Chem. Educ. 44, 84 (1967)

9.4: Multisubstrate Systems

The Michaelis –Menten model of enzyme kinetics was derived for single substrate reactions. Enzymatic reactions requiring multiple substrates and yielding multiple products are more common and yielding multiple products are more common than single-substrate reaction. In these types of reactions, the all the substrates involved are bound to the enzyme before catalysis of the reaction takes place to release the products. Sequential reactions can be either ordered or random. In contrast to the Michealis-Menton kinetics where a binary Enzyme-Substrate complex is generated in the mechanism ($[ES]$, in bisubstrate enzyme reactions, a ternary complex of the enzyme and two substrates is generated:

$$A + B \overset{E}{\rightleftharpoons} P + Q \tag{9.4.1}$$

Bisubstrate reactions account for ~ 60% of the known enzymatic reactions. Multi-substrate reactions follow complex rate equations that describe how the substrates bind and in what sequence. The analysis of these reactions is much simpler if the concentration of substrate A is kept constant and substrate B varied. Under these conditions, the enzyme behaves just like a single-substrate enzyme and a plot of v by $[S]$ gives apparent K_M and V_{max} constants for substrate B. If a set of these measurements is performed at different fixed concentrations of A, these data can be used to work out what the mechanism of the reaction is. For an enzyme that takes two substrates A and B and turns them into two products P and Q, there are two types of mechanism: ternary complex and ping–pong.

How do you resolve the enzymes kinetics of these more complicated systems? The answer is fairly straightforward. You keep one of the substrates (B, for example) fixed, and vary the other substrate (A) and obtain a **series** of hyperbolic plots of v_o vs A at different fixed B concentrations. This would give a series of linear $1/v$ vs. $1/A$ double-reciprocal plots (Lineweaver-Burk plots) as well. The pattern of Lineweaver-Burk plots depends on how the reactants and products interact with the enzyme.

Sequential Mechanism

In this mechanism, both substrates must bind to the enzyme before any products are made and released.The substrates might bind to the enzyme in a random fashion (A first then B or vice-versa) or in an ordered fashion (A first followed by B). An abbreviated notation scheme developed by W.W. Cleland is shown in Figure 9.4.1 for the sequential random and sequential ordered mechanisms (Figure 9.4.1). For both mechanisms, Lineweaver-Burk plots at varying A and different fixed values of B give a series of intersecting lines. Derivative curves can be solved to obtain appropriate kinetic constants. In ordered sequential reactions, all the substrates are first bound to the enzyme in a **defined order or sequence**. The products, too, are released after catalysis in a defined order or sequence.

An example is the lactate dehydrogenase enzyme, which is a protein that catalyzes glucose metabolism. In this ordered mechanism, the coenzyme, NADH, always binds first, with pyruvate binding afterward. During the reaction, the pyruvate is reduced to lactate while NADH is oxidized to NAD^+ by the enzyme. Lactate is then released first, followed by the release of NAD^+.

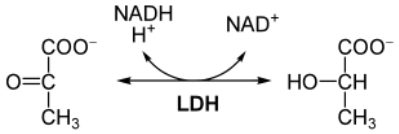


Figure 9.4.2*: Reaction of the lactate dehydrogenase: pyruvate (left) is oxidized to lactate (right) by expense of NADH*

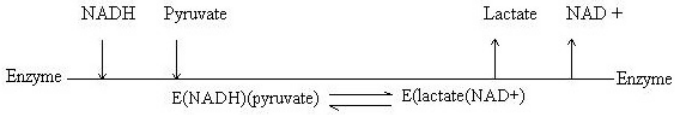


Figure 9.4.3*:* Ordered Sequential Mechanism for the lactate dehydrogenase enzyme

This is a characteristic of a ternary complex, which consists of three molecules that are bound together. Before catalysis, the substrates and coenzyme are bound to the enzyme. After catalysis, the complex consists of the enzyme and products, NAD^+ and lactate.

In random sequential reactions, the substrates and products are bound and then released in no preferred order, or "random" order (Figure 9.4.1). An example is the creatine kinase enzyme, which catalyzes creatine and ATP (the two substrates) to form phosphocreatine and ADP (teh Products) in Figure 9.4.4. In this case, **either** substrates may bind first and either products may be

released first. A ternary complex is still observed for random sequential reactions. Before catalysis, the complex is generated that includes the enzyme, ATP, and creatine. After catalysis, the complex consists of the enzyme, ADP, and phosphocreatine.

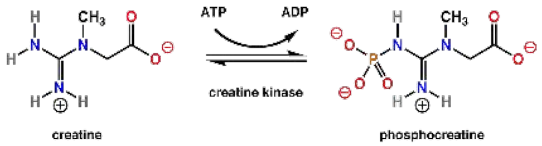


Figure 9.4.4: The metabolism of creatine follows a random sequential mechanism.

Ping-Pong Mechanism

In this mechanism, one substrate bind first to the enzyme followed by product P release. Typically, product P is a fragment of the original substrate A.The rest of the substrate is covalently attached to the enzyme E, which is designated as E'. Now the second reactant, B, binds and reacts with the enzyme to form a covalent adduct with the A as it is covalentattached to the enzyme to form product Q. This is now released and the enzyme is restored to its initial form, E. This represents a ping-pong mechanism. An abbreviated notation scheme is shown below for the ping-pong mechanisms. For this mechanism, Lineweaver-Burk plots at varying A and different fixed values of B give a series of parallel lines. An example of this type of reaction might be low molecular weight protein tyrosine phosphatase against the small substrate p-initrophenylphosphate (A) which binds to the enzyme covalently with the expulsion of the product P, the p-nitrophenol leaving group. Water (B) then comes in and covalently attacks the enzyme, forming an adduct with the covalently bound phosphate releasing it as inorganic phosphate. In this particular example, however, you cannot vary the water concentration and it would be impossible to generate the parallel **Lineweaver-Burk** plots characteristic of ping-pong kinetics.

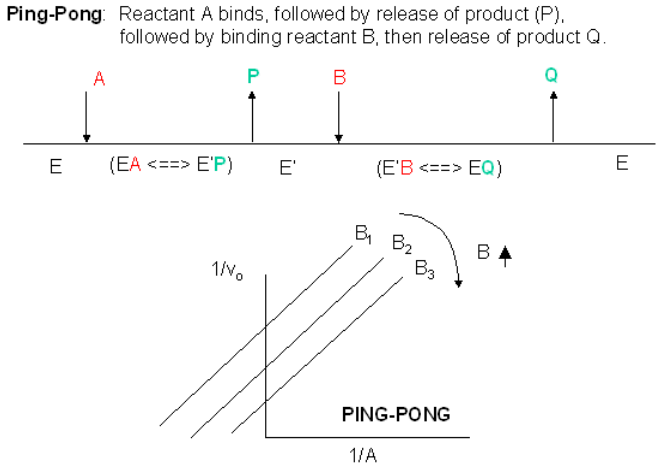


Figure 9.4.5): The Ping-Pong mechanism and associated **Lineweaver-Burk**.

Contributors

- Prof. Henry Jakubowski (College of St. Benedict/St. John's University)

9.5: Enzyme Inhibition

Enzymes can be regulated in ways that either promote or reduce their activity. There are many different kinds of molecules that inhibit or promote enzyme function, and various mechanisms exist for doing so. In some cases of enzyme inhibition, for example, an inhibitor molecule is similar enough to a substrate that it can bind to the active site and simply block the substrate from binding. When this happens, the enzyme is inhibited through competitive inhibition, because an inhibitor molecule competes with the substrate for active site binding. On the other hand, in noncompetitive inhibition, an inhibitor molecule binds to the enzyme in a location other than an allosteric site and still manages to block substrate binding to the active site.

Elucidating Mechanisms for the Inhibition of Enzyme Catalysis

When an **inhibitor** interacts with an enzyme it decreases the enzyme's catalytic efficiency. An irreversible inhibitor covalently binds to the enzyme's active site, producing a permanent loss in catalytic efficiency even if we decrease the inhibitor's concentration. A reversible inhibitor forms a noncovalent complex with the enzyme, resulting in a temporary decrease in catalytic efficiency. If we remove the inhibitor, the enzyme's catalytic efficiency returns to its normal level.

There are several pathways for the reversible binding of an inhibitor to an enzyme, as shown in Figure 9.5.1. In **competitive inhibition** the substrate and the inhibitor compete for the same active site on the enzyme. Because the substrate cannot bind to an enzyme–inhibitor complex, EI, the enzyme's catalytic efficiency for the substrate decreases. With **noncompetitive inhibition** the substrate and the inhibitor bind to different active sites on the enzyme, forming an enzyme–substrate–inhibitor, or ESI complex. The formation of an ESI complex decreases catalytic efficiency because only the enzyme–substrate complex reacts to form the product. Finally, in **uncompetitive inhibition** the inhibitor binds to the enzyme–substrate complex, forming an inactive ESI complex.

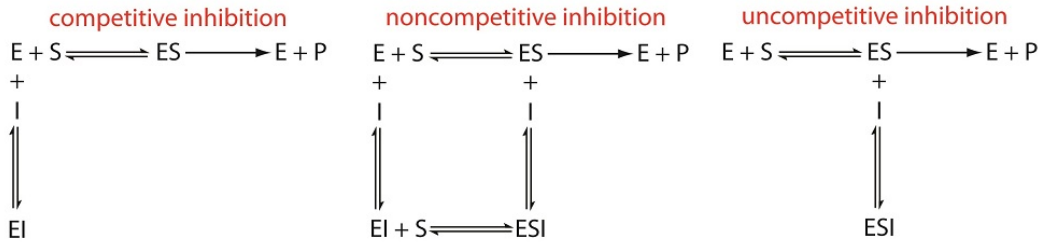


Figure 9.5.1: *Mechanisms for the reversible inhibition of enzyme catalysis. E: enzyme, S: substrate, P: product, I: inhibitor, ES: enzyme–substrate complex, EI: enzyme–inhibitor complex, ESI: enzyme–substrate–inhibitor complex.*

We can identify the type of reversible inhibition by observing how a change in the inhibitor's concentration affects the relationship between the rate of reaction and the substrate's concentration. As shown in Figure 13.14, when we display kinetic data using as a Lineweaver-Burk plot it is easy to determine which mechanism is in effect. For example, an increase in slope, a decrease in the *x*-intercept, and no change in the *y*-intercept indicates competitive inhibition. Because the inhibitor's binding is reversible, we can still obtain the same maximum velocity—thus the constant value for the *y*-intercept—by adding enough substrate to completely displace the inhibitor. Because it takes more substrate, the value of K_m increases, which explains the increase in the slope and the decrease in the *x*-intercept's value.

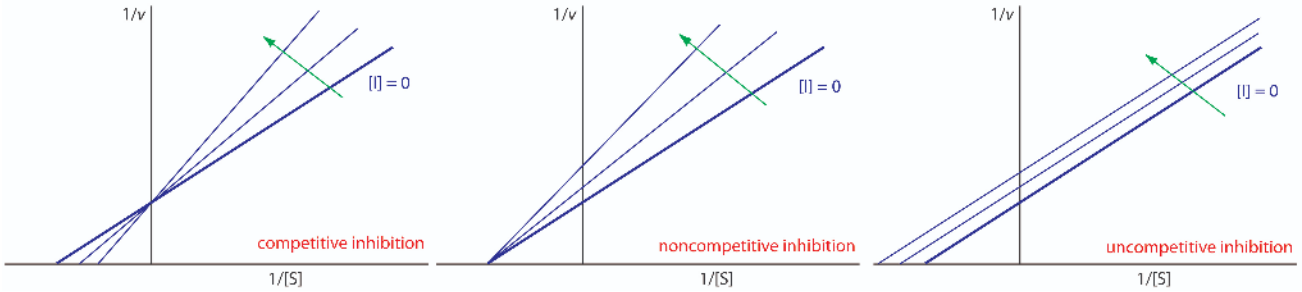


Figure 9.5.2: Linweaver–Burk plots for competitive inhibition, noncompetitive inhibition, and uncompetitive inhibition. The thick blue line in each plot shows the kinetic behavior in the absence of inhibitor, and the thin blue lines in each plot show the change in behavior for increasing concentrations of the inhibitor. In each plot, the inhibitor's concentration increases in the direction of the green arrow.

Example 9.5.1

Practice Exercise 13.3 provides kinetic data for the oxidation of catechol (the substrate) to *o*-quinone by the enzyme *o*-diphenyl oxidase in the absence of an inhibitor. The following additional data are available when the reaction is run in the presence of *p*-hydroxybenzoic acid, PBHA. Is PBHA an inhibitor for this reaction and, if so, what type of inhibitor is it?

[catechol] (mM)	0.3	0.6	1.2	4.8
rate (ΔAU/min)	0.011	0.019	0.022	0.060

Solution

Figure 9.5.3 shows the resulting Lineweaver–Burk plot for the data in Practice Exercise 13.3 and Example 13.7. Although the *y*-intercepts are not identical in value—the result of uncertainty in measuring the rates—the plot suggests that PBHA is a competitive inhibitor for the enzyme's reaction with catechol.

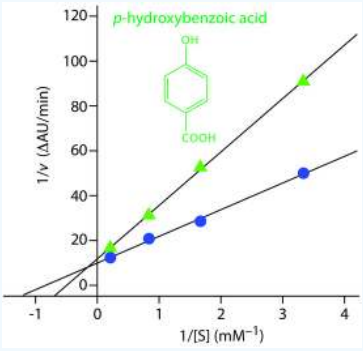


Figure 9.5.3: Lineweaver–Burk plots for the data in Practice Exercise 13.3 and Example 13.7.

Exercise 9.5.1

Practice Exercise 13.3 provides kinetic data for the oxidation of catechol (the substrate) to *o*-quinone by the enzyme *o*-diphenyl oxidase in the absence of an inhibitor. The following additional data are available when the reaction is run in the presence of phenylthiourea. Is phenylthiourea an inhibitor for this reaction and, if so, what type of inhibitor is it? The data in this exercise are adapted from jkimball.

[catechol] (mM)	0.3	0.6	1.2	4.8
rate (ΔAU/min)	0.010	0.016	0.024	0.040

Click here to review your answer to this exercise.

Contributors

- David Harvey (DePauw University)

9.6: Allosteric Interactions

Learning Objectives

- When a substrate binds to one enzymatic subunit, the rest of the subunits are stimulated and become active. Ligands can either have non-cooperativity, positive cooperativity or negative cooperativity.

A significant portion of enzymes function such that their properties can be studied using the Michaelis-Menten equation. However, a particular class of enzymes exhibit kinetic properties that cannot be studied using the Michaelis-Menten equation. The rate equation of these unique enzymes is characterized by an “S-shaped” sigmoidal curve, which is different from the majority of enzymes whose rate equation exhibits hyberbolic curves. Allosteric regulation is the regulation of an enzyme or other protein by binding an effector molecule at the protein's allosteric site (that is, a site other than the protein's active site). Effectors that enhance the protein's activity are referred to as allosteric activators, whereas those that decrease the protein's activity are called allosteric inhibitors. The term allostery refers to the fact that the regulatory site of an allosteric protein is **physically distinct** from its active site. Allosteric regulations are a natural example of control loops, such as feedback from downstream products or feedforward from upstream substrates. Long-range allostery is especially important in cell signaling.

Allosteric Modulation (Cooperativity)

Cooperativity is a phenomenon displayed by enzymes or receptors that have multiple binding sites where the affinity of the binding sites for a ligand is increased, positive cooperativity, or decreased, negative cooperativity, upon the binding of a ligand to a binding site. We also see cooperativity in large chain molecules made of many identical (or nearly identical) subunits (such as DNA, proteins, and phospholipids), when such molecules undergo phase transitions such as melting, unfolding or unwinding. This is referred to as subunit cooperativity (discussed below).

An example of positive cooperativity is the binding of oxygen to hemoglobin. One oxygen molecule can bind to the ferrous iron of a heme molecule in each of the four chains of a hemoglobin molecule. Deoxy-hemoglobin has a relatively low affinity for oxygen, but when one molecule binds to a single heme, the oxygen affinity increases, allowing the second molecule to bind more easily, and the third and fourth even more easily. The oxygen affinity of 3-oxy-hemoglobin is ~300 times greater than that of deoxy-hemoglobin. This behavior leads the affinity curve of hemoglobin to be sigmoidal, rather than hyperbolic as with the monomeric myoglobin. By the same process, the ability for hemoglobin to lose oxygen increases as fewer oxygen molecules are bound.

Negative allosteric modulation (also known as allosteric inhibition) occurs when the binding of one ligand **decreases** the affinity for substrate at other active sites. For example, when 2,3-BPG binds to an allosteric site on hemoglobin, the affinity for oxygen of all subunits decreases. This is when a regulator is absent from the binding site.

Another instance in which negative allosteric modulation can be seen is between ATP and the enzyme Phosphofructokinase within the negative feedback loop that regulates glycolysis. Phosphofructokinase (generally referred to as PFK) is an enzyme that catalyses the third step of glycolysis: the phosphorylation of Fructose-6-phosphate into Fructose 1,6-bisphosphate. PFK can be allosterically inhibited by high levels of ATP within the cell. When ATP levels are high, ATP will bind to an allosteric site on phosphofructokinase, causing a change in the enzyme's three-dimensional shape. This change causes its affinity for substrate (fructose-6-phosphate and ATP) at the active site to decrease, and the enzyme is deemed inactive. This causes glycolysis to cease when ATP levels are high, thus conserving the body's glucose and maintaining balanced levels of cellular ATP. In this way, ATP serves as a negative allosteric modulator for PFK, despite the fact that it is also a substrate of the enzyme.

Sigmoidal kinetic profiles are the result of enzymes that demonstrate positive cooperative binding. cooperativity refers to the observation that binding of the substrate or ligand at one binding site affects the affinity of other sites for their substrates. For enzymatic reactions with multiple substrate binding sites, this increased affinity for the substrate causes a rapid and coordinated increase in the velocity of the reaction at higher $[S]$ until V_{max} is achieved. Plotting the V_0 vs. $[S]$ for a cooperative enzyme, we observe the characteristic sigmoidal shape with low enzyme activity at low substrate concentration and a rapid and immediate increase in enzyme activity to V_{max} as $[S]$ increases. The phenomenon of cooperativity was initially observed in the oxygen-hemoglobin interaction that functions in carrying oxygen in blood. Positive cooperativity implies allosteric binding – binding of the ligand at one site increases the enzyme’s affinity for another ligand at a site different from the other site. Enzymes that demonstrate cooperativity are defined as allosteric. There are several types of allosteric interactions: homotropic (positive) and heterotropic (negative).

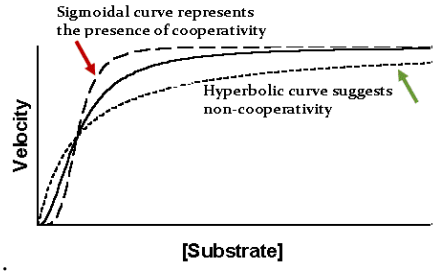


***Figure 1**: Rate of Reaction (velocity) vs. Substrate Concentration.*

Positive and negative allosteric interactions (as illustrated through the phenomenon of cooperativity) refer to the enzyme's binding affinity for other ligands at other sites, as a result of ligand binding at the initial binding site. When the ligands interacting are all the same compounds, the effect of the allosteric interaction is considered homotropic. When the ligands interacting are different, the effect of the allosteric interaction is considered heterotropic. It is also very important to remember that allosteric interactions tend to be driven by ATP hydrolysis.

The Hill Equation

The degree of cooperativity is determined by Hill equation (Equation 9.6.1) for non-Michaelis-Menten kinetics. The Hill equation accounts for allosteric binding at sites other than the active site. n is the "Hill coefficient."

$$\theta = \frac{[L]^n}{K_d + [L]^n} = \frac{[L]^n}{K_a^n + [L]^n} \tag{9.6.1}$$

where

- θ is the fraction of ligand binding sites filled
- $[L]$ is the ligand concentration
- K_d is the apparent dissociation constant derived from the law of mass action (equilibrium constant for dissociation)
- K_a is the ligand concentration producing half occupation (ligand concentration occupying half of the binding sides), that is also the microscopic dissociation constant
- n is the Hill coefficient that describes the cooperativity

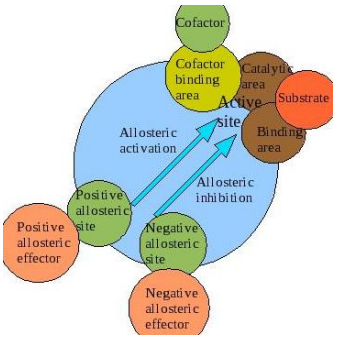
Taking the logarithm of both sides of the equation leads to an alternative formulation of the Hill Equation.

$$\log\left(\frac{\theta}{1-\theta}\right) = n\log[L] - \log K_d \tag{9.6.2}$$

- when $n < 1$, there is negative cooperativity
- when $n = 1$, there is no cooperativity
- when $n > 1$, there is positive cooperativity

Allosteric Models

Currently, there are 2 models for illustrating cooperativity: the concerted model and the sequential model. Most allosteric effects can be explained by the concerted MWC model put forth by Monod, Wyman, and Changeux, or by the sequential model described by Koshland, Nemethy, and Filmer. Both postulate that enzyme subunits exist in one of two conformations, tensed (T) or relaxed (R), and that relaxed subunits bind substrate more readily than those in the tense state. The two models differ most in their assumptions about subunit interaction and the preexistence of both states.



The Concerted model

The concerted model of allostery, also referred to as the symmetry model or MWC model, postulates that enzyme subunits are connected in such a way that a conformational change in one subunit is necessarily conferred to all other subunits. Thus, all subunits must exist in the same conformation. The model further holds that, in the absence of any ligand (substrate or otherwise), the equilibrium favours one of the conformational states, T or R. The equilibrium can be shifted to the R or T state through the binding of one ligand (the allosteric effector or ligand) to a site that is different from the active site (the allosteric site).

The Sequential model

The sequential model of allosteric regulation holds that subunits are not connected in such a way that a conformational change in one induces a similar change in the others. Thus, all enzyme subunits do not necessitate the same conformation. Moreover, the sequential model dictates that molecules of substrate bind via an induced fit protocol. In general, when a subunit randomly collides with a molecule of substrate, the active site, in essence, forms a glove around its substrate. While such an induced fit converts a subunit from the tensed state to relaxed state, it does not propagate the conformational change to adjacent subunits. Instead, substrate-binding at one subunit only slightly alters the structure of other subunits so that their binding sites are more receptive to substrate. To summarize:

- subunits need not exist in the same conformation
- molecules of substrate bind via induced-fit protocol
- conformational changes are not propagated to all subunits

Note: Allosteric database

Allostery is a direct and efficient means for regulation of biological macromolecule function, produced by the binding of a ligand at an allosteric site topographically distinct from the orthosteric site. Due to the often high receptor selectivity and lower target-based toxicity, allosteric regulation is also expected to play an increasing role in drug discovery and bioengineering. The AlloSteric Database (ASD, provides a central resource for the display, search and analysis of the structure, function and related annotation for allosteric molecules. Currently, ASD contains allosteric proteins from more than 100 species and modulators in three categories (activators, inhibitors, and regulators). Each protein is annotated with detailed description of allostery, biological process and related diseases, and each modulator with binding affinity, physicochemical properties and therapeutic area. Integrating the information of allosteric proteins in ASD should allow the prediction of allostery for unknown proteins, to be followed with experimental validation. In addition, modulators curated in ASD can be used to investigate potential allosteric targets for a query compound, and can help chemists to implement structure modifications for novel allosteric drug design.

Summary

Allosteric enzymes are an exception to the Michaelis-Menten model. Because they have more than two subunits and active sites, they do not obey the Michaelis-Menten kinetics, but instead have sigmoidal kinetics. Since allosteric enzymes are cooperative, a sigmoidal plot of V_0 versus $[S]$ results: There are distinct properties of Allosteric Enzymes that makes it different compared to other enzymes.

1. One is that allosteric enzymes do not follow the Michaelis-Menten Kinetics. This is because allosteric enzymes have multiple active sites. These multiple active sites exhibit the property of cooperativity, where the binding of one active site affects the affinity of other active sites on the enzyme. As mentioned earlier, it is these other affected active sites that result in a sigmoidal curve for allosteric enzymes.

2. Allosteric Enzymes are influenced by substrate concentration. For example, at high concentrations of substrate, more enzymes are found in the R state. The T state is favorite when there is an insufficient amount of substrate to bind to the enzyme. In other words, the T and R state equilibrium depends on the concentration of the substrate.
3. Allosteric Enzymes are regulated by other molecules. This is seen when the molecules 2,3-BPG, pH, and CO2 modulates the binding affinity of hemoglobin to oxygen. 2,3-BPG reduces binding affinity of O2 to hemoglobin by stabilizing the T- state. Lowering the pH from physiological pH=7.4 to 7.2 (pH in the muscles and tissues) favors the release of O_2. Hemoglobin is more likely to release oxygen in CO_2 rich areas in the body.

There are two primary models for illustrating cooperativity. **The concerted model** (also called the **Monod-Wyman-Changeux** model) illustrates cooperativity by assuming that proteins have two or more subunits, and that each part of the protein molecule is able to exist in either the relaxed (R) state or the tense (T) state - the tense state of a protein molecule is favored when it doesn't have any substrates bound. All aspects, including binding and dissociation constants are the same for each ligand at the respective binding sites. **The sequential model** aims to demonstrate cooperativity by assuming that the enzyme/protein molecule affinity is relative and changes as substrates bind. Unlike the concerted model, the sequential model accounts for different species of the protein molecule.

Contributors

Template:ContribWorldPublicLibrary

9.7: The Effect of pH on Enzyme Kinetics

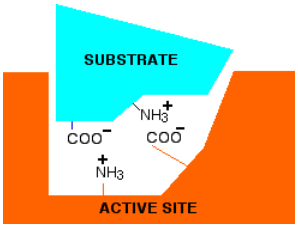
In the same way that every enzyme has an optimum temperature, so each enzyme also has an optimum pH at which it works best. For example, trypsin and pepsin are both enzymes in the digestive system which break protein chains in the food into smaller bits - either into smaller peptide chains or into individual amino acids. Pepsin works in the highly acidic conditions of the stomach. It has an optimum pH of about 1.5. On the other hand, trypsin works in the small intestine, parts of which have a pH of around 7.5. Trypsin's optimum pH is about 8.

Table 9.7.1: pH for Optimum Activity

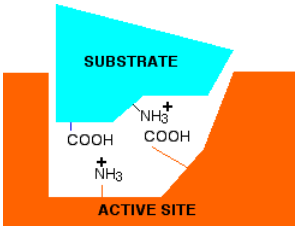
Enzyme	Optimal pH	Enzyme	Optimal pH
Lipase (pancreas)	8.0	Invertase	4.5
Lipase (stomach)	4.0 - 5.0	Maltase	6.1 - 6.8
Lipase (castor oil)	4.7	Amylase (pancreas)	6.7 - 7.0
Pepsin	1.5 - 1.6	Amylase (malt)	4.6 - 5.2
Trypsin	7.8 - 8.7	Catalase	7.0
Urease	7.0		

If you think about the structure of an enzyme molecule, and the sorts of bonds that it may form with its substrate, it isn't surprising that pH should matter. Suppose an enzyme has an optimum pH around 7. Imagine that at a pH of around 7, a substrate attaches itself to the enzyme via two ionic bonds. In the diagram below, the groups allowing ionic bonding are caused by the transfer of a hydrogen ion from a -COOH group in the side chain of one amino acid residue to an $-NH_2$ group in the side chain of another.

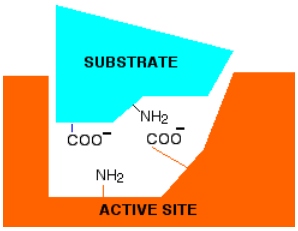
In this simplified example, that is equally true in both the substrate and the enzyme.



Now think about what happens at a lower pH - in other words under acidic conditions. It won't affect the $-NH_3^+$ group, but the $-COO^-$ will pick up a hydrogen ion. What you will have will be this:



You no longer have the ability to form ionic bonds between the substrate and the enzyme. If those bonds were necessary to attach the substrate and activate it in some way, then at this lower pH, the enzyme won't work. What if you have a pH higher than 7 - in other words under alkaline conditions. This time, the $-COO^-$ group won't be affected, but the $-NH_3^+$ group will lose a hydrogen ion. That leaves . . .



Again, there is no possibility of forming ionic bonds, and so the enzyme probably won't work this time either. At extreme pH's, something more drastic can happen. Remember that the tertiary structure of the protein is in part held together by ionic bonds just like those we've looked at between the enzyme and its substrate. At very high or very low pH's, these bonds within the enzyme can be disrupted, and it can lose its shape. If it loses its shape, the active site will probably be lost completely. This is essentially the same as denaturing the protein by heating it too much.

Kinetics

The rates of enzyme-catalysed reactions vary with pH and often pass through a maximum as the pH is varied. If the enzyme obeys Michaelis-Menten kinetics the kinetic parameters k_0 and k_A often behave similarly. The pH at which the rate or a suitable parameter is a maximum is called the *pH optimum* and the plot of rate or parameter against pH is called a *pH profile*. Neither the pH optimum nor the pH profile of an enzyme has any absolute significance and both may vary according to which parameter is plotted and according to the conditions of the measurements.

If the pH is changed and then brought back to its original value, the behavior is said to be *reversible* if the original properties of the enzyme are restored; otherwise it is *irreversible*. Reversible pH behavior may occur over a narrow range of pH, but effects of large changes in pH are in most cases irreversible. The diminution in rate as the pH is taken to the acid side of the optimum can be regarded as inhibition by hydrogen ions. The diminution in rate on the alkaline side can be regarded as inhibition by hydroxide ions. The equations describing pH effects are therefore analogous to inhibition equations. For single-substrate reactions the pH behavior of the parameters k_0 and k_A can sometimes be represented by an equation of the form

$$k = \frac{k_{opt}}{1 + \frac{[H^+]}{K_1} + \frac{K_2}{[H^+]}} \tag{9.7.1}$$

in which k represents k_0 or k_A, and k_{opt} is the value of the same parameter that would be observed if the enzyme existed entirely in the optimal state of protonation; it may be called the *pH-independent* value of the parameter. The constants K_1 and K_2 can sometimes be identified as acid dissociation constants for the enzyme. substrates or other species in the reaction mixture. The identification is, however, never straight forward and has to be justified by independent evidence. The behavior is frequently much more complicated than represented by Equation 9.7.1.

It is not accidental that this section has referred exclusively to pH dependences of k_0 and k_A. The pH dependence of the initial rate or, worse, the extent of reaction after a given time is rarely meaningful; the pH dependence of the Michaelis constant is often too complex to be readily interpretable.

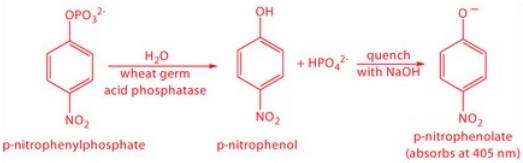
> *The pH dependence of the Michaelis constant is often too complex to be readily interpretable.*

Quenching Enzyme Activity

When using Representative Method 13.1 to determine the concentration of creatinine in urine, we follow the reactions kinetics using an ion selective electrode. In principle, we can use any of the analytical techniques in Chapters 8–12 to follow a reaction’s kinetics provided that the reaction does not proceed to any appreciable extent during the time it takes to make a measurement. As you might expect, this requirement places a serious limitation on kinetic methods of analysis. If the reaction’s kinetics are slow relative to the analysis time, then we can make our measurements without the analyte undergoing a significant change in concentration. When the reaction’s rate is too fast—which often is the case—then we introduce a significant error if our analysis time is too long.

One solution to this problem is to stop, or **quench** the reaction by adjusting experimental conditions. For example, many reactions show a strong pH dependency, and may be quenched by adding a strong acid or a strong base. Figure 13.7 shows a typical example

for the enzymatic analysis of *p*-nitrophenylphosphate using the enzyme wheat germ acid phosphatase to hydrolyze the analyte to *p*-nitrophenol.



The reaction has a maximum rate at a pH of 5. Increasing the pH by adding NaOH quenches the reaction and converts the colorless *p*-nitrophenol to the yellow-colored *p*-nitrophenolate, which absorbs at 405 nm.

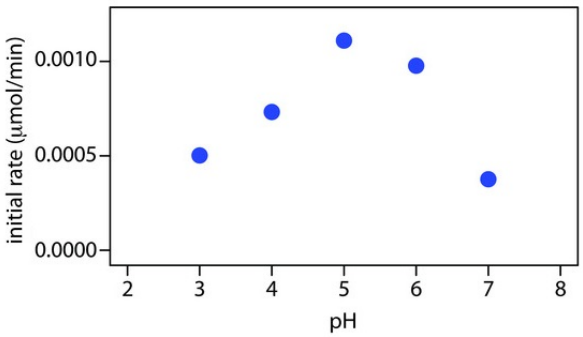


Figure 13.7: Initial rate for the enzymatic hydrolysis of p-nitrophenylphosphate using wheat germ acid phosphatase. Increasing the pH quenches the reaction and coverts colorless p-nitrophenol to the yellow-colored p-nitrophenolate, which absorbs at 405 nm. The data are adapted from socrates.hunter.cuny.edu.

Contributors

- Template:Clark

9.8: The Effect of Temperature on Enzyme Kinetics

Enzymes are generally globular proteins, acting alone or in larger complexes. Like all proteins, enzymes are linear chains of amino acids that fold to produce a three-dimensional structure. The sequence of the amino acids specifies the structure which in turn determines the catalytic activity of the enzyme. Although structure determines function, a novel enzyme's activity **cannot** yet be predicted from its structure alone. Enzyme structures unfold (denature) when heated or exposed to chemical denaturants and this disruption to the structure typically causes a loss of activity.

Protein folding is key to whether a globular protein or a membrane protein can do its job correctly. It must be folded into the right shape to function. But hydrogen bonds, which play a big part in folding, are rather weak, and it does not take much heat, acidity, or other stress to break some and form others, denaturing the protein. This is one reason why tight homeostasis is physiologically necessary in many life forms.

Denaturation

Denaturation is a process in which proteins or nucleic acids lose the quaternary structure, tertiary structure and secondary structure which is present in their native state, by application of some external stress or compound such as a strong acid or base, a concentrated inorganic salt, an organic solvent (e.g., alcohol or chloroform), radiation or heat. If proteins in a living cell are denatured, this results in disruption of cell activity and possibly cell death. Denatured proteins can exhibit a wide range of characteristics, from conformational change and loss of solubility to aggregation due to the exposure of hydrophobic groups.

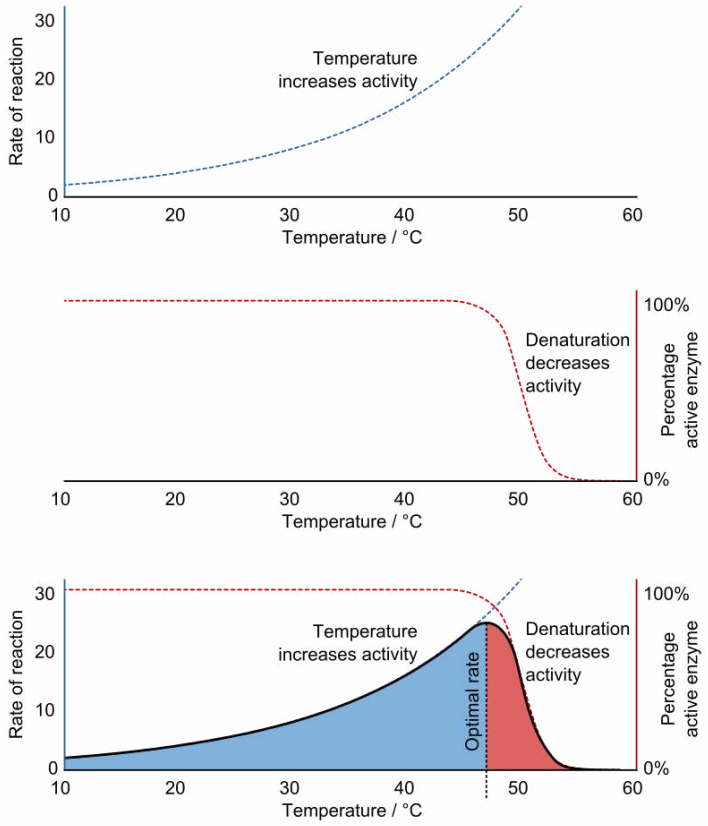


Figure* 9.8.1*: *Enzyme activity initially increases with temperature until the enzyme's structure unfolds (denaturation), leading to an optimal rate of reaction at an intermediate temperature. (CC-BY-SA-4.0;* Thomas Shafee)

Enzyme denaturation is normally linked to temperatures above a species' normal level; as a result, enzymes from bacteria living in volcanic environments such as hot springs are prized by industrial users for their ability to function at high temperatures, allowing enzyme-catalyzed reactions to be operated at a very high rate.

Contributors

- Wikipedia

9.E: Exercises

Numerical Solutions to Rate Laws

Numerical Solutions

```
%%python3

import numpy as np
from scipy.integrate import odeint
import matplotlib.pyplot as plt

def rxn1st(C,t,*k):

  r1=k[0]*C[0] #k[0]*(concentation of A)

  dAdt=-r1 #rate of change of A decreased by forward reaction and increased by revers
  dBdt=r1 #rate of change of B increased by forward reaction and decreased by reverse

  return(dAdt,dBdt)

t=np.linspace(0,10,101) #the first number is the beginning point, the second number i
C0=[1,0] #initial concentrations of A and B
k1=1
k2=0
k=[k1]
C=odeint(rxn1st,C0,t,(k1,k2))

cA=C[:,0] #define cA to give the concentration from the first (zeroth) column of the
cB=C[:,1] #define cB to give the concentration from the second column of the C array
```

Index

Glossary

Sample Word 1 | Sample Definition 1

Detailed Licensing

Overview

Title: Kinetics

Webpages: 95

All licenses found:

- CC BY-SA 4.0: 60% (57 pages)
- Undeclared: 40% (38 pages)

By Page