The Concept of Entropy: Relationship to the Energy Level Spectrum of a System

- The first law of thermodynamics allows the definition of heat and work in terms of energy changes.
- The first law of thermodynamics also restricts the class of possible processes to those that conserve energy.
- But processes that conserve energy do not necessarily occur.
- The first law does not predict whether a process will actually occur.
- For example, consider a box with two compartments.
 - Each compartment is filled with a different gas.
 - A value is opened between the compartments.

- Continued:
 - The gases mix and energy is conserved.
 - But demixing of the gases is also consistent with the first law.
 - Yet it is never observed.
- Consider an adiabatic bomb calorimeter filled with H₂ and O₂.
 - By the first law, energy is conserved since energy is unable to leave the system as work or heat.
 - The reaction:

$$\mathsf{H}_2 + \frac{1}{2} \mathsf{O}_2 = \mathsf{H}_2 \mathsf{O}$$

is vigorously exothermic and can be initiated with a slight perturbation (such as a spark).

- Continued:
 - The first law is unable to predict whether this reaction will be occur in a given set of conditions.
- Consider two systems in contact through a diathermal wall, but otherwise isolated.
 - Such systems are observed to spontaneously reach a common temperature.
 - Although it is consistent with the first law, it is never observed that the systems spontaneously change such that the warmer system becomes hotter and the cooler system cools further.
- Knowing the energy change of the system associated with a process is insufficient to predict the direction of the process.

- There are many spontaneous processes that are either exothermic or endothermic.
 - Mixing of water with sulfuric acid is exothermic.
 - Mixing of water with ammonia is endothermic.
- Criteria in addition to the first law is required to predict spontaneity.
 - The thermodynamic function of entropy is needed.
 - The second law of thermodynamics describes equilibrium in terms of entropy.
 - Derivatives of entropy with respect to the parameters that describe the constraints of the system determine the stability of equilibrium.
 - If the energy is constant, entropy determines whether or not a process is spontaneous.

- The second law of thermodynamics is considered from two points of view.
 - One approach will be with respect to the redistribution of a population of molecules over energy levels as energy enters or leaves the system as heat.
 - How much information about the microscopic states of a system contained by the specification of the macroscopic parameters will be considered.
 - Also considered will be degeneracy of the energy spectrum. (Degeneracy is related the number of distributions that have the same energy.)
 - Statistical mechanics deals with the most probable distributions that are consistent with a particular energy.

- Continued:
 - Statistical mechanics shows that in most cases the most probable distribution is sharply peaked.
 - It also shows that the macroscopic properties are somewhat insensitive to the details of the energy spectrum.
- The second law will then be considered from a macroscopic perspective.

The Relationship between Average Properties and Molecular Motion in an N-Molecule System: Time Averages and Ensemble Averages

Macroscopic properties have been related to averages of microscopic properties.

- Three questions arise:
 - Why is averaging used?
 - Which type of average is correct?
 - Do different averaging processes give the same description of the system?
- Averaging is used because a thermodynamic system of many molecules is described by a few macroscopic variables.
 - The thermodynamic description is consistent with some type of averaging.
 - The thermodynamic description is coarser than the microscopic description.

- To discern the type of averaging to be used and whether the type of average affects the result requires careful consideration of how measurements are made.
- Already considered have been:
 - The treatment of all molecules as independent particles.
 - The momentum transferred by one molecule per unit area per unit time.
 - The sum of the contributions all molecules moving in the appropriate direction to momentum transport per unit time per unit area.
 - To do this, the distribution function $f(\mathbf{v})$ is assumed to be independent of time.

- Assuming that the distribution function $f(\mathbf{v})$ is independent of time is equivalent to following the trajectory of a single molecule for a very long time and averaging the momentum transported across the plane each time it crosses the plane.
- But, what exactly does this type of equivalency mean?
- Successive crossings of the plane by different molecules are not uniform in time.
 - Instantaneous rates of momentum vary with time.
 - But at equilibrium, the average rate of crossing the plane must be independent of time.
- Consider the time average of momentum transfers occurring in some long time interval S.

- Continued:
 - If ℑ is short (i.e. the interval between crossing is comparable to ℑ, then averages over that time interval will fluctuate with that time scale.
 - If ℑ is large (i.e. the interval between crossings is small relative to ℑ), then the average over ℑ is independent of ℑ.
- Now consider following the trajectory of a single molecule in a gas.
 - If it is followed long enough, it will approach arbitrarily close to each point within the volume.
 - It will also approach arbitrarily close to each velocity in the velocity distribution.
 - This is the quasi-ergodic hypothesis.

- If the gas is dilute, collisions between molecules are so infrequent that they may be ignored.
 - In this case it is feasible to follow the trajectory of a particular molecule sufficiently long to compute the time average of a mechanical quantity.
 - The distance that the molecule travels in time t grows linearly with t.
 - The contributions to momentum transport across a plane could be summed and then divided by t.
 - Then taking the limit as t goes to infinity becomes the equivalent to averaging over a distribution function.
- If the gas is not dilute, then the molecule is almost always within the force field of another molecule.
- The trajectory is very complicated.

- In such cases, calculation of the trajectory is computationally impossible.
- An alternative method is needed to relate molecular properties to macroscopic properties.
- The alternative approach was proposed independently by Gibbs and by Einstein.
- Recall that for a given macroscopic state of the system, there are many microscopic states that are consistent with the macroscopic state.
- Instead of one system in a microscopic state consistent with the macroscopic state, many such systems in different microstates are constructed.
 - This collection of systems encompasses the entire distribution of microstates that match the macrostate.

- Continued:
 - Each member system of the collection is a replica system.
 - The entire collection is called an *ensemble*.
- Averages may be calculated by considering the quantity of interest for each of the replicas, then averaging over the entire ensemble.
- Statistical mechanics takes as a postulate that the ensemble average is equal to the time average in the limit of infinitely long time.
- Ensemble averages are easier to calculate than are time averages.
- This postulate follows from the quasi-ergodic hypothesis.

- To understand the quasi-ergodic hypothesis, consider a geometric visualization of molecular dynamics.
 - Classical mechanics will be used.
 - Assume that the molecules have no internal structure.
 - Consider a many dimensional coordinate system that describes the position and momentum of N particles in volume V.
 - \bullet 6*N* axes are required.
 - There are 3N position coordinates

 $x_1, y_1, z_1, \ldots, x_N, y_N, z_N$.

There are 3N momentum coordinates

 $p_{x_1}, p_{y_1}, p_{z_1}, \ldots, p_{x_N}, p_{y_N}, p_{z_N}$

This space is called the *phase space* of the N molecule system.

Phase space is assumed to be Cartesian with distances between two points given by:

$$D(w, w^*) = \left[\sum_{j=1}^{6N} (w_j - w_j^*)^2\right]^{1/2}$$

where w_j and w_j^* are the *j* elements of the vector of coordinates and momentum describing point *w* or w^* in 6N dimensional phase space.

- For a given value of energy, the possible values of coordinates and momentum will describe a surface in phase space.
 - For a perfect gas, energy depends only on the momenta of the particles.

- Continued:
 - The total energy is:

$$E = \frac{(p_{x_1}^2 + \dots p_{z_N}^2)}{2m}$$

which describes a sphere in 3N dimensions.

- If the molecules interact, then the energy depends also on the instantaneous positions.
 - The total energy expression still describes a surface, but that surface is much more complicated.
 - If a trajectory on an isolated system is considered, then the trajectory moves on the surface in phase space that is defined by energy conservation.

- Consider the difference between time averaging and ensemble averaging.
 - The trajectory moving over the surface defined by energy conservation will sample coordinates and momenta consistent with the total energy.
 - The time average will average over these.
 - For an ensemble, each replica system has the same value of E, V and N.
 - Therefore the momenta and coordinates will be a representative point on the same surface considered for the time averaging.
 - Different replicas will will have different representative points, but all points lie on the same surface.
 - The possible points will densely cover the surface.

- Unless the time average covers this surface in the same way as the ensemble, the averages will be different.
 - There exist conditions for which they are equivalent.
- The ensemble average is the fundamental concept in the description of equilibrium.
 - By definition, a macroscopic property of a system is to be identified with the ensemble average of the corresponding microscopic dynamical property.
 - This definition can be justified by the comparison of theory and observation.
- Ergodic theory as applied to statistical mechanics explores the equivalence of time and ensemble averages.

- The ergodic hypothesis as originally formulated by Boltzmann is incorrect.
 - Boltzmann postulated that, given enough time, a trajectory representing the time evolution of the system would pass through every point on the energy surface.
 - This cannot occur because:
 - A solution to an equation in mechanics has a unique solution which never crosses itself.
 - The trajectory is of lower dimensionality than the surface and so can thus never fill the higher dimensional energy system.
- The quasi-ergodic hypothesis makes a weaker statement about the trajectory on the energy surface.

- Birkhoff (1931) proved that if a boundary on an energy surface could not be constructed such that a trajectory originating from a representative point would lie entirely on one side of the boundary then the trajectory will pass arbitrarily close to every point on the energy surface.
- The quasi-ergodic hypothesis is valid only in classical mechanics.
- It cannot be rigorously proven,
- There are other difficulties in basing a description of equilibrium on ergodicity.
 - Any microscopic description of a system moving to equilibrium should have a time-independent description of the equilibrium state.

- That a system is quasi-ergodic is necessary, but not sufficient, to give a time independent average.
 - A quasi-ergodic trajectory can also be quasi-periodic, such as in the case of couples oscillators with an irrational ratio of frequencies.
- Two trajectories can start at points close together on the energy surface.
 - As they evolve, they remain close together.
 - The distance between them grows smoothly with time.
 - If reversibility is to be avoided, more turbulent trajectories are needed.
 - These turbulent trajectories are associated with mixing and diverge exponentially with time.

- Continued:
 - The trajectory is chaotic and does not return to the initial state.
 - Thus mixing implies quasi-ergodicity.
 - Simulation studies have demonstrated mixing and quasi-ergodicity.
 - For general mechanical systems, there exist energy regions which exhibit dynamical chaos.
 - In the limit of thermodynamic size ($N \approx 10^{23}$), ensemble averages satisfactorily describe equilibrium properties.
- To use ensemble averages in place of time averages, a number of issues must be considered.

- Continued:
 - The mixing property has been established only for classical mechanics.
 - Quantum ergodic theory is not well developed.
 - The use of ensemble averages is valid both quantally and classically.
 - Equilibrium is defined as being invariant with time.
 - This neglects fluctuations about equilibrium.
 - Ensemble averages permit the study of these fluctuations.
 - The type of the equilibrium that exists depends on the constraints.
 - Constraints are more easily implemented in ensembles than in trajectories.

Ensembles and Probability Distributions

- Consider tossing a coin:
 - It can land either head or tails.
 - Which it will be cannot be predicted with certainty.
 - If the coin is balanced, then head and tails have equal probability.
 - For a large number of tosses, the number of heads should be approximately equal to the number of tails.
 - But other outcomes are possible, although not as likely.
- For a single toss of the balanced coin, the probability of heads is $\frac{1}{2}$.
- The probability of tails is also $\frac{1}{2}$.

- Consider *n* tosses.
 - The probability that n_1 will be heads is:

$$\frac{n!}{n_1!(n-n_1)!} \left(\frac{1}{2^n}\right)$$

where $n!/n_1!(n-n_1)!$ is the number of ways that n_1 heads could occur regardless of order and 2^n is the number of possible sequences of heads and tails.

- For 4 tosses this predicts:
 - 1/16 will be all heads and 1/16 will be all tails.
 - 4/16 will be 3 heads and 1 tail and 4/16 will be 1 head and 3 tails.
 - 6/16 will be 2 heads and 2 tails

- For 400 tosses, the probability that 200 will be heads is 5×10^{22} more likely than 100 of 400 tosses being heads.
- Now consider this in the context of molecules in a gas.
 - Molecules of a gas are in motion.
 - The number of molecules in a fixed volume element will vary from one moment to the next.
 - Consider a perfect gas of N molecules in a volume
 V.
 - Now consider some volume element ω within the container.
 - ω/V is the fraction of the total volume occupied by ω .
 - Assume that the gas is uniformly distributed.
 - The probability of finding a particular molecule within ω is ω/V

- The probability of all N molecules being within ω is $(\omega/V)^N$.
- If ω is small with linear dimensions comparable to molecular separations, then the number of molecules in ω would most likely be 0,1, or 2.
- If ω was 0.1 of the total volume, then N/10 molecules (with small fluctuations) would be most likely to be found within ω .
- $P(N_1)$ is the probability distribution that describes the probability of finding N_1 in ω and $N N_1$ in $V \omega$.

$$P(N_1) = \frac{N!}{(N-N_1)!N!} \left(\frac{\omega}{V}\right)^{N_1} \left(\frac{V-\omega}{V}\right)^{N-N_1}$$

- The relative fluctuation in the number of molecules in ω decreases as ω increases and is given by $(N\omega/V)^2$.
 - The relative deviation is a volume of 1 mm³ is of the order of 10^{-8} at room temperature and pressure.
 - Macroscopic density does not vary perceptibly from point to point in a gas at equilibrium.
- In applying probability to the study of molecules the following must be considered:
 - The relationship between the definition of the probability distribution and the class of events being described.
 - The relationship between the probability and the fundamental classical or quantum description of the system.

- The relationship between an ensemble and the probability distribution must be considered.
 - An ensemble is a collection of replica systems.
 - The following must be specified:
 - The nature of the collection
 - The set of possible states for the objects composing the collection
 - Any other relevant details.
 - The probability of finding a characteristic is the fraction of replicas that possess the characteristic.
 - The probability depends on the nature of the ensemble.
 - Replicas composing the ensemble are consistent with the constraints.

- Consider an ensemble composed of pennies:
 - Whether they are heads up or tails up does not need to be specified since they are included in the properties of a penny.
- Consider an ensemble of containers of gas having rigid adiabatic walls and the same density of molecules.
 - Do not need to state whether the pressures in all containers are the same.
- There is a relationship between the external constraints defining a macroscopic system and the nature of the corresponding ensemble.
 - This allows the establishment of correspondence between macroscopic and microscopic approaches to the behaviour of matter.

- Consider the number of replicas \mathcal{N} within the ensemble to be sufficiently large that that the limit $\mathcal{N} \to \infty$ may be taken.
 - This means that the probability distribution function may be treated as continuous and may be considered in terms of a probability density.
 - ✓ For a discrete probability function, the probability of x occurring is considered.
 - For a continuous probability function, the probability that x is between a and b is considered.
 - This probability is expressed in terms of an integral over a probability density, $\mathcal{P}(x)$:

$$P(x_1 \le x \le x_2) = \int_{x_1}^{x_2} \mathcal{P}(x) dx$$

• If $x_2 = x_1 + dx$ then:

$$P(x_1 \le x \le x_1 + dx) = \mathcal{P}(x)dx$$

• Note that P(x) is dimensionless, but \mathcal{P} has units x^{-1} .

It is possible to relate this to a discrete distribution by dividing x into arbitrarily small intervals Δ and labelling each with an index i. Then:

$$P(x_i) = \mathcal{P}(x_i)\Delta$$

This is valid as long as Δ is small enough for \mathcal{P} to be treated as a constant over the interval.

Analogously to continuous distribution functions, averages can be determined using discrete probability functions.

- Let x be a discrete variable from the set x_1, \ldots, x_n with probabilities $P(x_1), \ldots, P(x_N)$.
- The average value of x (also known as the arithmetic average or the mean or expectation value) is:

$$\langle x \rangle = \frac{x_1 P(x_1) + x_2 P(x_2) + \dots + x_N P(x_N)}{P(x_1) + P(x_2) + \dots + P(x_N)}$$

$$=\frac{\sum_{i} x_i P(x_i)}{\sum_{i} P(x_i)}$$

• The mean value of f(x) is:

$$\langle f(x) \rangle = \frac{\sum_{i} f(x_i) P(x_i)}{\sum_{i} P(x_i)}$$

Continued:

- If $(f(x) = x^n$, the $\langle f(x) \rangle$ is the *n*th moment of P(x)
- For a continuous variable x and continuous f(x), the equations become:

$$\langle x \rangle = \frac{\int x \mathcal{P}(x) dx}{\int \mathcal{P}(x) dx}$$
$$\int f(x) \mathcal{P}(x) dx$$

$$\langle f(x) \rangle = \frac{\int f(x) \mathcal{P}(x) dx}{\int \mathcal{P}(x) dx}$$

From this follows relations:

$$\langle f(x) + g(x) \rangle = \langle f(x) \rangle + \langle g(x) \rangle$$

 $\langle cf(x) \rangle = c \langle f(x) \rangle$

Probability distributions are often normalized:

$$\sum_{i} P(x_i) = 1$$
$$\int \mathcal{P}(x) dx = 1$$

- Normalization is not necessary if only relative probabilities are of interest.
- The spread of a probability density distribution $\mathcal{P}(x)$ is given by $\langle x^2 \rangle \langle x \rangle^2$
 - Except when $\mathcal{P}(x)$ is a delta function, $\langle x^2 \rangle > \langle x \rangle^2$ where $\langle x^2 \rangle = \int x^2 \mathcal{P}(x) dx$

428 • The Concept of Entropy: Relationship to the Ene



Figure 15.1 (a) A typical probability density distribution with $\langle x \rangle \neq 0$. (b) A typical symmetric probability density distribution with $\langle x \rangle = 0$.
- One implication of a distribution having width is the possibility of fluctuations.
 - There is a nonzero probability of finding a value of x that differs from the average value of x.
 - The probability of a fluctuation depends on its deviation from the mean value.
 - These fluctuations show up in fluctuations of the macroscopic variables.
- Another property of probability distributions that is important is the concept of joint probability.
 - Consider two variables, x and y, with probability density distributions $\mathcal{P}_1(x)$ and $\mathcal{P}_2(y)$ respectively.
 - The joint probability distribution is $\mathcal{P}(x, y)dxdy = \mathcal{P}_1(x)\mathcal{P}_2(y)dxdy$

- Continued:
 - This gives the probability of finding x between x and x + dx and y between y and y + dy if $\mathcal{P}_1(x)$ and $\mathcal{P}_2(y)$ are independent of each other.
- If $\mathcal{P}_1(x)$ and $\mathcal{P}_2(y)$ are not independent but rather are correlated, then the joint probability includes a correlation function:

$$\mathcal{P}(x,y)dxdy = \mathcal{P}_1(x)\mathcal{P}_2(y)g(x,y)dxdy$$

or

$$\mathcal{P}(x,y)dxdy = \mathcal{P}_1(x)\mathcal{P}_2(y)[1+h(x,y)]dxdy$$

where g(x, y) and h(x, y) are the correlation functions.

• When x and y are independent, g(x, y) = 1 and h(x, y) = 0

Some Properties of a System with Many Degrees of Freedom: Elements of the Statistical Theory of Matter at Equilibrium

- Consider a sample containing a large number of molecules, N, in volume V.
- The energy of the system is between E and E + dE with dE/E << 1.
- An energy spectrum exists for the system.
 - The details of the spectrum do not need to be known.
- There are an enormous number of quantum states consistent with the macroscopic constraints of E, V, N.
- There are many distributions over the energy levels that are consistent with E, V, N.

- The equations of motion for N molecules cannot be solved exactly (classically or quantally).
 - Therefore the exact mechanical behaviour of the system is not determined and is not available for a macroscopic system.
 - This is not of interest in macroscopic thermodynamics.
- $\Omega(E, V, N)$ is the number of states with energy between *E* and E + dE.
 - Each solution of the N molecule Schrödinger equation gives a small number of the $\Omega(E,V,N)$ states.
 - The properties of the macroscopic variables depend only on $\Omega(E, V, N)$, not the microscopic details.

The fundamental postulate of equilibrium statistical mechanics is:

"All possible quantum states of an isolated system consistent with a given set of macroscopic parameters of constraint are to be considered as equally probable".

- This is also known as the postulate of equal a priori probability.
- This is an assumption which has been validated by the success of statistical mechanics in producing results consistent with observations.
- To apply this postulate an ensemble of systems is considered.
- A large number of replica systems is constructed, each consistent with the macroscopic variables E, V, N.

- All other properties are allowed to vary freely.
- Suppose that all possible quantum states consistent with the definition of the ensemble are represented in the ensemble.
 - The relative frequency definition of probability is in agreement with the assumption of equal *a priori* probability of all quantum states consistent with *E*, *V*, *N*.
 - Therefore the probability of a random choice of a replica from the ensemble being the same as an arbitrarily chosen member of the ensemble is 1/N when there are N replicas in the ensemble.
- Consider now the constraints used to define the ensemble.

- Continued:
 - The equilibrium macroscopic properties of an isolated system are independent of time.
 - Therefore the distribution over the energy levels is independent of time.
 - Constants of motion include total energy, total linear momentum, and total angular momentum of the system.
 - In an isolated system, no energy is transferred across the borders and thus no forces act upon the system.
 - Therefore there is no change in either linear or angular momentum.
 - The constraints make it possible to define an appropriate probability distribution.

- Continued:
 - Often it is sufficient to use only one of the constants of motion, typically the total energy (since it is used to define the ensemble).
 - The other constants of motion are allowed to vary freely.
- Consider an ensemble of replica systems, each with volume V and number of molecules N.
- The ensemble is specified by the requirement that the energy be between E and E + dE.
 - The number of quantum states with energy less than E is $\Gamma(E, V, N)$.
 - Therefore the number of quantum states with energy between E and E + dE is:

$\Omega(E, dE, V, N) = \Gamma(E + dE, V, N) - \Gamma(E, V, N)$

- $\Omega(E, dE, V, N)$ describes the energy spectrum including degeneracy and dependence on volume.
 - The dependence on V is due to the effect of the boundary conditions on the energy spectrum of the isolated system.
- In large systems, the dependence of Ω on the energy interval dE is insignificant.
 - As the system becomes larger the dependence of Ω on dE is increasingly insignificant.
 - This is a consequence of the Schrödinger equation, the solution of which depends on the form of the potential and the boundary conditions.

- The solutions to the Schrödinger equation for particles in a box are of the form $\sin kx$ and $\cos kx$.
 - The boundary conditions assume that the walls of the box are infinitely repulsive (there is zero probability of the particle being outside the box).
 - These boundary conditions force k and consequently the energy to have finite values.
 - The larger the box, the greater number of finite energies allowed in the spectrum.
 - In the limit of an infinite box (i.e. a free particle), the allowed values of k become continuous and the energy spectrum becomes a continuum.
- The fact that a macroscopic system contain a large number of molecules allows the analysis to be simplified.

- Divide the energy scale into intervals of dE with dE << E
 - For a large system the interval dE contains many energy levels.
 - The number of states between *E* and E + dE is $\Omega(E)$.
 - $\Omega(E)$ varies with *E* and obeys the relation:

$$\Omega(E) = \left(\frac{\partial\Gamma}{\partial E}\right)dE = \omega(E)dE$$

- $\omega(E)$ is the density of states and is characteristic of the system.
- The sensitivity of $\Omega(E)$ and $\omega(E)$ to the energy *E* of the system will now be considered.

- It can be shown that $\Omega(E)$ is insensitive to the details of the nature of the energy spectrum in the limit of a large system.
 - This is a characteristic of a macroscopic parameter.
 - $\Omega(E)$ connects the microscopic properties of the system to the macroscopic system.
- Let a system of energy E be described by ν quantum numbers.
- The energy per degree of freedom is:

$$\epsilon = \frac{E}{\nu}$$

• For a perfect gas with 3N degrees of freedom:

$$\epsilon_{\text{perfect gas}} = \frac{1}{2}k_BT$$

- In classical mechanics, the total energy is divided among the degrees of freedom such that the average energy per degree of freedom is $\frac{1}{2}k_BT$.
- In quantum mechanics, this is true if k_BT is large relative to the energy spacings, but not when k_BT is small relative to the energy spacings.
 - For real molecules at room temperature, k_BT is often small relative to the spacing of vibrational levels.
- The equipartition of energy ($\epsilon = E/\nu$) will be used as an approximation to further explore the properties of $\Omega(E)$ for an arbitrary system with ν degrees of freedom.

- Consider $\Gamma(E)$, the total number of quantum states with energy less than *E*.
 - The energy is separable into contributions from ν degrees of freedom.
 - Each of the possible contributions contributes a state to $\Gamma(E)$.
 - E_i is the energy of the *i*th degree of freedom.
 - $\Gamma(E_i)$ is the number of states of the *i*th degree of freedom with energy less than E_i .
 - Then $\prod_{i=1}^{\nu} \Gamma_i(E_i)$ is the total number of states with energy less than *E* for a particular subdivision of energy.
 - For all possible subdivisions of E, the total number of states with energy less than E is:

$$\Gamma(E) = \sum_{E_i} \prod_{i=1}^{\nu} \Gamma_i(E_i)$$

where $E = \sum_{i} E_{i}$.

- Note that the summation over E_i is over all possible ways of subdividing the energy that are consistent with the constraint of conservation of energy.
- To estimate $\Gamma(E)$ require that each of the E_i be the average energy ϵ .
 - This is equivalent to choosing the one term in the sum over all the subdivisions of energy that has uniform distribution of energy over all the degrees of freedom.
 - This gives:

$$\Gamma(E) = \prod_{i=1}^{\nu} \Gamma_i(\epsilon) = \left[\Gamma_1(\epsilon)\right]^{\nu}$$

- Although crude, this can give an adequate order of magnitude estimate.
- It is assumed that all degrees of freedom are sufficiently alike that the same average behaviour may be ascribed to them.
- Expansion in a Taylor series gives:

$$\Omega(E) = \frac{\partial \Gamma}{\partial E} dE = \frac{\partial}{\partial E} \left\{ \left[\Gamma_1(\epsilon) \right]^{\nu} \right\}$$

$$=\nu\Gamma_1^{\nu-1}\frac{\partial\Gamma_1}{\partial(\nu\epsilon)}dE = \Gamma_1^{\nu-1}\frac{\partial\Gamma_1}{\partial\epsilon}dE$$

- $\Gamma_1(\epsilon)$ increases as *E* increases.
 - Since ν is very large in a macroscopic system, $\Omega(E)$ increases rapidly with *E*.
 - The rate of increase is evident when $\Omega(E)$ is rewritten as:

$$\ln \Omega(E) = (\nu - 1) \ln \Gamma_1(\epsilon) + \ln \left(\frac{\partial \Gamma_1}{\partial \epsilon} dE\right)$$

- Recall that dE is much larger than the separation of the energy levels.
- This means that $\ln (\partial \Gamma_1 / \partial \epsilon) dE$ is negligible relative to the first term of the expression.
 - Since $\Gamma_1(\epsilon)$ is proportional to ϵ , $(\partial \Gamma_1/\partial \epsilon)$ is $\Gamma_1(\epsilon)/\epsilon$.

- Continued:
 - For simple model systems, $\Gamma_1(\epsilon)$ is proportional to ϵ^{α} where α is small.
 - For a harmonic oscillator, $\alpha = 1$
 - For a particle in a box, $\alpha = 1/2$.
 - This means that $\ln (\partial \Gamma_1 / \partial \epsilon) dE$ is a factor of ν smaller than $(\nu 1) \ln \Gamma_1(\epsilon)$
- The magnitude of $(\nu 1) \ln \Gamma_1(\epsilon)$ can be estimated.
 - Δ is the average spacing between energy levels.
 - Therefore $\Gamma_1(\epsilon)$ is of the order of ϵ/Δ and

$$\ln \Omega(E) = \nu \ln \Gamma_1(\epsilon)$$

• This means that Ω is of order magnitude of ν or better for *E* well above the ground state.

The Influence of Constraints on the Density of States

- The role of constraints in defining the density of states $\omega(E)$ or the number of states $\Omega(E)$ between *E* and E + dE is now examined.
- Consider a perfect gas, described as N particles in a cubical box of volume $V = a^3$
 - The wave function for a single particle in a cubical box is:

$$\Psi_j = A \sin \frac{n_{jx} \pi x}{a} \sin \frac{n_{jy} \pi y}{a} \sin \frac{n_{jx} \pi z}{a}$$

for 0 < x < a, 0 < y < a, 0 < z < a, where
j = 1,..., N labels the particles,
A is the normalization constant.

• The quantum numbers n_{jx} , n_{jy} , n_{jz} satisfy:

$$n_{jx}^{2} + n_{jy}^{2} + n_{jz}^{2} = \frac{2m\epsilon_{j}}{\pi^{2}\hbar^{2}}a^{2}$$

where ϵ_j is the energy of particle *j*.

- The values of n_{jx}, n_{jy}, n_{jz} are restricted to positive energies.
- The total energy of N particles is:

$$E = \sum_{j=1}^{N} \epsilon_j$$

This may be rewritten as:

$$\sum_{j=1}^{N} (n_{jx}^2 + n_{jy}^2 + n_{jz}^2) = \frac{2mE}{\pi^2\hbar^2}a^2$$

- The total number of states with energy below E can be determined by the number of states with integer coordinates lying within the 3N dimensional hyperspherical surface with radius R_s of $(2mEa^2/\pi^2\hbar^2)^{1/2}$
 - The dimensionality of 3N comes from the allowed values of the quantum numbers n_{jl} where

j = 1, ..., N and l = x, y, z.

When the box is large, the spacing between energy levels is small and the volume enclosed by the hyperspherical surface is:

$$\frac{\pi^{3N/2}}{(3N/2)!} R_s^{3N} = \frac{\pi^{3N/2}}{(3N/2)!} \left(\frac{2mEa^2}{\pi^2\hbar^2}\right)^{3N/2}$$

Since only positive values of n_{jl} are allowed, the volume of interest is the positive orthant which is $(\frac{1}{2})^{3N}$ of the total volume. Therefore:

$$\Gamma(E) = \frac{1}{(3N/2)!} \left(\frac{mEa^2}{2\pi\hbar^2}\right)^{3N/2}$$

• The number of states between E and E + dE is:

$$\Omega(E) = \left[\frac{\partial\Gamma(E)}{\partial E}\right] dE = \frac{E^{3N/2-1}}{((3N/2) - 1)!} \left(\frac{ma^2}{2\pi\hbar^2}\right)^{3N/2} dE$$

Therefore, the density of states is:

$$\omega(E) = \frac{E^{3N/2-1}}{((3N/2) - 1)!} \left(\frac{ma^2}{2\pi\hbar^2}\right)^{3N/2}$$

- Note that $(a^2)^{3N/2}$ is V^N and it is the boundary condition that defines the energy spectrum.
- In more general cases, other macroscopic variables (besides V) can affect the energy spectrum.
- Consider a macroscopic system described by the set of variable:

$$y_1 = \alpha_1, y_2 = \alpha_2, \dots, y_n = \alpha_n$$

 There exists a corresponding dependence of the energy level spectrum on these variables. In this case, the number of states between E and E + dE is:

$$\Omega = \Omega(E, y_1, y_2, \dots, y_n)$$

when y_i is between y_i and $y_i + dy_i$

If a constraint is removed, then the number of states is increased:

$$\Omega(E, y_1, y_2, \dots, y_{n-1}) > \Omega(E, y_1, y_2, \dots, y_n)$$

- In most physical processes, constraints are altered instead of being removed.
 - How does $\Omega(E, y)$ change in such cases?
 - Consider N particles in a box. The constraint is the volume of the box.

By definition, relaxation of a constraint always leads to:

 $\Omega_f(E, y) > \Omega_i(E, y)$

- Consider an isolated system at equilibrium with a particular set of constraints.
 - This system has Ω_i equally accessible states.
- Some constraints are removed or relaxed by an external change involving no work.
- The new system thus created has Ω_f equally accessible states.
- Therefore:

$$\Omega_f \ge \Omega_i$$

- Consider an ensemble of systems that are in equilibrium with the constraints.
 - All Ω_i accessible states are equally probable.
- Let one constraint be relaxed or removed.
 - At the instant that the constraint is changed, the system is still in one of the Ω_i states.
 - The system is not at equilibrium.
- The approach to equilibrium requires that the Ω_f states be occupied with equal probability.
- The system changes until a new equilibrium is achieved, consistent with the remaining constraints.
- Now consider the case in which an additional constraint is imposed.

- The system cannot go to an equilibrium consistent with these constraints unless work is done on the system.
- Consider the particle in a box model of the perfect gas.
- As shown in Figure 15.2(a), the gas occupies half the volume of an isolated system.
 - The system is separated into two equal volumes with the other half of the system is evacuated.
- The barrier is removed by the expenditure of a negligibly small amount of work.
- The gas expands to fill the entire volume $V = 2V_i$.
- For this process:

$$\Omega_f \ge \Omega_i$$



Figure 15.2 Schematic illustration demonstrating that restoration of a previously relaxed constraint without execution of work does not regenerate the initial state of the system.

- Now consider reinsertion of the barrier.
- The system does not return spontaneously to its initial state.
- In this case the relationship between Ω_f and Ω_i is not the inverse foe the previous case.
- Recall:

$$\Omega = \left[\frac{\partial\Gamma(E)}{\partial E}\right] dE = \frac{E^{3N/2-1}}{((3N/2) - 1)!} \left(\frac{ma^2}{2\pi\hbar^2}\right)^{3N/2} dE$$

Therefore

$$\Omega(E, V, N) = CV^N \left(\frac{E}{N}\right)^{3N/2} dE$$

where C is a constant independent of E and V.

This was derived using the Stirling Approximation.

$$N! \approx N^N \exp(-N)$$

- For the initial state for 15.2b, the particles are anywhere within volume V.
- For the final state for 15.2, N/2 particles are in each of the two subvolumes V/2.
- The energy per particle has not changed in this process.
- There are [N!/(N/2)!(N/2)!] ways that N/2 particles can each be placed in the subvolumes.



$$\Omega_f(E,V,N) =$$

$$C \frac{N!}{\left(\frac{N}{2}\right)! \left(\frac{N}{2}\right)!} \left(\frac{V}{2}\right)^{N/2} \left(\frac{E/2}{N/2}\right)^{3N/4} \left(\frac{V}{2}\right)^{N/2} \left(\frac{E/2}{N/2}\right)^{3N/4} dE$$

$$= CV^N \left(\frac{E}{N}\right)^{3N/2} dE$$

- This means that $\Omega_f = \Omega_i$ and the insertion of the barrier has not changed the system.
- To restore the system to the initial state given in 15.2a, the gas must be removed from one of the compartments.
 - Compress the gas into a subvolume by a sliding wall.

- This would involve energy being put into the system as work and would result in a temperature change.
- There is no way to get back to the original state of 15.2a without transferring work or heat across the boundary of the system.
- Once the barrier is removed and the molecules distribute themselves over the Ω_f states, the reimposition of the boundary condition does not cause the system to spontaneously return to the initial state.
 - This is an irreversible process.
- The groundwork has now been laid for the comparison of statistical mechanical and the thermodynamic descriptions of a spontaneous process.

- Consider a process in an isolated system that takes the system from an initial to a final state.
- If this final state is such that the imposition or relaxation of constraints without the requirement of external work cannot recreate the initial state, the process is irreversible.
- If the imposition or relaxation of constraints without external work recreates the initial state, the process is reversible.
- For processes at constant energy and constant volume.
 - If some of the constraints defining an isolated system are relaxed, $\Omega_f \ge \Omega_i$
 - If $\Omega_f = \Omega_i$, the systems of the ensemble are distributed over the same accessible states

- Continued:
 - before and after the constraint is relaxed. The system is in equilibrium at all steps of the process and the transition $i \rightarrow f$ is reversible.
 - If $\Omega_f > \Omega_i$, the transition $i \to f$ is irreversible.
- In Chapter 13, a reversible process was defined in terms of intensive variables being continuous across the boundary of the system.
 - This involves the system being in contact with some reservoirs.
 - How is a reversible process defined for an isolated system, which by definition is not in contact with any reservoirs?
- Consider the system illustrated in Figure 13.3.



21

50 HAND

- The system is subdivided into two subsystems, separated by a movable piston upon which weights rest confining subsystem 1.
- The weights can be moved by a pulley arrangement between the piston and a ledge at the same height.
 - Therefore moving the weight involves no net work.
 - The pulley arrangement is therefore a device to change constraints inside an otherwise isolated system.
- If the weights are infinitesimally small, the pressure will be continuous across the piston as subsystem 1 expands and subsystem 2 is compressed.
 - Subsystem 2 may be considered the surroundings of subsystem 1.
- The thermodynamic definition of a reversible process requires that every point on the path be infinitesimally close to an equilibrium state of the system.
 - At every point along the path, the relevant intensive variables are continuous across the boundary of the system.
 - Therefore the extensive variables of the system must also be infinitesimally close in value in two states that are infinitesimally close.
 - This is equivalent to $\Omega_f = \Omega_i$ since the extensive variables usually describe the constraints or boundary conditions defining Ω .

The Entropy: A Potential Function for the Equilibrium State

- Consider the properties of the number of accessible states for a composite system.
 - The total energy may be distributed over parts of the system.
 - The most probable distribution of the total energy over the system will be shown to be equivalent to systems in equilibrium having equal temperatures.
- This will be the basis of the definition of entropy.
 - It will be shown that it is related to work and internal energy.
 - Entropy may be considered as providing a relationship between microscopic and macroscopic descriptions of the system.

- Consider two macroscopic systems, A and B with energies E_A and E_B respectively.
- Let $\Omega_A(E_A)$ and $\Omega_B(E_B)$ be the number of states between E_A and $E_A + dE_A$ and between E_B and $E_B + dE_B$ respectively.
- *N_A* and *N_B* are the numbers of *A* and *B* molecules and are fixed.
- The ensemble is specified by the requirement that $E_T = E_A + E_B$ is constant.
- The ensemble consists of pairs of systems, with the replicas having different distributions of E_T between A and B.
- The volumes V_A and V_B are independently varied as are any other constraints.

- The ensemble is defined by the value of E_T .
- If system A of a replica pair has energy E_A , then system B has an energy of $E_T E_A$.
- When system A has energy E_A it can be in any one $\Omega_A(E_A)$ equally likely state while system B can simultaneously be in any one of $\Omega_B(E_T E_B)$ states.
- Every possible state of A can be combined with every possible state of B, so the total number of states is:

$$\Omega_T(E_T, E_A) = \Omega_A(E_A)\Omega_B(E_T - E_B)$$

- Since E_T is constant, $\Omega(E_T, E_A)$ is a function of E_A .
- Now consider the contribution to $\Omega_T(E_A) = \Omega_T(E_T, E_A)$ from different subdivisions of E_T between A and B.

• The probability of finding A + B with the energy of A between E_A and $E_A + dE_A$ is:

$$\frac{\Omega_T(E_A)}{\sum_{E_A} \Omega_T(E_A)}$$

where $\Omega_T(E_A)$ is the total number of states accessible to the total system for a given E_A taken over all possible values of E_A .

🥒 Let

$$\sum_{E_A} \Omega_T(E_A) \equiv C^{-1}$$

• Therefore the probability of finding A + B in a state with energy between E_A and $E_A + dE_A$ is:

$P(E_A) = C\Omega_A(E_A)\Omega_B(E_T - E_B)$

- As E_A increases, $\Gamma_A(E_A)$ increases but $\Gamma_B(E_T E_A)$ decreases.
- However $P(E_A \text{ has a maximum at some } E_A^*$.
 - Because $\Gamma_A(E_A)$ and $\Gamma_B(E_T E_A)$ are rapidly varying functions, the maximum will be sharp.
- If Γ_A and Γ_B are treated as continuous functions, then the maximum may be found by differentiation:

$$\left(\frac{\partial P(E_A)}{\partial E_A}\right)_{V,N} = 0$$

and

$$C\Omega_B(E_T - E_B)\left(\frac{\partial\Omega_A}{\partial E_A}\right) + C\Omega_A(E_A)\left(\frac{\partial\Omega_B}{\partial E_A}\right) = 0$$

• Dividing by $P(E_A)$ and using $E_A = E_T - E_B$,

$$\left(\frac{\partial \ln \Omega_A}{\partial E_A}\right)_{V,N} = \left(\frac{\partial \ln \Omega_B}{\partial E_B}\right)_{V,N}$$



$$\beta = \left(\frac{\partial \ln \Omega}{\partial E}\right)_{V,N}$$

- Therefore $\beta_A = \beta_B$ (which has the same properties as temperature).
 - It can be shown that $\beta = (k_B T)^{-1}$.

By definition

$$S = k_B \ln \Omega$$

where S is entropy.