Chapter 26

THE ORIGIN OF THE BASIC ASSUMPTION

The development of statistical mechanics is very strange. As mentioned in th preface, it is rather like an inverted pyramid. Starting from the basic assumptio the development is upwards. Specialists of physics, chemistry and material science, etc. apply it to various models, successfully explaining many phenomena Mathematicians make it elegant and abstract and use rigorous proofs to deriv some results. The development downwards is less remarkable. Physicists ar certainly interested in the origin of the basic assumption, but are on the whol not enthusiastic in pursuing this problem because it is too difficult.

Physicists do not have the habit of avoiding difficult problems. Perhaps the are content with the success of the various applications and thus overlook th importance of understanding the origin. So the quest for the origin of the basi assumption, though certainly within the realm of physics, tends to fall into th hands of some great mathematicians. However, most mathematicians lik elegant mathematics and they sometimes neglect some of the requirements o physics.

In fact the understanding of the basic assumption is not only a physic problem, but also an understanding of irregular phenomena in general. How d the regular rules of motion cause random results? This is a very commo problem.

Matters of principle are difficult, but applications are not easy eithe Although the applications are quite successful, most calculations involve approx mations that may not be reliable. From the experience of the above chapter whenever the interaction cannot be neglected, the problem becomes extremel difficult. This difficulty is linked to that encountered in the basic assumptio

In this chapter, we shall summarise our views on the basic assumption, i.e. our views on the bases of statistical mechanics. We then discuss the concepts of ergodicity, ensembles and mixing. Then we look into and criticise some conventional views. The instability of the trajectory is also introduced here.

The discussion in this chapter cannot be said to be an objective view of statistical mechanics, but is the subjective opinion of the author. These ideas, scattered in previous chapters, are summerised here.

26.1. The Basic Assumption

We now review the basic assumption as discussed in the previous chapters. We classify the variables of motion into two types: (1) the invariants, i.e. variables unchanged during the observation time, and (2) the rapidly changing variables. Then we define the region of motion as a set including all the configurations of the changing variable, i.e. the set of configurations satisfying the conditions of invariance. We have emphasised repeatedly that invariance means unchanging quantity during the observation time. What variables are invariant depends on the time scale and in some cases it is not obvious how to classify the variables. The strict definition of the region of motion depends on the detailed trajectory of motion.

Let Γ be the volume of the region of motion, i.e. the total number of configurations of the region. The basic assumption says that entropy is the logarithm of Γ:

$$S = \ln \Gamma \quad . \tag{26-1}$$

This assumption must be supplemented by the following condition

$$S = O(N) \quad , \qquad \text{i.e. } \Gamma = O(e^N) \quad , \tag{26-2}$$

where N is the particle number or the total number of variables. From (26-1) and (26-2) the entire theory of thermodynamics can be deduced and all the equilibrium properties can be calculated. This is the conclusion of Chapter 2 and the content of Chapter 23. Notice that condition (26-2) is most important. Without it there would be no thermodynamics. After Chapter 4 we have come to regard this condition as that establishing the independence of the various variables of motion. Hence (26-2) is the hypothesis of independence. In discussing the central limit theorem (Chapter 12, and the end of Chapter 6), we pointed out that the time average of large value ($O(N)$) variables is the same as the average value in the region of motion. Secondly, the fluctuations

of these variables are of $O(\sqrt{N})$ and normally distributed. This conclusion can be written as

$$P(R) \equiv \frac{\mathcal{T}(R)}{\mathcal{T}} = \frac{\Gamma(R)}{\Gamma} \quad , \tag{26-3}$$

where $\mathcal{T}$ is the total observation time and R is a subset whose volume $\Gamma(R)$ is comparable with Γ:

$$\Gamma(R) = O(e^N) \quad . \tag{26-4}$$

$\mathcal{T}(R)$ is the time that the trajectory stays in R. These also follow from (26-2). (See Chapter 12.) $P(R)$ can be interpreted as the probability of the configurations being in R. Notice that $\mathcal{T}$ is a finite time and cannot be regarded as a large number of $O(e^N)$. R must be a large subset. Otherwise (26-3) is meaningless. (See Sec. 12.9.) The fluctuation of a large variable is directly related to the differentials of the thermodynamical potentials. The term "large variable" here denotes the sum of the variables describing various parts of the system, e.g. the total magnetic moment is the sum of the various magnetic moments. Its average value may be zero. This type of variable also includes variables like

$$\int d^3r \, \rho(r) \, e^{-i\mathbf{k}\cdot\mathbf{r}} \quad .$$

It is the sum, over various parts of the system, of the product of the density with $e^{-i\mathbf{k}\cdot\mathbf{r}}$. To sum up, the above basic assumption is sufficient to analyse all experiments of thermodynamics and scattering.

Hence we see that (26-1) is a daring assumption and it links entropy with the microscopic motion through Γ. From the deductive viewpoint, (26-2) is the origin of this assumption. To understand statistical mechanics, we must first understand (26-2).

We now mention here ergodicity and ensembles, two most important concepts in the development of statistical mechanics.

26.2. Ergodicity and Ensembles

The above basic assumption is not too harsh a condition because we discuss only large variables and large regions. The reader may notice that this basic assumption may appear in different forms in the literature. When Boltzmann wrote down this assumption, his thinking went like this: the trajectory will pass through every configuration in the region of motion, i.e. it is ergodic, and hence the infinite time average is equal to the average in the region. This argument is

wrong, because this infinite long time must be much longer than $O(e^N)$, while the usual observation time is $O(1)$. To require the time average of every variable to equal the average in the region is a harsh condition. We do not require this in this book, but we require that the time average of large variables equal the average over the region. This condition is not too harsh because the values of large variables are nearly the same everywhere in the region of motion and the trajectory need not pass through every configuration, i.e. it need not be truly ergodic.

The ergodicity of the trajectory is a major problem in mechanics.[a] Many mathematicians have devoted much effort in proving the ergodicity of some models. However these results are not too helpful to statistical mechanics.

The reader must have noticed that the definition of terms is not uniform in the literature. The term "ergodicity" may have different definitions in different papers. In some places the definition is as above. In other places it means that the trajectory "nearly" passes through every point. Others may say that whenever the basic assumption holds the motion is ergodic. This is a mess. The reader should be clear about the definition in any paper before delving into it.

The ergodic theory of Boltzmann met with immediate criticism and doubt was cast on his theory including the H-theorem. Boltzmann perhaps died of the resulting frustration.

The theory of Boltzmann, reformulated by Gibbs in terms of "ensembles",[b] forms the main trend of modern statistical mechanics.

The theory of ensemble says that the equilibrium states of the system form an ensemble, i.e. a system with infinitely many similar structures. Each configuration of the system is a point in phase space. These points distributed in phase space can be thought of as an ideal gas (each point is a "molecule"). We need at least $\sim e^N$ molecules to define the distribution.

Properties of the equilibrium states are average properties of this ensemble. This abstract way of thinking has many advantages, because we have a rather strong intuition about the flow and distribution of a gas. The ideas of Gibbs produce many useful formulas (equivalent to the calculational rules of Chapter 7) and applications. The ensemble theory has become the traditional basic concept.

Nevertheless the ensemble is an abstract artifact. In reality there is one system, not infinitely many. This was pointed out by Gibbs in his book, but is usually neglected nowadays. Ensemble is a mathematical concept and it cannot

[a] See Arnold (1968), Chap. 6 of Yang (1978) and other books on Ergodic Theory.

[b] Gibbs (1960).

solve the physical problem of Boltzmann's assumption. If the ergodic property proposed by Boltzmann is said to be unrealistic then ensemble theory is perhaps more so. The content of the ensemble theory is identical to the assumption of Boltzmann. It distributes the ensemble in the region of motion. The ensemble average is the average over the region of motion.

After Gibbs, many scholars have tried to axiomatise this theory of ensembles or at least give it a more definite meaning. Today the most commonly held idea is to define the ensemble as the knowledge of a system by the experimenter,[c] or "the degree of uncertainty". This degree of uncertainty is regarded as an *a priori* probability; that is to say, before the observation, the configurations of the system have a probability distribution. We have criticised this view in the above chapters (especially Chapter 24). In the next section we shall compare and discuss this view with that in our book. Here we first mention some basic concepts of the ensemble theory.

The ensemble is like a cloud in phase space and its density can be represented by a distribution function ρ. Now phase space is a $6N$ dimentional space of N particles all of which obey Newton's laws. Gibbs discussed the change of ρ and defined entropy as

$$S_G = -\int \rho \ln \rho \, d\Gamma \quad , \tag{26-5}$$

where ρ is normalised $\int \rho \, d\Gamma = 1$. The integral in (26-5) is over the whole of phase space.

The motion of each system in the ensemble obeys Newton's law, tracing out a trajectory. The trajectories of different systems are non-overlapping. (If there is intersection, that would mean different trajectories can come from the same initial conditon.) The reader has probably come across Liouville's theorem in mechanics, i.e. ρ is unchanged along the trajectory. Let us fix a region $R(0)$ in phase space and let every point in this region move according to the laws of mechanics. After a time t they form another region $R(t)$. According to Liouville's theorem the volumes of $R(0)$ and $R(t)$ are the same; only the shape changes. From this we can prove that the entropy defined by Gibbs in (26-5) is an invariant. Obviously (26-5) cannot be used to discuss phenomena of entropy increase. Many people tried to explain and patch up this unreasonable conclusion of constant entropy. Suppose $R(0)$ is a region like a lump of dough. After a long time, $R(t)$ will be stretched out like intertwining noodles made from the dough. If at $t = 0$, ρ is nonzero only in $R(0)$, then ρ is nonzero

[c] See Tolman (1962).

only in the noodles at time t. The expression in (26-5) is the volume of the dough or noodles, but it does not represent the randomness of these noodles or threads. We do not discuss the correction to (26-5) because it is built upon an incorrect theory. The equilibrium state is not an instantaneous concept and entropy needs a time scale to be defined. This point is emphasised from the beginning of this book. The ensemble theory attempts to use the instantaneous properties of infinitely many systems to represent equilibrium. This is an unreasonable approach. We will discuss it carefully in Sec. 26.5.

26.3. Mixing and Independence

Many scholars in the study of modern statistical mechanics devote themselves to the analysis of mixing. Mixing implies that the noodles or threads $R(t)$ will distribute uniformly in the region of motion. If $t \to \infty$, then

$$\lim_{t\to\infty} \frac{\Gamma(R(t)\cap R')}{\Gamma} = \frac{\Gamma(R(0))}{\Gamma} \cdot \frac{\Gamma(R')}{\Gamma} \quad , \tag{26-6}$$

where R' is any subset of the region of motion, Γ is the total volume of the region of motion, and $\Gamma(R')$ is the volume of R'. That is to say, the volume of $R(t)$ in R' is proportional to the volume R'.

The property of mixing is more reasonable than that of ergodicity. It requires that any region $R(0)$ be transformed to a patch of fine threads spreading uniformly over the region of motion. The analysis of mixing is a complicated mathematical problem and the author is no expert in this field. There are some recent books on this aspect.[d]

The theory of mixing is commonly accepted now. It is more advanced than the theory of Boltzmann or Gibbs, and it can be said to have descended from Gibbs' theory.

Now we rewrite the view adopted in this book in a form similar to that of mixing and then make a general discussion. The view of this book is: the basic assumption is an assumption of independence, i.e. an assumption on the correlation time and the correlation length.

In Sec. 26.1 we pointed out that the independence property (26-2) is the root of the basic assumption. The calculation of entropy from the trajectory discussed in the last chapter shows that the region of motion must be defined and understood from the trajectory, and the independence property is a necessary property for defining entropy from the trajectory. This independence property

[d] See Sinai (1979) and Krylov (1979).

is the independence beyond long time scales and long length scales. That is to say, if we have correlation time τ and correlation distance ξ and $|t_1 - t_2| >> \tau$ or $|\mathbf{r}_1 - \mathbf{r}_2| >> \xi$, then

$$\langle A(\mathbf{r}_1, t_1)\, B(\mathbf{r}_2, t_2)\rangle$$
$$\equiv \frac{1}{\mathcal{T}} \int_0^{\mathcal{T}} dt\; A(\mathbf{r}_1, t+t_1)\, B(\mathbf{r}_2, t+t_2) \longrightarrow \langle A(\mathbf{r}_1)\rangle\, \langle B(\mathbf{r}_2)\rangle \quad . \tag{26-7}$$

The average value is the time average over the observation time $\mathcal{T}$. A small change of $\mathcal{T}$ will not affect the average value. The variables A and B are local variables of motion near $\mathbf{r}_1$ and $\mathbf{r}_2$; "near" means distance smaller than ξ, and "local" means related to a small number of variables of motion.

If we can start from the principles of mechanics to derive (26-7), then the region of motion can be determined, (26-2) is established, and the whole basic assumption is built up. In the above, we have not supplied the mathematical steps to obtain (26-2) from (26-7). These steps are probably not easy, or may require other conditions. The main points here is to understand (26-7), i.e. how the independence property follows from mechanics.

What is the difference between mixing and independence? We look at a simple example here.

Consider two one-dimensional models: (A) a particle restricted in a rigid box between $0 < x < L$ and (B) a particle attached on a spring and executing simple harmonic motion with frequency ω. Let us look at the property of mixing in these two examples.

The black disc region in Fig. 26-1(a) becomes the black threaded region in Fig. 26-1(b) after a long time. The longer the time, the finer and denser the threaded region becomes. The total area of the black region is unchanged, but the black disc is stretched to threads. Obviously these black threads are distributed uniformly in the region

$$a < |p| < b \quad , \qquad 0 < x < L \quad , \tag{26-8}$$

where $\Delta p = b - a$ is the diameter of the black disc. This uniform distribution occurs when $t >> \tau$,

τ = time interval to traverse length of box .

We can say that this model has the mixing property.

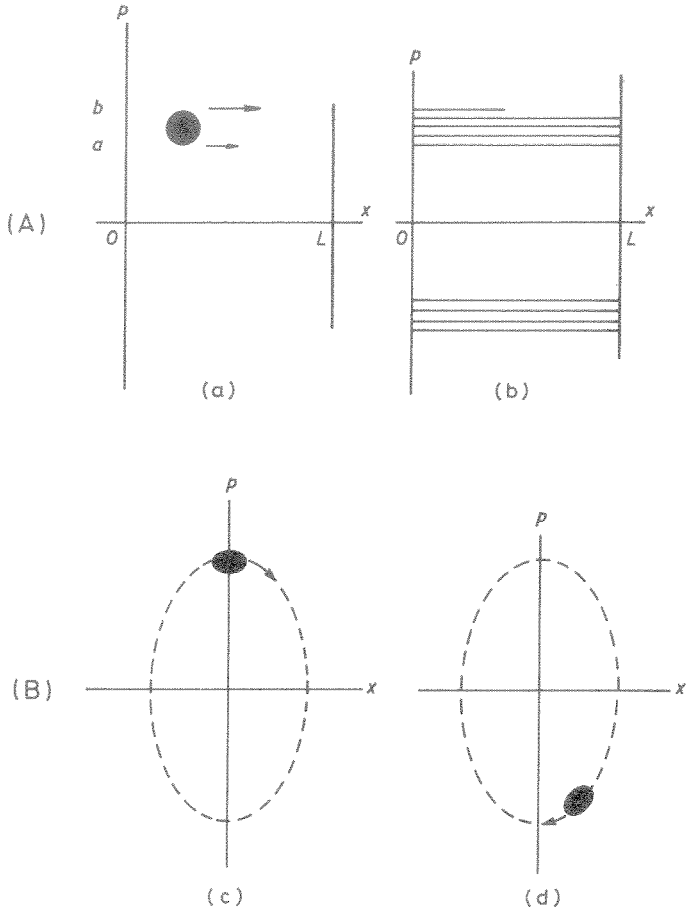

Fig. 26-1

Model (B) is completely different. The black disc in Fig. 26-1(c) goes in an elliptical path with slightly changing shape but never stretches into threads. No matter how long the time is, situations such as Fig. 26-1(b) cannot occur. (See Fig. 26-1(d).) Therefore, this model has no mixing.

So, from the point of mixing, these two models are completely different. But, from the point of independence these two models have no great difference, i.e. neither shows the independence property. Both are periodic motion without chaos.

The viewpoint of this book is not to consider the initial state as a group of points. We consider only one initial configuration. The period of motion of model (A) is determined by $|p|$ to be a constant.

The above examples are obviously too simple. We do not discuss whether (26-6) and (26-7) are different in more complicated models, but at least these examples point out certain basic differences of viewpoints.

26.4. Probability and Experiments

In current literature probability usually denotes something like the Gibbs' ensemble. The discussions of equilibrium and the increase of entropy necessarily involve the evolution of probability distributions in order to see how the probability evolves from an initial distribution (or *a priori* probability) and gradually disperse to fill up the region of motion. Correlation time refers to the time of dispersion for this probability distribution.

In the author's opinion, the meaning of this probability is very unclear. Let us forget about quantum mechanics for the moment. The concepts of statistical mechanics are derived from classical mechanics. If we cannot elucidate the situation in classical mechanics, the case of quantum mechanics would be hopeless.

In one point of view, probability expresses the knowledge of the observer. If he knows more about the system, the probability distribution is more concentrated. This is obviously incorrect. The motion of a system is independent of the psychological condition of the observer. According to this view, different persons observing the same system will have the same knowledge. This is unacceptable.

Another way is to say that this probability represents the outcome of experiments. This seems to be a more reasonable view, since experiments give us a sense of objectivity. But this proposition is more or less the same as the previous one. The question is: What is the experiment? How many experiments must be performed? The answers of "macroscopic experiments" or "feasible experiments" are incorrect. In the discussion of fixed impurities and the analysis of echo phenomena in Chapter 24, the position of each impurity atom is known. Is this a feasible experiment?

The experiments discussed in the literature on statistical mechanics are usually unrelated to actual experiments. For example, the canonical ensemble represents a fixed temperature for which the experimenter should be able to measure energy to an accuracy of $O(1/\sqrt{N})$. The other large variables are also determined to $O(1/\sqrt{N})$. In practise $N \sim 10^{22}$. Therefore $N^{-\frac{1}{2}} \sim 10^{-11}$. No experiments have been performed to such accuracy. In addition if the probability

tends to zero, then the event is explained to be not happening or not observed in experiments. This too is untenable, e.g. any metastable states such as diamond have a probability tending to zero. Yet the systems encountered in the laboratory are usually metastable.

One main weakness in the literature is to regard the region of motion or the equilibrium probability distribution as something known. The total energy function H is fixed, and then the equilibrium probability can be simply determinded, treating every configuration with the energy E as having the same probability. This manner of analysis is a habit acquired from analysing the gaseous state. (This book has repeatedly emphasised that the determination of the region of motion is not so simple and it depends on the actual distribution of the trajectories.) This habit is a hindrance to the analysis of the logical roots of statistical mechanics.

The author thinks that Gibbs' ensemble and probability have no relation with mechanics. From the viewpoint of statistical mechanics, they are both unnecessary and unrealistic. These concepts are helpful for our abstract analysis, but they are unrelated to any discussion on the logical roots of statistical mechanics. If we want a correct analysis of these roots, we have to get rid of this concept of ensemble probabiltiy.

Modern computers can calculate the trajectory of some simple models and can calculate the various thermodynamical properties from the trajectory. (See Chapter 22 and Chapter 25.) So the problem of experiments can be separated from that of the logical roots. These computer calculations are purely objective. Of course, the understanding of these simple models does not form the basis of statistical mechanics. But if we cannot even understand these models, the future does not look bright. The numerical solution requires Newton's laws but not the concepts of ensemble or probability.

The definition of probability given in this book is quite different. Probability is looked upon as a tool for arranging the information and not a physical concept. It is not a necessity but a convenience. We must first have numbers and information before we can do statistics, classify the states, make ensembles, calculate the probability and discuss equilibrium or nonequilibrium. All the properties must come from the trajectory.

We have performed a critical analysis of the concepts of ensemble and probability as presented in the literature. Our aim is to clarify the most basic concept. But the above criticism is not meant to denigrate the achievements attained in the literature. As mentioned above, we have a strong intuition about space and the flow and dispersion of a cloud in space. The concept of ensemble gives us a new tool and quite an elegant one. We cannot avoid assumptions of mathematical

concepts, but we have to clarify our concepts and make effective use of it.

26.5. Instability of the Trajectory

To understand mixing or independence from the laws of mechanics is a very difficult task. Equations (26-6) and (26-7) seem to be rather remote from the details of the trajectory. We now discuss instability and its details. The meaning of instability is as follows. Let $s(t)$ be the trajectory. If the initial state $s(0)$ is changed by a bit to $s(0) + \delta s(0)$, then the change of the trajectory is $\delta s(t)$. If

$$\delta s(t) \propto e^{\alpha t} \quad , \tag{26-9}$$

with $\alpha > 0$, the trajectory is said to be unstable. If there are many different $\delta s(0)$, forming a small cloud surrounding $s(0)$, then this cloud will disperse, conforming to the requirement of mixing.

In model (A) of Sec. 26.3, if initially momentum is increased by δp, then

$$\delta x(t) = \frac{\delta p}{m} t \quad . \tag{26-10}$$

There is also instability, but not as serious as (26-9).

Now let us look at a more complicated example. Suppose a particle collides with a number of fixed hard spheres, as in Fig. 26-2. We consider only a model with planar motion. The configuration of this particle is determined by its position and velocity. The magnitude of the velocity is unchanged, so we only consider the direction of the velocity. Starting from the origin, if the direction makes a small change $\delta\theta(0)$ the trajectory will change, and $\delta\theta$ changes for every collision. Assume that the density of the hard spheres is very low and the mean free distance λ of the particle is much larger than the radius a of the spheres. Figure 26-2 shows the situation after two collisions. As seen from the figure

$$\delta\theta' = 2\delta\alpha = 2\lambda\delta\theta/a\cos\alpha \quad . \tag{26-11}$$

Therefore $|\delta\theta'| \geq (2\lambda/a)|\delta\theta|$, and after n collisions

$$|\delta\theta(t)| \geqslant |\delta\theta(0)| e^{t/\tau} \quad ,$$

$$\tau \equiv \lambda \Big/ \left(v \ln \frac{2\lambda}{a} \right) \quad , \tag{26-12}$$

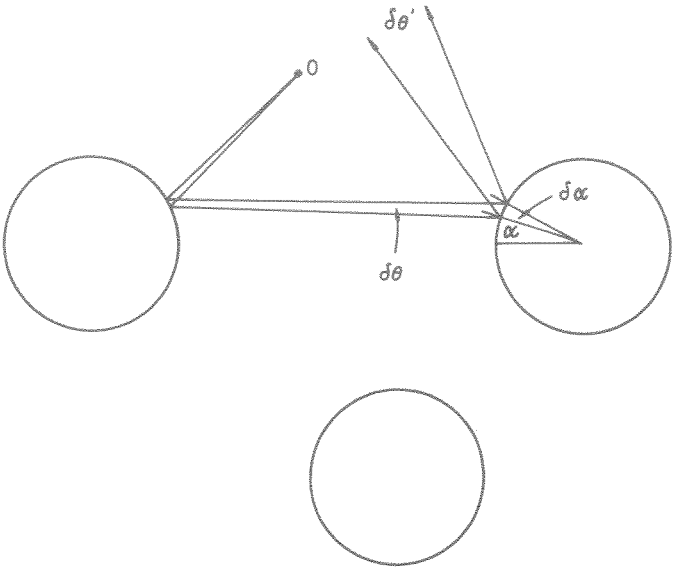

Fig. 26-2 Collision of a particle with fixed hard spheres.

where v is the speed of the particle. This result is the same as (26-9). If the time is sufficiently long, i.e.

$$t \geqslant \tau \ln(1/|\delta\theta(0)|)$$

$$= \frac{\lambda}{v} \frac{\ln|1/\delta\theta(0)|}{\ln 2\lambda/a} \quad , \tag{26-13}$$

then $|\delta\theta(t)| \sim 1$, i.e. the direction is randomised. The above analysis can be extended to the gas model with similar results. We mention it briefly here.

Suppose we have N small hard spheres of radius a and the density of the molecules is very low. Let $\mathbf{r}_i$, $i = 1, 2, 3, \ldots, N$ be the positions of the spheres. Let

$$\mathbf{r} \equiv (\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_N) \quad , \tag{26-14}$$

where $\mathbf{r}$ lies in a $3N$ dimensional space. In this space there are $\frac{1}{2}N(N-1)$ cylinders:

$$|\mathbf{r}_i - \mathbf{r}_j| \leqslant a \quad , \qquad i, j = 1, 2, \ldots, N \quad , \tag{26-15}$$

and $\mathbf{r}$ is excluded from the cylinders because they are hard spheres. So the motion of these molecules is represented by $\mathbf{r}(t)$, which goes in a straight line and is reflected whenever it hits a cylinder. This situation is similar to that of the above example. We have to regard the black disc of Fig. 26-2 as the cylinder in the $3N$ dimensional space. Notice that $d\mathbf{r}/dt = \mathbf{v}$ is the velocity and

$$v^2 = \sum_{i=1}^{N} \left(\frac{d\mathbf{r}_i}{dt} \right)^2 \quad , \tag{26-16}$$

is unchanged because v^2 is proportional to the total energy.

If we follow $\mathbf{r}(t)$, we obtain a trajectory. Now alter the direction of $\mathbf{v}(0)$ slightly to $\mathbf{v}(0) + \delta\mathbf{v}(0)$, and see how $\delta\mathbf{v}(t)$ changes. It is the same as the above example except that $\delta\mathbf{v}$ is more complicated than the above $\delta\theta$. The infinitesimal velocity $\delta\mathbf{v}$ has $3N$ components. One good feature is that the collisions are binary. In the centre-of-mass coordinate system of two molecules, the analysis of the angle is similar to the above example. Approximately, after each collision $|\delta\mathbf{v}|$ will increase by a factor λ/a, where λ is the mean free distance. But collisions with different molecules will change the direction of $\delta\mathbf{v}$. The changes of $\delta\mathbf{v}$ in collisions with different molecules are mutually perpendicular. For example, in a collision of molecules 1 and 2, there are changes of $\mathbf{v}_1$ and $\mathbf{v}_2$, so

$$\delta\mathbf{v}_{12} = (\delta\mathbf{v}_1, \delta\mathbf{v}_2, 0, 0, \ldots, 0) \quad . \tag{26-17}$$

If molecules 7 and 8 collide, then only $\delta\mathbf{v}_7$ and $\delta\mathbf{v}_8$ are nonzero.

$$\delta\mathbf{v}_{78} = (0, 0, \ldots, \delta\mathbf{v}_7, \delta\mathbf{v}_8, 0, 0, \ldots, 0) \quad . \tag{26-18}$$

Therefore,

$$\delta\mathbf{v}_{12} \cdot \delta\mathbf{v}_{78} = 0 \quad . \tag{26-19}$$

This is similar to the analysis in Sec. 5.9.

So, after n collisions $\delta\mathbf{v}(t)$ becomes the sum of n vectors. These n vectors are mutually perpendicular and

$$n \sim \frac{t}{\tau_0} N \quad ,$$

$$\tau_0 \equiv \lambda / v \quad , \tag{26-20}$$

where v is the average velocity of the molecules. The magnitude of each vector is about

$$\delta \mathrm{v}_i(t) \sim \left(\frac{\lambda}{a}\right)^{t/\tau_0} \delta \mathrm{v}_i(0) \quad . \tag{26-21}$$

Notice that t/τ_0 is the number of collisions of each molecule. The correlation time can be estimated as $\tau_0 / \ln(\lambda/a)$ similar to (26-12). As the vectors are mutually perpendicular, they form a volume

$$\prod_i \delta \mathrm{v}_i(t) \sim \left(\prod_i \delta \mathrm{v}_i(0)\right)\left(\frac{\lambda}{a}\right)^{3Nt/\tau_0} \quad . \tag{26-22}$$

The increase of this volume is exponential:

$$\left(\frac{\lambda}{a}\right)^{3Nt/\tau_0} = \exp\left[3N(t/\tau_0)\ln(\lambda/a)\right] \quad . \tag{26-23}$$

The space of v is momentum space $\mathbf{p} = (\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_N)$ as $\mathbf{p} = m\mathbf{v}$. The above analysis shows how the volume expands. Notice that (26-23) is a large number of magnitude e^N. The above is only a crude analysis.[e]

The above analysis explains approximately the meaning of instability. If the trajectory is unstable, then the trajectory is not easily controlled and extremely sensitive. If the initial state differs a bit, the final outcome will be completely different. If the trajectory has the property of instability, there will be no problem in mixing. Notice that instability is a result of mechanics and does not involve the concept of probability. To use it to discuss mixing we must first assume that there is an initial probability distribution.

In this kind of analysis it is very easy to allow the assumptions of mixing and independence to creep in unnoticed. For example, in (26-13) we assume that after $|\delta\theta(t)|$ equals 1, it will be "unrelated" to $\delta\theta(0)$ and hence we can disregard $\delta\theta(0)$. In fact the above analysis does not say how $\delta\theta(t)$ changes when it becomes very large. Equation (26-12) is appropriate only when $\delta\theta(t)$ is very small. When $\delta\theta(t)$ becomes large, we are ignorant and naturally there is no reason to draw any conclusion. This analysis cannot be used as a proof of mixing or independence. Equation (26-12) or (26-13) can only be regarded as an estimate, assuming that $\delta\theta(t)$ will be unrelated to the situation at $t = 0$ or independent when $\delta\theta(t)$ grow large. However, this way of analysing enables us to have a deeper understanding of the characteristics of the trajectory.

[e] For details, see Krylov (1979), p. 193. It is detailed but not quite rigorous.

Part of the contents of this section came from inspiration while readir Krylov's book. The discussion by Krylov is very penetrating but the views c the author do not quite agree with those of Krylov.

26.6. The General Problem of Independence

To summarise: From the point of view of physics, to understand statistic mechanics we must understand

(a) the existence of a correlation time $\tau = O(1)$. That is to say, the correl tion time must not only be finite, but of $O(1)$ and not a large number of $O(e^N$

(b) the existence of a correlation distance $\xi = O(1)$. That is to say, a syste can essentially be regarded as composed of independent parts. The correlatio distance ξ must also be of $O(1)$ and cannot be a large number of order Λ

These two points are the contents of the independence condition (26-7

In recent years, the research of mechanics in the direction of random traje tory has made great advances. Many simple models (N very small, no mor than 3 or 4 variables) exhibit random (i.e. independent in time) motion. Thes works show that random motion is a result of Newton's laws. However, up t now, the values of N investigated is still very small and is not sufficient t discuss the independence of the parts of a system, i.e. the existence of th correlation length ξ. When N is very small, the study of random motion is i the field of numerical simulation and has not reached a profound understandin of the problem. The author thinks that the research should be directed to th influence of the value of N, especially when N is around 10. Large values o $N \geqslant 10$ may produce random motion different from that of small N.

Indeed, our understanding of the basic assumption is still very primitive

Although to derive (26-7) from the principles of mechanics seems extremel difficult, this property of independence is actually not unexpected from ou daily experience. Independence is "randomness" or "chaos". This has bee discussed in Chapter 3. The property of independence is so common that i should not follow only from Newton's laws. For example, if we can understan random sequences we can understand more about the basic assumption. Th random sequence is generated by a computer programme and should be easie to understand.

Why a random sequence is random at all is still unknown. From the molecula motion to the throwing of a die, to the random sequence and the stock market all these random phenomena are not understood. The cause of chaos seems t relate to the degree of complexity, but simple models can still give rise to chaos These problems are waiting for an answer.

26.7. Difficulties Encountered in Applications

From the examples of the above sections we see that the application of statistical mechanics is an extremely difficult mathematical problem. If we cannot transform a model to one similar to the ideal gas, we cannot solve it. That is to say, we have to separate the variables of motion into mutually independent parts and then solve each part separately. In an ideal gas, each molecule is an independent part. In the crystal model, each elementary vibrating element (i.e. normal mode) is an independent part, etc.

If there are N variables, we have to look for the relation between them and put the correlated elements into a bunch, making the independence property obvious. In the various examples in this book, no matter how complicated the solution was (e.g. the exact solution of the two-dimensional Ising model), we had to separate out the independence property and the correlation property of the model, and then to solve it. At present the usual models are not tractable and have to be solved approximately or numerically.

The author thinks that the difficulty of solving problems is related to our lack of understanding of the independence property, i.e. as in the discussion of this chapter our lack of understanding of the origin of the basic assumption. The mutual interaction sets up the independence property, and is the origin of correlation. Correlation when mixed with independence makes the problem difficult to solve. From mechanics we cannot arrive at the independence property. But independence simplifies the application of the basic assumption.

At present, reliance on numerical solution is gradually increasing. As we said in Chapter 22, the numerical method is very important. But from the point of understanding the origin of statistical mechanics, this is perhaps not a good long-term solution.

Problems

1. As pointed out in Sec. 26.5, the correlation time of a gas is approximately equal to the mean free time of the molecules. This conclusion has some deficiencies. Try to find them from Prob. 23.1.

This problem shows that in some special cases the correlation time is very long. These special cases are usually overlooked.

2. Apply the method of Sec. 26.5 to the model of crystal vibration.

(a) Use a one-dimensional model; see Sec. 8.3. Use Eq. (8-41) as the interaction. First neglect the α' term and find the equation of motion.

(b) Add the α' term and discuss its effect, assuming α' to be small.

(c) Find the correlation time.