

2.1.1 Thermodynamic Probability and Entropy

In 1854, Rudolf J. E. Clausius introduced the *entropy* (S) as a parameter of phenomenological thermodynamics, and defined it as the heat, added to a system in a reversible way in relation to the temperature (see Eq. 3.10). Later, in 1894, Ludwig Boltzmann used this parameter in the framework of statistical thermodynamics. In these circumstances, the entropy, and in context to this, the second principle of thermodynamics becomes more imaginable. Entropy appears as a kind of measure of disorder, or as a degree of random distribution, i.e., of missing order. The correlation between order and probability and, as will be explained later – information – is of great importance for the understanding of the principles of biological organization.

Let us start with the assumption that the entropy is a measure indicating the degree of randomization of a given distribution. We will consider a system of maximal entropy as a system in maximal disorder. Furthermore, let us demand that the entropy be an *extensive* parameter. Therefore, like volume, or mass, but in contrast to the *intensive* parameters such as for example temperature or density, the entropies S_1 and S_2 of two systems can be added, if these systems come together:

$$S_1 + S_2 = S \quad (2.1)$$

How can we now define a parameter, which indicates the degree of randomization or, on the contrary, a degree of disorder? What does order of organization mean? Of course, our daily experience shows that an ordered system spontaneously transforms into a disordered one, but not vice versa. This, actually, is the consequence of the second principle of thermodynamics.

Let us consider a very simple structure, just the distribution of four distinguishable spheres on two compartments of a box (Fig. 2.1). Let each of these spheres, independently of the three others, just by chance fall into one or the other compartment of the box. All of the 11 possibilities of the distribution, as indicated in Fig. 2.1, therefore, have the same degree of probability, because the probability of each sphere individually, to fall into compartment 1 or into compartment 2 is equal. Summarizing the patterns of distribution shows that there is only one way to

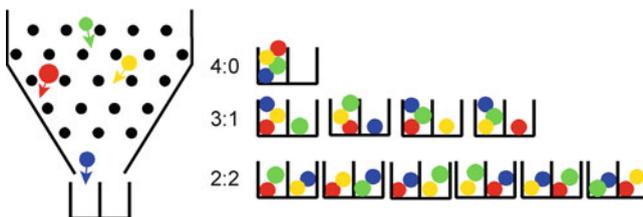


Fig. 2.1 All possibilities of the statistical distribution of four distinguishable spheres in two compartments of a box

realize the distributions 0:4 and 4:0. In contrast, there are four ways to realize the distributions 3:1 and 1:3, and, finally, six ways for equal distribution: 2:2.

Let us now ignore the fact that the spheres are distinguishable. Let us simply ask: How large is the probability that just by stochastic distributions one of the relations 4:0, 3:1, 2:2, 1:3, or 0:4 occur? Apparently, the probability of any kind of distribution will be larger, if it can be realized by a larger number of ways. The distribution mode 2:2, for example, is 6 times more likely, than the distribution 4:0, or 0:4. The number of ways which lead to the realization of a definite situation, in fact, seems to be a measure of the probability of the occurrence of it. We will designate this number of ways by the parameter W which we will call *thermodynamic probability*. The amount of W can be at least 1 and at maximum ∞ , in contrast to the mathematical probability (P), which we will use in Sect. 2.1.2, and which ranges between 0 and 1.

Now, we come to the following sequence of conclusions: If W really is a measure of the probability of getting a definite distribution, and if an increase of the degree of order is the most uncertain result of a stochastic distribution and finally, if the entropy (S) is a parameter, indicating the degree of disorder – than S should be a function of W . If two situations with relative probabilities W_1 and W_2 are connected together, then the probability of this combined situation results from the product ($W_1 \cdot W_2$). Using Eq. 2.1, this means:

$$S = f(W) = S_1 + S_2 = f(W_1) + f(W_2) = f(W_1 \cdot W_2) \quad (2.2)$$

This demand is met by the logarithmic function:

$$\ln A + \ln B = \ln(A \cdot B) \quad (2.3)$$

Hence entropy is proportional to the logarithm of W :

$$S = k \ln W \quad (2.4)$$

This is the Boltzmann equation of entropy. Boltzmann's constant k was defined as a universal constant later by Max Planck. It must have the same unit of measurement as entropy, and is as follows:

$$k = 1.380658 \cdot 10^{-23} \text{ J K}^{-1} = 8.6174 \cdot 10^{-5} \text{ eV K}^{-1}$$

This explanation was just based on the simplest experiment where four spheres were distributed randomly over two compartments. One step toward serious thermodynamics can be taken, considering for example the compartments of this box as molecules of a system, and the spheres, as quanta of energy, distributed among them. This complication, of course, means a transition of handling with larger numbers. If the number of elements and classes are increased, W cannot be

evaluated just by simple trial. It is possible to calculate this value using the following equation:

$$W = \frac{n!}{n_1! \cdot n_2! \cdot n_3! \cdot \dots \cdot n_m!} \quad (2.5)$$

where n is the total number of all elements in the system (in case of Fig. 2.1 – the total number of spheres: $n = 4$); n_i (for $i = 1 \dots m$) is the number of elements in each class of state (this means, the number n_1 in compartment 1 and n_2 in compartment 2); and m is the number of classes of state (namely: number of compartments in the box).

2.1.2 Information and Entropy

In 1948, C. E. Shannon introduced a parameter which in technical information theory has been proved as a useful measure of information content of a message. The information (I) of a message depends on the effort required to guess it by a highly systematic system of questions. Hence, information is some sort of degree of the actuality of a message.

It is not difficult to guess the result of the toss of a coin, since there are only two possibilities of equal probability. To guess a certain card in a full deck of playing cards is much more difficult. In this case, a much greater uncertainty factor has to be taken into account. Using a more systematic approach, a large number of yes-no questions have to be answered. Hence, the information content of a particular playing card is higher than that of a tossed coin. Should a deck consist of cards which are all the same, and should this be known to the challenged person, guessing will not make sense at all. The information content of each of these cards is zero. The probability by which possibilities are turned into reality, consequently, seems to become a measure of information.

In contrast to thermodynamics, in the information theory the mathematical term of probability (P) is used which is defined as follows:

$$P = \frac{\text{number of favorable cases}}{\text{greatest possible number of cases}}$$

On average, coins tossed a 100 times will land with heads up in 50 instances. Hence, the probability of heads facing up may be expressed by:

$$P = \frac{50}{100} = \frac{1}{2}$$

Conversely, the probability of throwing a “six” with some dice is only $P = 1/6$, whereas the probability of throwing one of the three even numbers would be higher: $P = 3/6 = 1/2$.

Whereas the thermodynamic probability (W) is always larger than 1, (cf. Sect. 2.1.1), the value of the mathematical probability lies between 0 and 1 ($0 \leq P \leq 1$). In this context, $P = 0$ means an impossibility, while $P = 1$ expresses absolute certainty.

The logical conclusions which led to the derivation of the Boltzmann equation (cf. Sect. 2.1.1) are the same as those on which the Shannon relation is based. Information (I) is a function of mathematical probability (P):

$$I = f(P)$$

The condition for the function f again, is satisfied by the logarithmic function, since here too, the multiplication rule for the calculation of probabilities must be valid (cf. Eq. 2.3). To obtain positive values of information, considering that $P \leq 1$, the negative value of the logarithmic function must be used. The information of a single event therefore is:

$$I = -K \ln P \quad (2.6)$$

This is the *Shannon equation* of information theory. The unit of I is determined by the unit of the always positive factor K . The bit (*binary digit*) is most commonly used. It expresses the number of binary yes-no decisions that are needed to determine a given message. For example, the one side of the coin can be guessed by one single decision, its information value, consequently, is 1 bit. Five binary decisions will be sufficient to guess a card from a deck. Hence, the information value of one card is 5 bits. The factor $K = 1/\ln 2 = 1.443$ must be used to calculate I in bits. In the information theory the logarithm to the base 2 (\log_2) is occasionally used:

$$I = -1.443 \ln P = -\log_2 P \quad (I \text{ in bit}) \quad (2.7)$$

A message usually consists of elements like symbols, words, or other structures of a set defined in advance. This concerns not only the letters of a word or the words in a written text, but the same approach can be applied also for various cases of biological information storage and transfer. So, for example, the information content of a DNA molecule is coded by the sequence of nucleic acids; the information content of a protein is given by their amino acids. In a similar way the information content of time-dependent structures like nervous signals can be evaluated too. In this case discrete patterns of neural spike trains are considered as elements of the message.

Assuming for example that a mammalian DNA molecule consists of 15,000 pairs of nucleotides and the four possible types of nucleoside bases have an equal probability of occurrence, then the information content of each single nucleotide

will, consequently, have a value of 2 bits. The information capacity of this DNA molecule therefore amounts to 30,000 bits.

In fact, however, in a DNA molecule the probability of the occurrence is not equal for all nucleotides. In this case, the information content of the whole message can be expressed by the sum of the evaluated probabilities (P_i) of the single elements (i) by the following equation:

$$I = -K \sum_i P_i \ln P_i \quad (2.8)$$

Estimates of information capacity in the whole genome of an organism in this way range up to about 10^{10} bit. The amount of actually stored information is even lower, if the redundancy in the storage of genetic information which is required for the protection of the information is taken into account.

It seems important to emphasize that the Shannon Information (I) as evaluated here, is the so-called *syntactic* information. It allows for example important conclusions about the maximum storage capacity of a DNA molecule of the whole genome. This Shannon Information (I), however, has nothing to do with the “information content” in the everyday use of this word, the so-called *semantic information*.

The difference between the syntactic and the semantic character of information can be illustrated best by the following example: A book, consisting just of stochastically jumbled letters, or even words according to Eq. 2.8 have a larger amount of syntactic information than such with a meaningful text, because in the stochastic sequence the probability of the occurrence of elements is lower than in an organized one. If P_i becomes smaller, its negative logarithm, and therefore I gets larger. Knowing some characters of a reasonable word, for example, the rest of them can usually easily be guessed – the probabilities of their occurrence therefore become larger, its content of syntactic information diminishes. If the same number of characters is mixed stochastically, guessing is more complicated – information is larger.

Transferred to the biological situation: The Shannon, i.e., the syntactic information of a polynucleotide with an arbitrary sequence is larger than that of a DNA with the same molecular length. In contrary to the *syntactic* information, the *semantic* information of a reasonable book, of course, is much larger than a book containing stochastically selected characters or even words. The same refers to the biological DNA in contrast to a polynucleotide with no biological function.

Does this mean that it is impossible to quantify biologically important information? Does it mean that the information concept is not applicable to biological systems at all? In fact, the semantic aspects of communication are irrelevant to most engineering problems and therefore not included in the definition of Shannon’s I . Despite many attempts, quantification of semantic information has not yet been achieved.

Nevertheless, even the use of the syntactic information properties of biomolecules and nervous processes has proved to be a useful tool in modern molecular biology.

The enormous increase of the amount of data on the primary structure of biological macromolecules, especially DNA sequences, as well as data of neurophysiological recordings nowadays required mathematical methods for their evaluation, comparison, and storage. For this, bioinformatics was developed as a special discipline, based on these considerations.

Formally, the Boltzmann's constant (k) can be used in the Shannon equation (Eq. 2.6). In this case information is obtained formally in entropy units (JK^{-1}). Therefore the term "entropy" as Shannon-Entropy in bioinformatics was frequently used in the sense of information. This formal interconnection of Shannon's definition of information with Boltzmann's entropy triggered off a discussion about the relation between these two parameters.

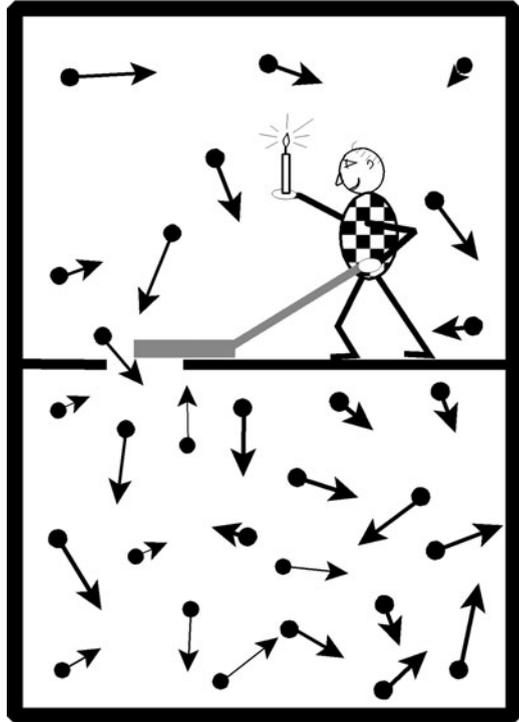
The starting point for this discussion was the second principle of thermodynamics, predicting that isolated systems spontaneously try to reach a state of maximum disorder. The growth and the existence of a living system on the contrary is possible only by decreasing or at least by the conservation of entropy. Erwin Schrödinger (1944) made the frequently quoted statement: "The living system feeds on negative entropy." This is the reason, why sometimes the term "negentropy" was used.

The interconnection of information with entropy in terms of thermodynamics may be illustrated best by a thought experiment conceived by James Clerk Maxwell in 1881 and still discussed today (Fig. 2.2). Maxwell proposed a room, which is connected with another by an opening. This opening can be closed by means of a slide. Both rooms are filled with a gas which is, in the beginning, in equilibrium, for example of equal pressure and temperature. An intelligent creature, later called "Maxwell's demon," is able to handle the slide between the two rooms with ease. This "demon" can observe accurately the direction and the velocity of the molecules in his room. Velocity and direction of these particles in the beginning are statistically distributed. If a particle in the upper room flies accidentally toward the opening, the demon opens the slide to let the particle pass. As a result of such sorting the pressure in the lower room would rise.

The demon could also take another approach. For example, he could separate fast from slow particles. In this case, a difference in the temperature between the two rooms would occur. In both cases, the entropy of the whole system would be reduced and energy might be generated by an appropriate turbine. The result would be a "*perpetuum mobile*," a perpetual motion machine of the second order, as it would contradict the second principle of thermodynamics.

This apparent contradiction subsequently was the subject of a large number of scientific and philosophical publications. The problem was finally resolved by the following consideration: The "demon" requires information to carry out the sorting. He collects this information by "watching" the molecules. In order to "see" he needs light. For this, the demon in Fig. 2.2 symbolically carries a lit candle. Yet, a body will only be able to emit light in a state of nonequilibrium relative to its environment. This, however, contradicts the equilibrium condition at the beginning of the experiment. The same would apply to any other approach to acquisition of information. This resolves the apparent contradiction to the second law of thermodynamics.

Fig. 2.2 Maxwell's demon



Why do we discuss this thought experiment here, if it is clear that it will not work as a perpetual moving machine? What has this to do with biophysics? In fact, independent of this discussion about the virtual violation of the second law of thermodynamics, Maxwell's demon demands particular interest in biophysics because of its analogy to various functions of living systems. The living cell, too, reduces its entropy at the expense of its environment, using information processing. Yet, in the latter instance it is not the energy of statistical fluctuations, which is used in this case. The biological system selects such molecules from its environment, which are rich in free Gibbs energy of formation and correspondingly, with a low content of entropy. It uses this energy and extrudes molecules with lower free energy and larger entropy. The basic information for this process of selection, in other words, the "software" for this process, is stored in the structure information of the proteins, which are responsible for the recognition of these molecules, and eventually for their metabolism. These proteins get this sort of semantic information during the process of synthesis via the RNA, from the DNA of the genome.

This example again raises the question of the relation between semantic and syntactic information: what is the threshold value of information that is required to control the processes of living systems? Or translated into the language of modern computer science: how large must the simplest software for this sort of a biological

Maxwell demon be? What is the threshold information that carries out not only the metabolic function of the primordial organism but additionally its replication? How did the first accumulation of information in a molecule occur? Purely accidental combination of monomers to build their first functional macromolecule must be ruled out. The probability for this occasion is too low by far. Today, so-called *probiotic evolution* is assumed, i.e., chemical selection of developing polymers even before the stage of the biopolymer (see also Sect. 5.2.5).

Further Reading

Shannon and Weaver 1962; Bioinformatics in molecular biology: Kauffman 1993; Strait and Dewey 1996; Yockey 2005; Lesk 2002; Maxwell's demon: Leff and Rex 1990.

2.1.3 Biological Structures: General Aspects

In the previous section we introduced expressions like *order*, *structure*, and *organization* and discussed them in context with entropy and information, as well as with the statements of the second law of thermodynamics. This touches on a set of questions which are of central interest in biophysics and which will be mentioned in many sections of this textbook. Therefore it is necessary at this point, to explain some basic definitions and principal ideas.

What, really, is a *structure*? To the biologist, the term “structure,” usually is related to the macroscopically or microscopically visible organization of an organism. This means, for example, the structure of an animal skeleton, structure of a cell, of a mitochondrion, etc. The term “molecular structure” already lies outside the limits of this view. It refers to a certain arrangement of atoms, without defined contours, which can be described just by means of wave mechanics. The same applies for the concentration profile of an electrical double layer (Fig. 2.43), and particularly for so-called “time structures,” namely special time courses, like oscillations of a biological system (Figs. 5.3 and 5.16), like the shape of an electrocardiogram (Fig. 3.38), or like the sonogram of a bat's cry (Fig. 4.19). This means that the definition of the term “structure,” which is used in biophysics has to be broader than that of the morphologists and cytologists. It must include these structures as well as those of metabolic networks, ecosystems, or others.

The best, and generalized definition of this term is given by the set theory of mathematics. Therefore: *a system is an aggregate of elements with certain interrelations between them*. The totality of these interrelations is called the *structure* of the system. This definition does not prescribe at all, what kind of elements, and what kind of interrelations these are. It is applicable to all kinds of systems including biological systems and structures. In biophysics, we are interested especially in *dynamic systems*, i.e., in such systems, the interrelations between their elements are *interactions*. In contrast to this, in *static systems* the elements have no interaction at all, but are just interrelated by formal *relations*. Examples for static

systems, are the system of natural numbers in mathematics, or the system of animal species or of plants in biology.

The elements of a metabolic network are the metabolites, and the interrelations between them, i.e., their interactions, are the steps of biochemical reactions. Correspondingly, an ecosystem is to be considered an interaction of individuals and populations depending on abiotic conditions.

Biological macromolecules can be considered as systems too. In this case, the following levels or organization occur depending on what will be assumed as their elements.

- *Primary structure* – linear sequence of monomers (= elements) in the linear molecular chain (= system). For example, a polypeptide chain: ...-serine-alanine-lysine-arginine-.
- *Secondary structure* – positioning in space of monomers (= elements) in a part of the molecule (= system) relative to each other. For example, the α -helix, or the β -sheet structure of an amino acid sequence in a protein.
- *Tertiary structure* – position in space of molecular regions of homogeneous secondary structures (= elements) in a molecule (= system). For example, intramolecular coordination of the position of several helical regions relative to each other or to a β -sheet.
- *Quaternary structure* – position in space of macromolecules (= elements) in a supramolecular complex (= system). For example, single proteins in a multienzyme complex.

When a salt is crystallized, a *periodic structure* forms, which is characterized by a periodic arrangement of their elements. Erwin Schrödinger (1944) called the biological structure an *aperiodic crystal*. This means a highly organized structure, the elements of which, are not simply repeated periodically. Sometimes one tries to evaluate this structural organization as *structural information*. As we pointed out in the previous Sect. 2.1.2, however, it is hard to quantify this parameter. Probably, the structural information should be measured as the effort which is necessary, to describe such a structure perfectly.

Consequently, the process of structure formation of biological systems, be it the development of life, its reproduction, or simply the biosynthesis of a macromolecule, all are accompanied by reduction of entropy. This appeared to be contrary to the second law of thermodynamics and has given rise to heated philosophical discussions in the past. The second law of thermodynamics actually postulates that in spontaneous processes occurring in isolated systems, the entropy strives towards a maximum. Yet, neither an organism nor its environment, i.e., the earth as a whole can be considered as an isolated system. The earth is constantly absorbing energy from the sun and is emitting this energy again. That continuous flow of energy maintains a permanent nonequilibrium state which manifests itself not only in a direct way in photosynthesis with subsequent metabolism of heterotrophic organisms, but also in the environment of life, for example, in flowing water, in alternation of light and darkness, and in changes in temperature, humidity, etc.

In fact, structures are also formed in inanimate nature under loss of entropy. Basically, a distinction must be made between two kinds of structure formation which at the same time explain the two ways of self-organization in living systems.

- *Equilibrium structures*, for example a salt crystal, formed spontaneously during evaporation of a solvent.
- *Nonequilibrium structures* (or: *dissipative structures*), for example an arrangement of cirrus clouds as the result of processes of air convection under special meteorological conditions.

The genesis of life which is based on the prebiotic formation of a first bio-macromolecule, and the subsequent evolution of all organisms, can be understood as the result of complicated processes that occur far from equilibrium. In this context, a major role is played by dissipative structures of the inanimate environment, such as periodic temperature variations, tides, etc. Substantial importance has to be attributed also to such nonequilibrium structures in the course of life itself. They are frequently represented by time structures, such as for example the heart rate, or other kinds of oscillations, which in some cases are associated with the so-called biological clock (for further explanation see Sects. 3.1.4, 5.2.4, and 5.2.5).

Equilibrium structures, such as inorganic salt crystals with a very primitive geometry, become very complex and variable in shape, when based on the sophisticated pattern of the primary structure of bio-macromolecules rather than on the relatively simple field of interaction of spherico-symmetric ions. The spontaneous folding of proteins and their arrangement to supramolecular structures, such as complex enzyme systems of even ribosomes, must be viewed from this aspect. Such processes are also referred to as *self-assembly*. More specific aspects relating to the formation of equilibrium and nonequilibrium structures will be given in subsequent chapters of this textbook.

Further Reading

Eigen 1971, 1992; Meinhardt 2008, 2009; Kauffman 1993; Strait and Dewey 1996.

2.1.4 *Distribution of Molecular Energy and Velocity at Equilibrium*

The Boltzmann equation of entropy (Eq. 2.4) as derived in Sect. 2.1.1, helps to illustrate the second law of thermodynamics, according to which, isolated systems spontaneously approach a state of maximum entropy. We demonstrated there that one can express this also in the following way: at equilibrium isolated systems reach a state of highest realization probability (maximum of W). Now, we will ask the question: what are the mean properties of the molecules at this equilibrium state? This is of fundamental importance for further considerations.

It is possible to answer this question on the basis of statistical thermodynamics, which helps to make predictions on the probability of energy distribution among the elements of a system in thermodynamic equilibrium. It should, however, be emphasized that a deterministic description of the properties of single particles lies outside the limits of this discipline. Nevertheless, it will be seen that even a statistical statement allows the drawing of important conclusions regarding reaction rates, stability of molecules, and many others.

Let us imagine a space that contains gas molecules of uniform mass (m), which at the beginning all have the same velocity (v), and consequently, the kinetic energy (E) which can be calculated by the following equation:

$$E = \frac{m}{2} v^2 \quad (2.9)$$

This equality of the kinetic energy of all molecules, in fact, is a highly improbable state. In this case, using Fig. 2.1, all molecules would belong to a single box, or a single class of properties n_i . Thus, according to Eq. 2.5: $W = n!/n_i! = 1$. This situation will change instantaneously. The molecules would exchange their energy by elastic collisions with each other, and soon a great number of energy states would be occupied. A simple mathematical example shows that W , and according to Eq. 2.4 also the entropy (S), will increase with a growing number of (m) of state classes, provided the following relation applies:

$$\sum_{i=1}^m n_i = n$$

Because of the law of conservation of energy (first principle of thermodynamics), the following condition must be satisfied at the same time:

$$\sum_{i=1}^m n_i E_i = \text{const} \quad (2.10)$$

The energy of this system, therefore, can be distributed randomly among all of its molecules. The total energy of the system, however, must always remain constant.

Now, let us look for a state with maximum probability, this means, with a maximum of entropy. Corresponding to the second principle of thermodynamics, this is actually the thermodynamic equilibrium. Applied to our example the question arises: How energy and, correspondingly, velocity will be distributed between the n particles, after a sufficiently long period of time if the system is isolated? Even under equilibrium conditions, of course, the energy of individual molecules will change permanently, but nevertheless, the statistical mode of energy distribution, or of the distribution of various degrees of molecular velocity at equilibrium becomes

stationary, i.e., time independent (for a detailed description of various kinds of stationary states see Sect. 3.1.4).

Considerations of this type lead to *Maxwell's equation of velocity distribution*:

$$\frac{dn(v)}{n_0 dv} = \frac{4}{\sqrt{\pi}} \left(\frac{m}{2kT} \right)^{3/2} v^2 e^{-\frac{mv^2}{2kT}} \quad (2.11)$$

The left part of this equation contains a relation which expresses the relative number (dn/n_0) of those molecules which are related to a particular velocity interval (dv). Its unit is $s\ m^{-1}$. This is a function expressing the probability of the distribution of the velocity, where m is the mass of a molecule (not to be confused with the numeral m in Eq. 2.5), and k is the Boltzmann constant. Multiplying in Eq. 2.11 the denominator and numerator of the expression within the brackets as well as that in the exponent with the Avogadro number ($N = 6.023 \cdot 10^{23}\ \text{mol}^{-1}$) one introduces molar, instead of molecular parameters:

$$M = N \cdot m \quad (2.12)$$

and

$$R = N \cdot k \quad (2.13)$$

(M – molar mass, $R = 8.314\ \text{J}\ \text{K}^{-1}\ \text{mol}^{-1}$ gas constant)

In Fig. 2.3, as an example, the velocity distribution of molecules of oxygen is depicted ($M = 0.032\ \text{kg}\ \text{mol}^{-1}$). This curve is not symmetrical. The mean velocity (weighted arithmetic mean value) in general will be higher than the maximum value. The following relation applies:

$$v_{\max} = \sqrt{\frac{2kT}{m}} = \sqrt{\frac{2RT}{M}}; \quad \bar{v} = \sqrt{\frac{8kT}{\pi m}} = \sqrt{\frac{8RT}{\pi M}} \quad (2.14)$$

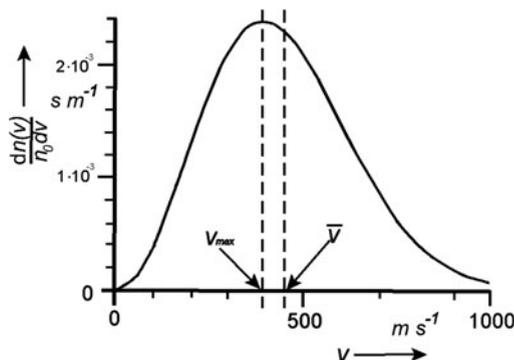


Fig. 2.3 Velocity distribution of O_2 molecules at 37°C corresponding to Eq. 2.11