Boltzmann was brilliant, undoubtedly a genius, far ahead of his time in theory. Of course he was not infallible. Most important for us moderns to realize, he was still very limited by the science of his era; dominant physical chemist and later Nobelist Ostwald named his estate “Energie” but did not believe in the physical reality of molecules nor in Boltzmann’s treatment of them.

Some interesting but minor details that are not too widely known: Even though Boltzmann died in 1906, there is no evidence that he ever saw, and thus certainly never calculated entropy values via the equation Planck published in a 1900 article, \( S = \frac{R}{N} \ln W \). It was first printed in 1906 in a book by Planck as \( S = k_B \ln W \), and subsequently carved on Boltzmann’s tombstone. Planck’s nobility in allowing \( \frac{R}{N} \) to be called ‘Boltzmann’s constant’, \( k_B \), was uncharacteristic of most scientists of that day, as well as now.

The important question is “what are the bases for Boltzmann’s introduction of order to disorder as a key to understanding spontaneous entropy change?” That 1898 idea came from two to three pages of a conceptual description, a common language summary, that follow over 400 pages of detailed theory in Brush’s translation of Boltzmann’s 1896-1898 “Lectures on Gas Theory” (University of California Press, 1964). The key paragraph should be quoted in full. (The preceding and following phrases and sentences, disappointingly, only expand on it or support it without additional meaningful technical details or indications of Boltzmann’s thought processes. I have inserted an explanatory clause from the preceding paragraph in brackets, and put in italics Boltzmann’s surprisingly naïve assumptions about all or most initial states as “ordered”.)

“In order to explain the fact that the calculations based on this assumption [“...that by far the largest number of possible states have the characteristic properties of the Maxwell distribution...”] correspond to actually observable processes, one must assume that an enormously complicated mechanical system represents a good picture of the world, and that all or at least most of the parts of it surrounding us are initially in a very ordered — and therefore very improbable — state. When this is the case, then whenever two of more small parts of it come into interaction with each other, the system formed by these parts is also initially in an ordered state and when left to itself it rapidly proceeds to the disordered most probable state.” (Final paragraph of #87, p. 443.)

[Pitzer’s calculation of a mole of any substance at near 0 K shows that none can be more ordered than having the possibility of \( 10^{26,000,000,000,000,000,000,000,000,000} \) different accessible microstates! (Pitzer, Thermodynamics, 3rd edition, 1995, p. 67.)]

Thus, today we know that no system above 0 K has any "order" in correct thermodynamic descriptions of systems of energetic molecules. The common older textbook comparison of orderly crystalline ice to disorderly liquid water is totally deceptive, It is a visual "Boltzmann error" not a proper thermodynamic evaluation. If liquid water at 273 K, with its \( 10^{1,991,000,000,000,000,000,000,000,000,000} \) accessible microstates (quantized molecular arrangements) is considered "disorderly", how can ice at 273 K that has \( 10^{1,299,000,000,000,000,000,000,000,000,000} \) accessible microstates be considered "orderly"? Obviously, using such common words is inappropriate in measuring energetic microstates and thus in discussing entropy change conceptually.

That slight, innocent paragraph of a sincere man — but before modern understanding of \( q_{rev}/T \) via knowledge of molecular behavior (Boltzmann believed that molecules perhaps could occupy only an infinitesimal volume of space), or quantum mechanics, or the Third Law — that paragraph and its similar nearby words are the foundation of all dependence
on “entropy is a measure of disorder”. Because of it, uncountable thousands of scientists and non-scientists have spent endless hours in thought and argument involving ‘disorder’and entropy in the past century. Apparently never having read its astonishingly overly-simplistic basis, they believed that somewhere there was some profound base. Somewhere. There isn’t. Boltzmann was the source and no one bothered to challenge him. Why should they?

Boltzmann’s concept of entropy change was accepted for a century primarily because skilled physicists and thermodynamicists focused on the fascinating relationships and powerful theoretical and practical conclusions arising from entropy’s relation to the behavior of matter. They were not concerned with conceptual, non-mathematical answers to the question, “What is entropy, really?” that their students occasionally had the courage to ask. Their response, because it was what had been taught to them, was “Learn how to calculate changes in entropy. Then you will understand what entropy ‘really is’.”

There is no basis in physical science for interpreting entropy change as involving order and disorder. The original definition of entropy (change) involves a transfer of heat from a thermal reservoir to a system via a virtually reversible energy flow process. Although Clausius described it and his equation of \( \frac{dq_{\text{rev}}}{T} \) or \( q_{\text{rev}}/T \) as a “Verwandlung” or “transformation”, he limited it and “disgregation” to discussions of fusion or vaporization where the “disgregation values” changed. Thus, Clausius was observing phase change, but he made no statements about “orderly crystalline substances” being transformed into “disorderly” liquids, an obvious claim for him to make from his observation. Unfortunately, Clausius did not see that his dq, an amount of ‘heat’ energy, initially relatively localized in a thermal reservoir, was transformed in any process that allowed heat to become more **spread out in space**. That is what happens when a warm metal bar is placed in contact with a similar barely cooler metal bar — or when any system is warmed by its slightly warmer surroundings. The final state of the “universe” in both of these examples is at equilibrium and at a uniform temperature. The internal energy of the atoms or molecules in the reservoir has become less localized, more dispersed in the greater final three-dimensional space than it was in the initial state. (More profoundly, of course, that energy has become more dispersed in phase-space, and spread over more energy levels in the once-cooler object than was its dispersal decreased in the once-hotter surroundings.)

That is also what happens when ideal gases A and B with their individually different internal energy content (\( S^0 \) values) but comparably energetic, constantly moving molecules in adjacent chambers are allowed access to one another’s chambers at 298 K. With no change in temperature, they will mix spontaneously because, on the lowest level of interpretation, the translational energy of the A and B molecules can thereby become more spread out in the larger volume. On a more sophisticated level, their energy is more widely distributed in phase-space. From the quantum-mechanical view of the occupancy of energy levels by individual molecules, each type of molecule has additional energy levels in the greater volume, because the energy levels become closer together. But the same causal description of energy spontaneously spreading out can be used as in the naïve view of seeing mobile molecules always moving to occupy newly available 3-D volume: the energy of the molecules is more dispersed, more spread out, now in terms of dispersal over more energy levels.

(Of course, this energy dispersal can best be described in terms of additional accessible microstates. The greater the number of possible arrangements of molecular energies over energy levels, the greater the entropy increase — because the system in any one arrangement at one instant has more probability of being in a different arrangement the next instant. The total energy of the system is unchanged over time, but there is a continuing ‘temporal dance’ of the system over a minute fraction of the hyper-astronomic number of accessible arrangements.)
The increase of entropy in either A or B can readily be shown to be equal to \( R \ln \frac{V_{\text{Final}}}{V_{\text{Initial}}} \), or more fundamentally, \(-nR(x_i \ln x_i)\). This result is not specific to gases, of course. What is shown to be significant by the basic equation is that any separation of molecules of one type from its own kind is an entropy increase due to spreading out of its intrinsic internal energy in greater space, both 3-D and phase-space. Further, this increased dispersal of energy is interpretable in terms of the increased number of accessible arrangements of the system's energy at any instant and thus, a greater number of chances for change in the next instant — a greater ‘temporal dance’ by the system over greater possibilities and a consequent entropy increase.

Boltzmann’s sense of “increased randomness” as a criterion of the final equilibrium state for a system compared to initial conditions was not wrong. What failed him (and succeeding generations) was his surprisingly simplistic conclusion: if the final state is random, the initial system must have been the opposite, i.e., ordered. “Disorder” was the consequence, to Boltzmann, of an initial “order” not — as is obvious today — of what can only be called a “prior, lesser but still humanly-unimaginable, large number of accessible microstates”

Clearly, a great advantage of introducing chemistry students to entropy increase as due to molecular energy spreading out in space, if it is not constrained, begins with the ready parallels to spontaneous behavior of kinds of energy that are well-known to beginners: the light from a light bulb, the sound from a stereo, the waves from a rock dropped in a swimming pool, the air from a punctured tire. However, its profound “added value” is its continued pertinence at the next level in the theoretical interpretation of energy dispersal in thermal or non-thermal events, i.e., when the quantization of molecular energies on energy levels, their distributions, and accessible microstates become the focus.

When a system is heated, and its molecules move more rapidly, their probable distributions on energy levels change so that previous higher levels are more occupied and additional high levels become accessible. The molecular energy of the heated system therefore has become more widely spread out on energy levels. The dispersal of energy on energy levels is comparable in adiabatic processes that some authors characterize as involving “positional” or “configurational” entropy. When a larger volume is made available to ideal components in a system — by expansion of a gas, by fluids mixing (or even by a solute dissolving) — the energy levels of the final state of each constituent are closer together, denser than in the initial state. This means that more energy levels are occupied in the final state despite no change in the total energy of any constituent.. Thus, the initial energy of the system has become more spread out, more widely dispersed on more energy levels in the final state..

The ‘Boltzmann’ equation for entropy is \( S = k_B \ln W \), where \( W \) is the number of different ways or microstates in which the energy of the molecules in a system can be arranged on energy levels. Then, \( \Delta S \) would equal \( k_B \ln \frac{W_{\text{Final}}}{W_{\text{Initial}}} \) for the thermal or expansion or mixing processes just mentioned. A most important \( \Delta S \) value in chemistry is the standard state entropy for a mole of any substance at 298 K, \( S^0 \), that can be determined by calorimetric measurement of increments of heat/T added reversibly to the substance from 0 K to 298 K. Any transition state or phase change/T is also added.

Obviously, therefore, considerable energy is ‘stored’ in any substance on many different energy levels when it is in its standard state. (A. Jungermann, J. Chem. Educ. 2006, 83, 1686-1694.) Further, for example, if energy from 398 K surroundings is spread out to a mole of nitrogen at 298 K, the nitrogen’s molecules become differently arranged on previous energy levels and spread to some higher levels. If, while at any fixed temperature, the nitrogen is allowed to expand into a vacuum or to mix with another gas, its energy levels in the new larger volume will be closer together. Even in fixed volume or steady temperature situations, the constantly colliding molecules in a mole of any gas are clearly not just in one unique arrangement on energy levels for more than an instant. They are continually changing from one
arrangement to another due to those collisions — within the unchanged total energy content at a given temperature and a distribution on energy levels consistent with a Boltzmann distribution. Thus, because $W_{\text{Initial}}$ at 0 K is arbitrarily agreed to be 1, nitrogen’s $S^0$ of 191.6 J/K mole = $1.380 \times 10^{-23}$ J/K $\ln W_{\text{Final}}$. Then, $W_{\text{Final}} = 10$ to the exponent of $6,027,000,000,000,000,000,000,000$ microstates, a number of possible arrangements for the nitrogen molecules at 298.15 K that is humanly beyond comprehension — except in the terms of manipulating or comparing that number mathematically.

It should be emphasized that these gigantic numbers are significant guides mathematically and physically and conceptually — i.e., that a greater or smaller such number indeed indicates a difference in real physical systems of molecules and we should sense that magnitude is significant. However, conceptually, we should also realize that in real time, it is impossible for a system to be more than a few quadrillion different states in a few seconds, perhaps spending most time in a few billion or million. It is impossible even in almost-infinite time for a system to explore all possible microstates or that even a tiny fraction of the possible microstates could be explored in a century. (It is even less probable that a gigantic number would be visited because the most probably and frequently occupied microstates constitute an extremely narrow peak in a ‘probability spectrum’.)

However, the concepts are clear. At one instant, all of the molecules are only in one energetic arrangement — an instantaneous ‘freeze-frame’ photo of molecular energies on energy levels (that is an abstraction derived from an equally impossible photo of the actual molecules with their speeds and locations in space at one instant.) Then, in the next instant, a collision between even two molecules will change the arrangement into a different microstate. In the next instant to another. Then to another. (Leff has called this sequence of instantaneous changes “a temporal dance” by the system over some of its possible accessible microstates.) Even though the calculated number of possible microstates is so great that there is no chance that more than a small fraction of that number could ever be explored or “danced in” over finite time, that calculated number influences how many chances there are for the system’s energy arrangement to be in at the next moment. The greater the number of microstates, the more chances a system has for its energy next to be in a different microstate. In this sense, the greater the number of microstates possible for a system, the less probable it is that it could return to a previously visited microstate and thus, the more dispersed is its energy over time. (In no way is this evaluation and conclusion a novel/radical introduction of time into thermodynamic considerations! It is simply that in normal thermo measurements lasting seconds, minutes or hours, the detailed behavior of molecules in maintaining a macro state is of no interest.) For example, heating a mole of nitrogen only one degree, say from the standard state of 298.15 K to 299.15 K, in view of its heat capacity of 29 J/K, results in an entropy increase of 0.097 J/K, increasing the microstates from $6,027 \times 10^{24}$ to $6,034 \times 10^{24}$. Thus, even a slight macro change in a system is signified by an equal change in the number of microstates — in the number of chances for the system to be in a different microstate in the next instant than in the previous moment. The greater the number of microstates possible for a system in a given state, the more probable is the dispersal of that system’s energy in the sense of its being in a different microstate in the next instant.

A greater dispersal of energy in a system means, in terms of its microstates, a ‘temporal dance’ over a greater number of possible microstates than if there were a smaller number of microstates.

The most frequently used example to show entropy increase as greater “disorder” in elementary chemistry texts for many years was that of ice melting to water. Sparkling orderly crystalline ice to disorderly mobile liquid water — that is indeed a striking visual and crystallographic impression, but appearance is not the criterion of entropy change. Entropy increase is dependent on the dispersal of energy — in three-dimensional space (an easily understandable generality for all beginning
students.) Then, more capable students can be led to see how entropy increase due to heat transfer involves molecular energies occupying more and higher energy levels while entropy increase in gas expansion and all mixing is characterized by occupancy of denser energy levels within the original energy span of the system. Finally, advanced students can be shown that any increase in entropy in a final system or universe that has a larger number of microstates than the initial system/universe as the ultimate correlation of entropy increase with theory, quantitatively derivable from molecular thermodynamics.

Crystalline ice at 273 K has an $S^0$ of 41.34 J/K mol, and thus, via $S = k_b \ln W$, there are 10 to an exponent of 1,299,000,000,000,000,000,000,000 possible accessible microstates for ice. Because the $S^0$ for liquid water at 273 K = 63.34 J/K mole, there are 10 to an even larger exponent of 1,991,000,000,000,000,000,000,000 accessible microstates for water. Does this clearly show that water is “disorderly” compared to crystalline ice? Of course not. For ice to have fewer accessible microstates than water at the same temperature means primarily — so far as entropy considerations are concerned — that any pathway to change ice to water will result in an increase in entropy in the system and therefore is favored thermodynamically.

Gibbs’ use of the phrase “mixed-upness” is totally irrelevant to ‘order-disorder’ in thermodynamics or any other discussion. It comes from a posthumous fragment of writing, unconnected with any detailed argument or logical support for the many fundamental procedures and concepts developed by Gibbs.

Finally, the idea that there is any ‘order’ or simplicity in the distribution of energy in an initial state of any real substance under real conditions is destroyed by Pitzer’s calculation of numbers for microstates in his “Thermodynamics” (Third edition, 1995, p. 67). As near to 0 K and thus, to as ‘practical’ a zero entropy as can be achieved in a laboratory, Pitzer shows that there must be $10^{26,000,000,000,000,000,000}$ of possible accessible microstates for any substance.

The divide between thermodynamic entropy and information theory’s ‘disorder’

Information theorists and thermodynamics text authors like Herbert B. Callen and Myron Tribus take Shannon’s valid theory of errors and disorder in communication and by ‘merely’ substituting $k_B$ for $k$ in $-k \sum p_i \ln p_i$ claim that “entropy is the quantitative measure of disorder in the relevant distribution of the system over its microstates.” [Callen, italics mine, to indicate a completely erroneous implication of the nature of a system’s occupancy of one microstate at one time in an equally erroneous connection of thermodynamic entropy with communications’ disorder.] Admittedly, both authors develop many pages of standard thermodynamics of systems, but a reader who scans only a page or hears only the Callen statement without an emphasis on how molecules behave spontaneously is seriously misled. This is not a normal substitution of one arbitrary mathematical constant for another. It constitutes insertion of the qualities and the behavior of all physical material systems into a purely mathematical equation without any consideration of the innate behavior of those physical systems. Although this may not be exactly like evaluating one’s obnoxious relatives in Chicago by some symbols in an equation from Einstein’s theory of general relativity, it is far more subtly erroneous because Shannon probability is a vital half of thermodynamic entropy, as described here:

“There are two requisites for entropy change in chemistry. An increase in thermodynamic entropy is enabled in a process by the motional energy of molecules (that, in chemical reactions, can arise from the energy released from a bond energy change). However, entropy increase is only actualized if the process results in a larger number of
arrangements (microstates) for the system’s energy, i.e., a final state that involves the most probable distribution for that energy under the new constraints.” [That probable distribution of Boltzmann is identical to Shannon’s.]

“The two requisites, energy and probability, are both necessary for thermodynamic entropy change; neither is sufficient alone. In sharp contrast, information theory (often misleadingly called information “entropy”) is alone sufficient for its mathematical purposes, depending only on $-k \sum p_i \log p_i$, where $k$ is an arbitrary constant that is not required to involve energy.”


Arbitrarily replacing k by $k_B$ — rather than have it arise from thermodynamic necessity via Boltzmann’s probability of random molecular motion — does violence to thermodynamics. The set of conditions for the use or for the meaning of $k_B$, of $R$ with $N$, are nowhere present in information theory. Thus, conclusions drawn from the facile substitution of $k_B$ for $k$, without any discussion of the spontaneity of change in a thermodynamic system (compared to the hundreds of information “entropies”) and the source of that spontaneity (the ceaseless motion of atoms and molecules) are doomed to result in confusion and error. There is no justification for this attempt to inject the overt subjectivity of “disorder” from communications into thermodynamic entropy.

Outside Links

http://entropysite.oxy.edu/boltzmann.html

Contributors

• Frank L. Lambert, Professor Emeritus, Occidental College