

On the unification of the classical laws of motion and thermodynamics

E. W. Dreiss

There exists an essential formal difference between the classical laws of motion as formulated by Newton and the laws of thermodynamics which evolved from empirical observations over the next three centuries. The development of the kinetic theory of gases brought this difference into sharp focus, particularly with regard to the second law of thermodynamics. The seeming conflict between Newton's laws and the second law is but one aspect of a more general discrepancy. Not only is there no law of motion similar to the second law, but there is no law of thermodynamics corresponding to Newton's law of force. Here we show that recognizing the second law as the thermodynamic counterpart of the law of inertia and adding a "fourth law" as the thermodynamic counterpart of Newton's law of force leads to a simplified set of physical laws which applies across both domains.

Modern classical (non-quantum) thermodynamics takes two forms, a macroscopic form which describes relationships between observables such as temperature, pressure and volume, and a microscopic form which deals with molecular interactions which cannot (with current technology) be directly observed. The two forms are linked by kinetic theory, which attempts to explain the macroscopic behavior in terms of the microscopic motion of molecules.

The macroscopic laws of thermodynamics were derived primarily by induction from empirical observations, and were related to practical problems, particularly the engineering of heat engines. For isolated systems, it was observed that gases always expand from regions of high pressure into regions of low pressure and that heat always flows from hot to cold. Such systems were observed to spontaneously evolve to a state of “thermodynamic equilibrium”, where the pressure and temperature appeared to be uniformly distributed throughout the system. These and other observations of irreversible behavior inspired various macroscopic formulations of the second law of thermodynamics¹.

Attempts to explain the second law in terms of kinetic theory revolved around the Maxwell velocity distribution. Boltzmann attempted to prove by mathematical deduction that an ideal gas for which the velocities are distributed in some arbitrary manner will, due to the randomizing effect of collisions between molecules and the “laws of probability”, always tend toward the Maxwell velocity distribution, terminating in a state of “statistical equilibrium”. Once this distribution is obtained, any chance deviations from this distribution will, by the same mechanism, quickly return to equilibrium with a high degree of probability. This logic led to Boltzmann’s probabilistic formulation of the second law².

While Boltzmann’s attempt to explain the irreversibility of the second law in terms of probability theory has been frequently criticized over the years, it is seldom pointed out that the problem he addressed was not the problem presented by the macroscopic second law. The tendency of pressure and temperature to equalize, the mixing of different gases, and most other observable

thermodynamic phenomena deal with the *spatial* distribution of molecules, not their *velocity* distribution. One would expect similar phenomena regardless of the shape of the velocity distribution, for example, if all of the molecules in a gas were moving at the same speed or if the speed distribution were uniform within a given range. In fact, as Boltzmann himself pointed out (in a somewhat different context), it is highly unlikely that we would ever find a velocity distribution in nature that would deviate appreciably from the Maxwell velocity distribution². It is equally difficult to see how such a distribution could be created or maintained in the laboratory. So it appears that Boltzmann solved a problem that doesn't exist and provided an explanation that is largely irrelevant to the practical question of why gases tend irreversibly toward spatial uniformity. A more direct approach would therefore seem advisable.

The second law and inertia

Kinetic theory explains pressure in terms of the momentum per unit area transferred to the walls of a container, and temperature as proportional to the average kinetic energy of the molecules. These observables are therefore directly linked to the motion of individual molecules which obey Newton's laws of motion. However, attempts to explain transitional phenomena such as diffusion and irreversibility in terms of kinetic theory have been less convincing. Describing the rate of diffusion of one gas into another as a function of the density gradient is like describing the rate at which water flows downhill in terms of the slope. While both accurately portray what happens, neither provide a causal explanation in terms of fundamental physical laws. Similarly, describing the approach to thermodynamic equilibrium in terms of

probability gradients provides no clue as to the physical processes involved.

A major conceptual barrier has been the assumption that the dynamical laws of motion are reversible with time, which seems to eliminate the possibility that irreversibility can be explained in terms of the dynamics of individual molecules, as is the case with pressure and temperature. To examine this assumption, we will consider the case of an ideal gas composed of identical particles which are perfectly spherical and elastic and strictly obey Newton's laws of motion. Since all motion is by definition deterministic, we must reject any line of reasoning that assigns a causal role to probability³. It follows that any application of probability theory to classical thermodynamic systems can only describe the state of our knowledge of its configuration, and has no bearing on the physical processes which drive its evolution⁴. Based on these assumptions, there is no reason not to believe that a collection of many of these particles will obey the same laws as a collection of just a few. Furthermore, there is no reason not to believe that a system composed of a large number of such particles would behave as dictated by the macroscopic versions of the second law formulated from empirical observation and consistent with everyday experience. In particular, we would expect that if such a gas were released into an empty container it would spread to fill the container and eventually settle into a steady state for which the gas molecules were more or less evenly distributed throughout the available space. Similarly, if we were to remove a barrier separating two different gases in an isolated container, we would expect these gases to mix spontaneously, once again approaching a state of relative uniformity. In neither case would we expect the

process to operate in reverse, spontaneously evacuating the container or unmixing the gases.

More specifically, in the first case we can imagine a container divided into two chambers, the first containing an ideal gas and the second a vacuum. There is a tiny hole in the barrier between chambers which can be opened or closed by a small door. If we open and close the door for just a moment, a small number of gas molecules (say two or three) will leak into the empty chamber. Since these were in motion prior to escaping through the hole, each will continue according to the law of inertia in the same direction and speed as it enters the empty chamber. As the molecules are extremely small compared to the volume of the chamber, it will be assumed that they do not interact. Furthermore, we expect the escaped molecules to be traveling in different directions, so that they will move apart after their emergence from the hole. At some point, each of the molecules will be reflected from the container wall and over time will trace a trajectory around the chamber independently of the other molecules. If the shape of the chamber is even slightly irregular or the chamber walls are not perfectly smooth, we would expect the trajectories of the molecules to eventually traverse most of the space enclosed by the chamber. For real systems this would almost certainly be the case since the atoms making up the walls of the container are roughly the same size as the gas molecules, eliminating the possibility of perfectly smooth walls³.

For this example it can be seen that molecules of gas have “expanded” into the second chamber solely as a consequence of the law of inertia. Inertia is also responsible for the molecules being distributed throughout the chamber, both directly and in

concert with the law of conservation of momentum which governs the collisions with the walls. If we then release a few more molecules into the second chamber, we will have more of the same. Interactions (collisions) between molecules will be negligible and the molecules will tend to “fill” the chamber due to inertia and the reflections off the chamber walls. Continuing this process, the density of molecules in the second chamber will increase and collisions between molecules will become more common. Eventually, enough molecules will have leaked into the second chamber to bring it into balance with the first chamber, in the sense that about as many molecules will move from the second chamber to the first when the door is opened as will move from the first to the second. Under these conditions, collisions between molecules are likely to be more common than collisions with the container walls. Nevertheless, the motion of the molecules will continue to be driven by the law of inertia.

In the case of the mixing of two different gases, we can imagine our container divided into two chambers separated by a barrier, with one of the gases initially confined to the first chamber and the other to the second. We will assume that for our initial trial, both gases are so dilute that the molecular interactions are negligible. When we remove the barrier, the molecules headed towards the barrier will continue to travel under the influence of inertia and, as a result, the gases will mix. Once again, the mixture will tend toward uniformity due to inertia and the effects of irregularities in the shape and interior surface of the container. Repeating the experiment with higher densities of molecules will increase the frequency of molecular collisions, but the inertial nature of the mixing process will not change.

We can see that expansion and mixing in an ideal gas can be understood exclusively in terms of Newton's laws of motion. The irreversibility of these processes is the natural consequence of the law of inertia acting in combination with the initial conditions imposed by the experimental arrangement. The evolution of the system toward a steady state is apparent whether the system is composed of two molecules or a trillion. In either case, inertia will cause the molecules to spread and mix to the extent allowed by the container. This pertains not only to very dilute gases but to those at room temperature and normal atmospheric pressure. Under these conditions the mean free path is on the order of 20 times the distance between molecules for air and between 10 and 100 for most other gases, so inertia remains the primary agent of molecular transport and mixing due to the interpenetration of the molecules.

The fourth law of thermodynamics

Where the molecules are non-interacting, it is clear that each will be moving in straight-line trajectories between reflections from the walls of the container, in accordance with the principle of least action. The action of the system will then be just the sum of the action of each of the molecules, so that minimizing the action of each of the molecules will also minimize the action of the system taken as a whole⁵. Following the same logic, the original derivations of the action principle by both Lagrange and Hamilton were for the general case of systems composed of multiple particles⁶. If the molecular density is increased so that the molecules collide, each molecule will continue to follow its path of least action, although its kinetic energy may increase or decrease with each collision. However, since energy is conserved, the total kinetic energy of an isolated system will remain constant.

The principle of least action is usually presented for the case of a closed system of particles where outside forces are present⁷. For each of the n molecules of the system, the energy at any point in time is

$$E_i(t) = KE_i(t) + PE_i(t) \quad (1)$$

where KE_i is the kinetic energy due to the motion of the i th molecule and PE_i is the potential energy due to the forces acting on the molecule. The action S_i for the molecule over the time interval from t_1 to t_2 is

$$S_i = \int_{t_1}^{t_2} (KE_i - PE_i) dt \quad (2)$$

Substituting for PE_i gives

$$S_i = \int_{t_1}^{t_2} (2KE_i - E_i) dt \quad (3)$$

This is equivalent to

$$S_i = (2\langle KE_i \rangle - \langle E_i \rangle)(t_2 - t_1) \quad (4)$$

where the angle brackets denote time averages over the interval t_1 to t_2 ⁸. Since the action is additive for a multi-particle system, the total action S_{Total} is

$$S_{\text{Total}} = \sum_n (2\langle KE_i \rangle - \langle E_i \rangle)(t_2 - t_1) \quad (5)$$

The total average energy of the system is

$$\langle E \rangle_{Total} = \sum_n \langle E_i \rangle \quad (6)$$

This is clearly true if the molecules are non-interacting, since the energy E_i for each molecule will remain constant over its trajectory. It is also true if the molecules interact, since energy conservation dictates that any kinetic energy lost by a molecule as the result of a collision will be transferred to the other molecule(s) involved. So the total action of the system can be written as

$$S_{Total} = (t_2 - t_1) \left(2 \sum_n \langle KE_i \rangle - \langle E \rangle_{Total} \right) \quad (7)$$

For an isolated system devoid of forces, the total energy of the system is equal to the total kinetic energy, so this equation reduces to

$$S_{Total} = (t_2 - t_1) \sum_n \langle KE_i \rangle \quad (8)$$

This last equation applies to the examples presented above, where all of the energy is related to the inertial motion of the molecules. According to kinetic theory, the thermal energy of an ideal gas is just the sum of the kinetic energies of the molecules, so that minimizing the thermal energy is equivalent to minimizing the action of the system.

For a closed system in a steady state, E_{Total} will be constant, and the minimum action will again correspond to the minimum kinetic energy. This is the “principle of the least dissipation of energy” as developed by Onsager for steady-state systems near “equilibrium”⁹. Prigogine interpreted this as the “theorem of minimum entropy production”, a principle he generalized to cover a variety of phenomena, including chemical reactions and the flow of current in electrical circuits¹⁰. Jaynes pointed out that minimum entropy only applies to isothermal systems, and that Kirchhoff’s “principle of minimum heat production” is the more general principle which “is surely true for fluxes of any kind of stable particles, or of anything else that is conserved (energy, momentum, etc.)”¹¹.

The principle of minimum heat production applies to open as well as closed systems. This is consistent with Eq. (7), since S_{Total} can be taken to encompass all of the n particles which are part of the system at any time during the interval between t_1 and t_2 . While Onsager and Prigogine address this principle only to steady-state systems, Eq. (7) does not require that E_{Total} be constant over time. The system defined by the number of particles n can be arbitrarily small or large, spanning the range from dynamics to thermodynamics.

From the above reasoning, we propose the following formulation of the fourth law of thermodynamics:

The flow of energy and matter through a system is such that the heat produced by the process is minimized.

This universal law follows logically from the principle of least action, since if the action of each particle in a multi-particle system is minimized, the sum of the kinetic energies of all the particles must also be a minimum, resulting in the minimum production of heat.

If a system is performing work, minimum heat production corresponds to maximum thermodynamic efficiency. The fourth law can then be stated as:

The flow of energy and matter through a system is such that the thermodynamic efficiency of the process is maximized, given the structural constraints on the system.

Classical thermodynamics can therefore be fully described with an energy conservation law (the first law) and a least action principle (the fourth law). This formulation is isomorphic to the least action version of the dynamical laws and is congruent in the sense that the laws of motion transform seamlessly into the laws of thermodynamics. The law of inertia and the second law of thermodynamics are merely special cases of these least action principles where no unbalanced forces are present.

Universality

The claim that least action constitutes a universal principle has a long pedigree. According to most reports, it was first proposed by Maupertuis, who was famously lampooned for this belief by Voltaire in *Candide*. Another distinguished proponent appears to have been Planck⁷. The primary philosophical objection to this generalization is that the principle of least action is teleological,

since it seems the particle must look ahead to compare all possible paths and then select the path for which the action is minimized^{5,7}. The presumption of clairvoyance follows from the integral formulations of the principle by Euler, Lagrange and Hamilton, which assume that both the initial and final conditions are specified and that the objective is to derive the trajectory connecting these conditions⁸. However, under the classical assumption of locality, the particle can only know its current physical state and has no way of anticipating where it will be in the future. It would appear that replacing the statistical formulation of the second law with a fourth law based on the principle of least action amounts to swapping one teleological law for another. However, Feynman pointed out that for the action over a path between two points to be a minimum, it must also be a minimum for any segment of the path. This segment can be taken as arbitrarily small, so that for an infinitesimal segment the integral principle of least action reduces to Newton's differential law of force⁵. This differential law is the only perspective available to the particle, and determines where it will be in the next instant. Therefore, while the least action formulations are clearly anthropomorphic, they are equivalent to the ontologically correct differential form. The infinitesimal least-action paths aggregate to the integral least-action path, which can then be analyzed from a human perspective. The locality implicit in classical dynamics must also apply to thermodynamics, so transcendent formulations of either the second or fourth law can only be metaphorical, not ontological.

Conclusion

On logical grounds, one would expect to be able to identify specific thermodynamic laws which correspond to the laws of

motion and *vice versa*. But this has clearly not been the case. However, applying the same logical criteria to both dynamics and thermodynamics leads to a symmetrical set of laws which is both consistent and complete. This is accomplished by first recognizing that the second law is a direct consequence of the law of inertia. Since the law of inertia cannot, by definition, constitute a law of force, a “fourth law” must be added to the laws of thermodynamics. The identification of the second law with the law of inertia and the addition of the fourth law result in two isomorphic sets of laws, one encompassing the laws of motion and the other the laws of thermodynamics. Each set is composed of the law of conservation of energy and a law derived from the principle of least action.

We conclude that thermodynamic systems must obey Newton’s laws of motion, which takes us back to our original assumption. In coming full circle we have shown that the asymmetrical nature of the second law is entirely due to the law of inertia acting within a system for which constraints on molecular motion have been relaxed. The process is irreversible unless work is applied from outside the system to restore the constraints. The extant versions of the second law represent heuristic models which may be useful in appropriate contexts, but should not be interpreted ontologically^{4,11}. Unlike these anthropomorphic formulations, the inertial second law is intrinsically passive and cannot be held accountable for either hindering or facilitating the evolution of complexity or the origin of life. This role is the proper domain of the law of force, which has been missing from the laws of thermodynamics. We therefore propose that Newton’s law of force be incorporated into thermodynamics as the fourth law.

1. Čápek , V. & Sheehan D. P. *Challenges to the Second Law of Thermodynamics* (Springer, Dordrecht, 2005), pp. 3-13.
2. Boltzmann, L. *Lectures on Gas Theory* (Dover, Mineola, NY, 1995), pp.36-55, 444.
3. Garrod, C. *Statistical Mechanics and Thermodynamics* (Oxford University Press, New York, 1995), p. 317, 352.
4. Jaynes, E. T. “Clearing up Mysteries – The Original Goal” in *Maximum-Entropy and Bayesian Methods* (J. Skilling, ed., Kluwer, Dordrecht-Holland, 1989), pp. 18-24.
5. Feynman, R. P., Leighton, R. B. & Sands, M. *The Feynman Lectures on Physics* (Addison-Wesley, Reading, MA, 1963), vol. II, p. 19-7, 19-9, 19-8.
6. Landau, L. D. & E. M. Lifshitz *Mechanics* (Third Edition, Elsevier, Oxford, 1976), pp. 8-10.
7. Yourgrau, W. & Mandelstam, S. *Variational Principles in Dynamics and Quantum Theory* (Dover, New York, 1968), pp. 24-44, 162-180.
8. Hanc, J., Tuleja, S. & Hancova, M. *Am. J. Phys.* **71**, 386 (2003).
9. Onsager, L. *Phys. Rev.* **37**, 405 (1931).
10. Kondepudi, D. & Prigogine, I. *Modern Thermodynamics* (Wiley, Chichester, West Sussex, 1998), pp. 392-404.

11. Jaynes, E. T. *Annu. Rev. Phys. Chem.* **31**, 579 (1980).