Chapter 9

The emergence and evolution of the Second Law of Thermodynamics

D. Wilkie, R.C. Dougal, M.W. Collins & A. Whitaker

Abstract

The results of measurements on thermal properties of common substances provided a backdrop to the successful development of heat engines in the eighteenth and nineteenth centuries. Early in the nineteenth century, Carnot’s theory of an idealised heat engine provided the stimulus to the ongoing studies, both practical and theoretical, of the properties of materials and of mechanical and thermal processes. By mid-century, the First and Second Laws of Thermodynamics had been established, in various equivalent forms. These clarified the role of energy and its transformations and led to the introduction of a new thermodynamic function, entropy, to complement energy and, from its thermodynamic properties, to clarify the limitations imposed by the Second Law. The understanding of thermal properties of matter from the laws and techniques of classical thermodynamics was enhanced by the introduction of and alignment with statistical thermodynamics. This provided an understanding of properties of assemblies of large numbers of molecules by incorporating the rules of probability theory, leading to a statistical interpretation of entropy. A merger with the well-known laws and processes of chemistry led to substantial advances in chemical thermodynamics, permitting an insightful subsequent reassessment of thermodynamics as a whole.

1 Introduction

Classical thermodynamics, emerging from studies of heat and bulk properties of matter, is underpinned by the First and Second Laws of Thermodynamics. The First Law is a special case of the principle of conservation of energy. The Second Law provides a criterion for the selection from processes permitted by the First Law only of those which occur naturally. These laws were formulated from a wide range of observations and measurements, and have a correspondingly wide range of applicability. In this chapter, we show how tactically the formulation of the Second Law benefitted from the formulation of the First Law, how both demonstrate the interplay of engineering and mathematics and how they led endeavours, which might have seemed peripheral, into the mainstream of science.
2 Heat and temperature: conceptualised and applied

2.1 Black and Watt

The science of energy began with Joseph Black (1728–1799) at Glasgow University in the early 1750s with his definition of the distinction between heat and temperature. Black stated that to assume two bodies at the same temperature would possess the same quantity of heat would be ‘taking a very hasty view of the subject. It is confounding the quantity of heat in different bodies with its general strength or intensity, though it is plain that these are two different things and should always be distinguished, when we are thinking of the distribution of heat’. Black showed that the temperatures of different substances increased at different rates when they were heated equally, implying different ‘capacities’ for absorbing heat. He was able to compile a comparative list of heat capacities per unit mass (‘specific heats’) for a number of common materials. He went on to introduce the concept of latent heat by heating ice and finding that no change in temperature occurred until all the ice had melted and likewise for water, which remained at boiling point until all the water had been converted to steam. He concluded that the heat supplied in these two events was latent in the melted ice and in the steam and was available for recovery in the reverse processes of freezing and condensation.

James Watt (1736–1819), born in Greenock, was engaged by Black as an instrument mechanic to repair and develop instruments. Watt had spent a year in London working for an instrument maker but his health suffered from overwork. As Watt had not served a full apprenticeship he was barred from obtaining employment in the trade, but the University had no such restrictions.

Watt repaired the department’s demonstration model Newcomen engine, which had been returned from London as beyond repair. This was an atmospheric rather than a steam engine, with the piston powered by the air above it, making it move down on a partial vacuum produced by the condensing of steam on the other side. During this work, he became aware of its extremely poor performance. This led him to invent a steam condenser separate from the cylinder thus avoiding the need to cool the piston on each stroke and re-raise steam resulting in a saving of time and fuel. It provided a fivefold improvement. In his early experiments with steam, Watt realised that raising steam required a lot of heat and deduced that the steam contained this heat.

2.2 Watt and Boulton

Black financed Watt’s development work initially but the industrialist John Roebuck (1718–1794), founder of the Carron ironworks (near Falkirk in Central Scotland), took over the financing in return for a two-thirds share in the patent rights. Roebuck became bankrupt in 1773 when his own businesses failed and Matthew Boulton (1728–1809), a versatile manufacturer from Birmingham, who had already met Watt and had been trying for some time to engage him, acquired Roebuck’s share. Watt went into partnership with Boulton in 1775.

In Watt’s engine, steam entered through an aperture in the side of the cylinder, above the piston. Early improvements included letting the steam enter alternately through opposite sides, making it a double-acting engine. Heat losses were reduced by the addition of a steam jacket, and insulation. A pressure indicator was modified by the works manager, John Southern
(c. 1755–c. 1825) to show the variation of pressure with piston displacement. From a plot of pressure versus piston displacement, a closed figure was produced, whose area represented the power developed by the engine. This allowed Watt to assess the effectiveness of developments to improve the power delivery. Among other things, Watt found that stopping steam entry to the cylinder before the end of the stroke improved the efficiency.

Being fully conversant with the operation of his steam engine Watt must have realised the importance of latent heat and therefore of stored energy. He also knew of the importance of attaining a maximum temperature difference, as did his extremely inventive assistant, William Murdock (1754–1839), who was responsible for many of the improvements to the Watt engine. For instance, he suggested using high pressure steam to increase the upper temperature, but this was dismissed by Watt because of the dangers of exploding boilers.

3 Towards the laws of thermodynamics

3.1 Carnot, Clapeyron and the original Carnot cycle

By the turn of the nineteenth century developments in the construction and operation of the steam engine had led to an appreciation of both its potential and its limitations. In 1824, the French engineer N.L. Sadi Carnot (1796–1832) stepped back from matters of detail of steam engines so as to identify their basic features and skilfully assemble them into a coherent whole. He then presented a description of the basic theory of an idealised steam engine, followed by a simpler, similarly idealised, air engine [1]. Carnot’s memoir, his only publication, was largely ignored until B.P. Emile Clapeyron (1799–1864) interpreted and re-presented it more analytically in 1834 [2,3] in a version that was the essential input to later researches of others, including William Thomson (1824–1907) [4], appointed as the professor of natural philosophy at Glasgow University in 1846. Clapeyron’s treatment led to a specification of ‘the Carnot cycle’. (Thomson did eventually, in 1848, get a copy of Carnot’s memoir from his friend, Lewis Gordon, Professor of Engineering at Glasgow University [5].

3.1.1 Carnot engines and Carnot’s theorem

The ‘structure’ of the idealised steam engine was chosen so that it could provide external mechanical work broadly in line with that from existing steam engines. This allowed Carnot to best place effort in identifying those features of the ideal engine which, once clarified, would allow an interpretation of relevance to other engines operating in repeated cycles. Optimum performance meant a tight control of heat flow: no waste from heat leaks down temperature gradients which would impair performance. Frictional effects were also ignored. He held to the caloric theory of heat, which obscured the link between heat flow, in and out, and external work done, per cycle, in the operation of the device, but he was still led to propose that – for the idealised engine – the operation of all stages in the cycle could be reversed. The performance of a reversible engine would then be optimum. He argued that if there was an ‘other’ engine having superior performance for the same heat flow, its output, being more than enough to drive the cycle of the reversible engine backwards, would result in net mechanical power from them connected as a composite system as from a perpetual motion machine, which was inadmissible. Carnot extended the argument to propose (Carnot’s theorem) that the output power from a reversible idealised cyclic engine had to be the maximum, or the process was at its optimum efficiency, whatever its working substance.
Carnot’s modelling (adapted) and analysis (essentially intact) were then transferred to an air engine, whose simpler working substance made it easier to visualise. In the resulting idealised cycle, heat \( Q_1 \) entered the working substance from heat source or reservoir (boiler, for steam) at a constant high temperature \( t_1 \) and caused the working substance to expand isothermally at that particular temperature and do mechanical work externally. This was followed by continued expansion with the working substance thermally isolated (later called adiabatic), and with its temperature dropping from \( t_1 \) to \( t_2 \). Two compression stages, now with work being done on the working substance, followed: isothermal at the temperature \( t_2 \) of a heat sink or reservoir (condenser) accompanied by a necessary release there of heat, \( Q_2 \), then a closing return adiabatic with temperature increasing to the initial \( t_1 \). The same net amount of mechanical work was delivered by the working substance each cycle. If the four stages of the closed loop are plotted clockwise on a \( p-v \) diagram, like Watt’s indicator diagram, the area of the loop is a measure of the work done by the working substance per cycle. The caloric theory of indestructible heat meant choosing \( Q_2 = Q_1 \).

For the steam engine, Carnot delayed addressing a potentially problematic return stage to the initial state, as described here, of the steam. Raising steam could have required heat input through finite temperature differences, for which reversibility would not pertain. This led him to consider an infinitesimal cycle for the working substance, working between an infinitesimal temperature difference. The loop area was also infinitesimal, meaning an infinitesimal work output per cycle. Undeterred, Carnot proposed that for an engine operating between a finite temperature difference \( t_1 - t_2 \), its cycle could be modelled as a composite in the form of a stack of infinitesimal cycles with the heat output from any one serving as the heat input to the next one down in the stack [3].

3.1.2 The Carnot function

All the stages of idealisation led to the independence of the efficiency of the idealised engine on everything except temperature. Engineering practice pointed to an improvement of mechanical output \( W \) with increasing temperature difference \( t_1 - t_2 \), moderated by some improvement for lower actual temperatures. Analogy with the mechanical output from a fall of water – proportionality to the height of the fall – suggested proportionality to \( t_1 - t_2 \) for a heat engine, moderated by a factor whose value increased as temperatures decreased. This is written here as \( C \), called the Carnot function. (Clapeyron used \( 1/C \) for the moderating factor, and he called this \( C \) the Carnot function.) Heat throughput \( Q \) could serve as the analogue of mass of water falling. For an infinitesimal cycle, with heat throughput \( dq \) and temperatures \( t + dt \) and \( t \), the Carnot function is easier to visualise, since some kind of averaging over temperature of it is not required. The work output \( dw \) can then be written \( dw = dq \times C \times dt \). The infinitesimal cycle and the Carnot function both had key roles in later developments of thermodynamics (Section 5.1), buttressing further Carnot’s remarkable insight.

3.1.3 Postscript

Carnot’s working notes after 1824 show that he moved from caloric towards accepting a mechanical theory of heat [3]. His premature death in 1832 prevented them from being recognised as a contribution to the formulation of the First Law of Thermodynamics: in 1824, there was no First Law, with little of direct relevance from the theoretical side on its precursor, transformations between heat and work. Carnot had, however, provided a convincing
association of heat flow with mechanical work which – however, tenuous – was supported – however, loosely – by a conservation principle: conservation of caloric thereby permitting the working substance to return to its starting state at the end of every cycle. It was Clausius who would provide significant additions to and replacements of – and consequential realignment of – the various pieces (Section 3.4); and eventually follow Thomson in accepting energy as the unifying concept (Section 3.5).

Remarkably, Carnot’s memoir is recognised as pointing the way to the more subtle Second Law of Thermodynamics (Section 4.2), with a universal Carnot function providing a clue to a fundamental interdependence of heat flow and temperature (Section 4.4) which was crucial to the establishment of that law.

3.2 Davy, Rumford, Mayer and Joule: transformations and conservation

Observations had been reported from just before 1800 onwards of a direct link between mechanics and heat, casting doubt on the caloric theory. Sir Humphrey Davy (1778–1829) produced heat from mechanical energy by melting the surfaces of contact between two blocks of ice by rubbing them together [6]. Count Rumford (1753–1814) observed the large amount of heat generated in the process of boring cannons which led him to deduce a relationship between mechanical work and heat which he published in 1798 [7]. Energy would appear later – in parallel with developments in language – as the enabling concept for the successful merging of heat and mechanics, building on to the notion that heat could be related to some kind of motion. In the 1840s, Robert J. Mayer (1814–1878) in Germany had speculatively broadened his interpretation, in chemical and thermal terms, of certain changes occurring in physiological processes. He suggested, in 1842, that conservation of some kind of ‘force’ was involved [8, 9], to parallel conservation of mass in chemistry. His treatment there of a relationship between mechanical work and heat prevented him from quantifying it convincingly, but history credits him with having recognised but not having driven forward adequately a principle of conservation of ‘energy’ (replacing ‘force’). In 1845, his more detailed account in a follow-up paper was rejected and he published it privately [10]. By that time measurements in England by James Prescott Joule (1818–1889), and his interpretation of them, had begun to dominate the field. In a sequence of remarkably precise direct measurements on the transformation from work to heat, he established an accurate value for a mechanical equivalent of heat including, early in the sequence, observations which had also shown a similar transformation of electrical energy [11]. The outcomes from these observations became crucial to the understanding of the parts of the Carnot cycle involving transformations between heat and work.

3.3 Thomson’s assessment of Carnot’s theory

In 1849, Thomson [4] presented an account of Carnot’s theory, putting a focus on the return of the working substance to its original state at the end of each cycle with heat gains by and heat losses from the working substance in balance. He noted Carnot’s own remarks that ‘to deny [it] would be to overturn the whole theory of heat, of which it is the fundamental principle’, followed by ‘several experimental facts appear nearly inexplicable in the actual state of this [caloric-based] theory’. Thus prepared, Thomson then gave an authoritative and sympathetic interpretation of the basis of Carnot’s theory, both steam and air engines, together with data from measurements, mainly at the laboratories of Henri Victor Regnault (1810–1878), on thermometry and associated thermal properties of materials, seeking compatibility. (Following graduation from Cambridge
University in 1845, Thomson had spent several months in Regnault’s laboratories in Paris. Results of measurements being made there provided a complementary addition to theoretical work [12] of J.B. Joseph Fourier (1768–1830) on heat transfer, already well known to him.)

In particular, he developed analytical expressions linking $M$ (our $W$) and $H$ (our $Q$): $M = H\int \mu \, dt$, with $\mu$ denoting ‘Carnot’s coefficient’. From experimental values of $\mu$ at one degree intervals from $1^\circ\text{C}$ to $230^\circ\text{C}$, he calculated and presented as a look-up table, the value of $M$ (in foot-pounds) per ‘thermic unit centigrade’ of heat transferring from ‘a body at any temperature less than $230^\circ\text{C}$ to a body at $0^\circ\text{C}$’. These show that Thomson’s $\mu$ corresponds to $J \times C$, where $J$, the mechanical equivalent of heat, converts thermal to mechanical values, and $C$ is the Carnot function (Section 3.1.2). He showed that there was good agreement between the few values (of $\mu$) quoted by Clapeyron and his own derived from Regnault’s observations. However, he cautioned that ‘[Carnot’s theory of the motive power of heat] and with it every other branch of the theory of heat, may ultimately require to be constructed upon another foundation, when our experimental data are more complete’.

### 3.4 Clausius’s extension of Carnot’s theory

In 1850, Rudolf J.E. Clausius (1822–1888) presented an analysis [13] of the assumptions in Carnot’s paper. He accepted as a fundamental principle ‘that whenever work is produced by heat and a permanent alteration of the body [working substance] in action does not at the same time take place, a certain quantity of heat passes from a warm body to a cold one’. However, he doubted that ‘the assertion, that in the production of work loss of heat never occurs, is sufficiently established by experiment’. He continued, ‘The difference between the two ways of regarding the subject has been seized with much greater clearness by W. Thomson, who has applied the recent investigations of Regnault, on the pressure and latent heat of steam, to the completing of the memoir of Carnot’.

Clausius sought to remove the conflict. Introducing a small alteration to Carnot’s model, he claimed ‘it is quite possible that in the production of work ... a certain portion of heat may be consumed, and a further portion transmitted from a warm body to a cold one; and both portions may stand in a certain definite relation to the quantity of work produced’. Clausius elaborated on the ‘portion of heat consumed’, treating it as the combination of heat retained (by a working substance) and work done internally (whilst the substance expands) and work done externally. Shortening Clausius’s argument, in a reversible cyclic process the heat retained and the internal work could be recovered when the working substance regains its initial state but not so the external work. The first two together could be regarded as a thermodynamic function of any two appropriate thermodynamic coordinates – say temperature $t$ and volume $v$ – just as, from the ‘gas laws’, the pressure $p$ of a gas can be related to $t$ and $v$ to a good approximation via the ideal gas equation. He gave it the symbol $U$, a function which limited whilst enabling the conversion of heat to external work. For small changes in heat $Q$, $U$ and volume $v$, he wrote $dQ = dU + Ap \, dv \ldots$ where $A$ denotes a constant which expresses the equivalent of heat for the unit of work’. He thus improved the tools with which to build a law suggested by the researches of Mayer and Joule taken together.

Throughout, Clausius was guided by ‘the general deportment of heat, which everywhere exhibits the tendency to annul differences of temperature, and therefore to pass from a warmer body to a colder one’. He was able to adapt Carnot’s reasoning on the composite of two...
reversible engines, one driving the other in reverse (Section 3.1.1), coupled this time so that the conjectured more efficient was tuned to provide, from a smaller throughput of heat, exactly the work needed to drive the Carnot engine in reverse. The resulting net flow of heat from cold to hot, unaccompanied by a net amount of work, was contrary to the general deportment of heat and the conjectured more efficient engine rejected.

3.5 Thomson and Clausius: the First Law of Thermodynamics stated

Within Thomson’s dynamical theory of heat, presented in 1851 [14], he associated heat with molecular motion (rather than caloric) and identified a mechanical energy, $e$, of a body as its means of participating in transformations and transmissions of heat. Thus, the increase $\Delta e$ in mechanical energy of a body resulting from heat input $\Delta Q$ and performance by it of external work $\Delta W$ could be written $\Delta e = \Delta Q - \Delta W$. Evaluating $e$ from a sufficiently extensive database of values of thermodynamic coordinates and properties was feasible. Later, Thomson referred to $e$ as intrinsic energy before it became widely accepted as internal energy and given Clausius’s symbol $U$. Much later, in a paper in 1865 on the dynamical theory of gases [15], Clausius, referred back to his introduction of $U$ 15 years earlier, and remarked that ‘since then, however, W. Thomson has proposed the term energy of the body for this magnitude, which mode of designation I have adopted as one very appropriately chosen’.

With heat $Q$ and external work $W$ both now accepted as manifestations of energy, the tracking of energy changes, without loss, led to a formulation of the First Law of Thermodynamics in the form of a tally. Following Clausius by starting with $\Delta Q$, it is: $\Delta Q = \Delta U + \Delta W$. He also gave there a succinct statement of the First Law: ‘the energy of the universe is constant’.

4 The Second Law of Thermodynamics

4.1 Interlude: an absolute temperature scale

Thomson [16] recognised the need for an absolute scale of temperature in the science of energy and conceived one which he published in 1848. He had identified a difficulty arising from the variation with temperature of thermal properties of materials. If such a property were used in the design of a thermometer, its responses to the same heat input at different temperatures would be different temperature rises as indicated by the increase in number of degrees or scale divisions of the chosen scale. The device would be unable to record heat input correctly (i.e. absolutely) at all temperatures.

In his search for an absolute thermometric scale, Thomson drew attention to the properties, following Carnot, of an ideal reversible heat engine operating between temperatures $t_1$ and $t_2$, (as described above). He wrote, using mechanical work output (the ‘mechanical effect’) as an intermediary: ‘The characteristic property of the scale which I now propose is, that all degrees have the same value; that is, that a unit of heat descending from a body A at the temperature $T˚$ of this scale, to a body B at the temperature $(T˚-1)$˚, would give out the same mechanical effect, whatever be the number $T$. This may justly be termed an absolute scale, since its characteristic is quite independent of the physical properties of any specific substance’. He explained how the scale could be realised, approximately, based on existing data from Regnault’s measurements on thermal properties of gases to access the mechanical effect, via the resulting values of his ‘Carnot coefficient’ $\mu$ (Section 3.3). He anticipated that more extensive data would improve
244 | Kelvin, thermodynamics and the natural world

the calculation of absolute temperature by correlation with temperature measured with an air
thermometer [17].

Thomson later simplified the specification of absolute temperature in ‘Fundamental Principles
of General Thermo-dynamics recapitulated’, his introduction to a paper in 1854 on thermo-
electric currents [18]. This time he summarised the argument: ‘the absolute values of two
temperatures are to one another in the proportion of heat taken in to the heat rejected in a perfect
thermo-dynamic engine working with a source and refrigerator at the higher and lower of the
temperatures respectively’. Thus, for a Carnot cycle working between heat reservoirs at absolute
temperatures \( T_1 \) and \( T_2 < T_1 \), and where the work output per cycle, \( W \), is equal to the difference,
\( Q_1 - Q_2 \), between heat in and heat out, \( Q_1/Q_2 = T_1/T_2 \). The efficiency for such Carnot engine,
\( W/Q_1 = (Q_1 - Q_2)/Q_1 = (T_1 - T_2)/T_1 \), proportional to temperature difference and with \( 1/T_1 \) as
the factor which moderates the efficiency in line with Carnot’s interpretation (Section 3.1.2).

4.2 Axioms and statements

In the course of clarifying Carnot’s theory by input from the First Law of Thermodynamics
(Section 3.4), Clausius had been led to propose in 1850 an overarching principle: heat flows
spontaneously only to reduce a temperature difference, not to increase it. For a Carnot engine
working in reverse, heat flow from cold to hot had to be accompanied by work input. Thomson
also identified a principle by a path that held closer to temperature-dependent restrictions on
the transformation of heat to mechanical work. He presented both, as axioms, in the first section
(March 1851) of his long paper [14] on a dynamical theory of heat.

Section 1 included a key paragraph comprising two propositions which together provided the
foundations for ‘the whole theory of the motive power of heat’. Briefly, ‘PROP. I’, acknowledging
Joule, was that the same amount of heat is used up (or generated) when equal amounts of
mechanical effect are produced (or lost) purely thermally. The two axioms were then presented
so as to validate the paper’s ‘PROP. II’, attributed to Carnot and Clausius, that a fully reversible
engine is the most efficient of all engines extracting heat from the same constant temperature
heat source whilst delivering heat to the same (lower) constant temperature heat sink:

**Clausius’s axiom:** It is impossible for a self-acting machine, unaided by any external agency, to convey
heat from one body to another at a higher temperature.

**Thomson’s axiom:** It is impossible, by means of inanimate material agency, to derive mechanical effect
from any portion of matter by cooling it below the temperature of the coldest of surrounding objects.

Thomson remarked that ‘It is easily shown, that … either [axiom] is a consequence of the other’.
This kind of exercise is still standard in the presentation of the two typical statements, below, of
the Second Law in textbooks on thermodynamics [19].

In Parts I and II of Section 1, on the ‘motive power of heat’, he amended Carnot’s treatment
replacing the caloric theory with interconversion of heat and mechanical work as demonstrat-
ed by Joule. The infinitesimal Carnot cycle retained the moderating factor, with \( \mu \) now called
Carnot’s function. Although it had the optimum efficiency, a large proportion of the heat input
was ‘irrecoverably lost to man, and therefore ‘wasted’, although not annihilated’. He gave an
analytical treatment of the relationship between work output and heat input for heat reservoirs
having a finite temperature difference, also modelling the system as a stack of infinitesimal cycles. The work output was less than would be got from Carnot’s theory, but tended to it as the temperature difference decreased. He also noted Joule’s suggestion that $\mu$ might be inversely proportional to absolute temperature. Part III demonstrated the ease with which the dynamical theory relationships between physical properties of substances – both those whose experimental values were well known and others – which had a bearing on thermodynamic processes. The single Part IV of Section 2 (April 1851) dealt with possible measurements on the relationship between work done and heat released during compression of gases, exploiting Joule’s proven instrumentation. Section 3 (December 1851) comprised Part V in which Thomson introduced the concept of the mechanical energy, $e$, of a fluid (Section 3.5).

The axioms provide the starting points for typical statements of the Second Law of Thermodynamics:

**The Clausius statement:** It is impossible to construct a device that, operating in a cycle, will produce no effect other than the transfer of heat from a cooler to a hotter body.

Thomson’s axiom was developed more substantially, following this re-statement of the law by Max Planck (1858–1947): *It is impossible to construct an engine which will work in a complete cycle, and produce no effect except the raising of a weight and the cooling of a heat-reservoir* [20]. It led to the compromise:

**The Kelvin–Planck statement:** It is impossible to construct an engine that, operating in a cycle, will produce no effect other than the extraction of heat from a reservoir and the performance of an equivalent amount of work.

### 4.3 Thomson’s insight

Thomson’s short paper [21] in 1852, just a few months after his major dynamical theory, was a response to Carnot’s proposition that there is ‘absolute waste of mechanical energy’ accompanying heat flow from a hot to a cold body by any process incompatible with the constraints satisfied by an ideal engine. As in his dynamical theory, he chose to replace waste by a transformation, rather than an annihilation of energy. That led him to consider the possibility of restoration ‘to its primitive condition’ of mechanical (in present day terms, potential and kinetic) energy that had, in part, been transformed to heat. He introduced the term dissipation of energy as a means of distinguishing irreversible from reversible processes.

Thomson set out four propositions which were ‘necessary consequences’ of his own axiom (Section 4.2). The first proposition put a condition on heat produced reversibly by transformation of mechanical energy. His choice of illustrating that the transformation could not be accounted for in isolation was: ‘When heat is created by a reversible process (so that the mechanical energy thus spent may be restored to its primitive condition), there is also a transference from a cold body to a hot body of a quantity of heat bearing to the quantity created a definite proportion depending on the temperatures of the two bodies’. The next three propositions, all involving dissipation of mechanical energy thereby ruling out its full restoration, were much more general, reflecting the abundance of irreversible processes. The examples were: any heat produced irreversibly (such as frictionally), heat diffused by conduction and the absorption of radiant heat or light.
Setting aside his few brief remarks on their relevance in general terms to biology and chemistry, his main conclusion from the propositions was that ‘there is at present in the material world a universal tendency to the dissipation of mechanical energy’. This overarching statement has to be acceptable as an informative, even if qualitative, guide to the Second Law of Thermodynamics.

4.4 Clausius’s initiative

By 1854, Clausius [22] had moved on from combining the Carnot cycle with his ‘first fundamental theorem’, essentially the First Law of Thermodynamics, extending his treatment from that set out four years before [13], which he included by way of introduction in a form which he claimed was ‘at once more general and more concise’. He claimed that the four-stage Carnot cycle was too restrictive. In particular, a higher temperature body (the boiler) had to provide both the heat to be converted to external work and the heat released to the lower temperature body (the condenser). It was, however, important to retain the link identified by Carnot between the conversion of heat to external work and the accompanying throughput of heat – both per cycle with all stages of the cycle reversible.

Clausius therefore introduced a six-stage reversible cycle. The first three stages are chosen here as in the Carnot cycle described above (Section 3.1.1) starting with an isothermal expansion at the higher temperature. The return adiabatic compression does not close the cycle. Instead it takes the working substance to a temperature of $t > t_1$, followed by an isothermal expansion at $t$ of such duration that the cycle can be closed by an adiabatic expansion, during which the temperature falls from $t$ to $t_1$. This provided Clausius with a combination of two mutually supportive heat transformations per cycle: heat flow $Q$ in at the temperature $t_1$ transformed to become the same amount $Q$ out at a lower temperature $t_2$; and heat flow $q$ at a temperature $t > t_1$ which, being the net heat in, was transformed completely into external work $w$. In completing the cycle, the working substance returned to its initial thermodynamic state.

He argued that these two particular transformations were ‘phenomena of the same nature’. Retaining the simplicity (dependence on temperatures only) of the expression for efficiency (Section 3.1.2), he gave them ‘equivalence values’: $Q \times F(t_1,t_2)$ for the transformation of $Q$, and $q \times f(t)$ for the ‘heat-to-work’ transformation of $q$. He noted that the transformation of $Q$ occurred in the direction of spontaneous heat flow, but that the transformation of $q$ was in the opposite direction to what could be chosen to be the spontaneous direction for transformations between heat and work, namely work to heat. He sought the implications of a balance between the two heat transformations, now defining the cycle by combining their equivalence values algebraically:

$$Q \times F(t_1,t_2) - q \times f(t) = 0.$$  

Clausius then showed how to extract a relationship between the two functions of temperature. The strategy was similar to that used by Carnot: combine two Carnot engines, one driving the other in reverse (Section 3.1.1). Clausius’s second cycle, performed in reverse, otherwise differed only in having the ‘third’ reservoir at a temperature $t' > t$, and $q' > q$. For it, $Q \times F(t_2,t_1) + q' \times f'(t') = 0$. Because $F(t_2,t_1) = - F(t_1,t_2)$, the two heat flows $Q$ cancelled in the composite system, leaving only $q$ at $t$ and $q'$ at $t'$, and $q' \times f'(t') = q \times f(t)$ which became an input to his analysis by the equivalence value method to a now two-reservoir system – essentially a Carnot cycle. The upshot was $F(t_2,t_1) = f(t_2) - f(t_1)$. The equivalence value for heat throughput could thus be replaced by the combination of two equivalence values; one, $Qf(t_1)$, for the release and the
other, \(Qf(t_2)\), for the absorption of heat, by reservoirs at the denoted temperatures. The algebraic sum of the equivalence values in his reversible six-stage cycle could now be rewritten \(Q\times f(t_2) - Q\times f(t_1) - q\times f(t) = 0\) without losing sight of the balance of the two processes involved. This was the basis of several immediately ensuing stages of generalisation and interpretation.

The return to the same thermodynamic state in each cycle suggested that the product \(f(t)Q\) could, like \(U\) (Section 3.4), form the basis of another thermodynamic function.

Then for any multistage reversible cycle with any number of transformations of heat \(Q\) at temperature \(t\), the summation, now written \(\Sigma f(t)Q\), could even be replaced by an integration \(\int f(t)\,dQ\) to allow for infinitesimal heat flows, each for its corresponding temperature. Clausius proposed that \(\Sigma f(t)Q = 0\) generalised to \(\int f(t)\,dQ = 0\) could be ‘the analytical expression for all reversible cyclical processes, of the second fundamental theorem in the mechanical theory of heat’.

Furthermore, the temperatures could equally well be taken to be those of the working substance during reversible, isothermal (the only) heat flows. This gave the option of a change of focus from heat reservoirs to working substance. Heat leaving a reservoir became heat entering the working substance but \(\Sigma f(t)Q = 0\) still pertained.

Clausius then introduced a descriptive term ‘compensation’ to distinguish reversible from irreversible processes. His selective list of the latter overlapped those in Thomson’s introduction of dissipation (Section 4.3). Clausius’s description then parallels Thomson’s. The sum of all the positive terms in \(\Sigma f(t)Q\) would be compensated by the sum of all the negative terms in a reversible cycle. In proceeding to address irreversibility, he proposed that heat thereby generated could only be included as an uncompensated transformation. Thus ‘for all cyclical processes, those that are reversible forming the limit: the algebraical sum of all transformations occurring in a cyclical process can only be positive’. In the notation being used here, \(\Sigma f(t)Q \geq 0\).

Finally, an appeal to the gas laws provided the path to the simplest explicit function of temperature for \(f(t)\), namely \(1/T\), with \(T\) the absolute temperature. A rare detailed account of Clausius’s 1854 paper is given by Cropper [23].

Clausius’s subsequent papers trace his systematic development of the issues and concepts he had recognised, removing the restriction to cyclic processes and extending a consideration of thermodynamic properties of substances. The final paper [11] in the sequence included writing \(dQ/T = dS\) and choosing \(S\) and the name entropy for this now well-established thermodynamic function. The more general version of his second fundamental theorem could be written \(dS \geq 0\), its brevity hiding its massive content. As he did with his statement of the First Law, he gave his succinct alternative statement of the Second Law of Thermodynamics: ‘the entropy of the universe tends to a maximum’.

4.5 Interpretation

One attraction of Clausius’s general statement (of the Second Law) is that it implies the importance of changes, if any, in the value of a thermodynamic function entropy. In introductory thermodynamics, a range of illustrative examples can be collated and a generalisation made:
entropy increase can be associated with a decrease of order that can be related to structure in its widest sense.

From the extreme of bulk properties of materials is the historically challenging change from an initial spatially structured (or ordered) temperature gradient across a metal bar to an equilibrium structureless (thereby less ordered) uniform temperature throughout, accompanied by no work output. This example serves to illustrate the increase in entropy in irreversible processes culminating in an equilibrium state. The alternative ‘disordered’ for ‘less ordered’ is not very helpful here.

Topics from condensed matter physics appear in the centre ground. Elementary examples are associated with differences in order below and above some transition temperature, itself determined by energy considerations. The 50/50 copper/zinc alloy, β-brass, shows large peak in its specific heat capacity around 730 K. Below that temperature, the copper and zinc atoms have an ordered arrangement on a body-centred-cubic, crystalline lattice. The arrangement can be visualised as having a zinc atom at the centre of each cube, and a copper atom at those points where eight contiguous cubes meet, so that each copper atom is shared equally amongst the eight cubes, one-eighth to each. Above the transition temperature, both sets of constituent atoms are randomly distributed, hence disordered, on the same complete set of lattice sites. The transition is accompanied by an increase in entropy [24]. For this topic-specific transition, the brief description ‘order-disorder’ is appropriate.

5 Thermodynamic properties and processes

5.1 Overview and Maxwell’s relations

In classical thermodynamics, thermodynamic properties of gases can be described in terms of those two coordinates, chosen from pressure, volume and temperature, which match them to reversible thermodynamic processes whose outcomes they determine. Internal energy \( U \) and the more sophisticated entropy \( S \) were augmented by other thermodynamic functions of classical thermodynamics such as enthalpy \( H \), free energy \( E \) and the Gibbs function \( G \). They are inter-related and their utility is topic-specific. Combined with inputs from engineering, there was a steady extension of the understanding – in breadth and depth – of basic thermodynamic properties. Some were essentially thermal, such as heat capacities for heat transfer and, for phase transitions, the heat storage and release in terms of latent heat. Others were essentially mechanical, building on to an empirical base of coefficients determining properties of solids such as compressibility and ductility. Inevitably, with heat regarded as energy and entropy change frequently acting as a kind of proxy for heat transfer, this theory provided thermal and mechanical interdependence.

Also in this period, descriptions of relationships amongst thermodynamic coordinates moved to the use of the language of differential calculus from that of geometry, the latter as exemplified in their thermodynamics books by Maxwell [25] and Tait [26]. For example, they chose a geometrical scenario of the extreme theoretical Carnot cycle (Section 3.1.2) with an infinitesimal difference in temperature \( \Delta t \) between heat source and heat sink, and the corresponding infinitesimal difference in pressure \( \Delta P \) between the two isothermal stages. The geometry of this cycle was nearly a rectangle of sides \( \Delta P \) and, also infinitesimal, \( \Delta V \) for the isothermal expansion
and compression stages. Two expressions for the efficiency were equated. One was largely a matter of definition: \( \Delta W / (J \times \Delta Q) \) for work output \( \Delta W \), heat input \( \Delta Q \) at the higher temperature with the factor \( J \), the mechanical equivalent of heat, converting heat input from thermal to mechanical units. The other was Carnot’s deduction that the efficiency was a function of temperature, independent of the choice of working substance, and was eventually rewritten \( C \times \Delta t \), with \( C \) representing the universal Carnot function.

The development, outlined in detail by Klein [27], centred on rearranging the factors. \( \Delta W \) was replaced by \( \Delta P \times \Delta V \), a good approximation to the area of the cycle plotted on a \( P-V \) diagram. Thus, \( \Delta P/\Delta t = J \times C \times \Delta Q / \Delta V \). Tait chose to put \( \Delta Q / \Delta V = M \), with the order of the three factors being Maxwell’s initials, JCM. In the three-way correspondence amongst Maxwell, Thomson and Tait, Maxwell became \( dp/dt \), with Thomson and Tait adopting the simpler \( T \) and \( T' \), respectively.

Returning to thermodynamics, that equation (together with another three, all similarly derived, concisely, in Maxwell’s book) was later re-presented. The absolute temperature \( T \) replaced \( t \), for consistency with the replacement \( 1/T \) for the Carnot function and \( J \) was presumed to have been merged with \( \Delta Q \). In the transition to differential calculus, the constant volume \( \Delta P/\Delta T \) was re-cast as \( (\partial P/\partial T)_V \) and the isothermal \( (\Delta Q / \Delta V) / T \), firstly written \( (\Delta S / \Delta V) \) with the entropy difference \( \Delta S \) replacing \( \Delta Q / T \), was re-cast as \( (\partial S / \partial V) \) to give the first of the four Maxwell’s relations:

\[
\begin{align*}
(\partial P / \partial T)_V &= (\partial S / \partial V)_T \\
(\partial T / \partial P)_S &= (\partial V / \partial S)_P \\
(\partial V / \partial T)_P &= -(\partial S / \partial P)_T \\
(\partial T / \partial V)_S &= -(\partial P / \partial S)_V
\end{align*}
\]

**5.2 Illustrative example: the Joule–Kelvin process**

This process consists of pushing a column of gas steadily through a large number of tiny holes in a barrier, settling beyond the barrier at a lower pressure. Its historical importance is twofold. It stems from a collaboration of Joule and Thomson on measuring the temperature change of a gas made to flow through a barrier consisting of a thermally insulated plug of either tightly packed cotton or a tightly rolled piece of silk. Its popular name, the ‘porous plug’ process, was introduced by Joule and Thomson [28]. A typical analysis, many years later [29], exemplifies some of the strategy of classical thermodynamics.

It is assumed that the gas can be tracked to the extent that during some arbitrary time interval an amount of mass \( M \) in volume \( V_1 \) at uniform pressure \( P_1 \) and temperature \( T_1 \) is pushed by the rest of the column further upstream to emerge to occupy a larger volume \( V_2 \) and so have the smaller uniform pressure \( P_2 \) and the common temperature \( T_2 \) of all the downstream gas. The barrier and the gas column are thermally insulated and so the difference in internal energies of the mass \( M \) before and after, \( U_2 - U_1 \), is equal to the work done on the mass \( M \) by the gas column also at pressure \( P_1 \) upstream (pressure × volume moved through) minus the corresponding work done by \( M \) pushing on the gas ahead of it downstream: \( P_1 V_1 - P_2 V_2 \). Hence, \( U_2 + P_2 V_2 = U_1 + P_1 V_1 \). The combination \( U + PV \) is called the enthalpy \( H \) (Section 5.1). Although \( H \) cannot
be tracked during the irreversible passage of the gas through the holes, the initial and final states of the mass $M$ are modelled as equilibrium states and for them the change in any other thermodynamic function, here the temperature $T$, can be calculated for a constant enthalpy process between the same end states, this being optimally chosen to ease the analysis. Thus, the temperature $T$ is chosen to be a function of pressure $P$ and enthalpy $H$. This leaves $T$ varying with $P$ only because $H$ does not change. The particular variation of $T$ with $P$ can therefore be written as $(\partial T/\partial P)_H = \mu_{JK}$, the Joule–Kelvin coefficient. For a change in pressure from $P_1$ to $P_2$, the change in temperature $T_2 - T_1 = \int \mu_{JK} dP$, integration being between the limits $P_1$ and $P_2$.

The now traditional tactics in thermodynamic analyses include replacing $H$ in terms of other thermodynamic functions and parameters. Start with $(\partial T/\partial P)_H = - (\partial T/\partial H)p(\partial H/\partial P)_T$. $(\partial T/\partial H)_P$ is readily shown to be equal to the reciprocal of the heat capacity, $C_P$, of the mass $M$ of gas at constant pressure. Entropy enters by incorporating a version of the First Law written $T dS = dU + P dV$, leading to $(\partial H/\partial P)_T = T(\partial S/\partial P)_T + V = - [(\partial V/\partial T)_P + V]$, using the Maxwell Relation – $(\partial S/\partial P)_T = (\partial V/\partial T)_P$. The outcome is $\mu_{JK} = [(T(\partial V/\partial T)_P - V)/C_P]$, which can be positive or negative, depending on the values of the thermodynamic coordinates and their mutual variation. From it, the choices making $T_2 - T_1$ negative point to the design details of devices using the effect for gas cooling and liquefaction, with the great advantage that, especially as the final stage of the cooling, it has no moving parts.

Joule and Thomson arrived at the same formula for $T_2 - T_1$ within their extensive papers. These merge a detailed account of Joule's meticulous measurements of temperature with a steady development of immediately applicable theory ‘founded on elementary thermodynamic principles’ over a 10-year period, together matching the inspired design of the apparatus. The set of named functions such as enthalpy and entropy, and the packaging of equivalences of properties as exemplified by Maxwell’s Relations, came later.

6 Statistical thermodynamics

6.1 An early sighting: dissipation, recoverability and Maxwell’s Demon

In his 1874 paper on the kinetic theory of dissipation of energy, Thomson [30] returned to the much discussed conduction of heat through a thermally insulated bar of metal initially warmer at one end than the other. The relevant details he extracted from a detailed treatment of a similar – but more easily visualised – process of heat diffusion between two halves, each initially in equilibrium, of a sample of gas, one half warmer than the other, in a closed, insulated vessel. As with the bar, the state of that system would change, tending to equilibrium. He concluded that for both bar and gas, a return to the initial temperature difference could only happen following input ‘by the agency of external sources of heat and cold’. This was the basis of his 1851 version of Clausius’s axiom clarified a year later by his introduction of dissipation, and subsequent unavailability, of energy (Section 4.3).

In the intervening years, interest had grown in acquiring an enhanced understanding of thermodynamics from the dynamics of gases treated as purely mechanical systems. Almost by way of expressing caution, Thomson began his paper by contrasting ‘abstract dynamics’ with ‘physical dynamics’. In the former, the motion of every part of a system could be reversed exactly – and still satisfy the laws of motion – if the velocity of each part were to be instantaneously reversed: ‘in mathematical language, any solution remains a solution when
The emergence and evolution of the Second Law of Thermodynamics

[time] \( t \) is changed into \(-t'\). He asserted that such perfect reversibility – by implication helpful in the study of some of the properties of gases – fails in the physical dynamics of responses to a wide variety of forces. These could be attributed to friction, imperfections in properties of materials in bulk – electrical and chemical as well as mechanical – and including absorption of radiation. Thomson addressed the properties of the simpler system, a gas consisting of a large number of identical molecules. He used much of the imagery of abstract dynamics. However, the impact of physical dynamics led him to replace certainties with probabilities.

He embedded in his discussion a thought experiment involving a large number of identical Maxwell Demons. Maxwell had pointed out the properties of such a ‘being’ in a letter to his lifelong friend Professor P.G. Tait (1831–1901), in Edinburgh [31] in December 1867 followed within a few weeks – in February 1868 – by a much briefer version to Thomson in Glasgow [32]. The context was Maxwell’s comments on the draft version of Tait’s textbook on thermodynamics. A detailed account is included in Chapter 5, ‘Kelvin, Maxwell, Clausius and Tait: the correspondence of James Clerk Maxwell’. Suitably provided with a sorting device and, significantly, with a complete knowledge of positions and velocities of each of a very large number of molecules constituting the sample, a demon (Thomson’s name for the being) could select molecules according to their energy and produce a non-uniform distribution of energy, and thus of temperature, within the sample. The letter to Tait had the demon in charge of a lightweight trapdoor in a partition so as not to disturb with the normal energies of the molecules, of which the more energetic could be chosen – by the demon briefly opening the trapdoor – to become more numerous on one side of the barrier and the number of less energetic made more numerous on the other (see also the Maxwell’s Demon Section in the Appendix). This procedure, continued, would enable a continuous transfer of heat from cold to hot in circumstances which violate the Second Law of Thermodynamics. Maxwell dispensed with such a structure in his letter to Thomson, who chose to extend Maxwell’s concept to his own approach to describing changes in, and stability of, energy distributions as might be monitored by distributions of temperature regarded as a bulk property, in Thomson’s phrase, ‘in the gross’. Outcomes from a three-stage thought experiment, also detailed in Chapter 5, included a much enhanced understanding of the equilibrium state, as well as the intended support for his theory of the dissipation of energy.

An army of Maxwell Demons was put in positions in a vessel containing a sample of gas in equilibrium so as to define the position of an eventual interface between two halves of the vessel. Acting to order, they allowed passage in either direction to enough molecules, identified by energy, to produce a disequalisation (Thomson’s word) in energy between the two halves. With the army of demons suddenly stood down, the dynamics of the interactions amongst the moving molecules would lead to an equilibrium state in some finite time interval. Then, with the velocities of all molecules suddenly reversed, the system would seek to recover the demons-produced disequalisation by having every molecule retrace its path. Thomson argued that this might begin but in the same finite time interval would be inhibited from happening. Thereafter, the system, still left to itself, would seek some equilibrium state.

Thomson moved to propose that, in reality, without external influence an equilibrium state would persist (essentially Clausius’s axiom), whilst subject continually to fluctuations about properties such as averages of distributions are amenable to description at the molecular level. Thomson used the language of probability theory to discuss these. The number of ways, from the mathematics of permutations, of making any allocation of molecules to locations in the
vessel had to be multiplied by the probability of occurrence of that allocation. This is an extension of predicting the numbers, H and T, of ‘heads’ and ‘tails’, respectively, in a repeated exercise of tossing a number, n, of identical, perfectly symmetrical coins. Different H’s and T’s are the more likely if n is small. If n is large, many repeats are needed for a significantly different H’s and T’s to occur. Translating to a very large number of molecules in the vessel, a prohibitively long time is needed for any substantial disequalisation to happen. Thomson confirmed this with calculations of ‘the number of chances against one’ that a mixture of oxygen and nitrogen molecules would redistribute themselves to different volumes in their containing vessel.

6.2 From Maxwell’s kinetic theory to the Maxwell–Boltzmann distribution

Maxwell’s kinetic theory [33] and his follow-up dynamical theory of gases (guided by experimental results) [34] established their roles in the description and interpretation of properties of fluids. It was not feasible to track the movements of all molecules in a sample of gas in equilibrium. However, it was possible to predict the average value of velocity, \( v \), and of \( v^2 \), of the molecules whose collisions with one another were governed by very short range intermolecular forces, but which otherwise moved freely.

For an equilibrium sample of \( N \) molecules, the number with velocity in the range \( v \) to \( v+dv \) is \( NF(v)dv \), where \( F(v) = \left( \frac{4}{a^3 \sqrt{\pi}} \right) v^2 \exp\left( -\frac{v^2}{a^2} \right) \). Here, \( a \) is a constant, whose value is determined by the molecular mass and the sample temperature. The analysis led to \( \frac{3}{2}a^2 \) for the average value of \( v^2 \). The function \( F(v) \) can be transformed to a function \( F(w) \) of kinetic energy \( w = \frac{1}{2}mv^2 \). Algebraic manipulation leads to \( F(w) = \left( \frac{2}{\sqrt{\pi}} \right) \frac{1}{kT} \left( \frac{w}{kT} \right) \exp\left( -\frac{w}{kT} \right) \). Now, the number of molecules with kinetic energy in the range \( w \) to \( w+dw \) is \( NF(w)dw \), where \( F(w) \) is usually called a Maxwell–Boltzmann distribution function. The factor \( kT \), a kind of reference energy, replaces \( \frac{1}{2}ma^2 \) and uses a link from the kinetic theory of gases between average molecular kinetic energy and temperature \( T \): \( \frac{1}{2}m(\frac{3}{2})a^2 = \frac{3kT}{2} \). ‘Boltzmann’s constant’ \( k = \frac{R}{N_0} \), where \( R \) is the universal gas constant and \( N_0 \) is Avogadro’s number.

6.3 Statistical mechanics: the Second Law and entropy

Klein [27] has provided an excellent commentary in his summary of early developments in the search for an account of principles of thermodynamics solely in terms of the dynamics of particles. It soon became apparent that statistical methods would have to be included, leading to ‘statistical mechanics’. From about 1865 onwards, the path of what amounted to a sequel to Maxwell’s kinetic theory was developed by Ludwig Boltzmann (1844–1906), with substantial contributions on matters of interpretation and substance from Clausius and others. The mathematics evolved around an exponential energy factor, structurally a generalisation of the kinetic energy factor in the Maxwell–Boltzmann distribution (Section 6.2). Once a gas has reached an equilibrium state, there could be continual reallocations of energy amongst the molecules whilst always retaining the equilibrium distribution. However, there was the option – of solutions in keeping with laws of motion – of having all molecular movements repeated in reverse. Thomson had discussed a very special case of this (Section 6.1) in the context of an inadequacy of ‘abstract mechanics’ to provide a complete description of changes in energy distributions, and was led to introduce probability arguments in support. Now, the issue centred on changes in entropy. The reversal of all molecular motions would move the system steadily away from equilibrium accompanied by an entropy decrease, thereby challenging the Second Law. Boltzmann countered this with an argument amounting to a general acceptance.
of a probabilistic strand in the Second Law. Simplifying, a system reaching its most probable state – the equilibrium state – would then never stray far from it.

### 6.4 Statistical entropy and chemical thermodynamics

The parameters specifying the state of a system at molecular level must be compatible with molecular properties, typically position and velocity. Guided by the mathematics of permutations and of probability, as was Thomson (Section 6.1), an analysis leads to that set of parameter values which will keep recurring most often during their continual reallocation amongst the molecules as a result of intermolecular collisions – subject to appropriate conditions. These are typically that the sum of the molecular energies is equal the total energy, and that the number of molecules is constant. That particular set is chosen to define the equilibrium distribution. Others that occur less often will not persist and will revert to the equilibrium one.

A convenient link to entropy is provided by seeking an analogy, between classical and statistical thermodynamics, in describing the approach to equilibrium of a gas sample. The resulting increase in classical entropy to an ultimate final value can be thought of as being matched – in statistical terms – by a path, through intermediate states, settling at the equilibrium state. It is the most probable by having the maximum number of ways \( \Omega \) of finding the molecules with the equilibrium set of parameter values across their ranges. [35]. The analogue of the state’s classical entropy is a statistical entropy defined as \( k \times \log \Omega \), where \( k \) is Boltzmann’s constant and \( \log \) denotes logarithm to the base e. It is a feature of statistical thermodynamics that \( \log \Omega \) acts as a valid, and mathematically very convenient, proxy for \( \Omega \). Throughout this period, there remained a gap between the developing statistical theory of molecular motions and the laws of thermodynamics. Professor J. Willard Gibbs (1839–1903), of Yale University, pointed the way to narrowing the gap with a paper in 1876 [36] dealing with the equilibrium of multicomponent systems including chemical effects explicitly. This led to the use of changes in entropy and energy in tracking chemical processes. A special thermodynamic function (to be called the Gibbs function \( G \)) was soon identified. Its constant value for reversible processes occurring at constant temperature and pressure was particularly useful in describing basic phase transitions. He led the subsequent systematic development of chemical thermodynamics jointly with the establishment of statistical mechanics as a new branch of theoretical physics.

### 7 Later developments: Keenan

In this section, we take a brief look at the work of Joseph Henry Keenan [37], an engineer who introduced entirely new ideas into the conceptual structure of thermodynamics. In particular, he endeavoured to construct a totally coherent account of the fundamental aspects of the subject, removing all the remaining ambiguities.

Keenan, who was born in 1900, was involved with every aspect of engineering, both in industry and in education. His first position, which he took up in 1922, was as a turbine design engineer for the General Electric Company in Schenectady. This was the start of a lifelong concern for the properties of steam and, of course, for thermodynamics. In 1928, he moved into academia with a position at the Stevens Institute of Technology. Here, he was a delegate to the First International Conference on the Properties of Steam in 1929; he was to maintain attendance at the following conferences on this area of engineering for the rest of his life. He also made a major contribution to the Steam Tables of 1930. In successive editions of 1936, 1939 and 1969, to which he also made
major contributions, these tables would be indispensable for generations of students and also for every practitioner in the steam-power industry for decades to come. In 1934, he moved to MIT, where he headed the Department of Mechanical Engineering for a number of years. Here, his work on thermodynamics became increasingly well known and he developed what is often called the Keenan School of thermodynamics or the MIT School of thermodynamics.

At MIT, he followed up his work on the Steam Tables by co-authoring the Air Tables and the Gas Tables, which were compulsory reading for those developing the gas-turbine industry. He also carried out research that was virtually always ground breaking. This included evaluation of the performance of steam-turbine nozzles, studies on the operation of the free-piston compressor in gas turbines, and the development of equipment to process coffee and cocoa, and for dust separation. He also worked on jet and rocket propulsion, carrying out research on the friction coefficients of air at supersonic speeds, and on injectors and on heat transfer at high speeds.

Keenan was equally devoted to teaching, where he was determined to achieve a simple approach to thermodynamics, which was nevertheless totally rigorous, and would eliminate those unnecessary and awkward features and ambiguities in the then current accounts that were a residue of historical struggles and misunderstandings. It was this desire in his teaching that led to his development of entirely new conceptual foundations of thermodynamics. In the 1930s, he adopted many of the views of J. Willard Gibbs, in particular, the use of the quantity of thermodynamic availability, sometimes called exergy. This is the maximum useful work that is possible during a process bringing a system into equilibrium with a heat reservoir. In the case that the reservoir is just the surroundings, the thermodynamic availability is the ability of the system to cause a change as it comes to equilibrium with its environment. It is effectively the energy that is available for use. The concept of thermodynamic availability is of great use in the steady-flow processes of engineering, being crucial in the allocation of fuel costs in processes that have many outputs in chemical and power-plant engineering.

His two books, *Thermodynamics* [38], published in 1941, and *Principles of General Thermodynamics* [39], published with George Hatsopoulos in 1965, have been hugely influential in the study of thermodynamics, and the second in particular has revolutionized our ideas of the subject. It shows that the second law of thermodynamics can be expressed in terms of the existence of stable equilibrium states. In fact, the authors claim that the first and second laws may be combined into a single law, which they call the law of stable equilibrium. Their statement is that: ‘A system having specified allowed states and an upper bound in volume can reach from any given state one and only one stable state and leave no effect on its environment’.

The law has been restated [40] as: ‘When an isolated system performs a process, after the removal of a series of internal constraints, it will always reach a unique state of equilibrium; this state of equilibrium is independent of the order in which the constraints are removed and is characterized by a maximum value of entropy’. This principle entails the Kelvin–Planck and Clausius statements of the second law but also allows the traditional definition of entropy to be extended to non-equilibrium situations. It describes the behaviour of systems that are some distance from equilibrium, and which, when certain constraints are removed, will move to a local equilibrium. It includes a formal statement of the second law that has all the attributes of the concept of entropy but does not use it in its definition. Thus, the statement implies a direction and an equilibrium end point for all processes [41].
It may also be noted that this restatement of thermodynamics has considerable potential for reaching out from its home of classical physics to the quantum realm, as well as coping with relativistic situations. The authors describe their intentions as follows: ‘Thermodynamics has been known as a science which is effective in getting directly to the heart of a problem with a minimum of interpretation. It was considered, however, to be limited in its application to non-relativistic non-fluctuating systems … The extensive developments in the book show that by relatively minor modifications of definitions and principles and by adoption of the Gibbs concept of an ensemble to describe a system, most of the limitations previously assumed cease to exist. Thermodynamics can now be applied to obtain quite easily familiar and unfamiliar relations from the special theory of relativity. It can also be applied to fluctuating systems to provide a firm new basis for what was previously a basically incompatible mechanical theory of heat. This application will lead quite naturally to other developments of which information theory is one. Thermodynamics appears to be a renascent science on the threshold of a new era’.

8 Myron Tribus’s thermodynamics: an alternative understanding

Myron Tribus [42], who was born in 1921, may be regarded as a polymath. In the second worldwar, he worked as a design-development officer for the United States Air Force, and after graduating with a PhD in 1949 he began work as a gas turbine design engineer at General Electric, but was unhappy in industry and moved back to education at the University of California at Los Angeles. In 1960, he moved to Dartmouth College as Dean of Engineering. Here, he developed an updated curriculum based on engineering design and entrepreneurship.

In 1969, he joined the Nixon administration as assistant secretary of Commerce for Science and Technology, but in the following year he became the senior vice principal for Research and Engineering in Xerox. Then, from 1974 to 1986 he was the director of the Centre for Advanced Engineering Study at MIT. He is a co-founder of Exergy, Inc., a company that specializes in designing high-efficiency power production systems. (As explained in the previous section, exergy is another word for thermodynamic availability.)

He has carried out research in a wide range of areas including heat transfer, fluid mechanics, probability theory, statistical inference, sea water demineralisation, aircraft heating and ice prevention on aircraft. He is a great advocate of Bayesian methods of probability and, in connection with his work on the economic considerations for converting sea water into fresh water, coined the term thermoeconomics.

As something of a guru, he has produced his famous Perversity Principle, which says that: ‘If you try to improve the performance of a system of people, machines, and procedures by setting numerical goals for the improvement of individual parts of the system, the system will defeat your efforts and you will pay a price where you least expect it’.

Like Keenan, Tribus regarded himself as a follower of J. Willard Gibbs, though in his case, the ideas of Gibbs were filtered through the work of Claude Shannon (1916–2001) and Edwin T. Jaynes (1922–1998). Shannon is regarded as the founder of information theory. His wartime experience in cryptography and communication theory led to his celebrated paper of
1948 [43] which initiated this topic. In this paper, he produced an expression for the amount of information contained in a sequence of numbers, and broadly it could be said that the more unlikely an event is, the higher the related information. Shannon was able to build up a rigorous mathematical account of the relationship between probability and information, perhaps better expressed as the relationship between uncertainty and information. In the course of this he coined the term information entropy, though the connection between this quantity and thermodynamic entropy was not immediately obvious.

The work of Jaynes [44] centred around the relationship between thermodynamics, statistical mechanics, probability and statistical inference, and in a paper of 1957 he introduced the so-called MaxEnt or maximum entropy interpretation of thermodynamics. This principle is based on inference techniques and is based on Shannon information theory, Bayesian probability and the principle of maximum entropy. While Jaynes’ work has been criticized by many, Tribus believed that it revolutionized the foundations of thermodynamics. In his famous book Thermodynamics and Statistical Mechanics [45], he wrote that the work of Gibbs was certainly correct but mathematically complicated.

In contrast, when he was directed to the work of Jaynes, he realized that it showed how, starting from the work of Shannon ‘one could obtain in an elementary way all the results given by Gibbs. It is difficult to overestimate the value to science education of this contribution. At one fell swoop the enormous conceptual difficulty of Gibbs’ approach has been swept away. It now becomes possible to begin by considering the properties of the smallest particles and by simple mathematical methods to deduce the properties of large systems. The results obtained by Joule, Mayer, Carnot, Maxwell, Boltzmann and Gibbs are seen in a new light. It is no longer necessary to cover the development of thermodynamics in the same way as the subject developed in history. New generations of students can, in shorter time, reach a state of understanding heretofore impossible for more than a small number of persons’.

9 Conclusions

The development of thermodynamics by the aforementioned formidable scientists was geared to the recognition and refinement of the First and Second Laws of Thermodynamics. These laws emerged from the empirical correlation of relevant observations and measurements via the postulation of equivalents (‘mechanical equivalent of heat’) to transformations. Once the overarching concept of energy was accepted, the quantitative content of the laws became clear. Tallying served the First Law well enough. However, some sophistication was needed for the Second Law. The interpretation of the energy tally had to be supported by an additional thermodynamic property, entropy, whose value always increased or, in the limit for reversible processes, remained constant. For various specialized processes other properties, such as enthalpy (H) and the Gibbs function (G), were identified. Mathematics provided links amongst the properties and its own increasing sophistication was essential. These supported the subject, theoretical and applied, as it expanded across the sciences, with a statistical component crucial to the successful incorporation of information theory, fortified by developments in computing generally. Thus, the last 50 years or so have completed a platform for current and future developments of which some will almost certainly continue the expansion into the realm of nanoscience.
References


