A History of Thermodynamics

Julian Barbour

Contents

1 Preface

This 'book' consists largely of text removed, on the very sensible recommendation of my editor TJ Kelleher, from a first draft of my *The Janus Point* published on 1st December 2020 by Basic Books in the United States and by Bodley Head on 3rd December in the United Kingdom. The history it contains would have taken up too much space in the book now published and tried the patience of readers wanting to get on to new ideas. However, colleagues urged me not to jettison the material and to seek to turn it into a separate book on the discovery of thermodynamics. In the preface to *The Janus Point* I announced my intention to do that and as an interim measure to put the material on my website, where you have now found it. In its present state, it is certainly not ready for publication, in part because it contains some material that is already included in *The Janus Point* and should be edited out and also because the present text needs the inclusion of more material.

For those readers who have or intend to read *The Janus Point*, Sadi Carnot's work is here, in the first chapter, much more fully covered, and the same is true of the next seven chapters. The chapter on Maxwell's demon and William Thomson on time contains material that, except for a small part, is not in the published book at all. The three chapters here with Boltzmann in their title have significant overlap with material in *The Janus Point*.

Julian Barbour, College Farm, South Newington.

Note added 16/08/2021. Since I posted this 'book' on my website in December 2020, Paul Sen's *Einstein's Fridge: The science of fire, ice and the universe* has been published. I recommend the book. It is very readable and contains much historical material that in part complements my own account. In particular Sen devotes much space to the outstanding work of J Willard Gibbs, who hardly appears below except for his critical comment about the need for confinement of a dynamical system if it is to be treated in thermodynamics and statistical mechanics. This issue, which I believe is critical if the system considered is the universe, is at the heart of my *The Janus Point* but is not considered by Sen. If I do manage to develop this text into a proper book, Gibbs must surely be given more prominence. I was recently able to acquire second-hand copies of his collected papers in two volumes originally published shortly after his death in 1903 and subsequently reprinted by Dover. Lucky the person who has them on their bookshelf.

I have also taken the opportunity to rejig the text, which I had inadvertently right aligned for the original posting. I, and you if you read this 'book', can, like me, be thankful to Kartik Tiwari, who is reading it and sent me an email yesterday asking "Why is the manuscript right-aligned? It seems like a very trivial and useless question but it was bugging me so much that I thought I should ask you."

2 Sadi Carnot and the Steam Engine

It can be argued that thermodynamics as a science began with a strange observation made by Joseph Black (1728-1799). A professor at Edinburgh who discovered magnesium and advised whisky distillers, he happened during his work to leave two buckets of water in a room. The first contained ice and water, the second only water. The water temperature in both buckets was at the freezing point. When he came back a few hours later, the first bucket contained less ice and correspondingly more water but, to Black's surprise, it was still at the freezing point. The water in the second bucket was noticeably warmer. Only when all the ice had melted did the water in the first bucket begin to get warmer.

Ambient heat in the room had obviously entered the contents of both buckets and brought about the changes. But why had it only melted some of the ice in the first bucket without heating the water as well? It seemed heat could disappear. Believing its amount should remain constant, Black suggested it had become *latent*, the Latin for hidden. He also showed that the addition of heat to boiling water did not raise its temperature, which also suggested it had become hidden. To this day, physics students learn about latent heat, but it now refers only to the phenomenon Black observed, not his interpretation of it. Latent heat is a measure of the amount of heat needed to bring about some transformation like the melting of ice.

Black's observation was a factor in the development by Antoine Lavoisier (1743-1794)[1] of the ice calorimeter, which he used to measure the heat released in chemical reactions. Together with Pierre-Simon Laplace (1749-1827) Lavoisier developed the theory of *caloric*. This was supposed to be an invisible, weighless and indestructible 'heat fluid' that could pass from one body to another, raising the latter's temperature and causing it to expand. Besides giving a simple explanation of Black's observations, caloric served as the foundation of Laplace's theory of the propagation of sound. Newton had argued this involved transfer of heat; Laplace's caloric-based theory assumed there was none and was a significant improvement on Newton's theory; it continued to make improvements to the theory of sound for a century.

I have recalled the theory of caloric for two reasons: first, it is a good example of a seemingly sound idea that, as we will see, remained unchallenged for a surprisingly long time. The turning points, when they do come, provide much of the fascination in the history of science. Second, caloric provides the background to this chapter on the work of Sadi Carnot.

There were, in fact, two Sadi Carnots, the one who first understood the essence of steam

[1]Known as 'the father of modern chemistry', Lavoisier is above all associated with its transformation, so decisive throughout the scientific revolution, from a qualitative to a quantitative science. He recognized and named oxygen and hydrogen and helped construct the metric system. There is a superb 1788 portrait of Lavoisier and his wife by the great painter Jacques-Louis David. The opulence of the setting gives a clue to Lavoisier's gruesome end. He was a tax-collecting administrator of the *Férme général*, which in Revolutionary France was a greatly hated part of the *Ancien Régime*, and from its profits he funded his scientific research and, no doubt, his private life. He was accused of selling adultered tobacco and guillotined. The appeal to spare his life so that he could continue his experiments is said to have been dismissed by the judge with the words "The Republic has no need of scientists (*savants*) or chemists." This may be apocryphal, but not the comment of the great mathematician Lagrange: "It took them only an instant to cut off this head, and one hundred years might not suffice to reproduce its like."

engines and the other his nephew, the president of France from 1887 until his assassination by an Italian anarchist 1894. The first Sadi's father was Lazare Carnot, a hugely significant historical figure on account of his role in revolutionary France. He appointed Napoleon to his first indepedent command and organized the fourteen French armies involved in the Napoleonic wars. He was the only one of Napoleon's generals to be undefeated. He was also a scientist of no mean ability. In 1803 he published a book on mechanical machines such as pulleys. This reveals an ability, inherited by his son, to see through details and grasp the essential. He saw clearly that the most important thing in machines is to maximize their efficiency. That seems obvious today but is in part due to Lazare and even more so his son.

Sadi, who was born in 1796 and was named for a Persian poet then in vogue, seems to have had a very attractive personality: honourable, sober, modest and an excellent violinist. There are some brief biographical details about him in the introduction by the editor, E. Mendoza, to the translation of Sadi's justly famous *Reflections on the Motive Power of Fire* published in 1824 (the year after Lazare died). For those who want to go into these things a bit more deeply, I strongly recommend the Dover publication of this work, which also includes translations of two papers, one in 1834 by Émile Clapeyron and the other in 1850 by Rudolf Clausius. These papers transformed Carnot's ideas into the two fundamental laws of thermodynamics and were extremely important, since Carnot's book went completely unnoticed by contemporary scientists and was only saved from oblivion by Clapeyron, who usefully sharpened the mathematical formulation of Carnot's insights and made the later work of Clausius and others possible.

As was so common until the modern age, Sadi's life was cut short by disease. He never lived to see the triumph of his work. As Mendoza reports, in June 1832, during anti-government riots, he saw "a drunken officer galloping down the street brandishing his sabre and knocking people down; Sadi dodged under the man's arm, toppled him off his horse and threw him in the gutter". Shortly afterwards, Sadi caught scarlet fever which turned into brain fever. His younger brother Hippolyte, the father of the later president of France, took him to the country to convalesce. There, when reading reports of the cholera epidemic then raging, he actually caught the disease. He died from it in a few hours. He was 36.

Coming now to his book, it is very much a product of its age. Many early scientists were inspired by pure curiosity—they wanted to understand the world and how it worked. Often, as with Newton, there was a strong religious component, the desire to discover God's design of the world and seek for its purpose. In contrast, the Revolution in France brought rivalry with other nations, above all England, to the fore. Whereas England's largely self-taught engineers had, as much by luck as judgement, made very important technical advances, above all through the invention of the steam engine, the French revolutionaries aimed at systematic advance of their country's interests through the foundation in 1794 of the École Polytechnique. Its many brilliant teachers and alumni did much to create the modern ethos and practice of science. Some of this clearly rubbed off on Sadi, as we see from this early passage in his book:

> If the honor of a discovery belongs to the nation in which it has acquired its growth and all its developments, this honor cannot be here refused to England. Savery, Newcomen, Smeaton, the famous Watt, Woolf, Trevithick, and some other English engineers, are the veritable creators of the steam-engine . . . Notwithstanding the work of all kinds

> done by steam-engines, notwithstanding the satisfactory condition to which they have been brought today, their theory is very little understood, and the attempts to understand them are still directed almost by chance.

Although he was clearly much too modest to make the claim for himself, it was he, the Frenchman, who changed that 'little understood' into 'much understood'. As so often with such matters, the long term consequences of first steps are hard to foresee. I doubt if Carnot had any inkling of the magnitude of the issue that would emerge from his work: why is the past so very different from the future? Why are we born only to die?

One of the interests in the history of thermodynamics is to see how the early practical issues in the development of the steam-engine still affect the way scientists seek to grasp the workings of nature and even the very universe. That will be a theme that runs through the next few chapters. We can start with a further comment of Carnot: "The most signal service that the steam-engine has rendered to England is undoubtedly the revival of the working of the coal mines, which had declined, and threatened to cease entirely, in consequence of the continually increasing difficulty of drainage, and of raising the coal." Carnot's formalized measure of the work that a steam-engine can perform, to be discussed shortly, can be traced to 'raising the coal'. He certainly leaves one in no doubt about its importance:

> To take away from England her steam-engines would be to take away at the same time her coal and iron. It would be to dry up all her sources of wealth, to ruin all on which her prosperity depends, in short to annihilate that colossal power. The destruction of her navy, which she considers her strongest defence, would perhaps be less fatal.

Let's now start with key aspects of Carnot's work. The rather quaint 'fire' in the title is a hangover from the time, not so distant from Carnot's, in which fire was still considered to be one of the four fundamental elements along with earth, air and water. Mendoza notes that, as late as 1783, Montgolfier said his hot-air balloon ascended because it was filled with fire, the lightest of the elements. For 'fire' in Carnot's title one should therefore understand *heat*, but heat itself is not Carnot's concern. His concern is what the title says, the motive power (*puissance motrice*) of heat. However, within little more than two decades, Carnot's work on that had become a major factor in concentrating intense interest on the nature of heat itself. Science was again driven more by pure curiosity than useful application. For his part, Carnot believed (at least when he wrote his book) that heat was caloric and hence indestructible.

Carnot is worth quoting above all for the breadth of his vision and novel approach. He highlights the vast motive power of heat, how it is responsible for

> so many movements which take place on the earth. It causes the agitations of the atmosphere, the ascension of clouds, the fall of rain ... the currents of water which channel the surface of the globe ... Even earthquakes and volcanic eruptions are the result of heat.

From this "immense reservoir" with which Nature has provided us "we may draw the moving force necessary for our purposes ...". The object of heat-engines is to "develop this power, to appropriate it to our purposes". He lists the great advantages of steam-engines compared with animal power, water falls, and air currents. It has the advantage of economy over the first and, over the other two, "the inestimable advantage that that it can be used at all times and places without interruption". What is more, "it causes rapid extension in the arts in which it is applied, and can even create entirely new arts".

What really distinguished Carnot was, first, his ability to make an all-encompassing survey of the relevant field and, second, identify the critical issues. Both can be seen in these two paragraphs:

> The production of motion by heat has not been considered from a sufficiently general point of view. We have considered it only in machines the nature and mode of action of which have not allowed us to take in the whole extent of application of which it is susceptible ... It becomes difficult to recognize its principles and study its laws.
>
> In order to consider in the most general way the principle of the production of motion by heat, it must be considered independently of any mechanism or any particular agent. It is necessary to establish principles applicable not only to steam engines but to all imaginable heat-engines, whatever the working substance and whatever the method by which it is operated.

Because it will become a main theme of the book, let me point out that Carnot's call, immensely positive for the development of thermodynamics, for "a sufficiently general point of view", was effective precisely because it was simultaneously circumscribed—to "the production of motion by heat". He describes in graphic detail the motive power of heat, but nowhere asks after the origin of heat itself or, more importanly as we will see, differences of temperature. That was totally justified for the purposes of his project; any attempt to go further in his day would have been hopelessly premature. However, it may not be premature in ours, and we should be careful not to transfer unthinkingly concepts developed to study steam in a box to an unconfined universe.

Immediately following the two quoted paragraphs, Carnot outlines his approach. Because he can draw on the already well developed theory of mechanical phenomena, he writes with confidence:

> Machines which do not receive their motion from heat, those which have for motor the force of men or of animals, a waterfall, an air current, etc., can be studied even to their smallest details by the mechanical theory.[2] All cases are foreseen, all imaginable movements are referred to these general principles, firmly established, and applicable under all circumstances. This is the character of a complete theory. A similar theory is evidently needed for heat-engines. We shall have it only when the laws of physics shall

[2] Because of Carnot's reference to mechanical theory, I will mention here things I call *gifts of nature*. Good examples are wood or metal, from which it is easy to make rulers to measure distances or goniometers to measure angles. Then there is the rotating earth, which gives us sidereal time and defines the (astronomical) day by passage of a chosen star across the meridian. All the observations, stretching over about three millenia, that enabled Kepler to discover the laws of planetary motion were of angles between celestial bodies at instants of sidereal time. Goniometers were ubiquitous, but the rotating earth alone had the accuracy and above all stability for the task. The terrestrial science to which Galileo gave such a boost was and still is unthinkable without wood, metal and the rotating earth. Their very existence plays a critical role in the story of time and the universe.

> be extended enough, generalized enough, to make known beforehand all the effects of heat acting in a determined manner on any body.

Commencing this ambitious undertaking, the creation of a complete theory, Carnot reminds his readers of the basic working of steam-engines. In the process he points out a universal feature of them that seems, up to his time, to have escaped the notice of theoreticians. It was that, to function at all, steam-engines must always transfer some heat from a hotter to a colder medium.

Anyone looking at the very first commercial steam-engine, Newcomen's of 1712 developed to pump water out of mines, would have this fact before their eyes. In it, the furnace heats cold water to generate steam. On the opening of a valve, the steam enters a cylinder that contains a piston and is open at the top to the atmosphere, which therefore presents no effective resistance as the piston is pushed up in its forward stroke. In the top position the steam inlet valve is closed and another valve opened to let in cold water that sprays and cools the steam, condensing it and thereby greatly reducing its volume. The external air pressure now forces the piston back down in its return stroke. Through mechanical linkage, a weight can be lifted at the same time. Work is done. In summary, heat from the furnace has generated steam which has done work after which heat has been transferred into the cold injected water.

At the start of some 4000 words that laid the foundations of thermodynamics, Carnot, having sketched the typical working of steam-engines such as Newcomen's, drew the conclusion that the production of motive power in steam-engines is due "not to actual consumption of caloric, but *to its transportation from a warm body to a cold body*." Carnot's italicized words bid fair to be the most fruitful false statement in the history of science. It's so easy to believe in something that is indestructible. And he made his words seem so plausible by an analogy:

> The motive power of a waterfall depends on its height and on the quantity of the liquid; the motive power of heat depends also on the quantity of caloric used, and on what we will call the *height of its fall*, that is to say, the difference of temperature of the bodies between which the exchange of the caloric is made. In the waterfall the motive power is exactly proportional to the difference of level between the higher and lower reservoirs. In the fall of caloric the motive power undoubtedly increases with the difference of temperature between the warm and the cold bodies; but we do not know whether it is proportional to this difference. We do not know, for example, whether the fall of caloric from 100 to 50 degrees furnishes more or less motive power than the fall of the same caloric from 50 to zero. It is a question which we propose to examine hereafter.

In his book, Carnot made only partial progress (for the case of an infinitesimal temperature difference) toward the answer to this profoundly important question, the beautiful answer to which will come later in the book. Before then, we will see how the waterfall analogy misled not only Carnot but also another great scientist.

Another passage I must quote is this:

> the production of heat alone is not sufficient to give birth to impelling power: it is necessary that there should also be cold; without it, the heat would be useless. And

> in fact, if we should find around us only bodies as hot as our furnaces, how can we condense steam?

Emphasizing the generality of his vision, he says

> Wherever there exists a difference of temperature, it is possible to have also the production of impelling power. All substances in nature can be employed for this purpose, all are susceptible of changes of volume, of successive contractions and dilatations, through the alternation of heat and cold. All are capable of overcoming in their changes of volume certain resistances, and of thus developing impelling power.

From this general standpoint, it's a small step to the next question: "Is the motive power of heat invariable in quantity, or does it vary with the agent employed to realize it as the intermediary substance?" Of course, such a question is meaningful only if one can quantify both heat and motive power. This was the least of Carnot's problems. Calorimetry was already a well-developed science. Carnot mentions the amount of heat needed to melt a kilogram of ice. As for motive power, that should be measured by the work it can do. The most common unit for that was lifting of a one-kilogram weight through one meter against the force of gravity at sea level. This obviously matches the basic task in mines: raising coal to ground level. Note that completely different units and phenomena are here involved. What has heating water to do with lifting coal? We'll see in the following chapters how Carnot's slim publication helped greatly to find the answer to this question.

Another observation of Carnot was this: difference of temperature can produce motive power but also motive power can generate difference of temperature. Is not "the friction of bodies actually means of raising their temperature, of making it reach spontaneously a higher degree than that of the surrounding bodies"? Also the temperature of gases is lowered by rarefaction and raised by compression.

Carnot uses this observation to introduce the idea which ensured for him a place in history. Some two decades after his untimely death, it became known as *the Carnot cycle*. The masterstroke was to conceive the design of an idealized heat-engine in which all losses apart from the inevitable transfer of some heat to the medium playing the role of the condenser are eliminated. The hardest part of this project was conceptual: how can one know that some particular machine is ideal, how can one tell that no heat has been lost except that which must unavoidably be lost if mechanical work is to be done?

In achieving this, one condition above all is absolutely critical. In formulating it, Carnot noted that if heat is allowed to pass from a hot to a colder medium without at the same time doing any useful work the potential motive power in the hot body will simply be squandered. It will have done nothing but heat the colder medium. Thus: "The necessary condition of the maximum is *that in the bodies employed to realize the motive power of heat there should not occur any change of temperature which may not be due to a change of volume."*

Of course, there must be some temperature difference if any motive power is to be generated. What Carnot realized was that, by careful arrangement, work can still be done however small the temperature difference—it can be made infinitesimally small and still allow the working medium to expand and do work. Although never fully realizable in actual practice, this conceptual picture enables one to determine what will be the maximal efficiency achievable under the most favourable conditions imaginable. Carnot introduced this very characteristic way of thinking into the theory of steam engines and beyond that into the nascent science of thermodynamics. What he thereby demonstrated was both surprising and illuminating. Let us follow him step by step to see how the optimal conditions can be achieved.

Carnot's gives two descriptions of the cycle, one with steam as the working medium, the other with ideal gases. I will describe the latter which has a double advantage: simplicity and the role, to be described in the following chapters, that ideal gases played in the discovery and interpretion of entropy.

First, a few words about ideal (or perfect) gases. Basically, they are gases that, when cooled, become, first, liquid and then solid well below room temperatures. The temperatures of these *phase transitions*, from gas to fluid and from fluid to solid, also depend on the pressure they are under. On account of the simplicity and universality of the laws that govern their behaviour, ideal gases played a critical role in the discovery of thermodynamics. When a given weight of such a gas is confined to a vessel of volume V it will settle down into a stable state—it will *equilibrate*. The resulting equilibrium state is completely determined by either its pressure P or its temperature T. In such a state, the temperature, which can be measured by a thermometer, is constant throughout the confined volume of gas, as is the pressure. The three quantities P, V, T are called *state functions*. If, for any given ideal gas, any two of the state functions are known, then the third can be found through the relation that connects them. It is called the *equation of state*. Equilibration plays a central role in thermodynamics and much of the discussion in this book will be about it.

To describe the Carnot cycle, I need to tell you about only one of the universal properties of ideal gases. In 1662, Robert Boyle found that if the temperature T of a confined ideal gas is kept constant, its pressure P and volume V are inversely proportional to each other.[3] This is expressed by the equation $PV = C$, where C is a constant that depends on the temperature, the gas and the amount there is of it. Typical curves relating P to V for the same gas at two different tempertures are the sections ab and cd in Fig. 1. They are *isotherms*: curves for the same (iso) temperature.

Figure 1 is in fact Clapeyron's graphical representation of the Carnot cycle. At the start of the cycle, a, ideal gas in a cylinder with a piston which has vacuum above it is at a temperature just below the temperature T_{high} of a furnace with which it can be brought into thermal contact. There can be lots of little lead weights piled on top of the piston that are just sufficient in number to equal the upward pressure of the gas. If a little heat is now allowed to flow into the cylinder, the gas will expand a bit, raising the weights. One of them can be taken off and placed secure at a slightly greater height than it had. The gas pressure will have decreased just a little and be in balance with the reduced weight it must balance. This process can be repeated many times. In principle, the lead weights can be made arbitrarily light and the amount of heat admitted at each step correspondingly infinitesimal. Theoretically, the most perfect heat engine corresponds to the mathematical limit in which there is no difference of temperature between the furnace and gas. It is also very important that the ideal process is perfectly reversible. If a lead weight is put back onto the piston, the gas will be compressed and get a bit hotter than the furnace due to the increased weight. The excess heat will flow back into the furnace.

[3]On the Continent, Boyle's law is often named for Edme Mariotte, who published it in 1676.

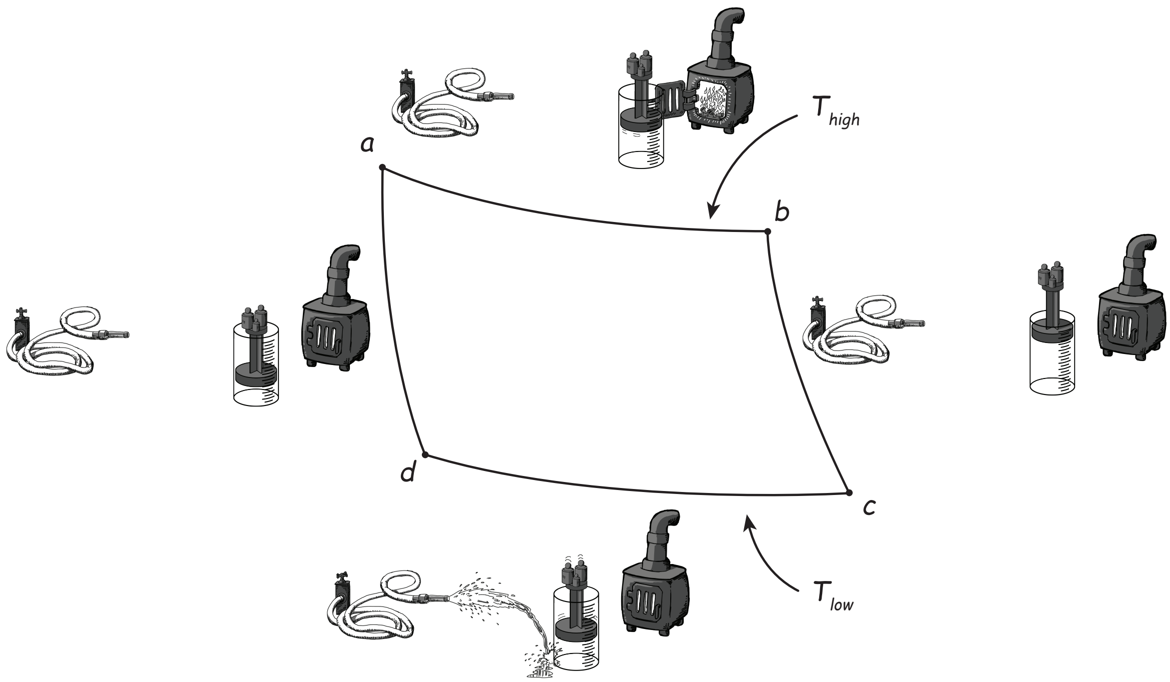


Figure 1: *The Carnot cycle.* The bounding curves in the sequence are Clapeyron's diagram. Flavio Mercati's vignettes indicate what the furnace, piston, and coolant do in each stage of the cycle. Only the piston is in action throughout the cycle; the furnace and coolant each act only once. The pressure of the gas increases vertically, the volume of the gas from left to right. While a brilliant idea, the Carnot cycle is completely impractical: a single cycle takes an infinite amount of time. See the main text for the full description.

The part of the Carnot cycle just described is the *isothermal* (same-temperature) stage. Through it a certain amount of work is done. More can be achieved without drawing any more heat from the furnace. One simply allows the gas to expand bit by bit by removing one of the little weights after another. Long after Carnot's death this part of the cycle, from b to c, came to be called the *adiabatic*[4] section. In it, the pressure falls more rapidly than in the isothermal stage, as is reflected in the steeper slope of the curve. At this stage, a certain amount of work has been 'banked'—lots of little weights have been raised through various heights and some are still sitting on top of the piston. The temperature of the gas in the cylinder is lower that it was on the isotherm ab.

Now comes a critical point. To be any use, a steam engine must run continuously. To achieve that, the working medium in the cylinder must be brought back to its orginal state at the start of the cycle. In Carnot's idealized cycle, this means from the point c to its original state at a. This could be done by simply reversing the process by putting back the weights to return the gas along the route cba. But this would exactly use up all the work that had been gained. It is clear that some work must be done to get back to a, but this need not be as much as was gained.

The procedure is as follows. At c, thermal contact is established with a heat reservoir, the refrigerator, whose temperature T_{low} is just below that of the gas and correspondingly significantly lower than that of the furnace. The remaining weights are sufficient, bit by bit, to compress the gas in the cylinder and change the state of the gas isothermally along cd. The important thing is that less work is needed to compress the colder gas under the conditions on cd than the hotter gas on ba. A glance at the figure shows that the pressure on the route back is at every stage lower than on the route out. This is the essential reason why a positive amount of work can be done during the cycle.

The final stage of the return cycle is again adiabatic: the gas is compressed along da and brought back to its original state. The steam engine can continue its labour.

Clapeyron's figure has two great virtues: first, it's a wonderfully transparent representation of the process. Second, the area enclosed by the closed cycle $abcda$ measures the net amount of work that has been done. A particularly nice feature of the cycle is that it can be run through in the reverse order: $adcba$. When this is done, work is expended to take heat from the refrigerator and transfer it to the furnace. This demonstrates the perfect reversibility of the Carnot cycle. Moreover, it is clear from the figure that the amount of work done depends on the two temperatures T_{high} and T_{low}. Neither Carnot nor Clapeyron could determine its quantity because, in contrast to the isotherms ab and bc, they did not know the form of the adiabats bc and da. As regards the temperatures for which heat engines can be operated, the properties of the furnace and the working medium fix T_{high}, while T_{low} is effectively that of water at the ambient air temperature (typically around $20°C$). These two temperatures fix the optimal cycle that in practice is possible.

I should emphasise that that my description in terms of heat flow corresponds to modern understanding. Both Carnot and Clausius thought in terms of the weighless, invisible and indestructible caloric. They supposed that in the section ab caloric flowed into the gas, causing it to expand. In the section bc the amount of added caloric, being indestructible, did not change. In the section cd all the caloric gained in ab passed, in their view, into the

[4] A coining in the late 19th century from Greek: *a-* (not) + *dia* (through) + *batos* (passable)

refrigerator causing its medium to expand but not get hotter.

This was a false conclusion but ironic; if Carnot had made an experiment to check for a temperature rise he wouldn't have found one. This is because a remarkably small amount of heat corresponds to a lot of work. It was about 40 years from Carnot's publication before improvement in experimental accuracy permitted a temperature rise in the refrigerator to be measured.

Fortunately, the incorrect notion of caloric did virtually nothing to reduce the value of Carnot's study. His observation that his idealized heat engine could be run in either direction was most important. Carnot said that in the forward, work-performing cycle caloric is transformed from the furnace to the refrigerator; in the reverse cycle the caloric is transferred back from the refrigerator to the furnace. We recall his words that motive power is due "not to actual consumption of caloric, but *to its transportation from a warm body to a cold body.*" The analogy with a waterfall, in which work really is done without any consumption of water, made this conclusion very plausible.

Carnot drew two key conclusions from his theory of idealized heat engines. Both were correct despite the false belief in caloric. In modern terms, the first was that not all of the heat suplied in a heat engine can be used to do useful work. Some inevitably finishes up in the cold reservoir (the condenser in a steam-engine). The correct deduction of this result when the notion of caloric was abandoned did not lead to any significant new idea. Quite different was what happened as a result of the correction of Carnot's second conclusion. This was that, no matter what working medium one used, all heat engines would give exactly the same maximal efficiency. This is Carnot's argument:

> Now if there existed any means of using heat preferable to those which we have employed, that is, if it were possible by any method whatever to make the caloric produce a quantity of motive power greater than we have made it produce by our first series of operations, it would suffice to divert a portion of this power in order by the method just indicated to make the caloric of the body B to return to the body A from the refrigerator to the furnace, to restore the initial conditions, and thus to be ready to commence again an operation precisely similar to the former, and so on: this would be not only perpetual motion, but an unlimited creation of motive power without consumption of caloric or any other agent whatsoever. Such a creation is entirely contrary to ideas now accepted, to the laws of mechanics and of sound physics. It is inadmissible. We should therefore conclude that *the maximum of motive power resulting from the employment of steam is also the maximum of motive power realizable by any means whatever.*

The search for perpetual-motion machines had long been a dream of mankind, but by Carnot's time repeated debunking of charlatans' claims to have created them had persuaded scientists their non-existence was a deep principle of nature. Given his belief in indestructible caloric, Carnot's argument is impeccable. Once the notion of caloric was abandoned, a subtle move that modified his argument but retained his conclusion led to the second law of thermodynamics. The revised argument is impregnable—only the creation of a perpetual-motion machine could break down its defences. I think this is a main reason why the claim, first made by Rudolf Clausius and to which we will come, that the entropy of the universe tends to a maximum is so widely accepted. Since I question the application of standard thermodynamic arguments to the universe (though, to be clear, not to systems like steam engines), to conclude this chapter I will highlight those aspects of Carnot's work that, brilliant in their context, nevertheless may be less helpful outside it.

For a start there's the remarkable robustness of Carnot's principles. In fact he created almost the entire conceptual framework of thermodynamics. It's as secure now as when he published his booklet in 1824. One way to see that, for the reader who acquires Carnot's booklet, is to follow it up with *Thermodynamics* by the Nobel Laureate Enrico Fermi (1901-1954) given as a lecture series in 1936.[5] The opening chapter defines a *thermodynamic system*. This core concept comes straight from Carnot's book: in the typical case it's a confined ideal gas whose pressure P, volume V and temperature T can be changed from without in infinitesimal *reversible* transformations through a succession of equilibrium states. These can take the system from an initial state A to a final state B but also, in a *cycle*, back from B to A. The work done in such a case is the area enclosed, as in Fig. 1, within the plot of P against V. It's pure Carnot. It should however be noted that thermodynamic systems are extreme idealizations. They don't exist in the universe.

Perhaps the greatest tribute to his work was something that Einstein, near the end of his life, said of thermodynamics: "It is the only physical theory of universal content which I am convinced that, within the framework of applicability of its basic concepts, will never be overthrown." Most of those basic concepts are Carnot's. However, note Einstein's caveat 'within the framework of applicability'. That will be critical later, but some things can be said now.

Besides the point about confinement in a box already made, it's worth noting that the interaction of thermodynamic systems with their environment is very 'one-sided'. By and large, things are done to them. They are not entirely passive but they cannot 'do their own thing'. They are reactive. The key concept of entropy can be defined for thermodynamic systems in large part because they can be confined and controlled. Who can do that for the universe?

I also find it relevant that a thermodynamic system is the acme of idealization. With the possible exception of black holes, whose behaviour (discussed in *The Janus Point*) follows the background arrow but does not seem to determine it, nothing remotely like thermodynamic systems exist in natural form in the universe. They can be realized in a good approximation for a long time in a laboratory, but in the pure form they exist only on paper and in the brains of theoreticians. And in those brains, as extensive reading of the professional and popular-science literature on time's arrows demonstrates, they do keep much thinking 'inside the box'.

I noted earlier that Carnot's "sufficiently general point of view" was effective precisely because it was simultaneously circumscribed to "the production of motion by heat". For very good reasons, Carnot made no attempt—it would have been hopeless if he had—to ask how it can be that hot and cold, the *sine qua non* for motive power, exist around us simultaneously. He would have faced a never ending series of questions relating, among other things, to the origin of the earth and its coal-bearing seams, the benificent sun that stimulates the growth of trees and the food that provides the fireman with the energy to shovel coal into the firebox, and so on. In short, it would have forced him to understand

[5]Both books are available online in pdf format.

the universe and its history. In fact, he did mention the universe just once, in a footnote in which he says a hypothetical perpetual machine would be

> capable of creating motive power in unlimited quantity, capable of starting from rest all the bodies of nature if they should be found in that condition, of overcoming their inertia; capable, finally, of finding in itself the forces necessary to move the whole universe, to prolong, to accelerate incessently its motion.

This calls up the quaint image of a machine 'pushing the whole universe' into ever faster motion. The picture I present in the Janus point is quite different—that of a universe, freed from a box, in which its parts move *relative to each other*. That doesn't make any of the parts into perpetual-motion machines, but it does open up interesting possibilities.

Related to this is another key aspect of thermodynamics inherited from the working of steam engines: the need to keep on bringing the working medium back to its initial state. This is the origin of the cycles that Fermi mentions. As we will see, they were crucial for the discovery of entropy in confined systems. But this raises a question if we want to define an entropy of the universe. For we certainly cannot bring it back to any earlier state. It seems determined to go its own way in eternal expansion. This may cast some doubt on the idea that the uinverse, at any instant, has a definite, and increasing, entropy.

Another thing is this. Carnot is renowned for observing that heat engines never operate at 100% efficiency. But why? It is their continuous operation. The working medium must, again and again, be brought back to its initial state, point a in Fig. 1. But in a single use one could stop at b. Then the heat would have done nothing but work. The wastage only comes through the recycling.

Carnot's viewpoint is antropocentric. His vision is all-encompassing, but his acuity is directed to the needs of mankind. His aim is to make the steam engine maximally efficient. He did not seek to understand the origin of coal but to maximize its utility. Too concerned with human needs, he may have missed the most important thing.

What is that, divorced from our mundane concerns, steam engines actually do? They lift coal from the depths of mines to the surface of the earth. In doing this, they change the shape of the universe.

3 The Mechanical Equivalent of Heat

The man known to science as Count Rumford was the American Benjamin Thompson (1753-1814). He was talented but had little success until he married a rich heiress who had inherited a property in Rumford (now Concord) in New Hampshire. Because he sided with the British in the Revolutionary War, he had to flee, abandoning his wife. He moved to London and became a British citizen knighted for his administrative talents. In 1784 he moved to Bavaria, where he was made the Army Minister and, in 1791, a Count of the Holy Roman Empire. In Bavaria, he did many admirable things, including the invention of Rumford's Soup for the poor and the introduction of potato cultivation. In 1789, he created, on behalf of the ruler of Bavaria, the magnificent Englischer Garten in Munich, one of the largest urban parks in the world. I personally, like millions more, am very grateful to him for that. In the time when I lived and studied in Munich from 1961 to 1966, I walked or bicycled through the park almost every day. It was in Munich that Rumford made the experiments which should have spelled the demise of caloric and spared Carnot his one mistake but somehow made little or no impact.

In 1798, Rumford published his paper "An Experimental Enquiry Concerning the Source of the Heat which is Excited by Friction" in the august *Philosophical Transactions of the Royal Society of London*. A German translation appeared at the same time. The paper gives a vivid impression of the man and is a beautiful example of a chance observation that, followed up, can lead to profound insights. He was overseeing work in the arsenal in Munich [6] and was

> amazed at the considerable heat that quickly appeared in a brass cannon during drilling and by the even greater heat of the splinters created by the drilling, which I found to be much greater than that of boiling water. The more I thought about these phenomena, the more remarkable and interesting they appeared to me. A thorough investigation of them even seemed to promise a deep insight into the hidden nature of heat and to enable us make a sensible conjecture about the existence or nonexistence of a fiery fluid, about which the opinions of natural philosophers have at all times been so divided.

Rumford was clearly a first rate experimentalist. He used two horses to turn a brass six-pounder against a blunt borer to create maximum friction and much heat. Bystanders were astonished "to see how, without fire, such a quantity of cold water could be heated and even brought to boiling." When drilling recommenced after a pause, just as much heat could be generated. Being manifestly inexhaustible, the heat excited by the friction "could not possibly be a material substance". There was also no sign of any chemical transformation; that too was ruled out as the heat source. Rumford quantified the effect of friction: the work of one horse during two and a half hours was sufficient to raise through 180° Fahrenheit 26.58 pounds of water. One pound heated by one degree was therefore "equivalent to 940 British units of work".

The great question was therefore the one that "has so often occupied the natural philosophers: What is heat? Is there a fiery fluid? Does there exist something that can actually be

[6] The site and its significance is marked by a plaque not far from the Chinese Tower in the Englischer Garten.

called heat substance?" Having shown that the heat excited by the friction is inexhaustible and could not possibly be a material substance, Rumford concluded: "It must therefore be *motion*." He made no conjectures about its precise nature, as heat "is a subject that scientists and philosophers have for so many thousands of years been vainly attempting to comprehend." But he did argue for the methods of science. Even if one cannot penetrate to the innermost depths of nature's workings, quantitative description of what one can observe guided by intuition of what might lie behind appearances can still take one very far.

How was it that, despite Rumford's experiments, the notion of caloric survived for half a century? Rumford had ample opportunity to make his discovery known. From 1799 to his death in 1814 he divided his time between Paris[7] and London, where he surely took part in many scientific discussions. Indeed, he and Sir Joseph Banks established the Royal Institution of Great Britain in 1799 with Sir Humphry Davy as the first lecturer.[8] In 1799, Davy reported an experiment in which he had melted two pieces of ice by rubbing them together, from which he concluded "The phenomena of repulsion are not dependent on a peculiar elastic fluid for their existence, or caloric does not exist." Thus, interaction with Rumford may have stimulated Davy to his experiment.

The historian of science Stephen Brush, whose books have been a considerable help to me, suggests that two factors may explain caloric's longevity. First, the study of heat was only one of many immensely exciting research topics in the first half of the 19th century, above all those connected with chemistry and also electricity and magnetism. In 1820, Oersted proved that an electric current could deflect a magnetic needle in its vicinity and in 1831 Faraday discovered electromagnetic induction. These discoveries pointed to a deep unity in the phenomena of nature and, as a second explanation of the survival of caloric, encouraged theories of underlying fluid-type mechanisms. Discoveries in the behaviour of light and radiant heat suggested that they could be transverse (perpendicular to the direction of wave propagation) vibrations in a fluid ether. By analogy, vibrations of caloric might still explain heat.

Given the already widespread recognition that science advances in large part through observation and measurement, it is still surprising that only in 1842 did the German Julius Mayer (1814-1878) became the first person after Rumford to relate mechanical work to heat quantitatively. He too was prompted by serendipity. As ship's physician on a Dutch three-master sailing to Jakarta in 1840, he noted that storm-whipped waves are warmer than calm sea. He started to think for the first time about physical laws, but his ideas were inchoate and largely expressed in terms redolent of medieval philosophy. However, he got help from professionals, including a sceptical professor at his alma mater Tübingen, who told him to shake water violently to prove that did actually increase its temperature. Mayer did do that and reported a positive effect, but another paper which, as the first after Rumford's, established a value for what came to be known as *the mechanical equivalent of heat*, was

[7] One of the many strange facts about Rumford's life was his second marriage, in 1804, to the widow of the guillotined Antoine Lavoisier.

[8] The institution flourished and became world renowned as a result of Davy's pioneering research. His assistant Michael Faraday greatly strengthened the Institution as a premier research laboratory. The tradition of his famous public lectures popularizing science continues to the present; the Royal Institution Christmas lectures attract large audiences through their TV broadcasts.

not based on actual measurements. It relied instead on an indirect argument related to the difference between the amounts of heat taken up by gases under conditions of constant volume or constant pressure.

Mayer's paper did not attract much early interest and I won't discuss it further because it was of much less quality and influence than results obtained by James Joule (1818-1889). He lived in Manchester, which at that time was the greatest industrial city in the world, attracting no less a person than Friedrich Engels to write there his famous 1845 book *The Condition of the Working Class in England.* Joule was the son of a wealthy brewer and as an adult his 'day job' was managing the brewery. Physics was a passionate hobby. He had been tutored as a boy by the great scientist John Dalton and, with his brother, developed a fascination for electricity. He and his brother used to give each other—and servants—electric shocks. He investigated the possibility that steam engines in the brewery could be replaced by electric motors, whose recent invention had followed Faraday's discovery of electromagnetic induction. Moreover, he was able to draw on the superb engineering skills available in Manchester for the construction of the exquisitely accurate apparatus used in his many important experiments.

Having the advantage of this expertise and his knowledge, both theoretical and practical, Joule was able to go far beyond Mayer in performing important experiments involving generation of heat not only by friction but also by electricity and magnetism. He performed direct experiments and published their results beginning with one in 1843. Joule's scientific papers were brief, the epitome of clarity and entirely free of Mayer's wordy philosophy. In his 1843 paper, he said "I shall lose no time in repeating and extending these experiments, being satisfied that the grand agents of nature are indestructible and that wherever mechanical force is expended, an exact equivalent of heat is always obtained."

Joule's most famous paper, of 1845, is a wonderful example of what can come of the simplest idea. He made the experiment that Mayer did not. He measured not only the rise in temperature but also the amount of 'violent shaking' that caused it. His apparatus consisted of a brass paddle-wheel working horizontally in a can of water. Motion could be communicated to this paddle by means of weights on pulleys. The paddle moved with great resistance in the water, so that the weights (each of four pounds) descended slowly at about one foot per second. The height of the pulleys from the ground was twelve yards. Joule rewound the pulley 16 times. His measurement of the resulting temperature increase led him to the conclusion that

> for each degree of heat evolved by the friction of water, a mechanical power equal to that which can raise a weight of 890 lbs to the height of one foot, had been expended. Any of your readers who are so fortunate as to reside amid the romantic scenery of Wales or Scotland, could, I doubt not, confirm my experiments by trying the temperature of the water at the top and at the bottom of a cascade. If my views be correct, a fall of 817 feet will of course generate one degree of heat; and the temperature of the river Niagara will be raised about one-fifth of a degree by its fall of 160 feet.[9]

Discussing other experiments made with air, Joule said that they "are inexplicable if heat be a substance but are, however, such as might have been deduced *à priori* from any

[9] In 1847, Joule went to the Niagara Falls with his wife on their honeymoon and made the appropriate measurement.

theory in which heat is regarded as a state of motion among the constituent particles of bodies." That the experiments of Rumford and Davy had not gone completely unnoticed is clear from Joule's further comment that his experiments "afford a new, and to my mind, powerful argument in favour of the dynamical theory of heat which originated with Bacon, Newton, and Boyle, and has been at a later period so well supported by the experiments of Rumford, Davy, and Forbes."

Joule concluded that "an enormous quantity of *vis viva* exists in matter" (*vis viva* was the expression then used for the energy associated with motion). In fact, Joule identified something that greatly affects our everyday lives, including the cost of heating our homes. What feels to us a very modest increase in warmth results from a great amount of mechanical energy (a fifth of a degree for the fall at Niagara). It is also disappointingly difficult to work off calories, which measure energy. The consolation is that a gallon of fuel in the automobile tank will take us a good long way. The mechanical equivalent of heat is the reason why Carnot could not have disproved the caloric theory by direct measurement with the accuracy he could have employed.

Not surprisingly, after a delay of five years, Joule's clearly written papers attracted far more attention than Mayer's verbose papers and are the reason why Joule, after whom a unit of energy is named, gained nearly all of the early credit for establishing the equivalence of heat and mechanical energy. The unfortunate Mayer was distraught when he discovered Joule had gained the credit. This and the death of two of his children in 1848 caused Mayer's mental health to deteriorate rapidly and in 1850 he attempted suicide. He was committed to a mental institution, from which he emerged a broken man. He achievement was at least recognized while he was still alive.

Joule, for his part, was in no doubt as to the significance of his work. He gave a popular lecture that appeared in the *Manchester Courier* in May of 1847. It brims with the confidence of a man who knows his experiments are helping to reveal the inner workings of nature and God's control of them. He argues that living force (*vis viva*) "is one of the most important qualities with which matter can be endowed, and, as such, that it would be absurd to suppose it can be destroyed". This is because "it is manifestly absurd to suppose that the powers with which God has endowed matter can be destroyed any more than that they can be created by man's agency".

Inviting his audience to behold "the wonderful arrangements of creation" Joule says that "we find a vast variety of phenomena connected with the conversion of living force and heat into one another, which speak in language which cannot be misunderstood of the wisdom and benificence of the Great Architect of nature." It is something we see "in our own animal frames, 'fearfully and wonderfully made'." Indeed,

> the phenomena of nature, whether mechanical, chemical, or vital, consist almost entirely in a continual conversion of attraction through space, living force, and heat into one another. Thus it is that order is maintained in the universe—nothing is deranged, nothing ever lost, but the entire machinery, complicated as it is, works smoothly and harmoniously. And though, as in the awful vision of Ezekiel, "wheel may be in the middle of wheel," and everything may appear complicated and in the apparent confusion of an almost endless variety of causes, effects, conversions, and arrangements, yet is the most perfect regularity preserved—the whole being governed by the sovereign

will of God.

Coming back down to earth, it's a nice thought that caloric came into science through an adviser to whisky distillers and went out through the efforts of a brewer.

4 The Conservation of Energy

Galileo initiated much in science. His vision was remarkable. In 1623 he wrote

> Philosophy is written in this immense book that stands ever open before our eyes (I speak of the Universe), but it cannot be read if one does not first learn the language and recognize the characters in which it is written. It is written in mathematical language, and the characters are triangles, circles and other geometrical figures, without the means of which it is humanly impossible to understand a word; without these philosophy is a confused wandering in a dark labyrinth.

.

His final book, smuggled out of Italy and printed at Leiden in 1638, contained something he had found many years earlier by rolling balls down gently inclined planes and using the amount of water flowing out of a tank to measure time: if the unit of distance traversed in the first unit of time is 1, in the next it will be 3, in the next 5, and so on. Galileo called this the *odd-numbers rule*. It was the first 'sentence' he had read in the immense book.

Galileo writes with rare confidence. Shortly after his statement of the odd-numbers rule he says

> It has been observed that missiles and projectiles describe a curved path of some sort; however no one has pointed out the fact that this path is a parabola. But this and other facts, not few in number or less worth knowing, I have succeeded in proving; and what I consider more important, there have been opened up to this vast and most excellent science, of which my work is merely the beginning, ways and means by which other minds more acute than mine will explore its remote corners.

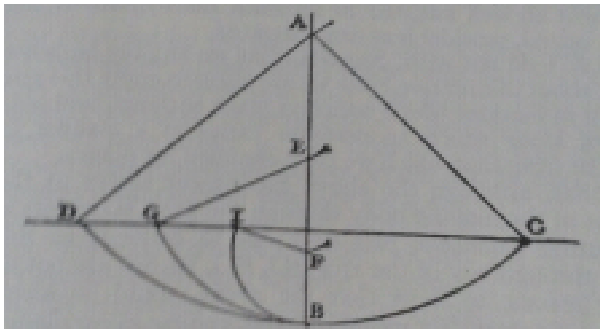
Some prophecy. Not a day passes now without another corner being explored. And surprises follow surprises, big and small.

Because free fall—of Newton's apple from the tree to the ground—would have been over too soon for Galileo to time, he invoked for it an indirect argument. It involved a pendulum and the argument you will find in the caption of Fig. 2. For my immediate purposes I don't need to give the steps from it to Galileo's law of free fall, which in its turn was a vital for Newton's great discoveries published in his *Mathematical Principles of Natural Philosophy* in 1687.[10] All I need is the pendulum argument, to which I will return a little later.

The key point is that, wherever a pin E or F is placed to thwart the full swing, the speed at B is always the same and the subsequent height of ascent after the passage through B is always the same. The same speed at B always allows the bob to climb to the same height. Having made this point, Galileo says there's no need to "trouble ourselves too much" about the matter of the pendulum. As we will see, it's an ironic comment. A little over two centuries later the phenomenon Galileo described was at the heart of physics and helped to usher in the principle that energy is conserved.

I might as well be anachronistic and anticipate the form the principle takes for the pendulum when expressed in modern terms. Suppose the bob has mass m, the thread to

[10]For a detailed account, see my *The Discovery of Dynamics.*



Without air resistance, the bob released at C would swing exactly up to D at the same height. A pin stuck in the backing bord at E or F would change the swing to G or I, both at the same height on CD. Imagining a mathematical ideal, here without friction, is a defining characteristic of the new science Galileo introduced.

Figure 2: Galileo's pendulum.

which it is attached is massless and the force of gravity is g. Suppose that at any instant the speed of the bob is v. Then its *kinetic energy*, T, is $mv^2/2$. Without the factor $1/2$, T is the *vis viva*, or living force,[11] referred to by Joule. In the gravitational field of the earth, the bob also has *potential energy*, V, which is equal to mgh, where h is the height above some nominal level and g the force of the earth's gravity. The fact that h is measured from some nominal value does not affect the statement of energy conservation, as I will explain in a moment. For simplicity let us first measure the height of the bob from point B in Fig. 2. Then in accordance with the principle of energy conservation, the *total energy*, E, is equal to

$$E = T + V = \frac{1}{2}mv^2 + mgh \tag{1}$$

and has the constant value E at all points of the motion. At B, the lowest point of the bob, $mgh = 0$ and all the energy is kinetic. At the highest point, anywhere on the line CD, the bob has come to rest and there is no kinetic energy. All the energy has become potential. In fact, I have not yet said anything of physical significance. If I had defined the kinetic energy T in any way such that $T = 0$ when $v = 0$ and the potential energy equally arbitrarily but so that it vanishes when $h = 0$, i.e., at B, the same statement would hold. The principle of energy conservation has real meaning because for all actual motions in which friction can be ignored the relation holds at all instants of the motion. In the swinging pendulum, there's a constant to and fro of energy passing back and forth between its two forms. If you change the reference height by adding the constant c to h then the right hand side of (1) becomes

$$\frac{1}{2}mv^2 + mg(h + c) = \frac{1}{2}mv^2 + mgh + mgc,$$

and since mgc is constant the constancy of the first two terms on the right is unaffected.

Suspecting that the secret of the conservation law (1) might be found within his pendulum argument was clearly beyond Galileo's ken and perhaps his ability to prove it. For over a century his diagram remained, like the suspension pin at A and the whole diagram, the tip of an iceberg below which much lay hidden. The first person to recognize its significance was the Dutchman Christiaan Huygens.

[11] *Vis viva* is an expression that Leibniz coined to reflect his belief in a universal vital force.

However, before I come to him I need describe another insight of Galileo that Huygens put together with the pendulum observation to draw an almost magical conclusion. The insight concerns Galileo's defence of the Copernican revolution.

The main thing about that was the motion around the sun that Copernicus attributed to the earth. That opened the door onto a new world. However, a secondary motion, seemingly less important, was forced upon Copernicus as a kind of corollary: in order to explain the observed diurnal sweep of the stars over the sky, the earth must spin on its axis, rotating once in 24 hours. Despite its secondary role, this postulated further motion also had momentous consequences. For about a century after Copernicus had published his book in 1543, many (if not most) educated people mocked his ideas as ridiculous. People said that things on a spinning earth just could not keep up with the colossal speed of the earth's surface as it revolved eastward. Church steeples would fall over backwards to the west, gales from east to west would howl past us, life would be impossible. The earth simply could not rotate at that speed.

One argument put forward, allegedly based on actual observation, was this. Suppose a ship at anchor in harbour. If somebody climbs the main-mast carrying a cannon ball and, from near the top, drops it, then, as is well known, it will hit the deck exactly next to the foot of the mast. Now suppose the same experiment is made when the ship is sailing smoothly at sea. In this case, it was argued, the cannon ball would not land at the foot of the mast but at some distance toward the stern, the distance being greater, the greater the ship's speed. Everybody believed this.

In 1632, Galileo published his *Dialogue Concerning the Two Chief World Systems*, which, as is well known, brought down upon him the wrath of the Inquisition. In his book, Galileo stated categorically that this claimed experimental fact was simply false. He said that, without even performing the experiment, he knew for certain that if performed when the ship is either under sail (if moving smoothly) or anchored in harbour the cannon ball would, in both cases, land at the foot of the mast. The first recorded test of Galileo's experiment was carried out in the harbour in Marseille in 1640, eight years after the *Dialogue* had been published. Galileo was totally vindicated. People also realized the experiment could be tested on horseback. Horsemen would ride at high speed holding a cannon ball in their hand and let it fall. Until it hit the ground, the rider could see the ball fall vertically below the hand from which it had been released. Experiences like that created a stir.[12]

[12] I cannot resist recalling two experiences I had about 350 years after Galileo had made his prediction. They brought home the difficulty his contemporaries had in accepting its truth. In the first I was on an express train in Lombardy heading at high speed from Milan to Venice and had cause to visit the toilet. You lifted the seat and found a verticle tube with diameter of about 10 cm open to the track, which thundered past a metre below its end. Men cannot resist aiming; my urine fell in a perfectly perpendicular stream until it reached the opening and was torn away in the direction to the back of the train; the stability of the fall seemed impossible given the speed of the train. People still find such experiences hard to accept. A few years later I flew to New York and with a random group of people was taken from JFK in a commercial minibus to New Haven to visit Lee Smolin, who was then at Yale. Hearing I was a theoretical physicist, the man next to me, with whom I had got into discussion, said there must be something wrong with the way the universe worked. On the flight, he had been to the restroom and while waiting for the current occupant had done a little test by making a small jump."There's got to be something wrong. That 747 was doing close on 600 miles in an hour. By rights, when I made that jump, I should have been shot clean out the back of the plane."

To persuade people that the earth could be rotating at high speed without church steeples falling down westwards, Galileo gave a famous argument that encompassed not only falling cannon balls but all natural processes. He identified a universal phenomenon. I give the relevant passage from the *Dialogue* in full; it's a gem:

> Shut yourself up with some friend in the main cabin below decks on some large ship and have with you there some flies, butterflies, and some other small flying animals. Have a large bowl of water with some fish in it; hang up a bottle that empties drop by drop into a wide vessel below it. With the ship standing still, observe carefully how the little animals fly with equal speed to all sides of the cabin. The fish swim indifferently in all directions; the drops fall into the vessel beneath; and, in throwing something to your friend, you need throw it no more strongly in one direction than another, the distances being equal; jumping with your feet together, you pass equal spaces in every direction. When you have observed all these things carefully (though there is no doubt that when the ship is standing still everything must happen in this way), have the ship proceed with any speed you like, so long as the motion is uniform and not fluctuating this way and that. You will discover not the least change in all the effects named, nor could you tell from any of them whether the ship was moving or standing still. In jumping, you will pass on the floor the same spaces as before, nor will you make longer jumps toward the stern than toward the prow even though the ship is moving quite rapidly, despite the fact that during the time you are in the air the floor under you will be going in a direction opposite to your jump. In throwing something to your companion, you will need no more force to get it to him whether he is in the direction of the bow or the stern, with you situated opposite. The droplets will fall as before into the vessel beneath without dropping toward the stern, although while the drops are in the air the ship runs many spans. The fish in the water will swim toward the front of the bowl with no more effort than toward the back, and will go with equal ease to bait placed anywhere around the edges of the bowl. Finally, the butterflies and flies will continue their flights indifferently toward every side, not will it ever happen that they are concentrated toward the stern, as if tired out from keeping up with the course of the ship, from which they will have been separated during long intervals by keeping themselves in the air. And if smoke is made by burning some incense, it will be seen going up in the form of a little cloud, remaining still and moving no more toward one side than the other. The cause of all these correspondences of effects is the fact that the ship's motion is common to all the things contained in it, and to the air in it.

Two and a half centuries after Galileo's *Dialogue* had been published, his cabin cameo led to the notion of an *inertial frame of reference*. It reflects the fact that observers in the cabin and on the quayside in Marseille would see processes unfolding in the cabin governed by identical laws. However, this would not be so for seagulls circling the ship. Motions that appeared unaccelerated in the cabin would appear accelerated to them. Most textbooks on dynamics define an inertial frame of reference as one in which Newton's laws hold.[13] If they hold in one such frame, they will hold in any frame moving relative to it with uniform velocity. This is the famous relativity principle, called Galilean relativity on account of the

[13] In fact, this 'dodges' a foundational issue in dynamics: what determines the inertial frames of reference? I will return to this at the end of this chapter; see also Chapter 8 of *The Janus Point*.

above passage and is supposed to apply to purely mechanical phenomena whereas Einstein extended the principle to include light and all electromagnetic phenomena in the famous paper in 1905 that led him to the many remarkable predictions concerning the behaviour of clocks and measuring rods for which he is so famous, including the equation $E = mc^2$.[14] I won't discuss Einstein's predictions here because my immediate concern is with a much earlier use of the relativity principle that bears directly on the mechanical interpretation of the laws of thermodynamics that followed soon after their formulation.

But first we must see what happened when Christiaan Huygens (1629-95), the greatest scientist between Galileo and Newton, contemplated the significance of Galileo's pendulum observation in conjunction with the relativity principle.

One of Huygens's many achievements was the discovery with his brother of Saturn's rings. Much more significantly, he played a crucial role in the development of the pendulum clock in 1656. This marked an important advance in the accuracy of time keeping. It's a curious fact—an example of a prophet not recognized in his own country—that Christiaan is hardly known in his native Holland. He is eclipsed by his father Constantijn, who was an important diplomat and admired poet. Constantijn also influenced Dutch architecture by building, in classicist style, his moated summer home Hofwijck near Den Haag. Christiaan inherited the house, which is now a museum. The Dutch visit it mainly on account of Constantijn but, as I learned from the curator,[15] most foreigners come because of their interest in his son.

In this chapter, I'm going to describe Huygens' discovery of the laws of elastic collision that underlie billiards and snooker. This beautiful work made critical use of Galileo's relativity principle and was done at about the same time as the work on the pendulum clock. It predated by 250 years the conclusions that Einstein, using the same principle, made so famous through his creation of the theory of special relativity. In a way, Huygens' insights were just as important, especially in the theoretical interpretation of the first and second laws of thermodynamics. This is why I discuss them now. However, Huygens' discoveries, though they greatly impressed his contempories, don't catch the modern imagination since they did not overturn beliefs about the nature of space and time taken for granted from time immemorial. You have to go back to Joule's article in the *Manchester Courier* and the decade 1840 to 1850, in which the principle of conservation of energy was discovered, to sense the impression Huygens's work made.

The background to it that I'm going to describe are claims that Descartes (1596-1650) made. At the age of 14, the precocious Huygens met the famous philosopher, who was

[14] I say 'supposed' because Einstein's extension is already implicit in Galileo's famous text. It's clear that he is thinking in universal terms and not just in effects manifested mechanically. Indeed, some flint would need to be struck to light the incense that is then "seen going up in the form of a little cloud". Moreover, Galileo did make a sophisticated but unsuccessful attempt to measure the speed of light. We now know that the movements of all animals rely critically on electrical processes in muscles. Galileo would surely, if pressed, have said his relativity principle must include all physical processes yet to be discovered. Otherwise his defence of the Copernican revolution would fail.

[15] Readers may like to view the film *Killing Time*, beautifully made by Dutch TV in the autumn of 1999. It's about the ideas I had just published in *The End of Time*. The film opens with a shot of me walking along the bank of the moat at Hofwijck; my arguments about the nature of time are presented in the hall of the house. I had the honour of being able to put forward objections to Newton's notions of absolute space and time much like the ones Huygens may well have formulated while reading Newton's *The Mathematical Principles of Natural Philosophy* in that very same room.

keeping safely out of reach of the Inquisition in Holland. Huygens greatly admired Descartes for the introduction of the mechanical theory of the universe. According to it, there is nothing in the universe but pieces of matter moving about in all possible ways in infinite space. In his influential *Principles of Philosophy* published in 1644, Descartes claimed to know the laws that would govern collisions between such pieces of matter.

The theory captured the imagination of Huygens, but he realized Descartes' laws were wrong in part and set out to find the correct ones. The simplest laws govern what happens when two identical perfectly elastic billiard balls collide. Much more subtle are the ones that govern the collision of balls that are not identical, having different masses.[16] These laws are important for this book since, first, they were the basis of the mechanical theory of heat and the attempts to understand the true nature of entropy.

Huygens began his considerations, which were almost entirely theoretical, by asking what one should expect to happen if two identical balls collided head on with equal but opposite speeds. The symmetry of the situation strongly suggests the only possible answer: the balls would spring back from each other with the magnitude of their pre-collision velocities unchanged but directions reversed. Although billiards had been played for about two centuries, perfectly symmetric collisions do not generally occur in the game, so perhaps Huygens' correct assumption had passed unnoticed. However, most people would surely have assented to it if asked. It shares with Euclid's axioms the appearance of self-evident truth. Anyway, the symmetrical outcome, the simplest example of a time-reversal symmetric law, was the hypothesis Huygens made.

He then asked this question: what happens if two identical balls meet head on with unequal speeds? Here he could not invoke symmetry arguments. Instead, he used Galileo's cabin argument, transposed from a sailing ship in the Mediterranean to a boat on a river in Holland. Galileo's verbal skill was so great he had no need of pictures to make his points, which were in fact about everyday actions familiar to children. Huygens was considering precise and rather abstract matters and clearly felt the need of pictures to make his points. His *De Motu Corporum* (On the Motion of Bodies) published posthumously gives the derivation of results that he had simply stated in 1673 in his influential masterpiece, the book *Horologium Oscillatorum* on pendulum clocks.

In his woodcut, which I reproduce from the book in Fig. 3, we see a man in the boat holding two suspended balls. As the boat moves at uniform speed from right to left past the man standing on the bank, the boatman brings the balls together with equal but opposite speeds, $-u$ and u, relative to his frame of reference, the boat. Huygens wishes to make it clear that one should not think the collision takes place *in the boat*. It is not tied to the boat, it could be regarded just as well as taking place *on the bank*. He makes this point as follows. With outstretched arms, the man on the boat is initially holding the balls suspended at rest. Exactly at the moment he comes opposite the man on the bank, the two link hands and hold the strings jointly. The man on the boat moves his hands so that relative to it the balls collide with equal and opposite speeds. For the man on the bank, who moves his hands with

[16] The notion of mass only emerged with the work of Newton. Like his contemporaries, Huygens thought in terms of the size of the bodies or also their weight. In one of his great insights, Newton realized that mass and weight are not the same thing but that in any body they have a definite ratio determined by the local strength of gravity. This fact later came to play a key role in Einstein's theory of gravitation.

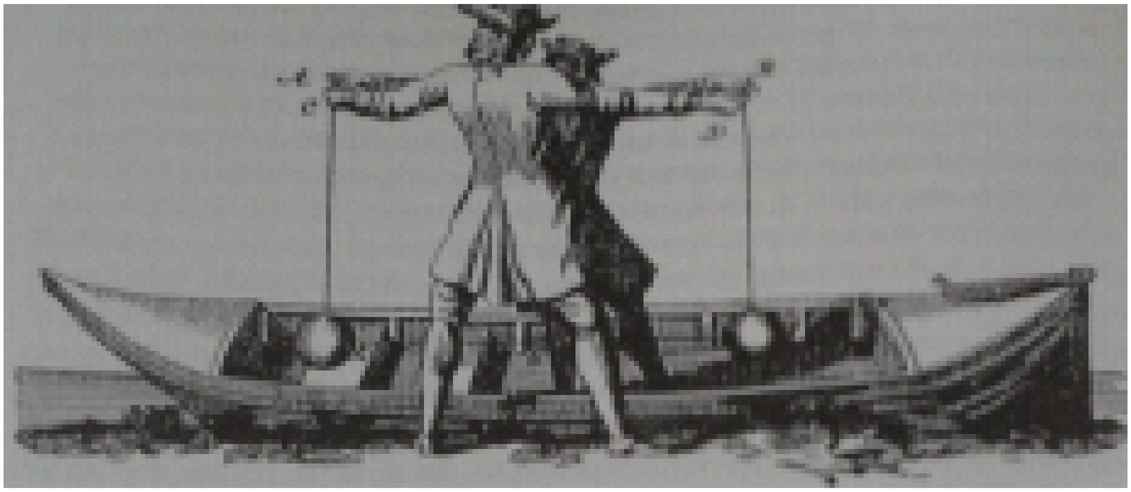


Figure 3: Huygens' woodcut with two men linking hands as the boat moves from right to left. The balls cannot 'know' which man is causing them to collide.

those of the man on the boat, the balls collide with unequal speeds.

Now the balls clearly cannot 'know' which of the two men is causing them to collide. As far as they are concerned, a single event is taking place: they are colliding head-on with a certain relative velocity. The single event is being watched by two different observers—one on the boat, the other on the bank.

Moreover, by Huygens' original hypothesis, we know what would happen if the man on the bank makes the balls collide with equal and opposite speeds. The balls must bounce back with equal but opposite speeds relative to the bank. By Galilean relativity, that must also happen if the balls are brought together with equal but opposite speeds relative to the uniformly moving boat.

This argument leads to a prediction. Suppose the boat is moving to the left in the figure with speed v relative to the bank. Then, seen from the bank, the unequal speeds before the collision are

$$-u + v \quad \text{and} \quad u + v. \tag{2}$$

By the relativity principle, we know that relative to the boat the balls spring back with equal and opposite speeds. Therefore, seen from the bank the speeds after the collison are

$$u + v \quad \text{and} \quad -u + v. \tag{3}$$

Huygens has been able to deduce, under the assumption that Galileo's relativity principle holds, what will happen in a collision with unequal speeds. Moreover, since the speed of the boat is arbitrary, Huygens can use an assumption about one single collision together with the relativity principle to predict what will happen in infinitely many: all those in which the boat's speed is varied from 0 to ∞. In fact, from (2) and (3) he can predict a two-fold infinity of solutions, since the relative speed $2v$ with which the balls approach each other can also range from 0 to ∞.

Huygens then attacks what looks like a much tougher problem. The original assumption of a symmetrical bounce was so plausible because the balls were assumed to be identical, to have the same mass. But what will happen if the balls have unequal masses?

Before we continue, let me remind you of the concept of the centre of mass (or centre of gravity), which is very familiar from the operation of a lever. If two weights of masses a and b are placed either side of the fulcrum at distances x and y from it, then they will balance if $ax = by$. Thus, if a is much less than b, there will still be balance if x is much greater than y. This, of course, is how a small child can balance a parent on a see-saw and how a great weight can be lifted with little effort provided the lever arm is adequate. The point at which the condition $ax = by$ holds is the centre of mass. It is defined with or without a fulcrum. It can also be defined for any number of point masses distributed in space. At the centre of mass, there is 'balance' about all three spatial directions when one adds up the contributions of all the masses.

Let me now describe how Huygens solved the problem of unequal-mass collisions. This is where my suggestion that the point of suspension of Galileo's pendulum (Figure 2, p. 000) is the tip of an iceberg starts to make sense. Huygens notes that the speed which the pendulum bob has at its lowest point is always sufficient, provided air resistance and friction can be ignored, to carry the bob up to the same height whatever the curve through which it swings. Moreover, it cannot go any higher.

Huygens then supposes that two balls of unequal masses m_1 and m_2 are let fall from heights h_1 and h_2, deflected into the horizontal with the acquired speeds $u_1 = \sqrt{2gh_1}$ and $u_2 = \sqrt{2gh_2}$ (g is the acceleration due to gravity), allowed to collide and then to ascend inclined planes to heights $\bar{h}_1$ and $\bar{h}_2$, where they are held. The new heights will not be the same as the initial heights since the speeds will have been changed by the collision.

Huygens then makes the assumption that the centre of gravity of the two balls *is not higher than it was at the start of process*. For if it were, one could make a perpetual motion machine: by lowering the weights to their original heights, one could raise other weights by a certain amount. Repeating the collision process from the new position and again coming to a state in which the post-collison centre of gravity was higher, one could do further lifting work and so *ad infinitum*. Huygens resolved another mechanical problem by the same kind of argument in *Horologium Oscillatorum* and with it caused a considerable stir. People began to realise that the ancient dream of a perpetual motion machine could not be realized. One cannot get something for nothing. Huygens's deductions, together with all the failed attempts to build perpetual-motion machines, were the background to Carnot's invocation of their non-existence in laying the foundation of his theory of steam engines. We have yet to see, in the next chapter, the modification of Carnot's argument that led to the second law of thermodynamics.

In fact, Huygens did anticipate the first law, which says that energy is conserved, by making a more precise assumption suggested by Galileo's observation of what happens with the pendulum. Its bob ascends to the height from which it had fallen. Huygens, making the kind of idealization that is so characteristic of great theoretical advances, assumed that, in the absence of friction and air resistance, the centre of gravity would reascend to *exactly* its original height. Unlike the behaviour of the bob, which a child can see and grasp without difficulty, this extension is far reaching. It involves observation of *two* balls and calculations that few adults would care to undertake. Although the extension is guided by a clear idea, one is dealing with something much less obvious. It's worth mentioning that we have here a classic example of the power that the simplest possible insights have in science. One starts with the bare minimum: a single body and a transparent process. From it, one hypothesizes by extension a law that governs two, three, four ... indeed any finite number of bodies. It is this organizing and controlling principle that Joule called "the powers with which God has endowed matter" so that "order is maintained in the universe—nothing is deranged, nothing ever lost".

It's worth mentioning here that energy is often thought, even by some scientists, to be some kind of indestructible substance not totally unlike caloric. Nothing could be further from the truth. The principle of energy conservation is much more like a book keeping rule. We saw that in the to and fro of Galileo's pendulum. Energy is transferred back and forth between the kinetic and work accounts. It's rather like currrency exchange. You can still buy the same number of oranges with euros or dollars. In fact, the hopefully modest exchange charges you incur are rather like the effect of friction. Just as that generates heat, the charges swell the profits of banks. As Joule said, "nothing is ever lost".

Although they are relatively simple, I will spare you the calculations, but I have still to tell you what were the laws of collision Huygens was able to deduce. They govern collisions with arbitrary pre-collision speeds u_1, u_2 and are

$$m_1u_1^2 + m_2u_2^2 = m_1v_1^2 + m_2v_2^2 \tag{4}$$

and

$$m_1u_1 + m_2u_2 = m_1v_1 + m_2v_2, \tag{5}$$

where v_1, v_2 are the post-collision speeds. In the modern terms I already introduced in equation (1) when discussing Galileo's pendulum, the quantities in equation (4) quadratic in the speeds, for example $m_1u_1^2$, are, with the factor 1/2 added,[17] the kinetic energies of the particles. In Huygens time and long after, as I mentioned in footnote 11, one used for them Leibniz's expression *vis viva* (living force), which reflected his belief (long discarded by the great majority of scientists) that the whole universe is animate. Since no potential energy is present in (4), it is special form of energy conservation that holds when only (frictionless) collisions are involved. The quantities linear in the speeds in (5) are the momenta of the particles and the equation expresses the law of conservation of momentum.

It is truly remarkable that Huygens, using clear principles and such simple arguments, deduced not one but two fundamental laws. The laws of conservation of energy and momentum are among the most important in science and will play a central role later in the book.

They will do that in a form in which Huygens, in a further important result, identified what one might call the intrinsic form of collisions. The basis for this was his proof that the centre of mass of the two bodies moves uniformly throughout the motion: before, during and after the collision. By the relativity principle, this means one can describe the collision process in a frame of reference in which the centre of mass is at rest throughout. The process then takes a particularly simple form: the two particles are always situated relative

[17] The 1/2 has no physical significance. It was added for mathematical convenience soon after the laws of thermodynamics were discovered so that the derivative of the kinetic energy with respect to v the same as the magnitude of the momentum mv.

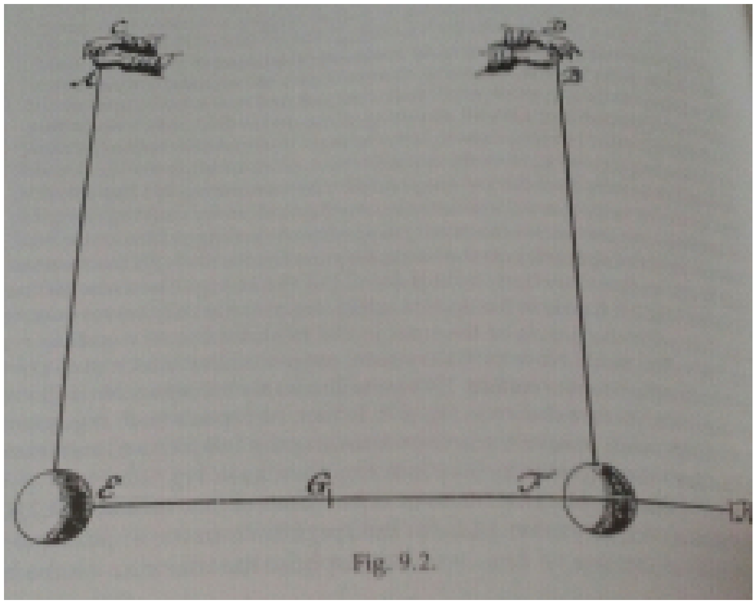


Figure 4: Huygens' disemboded hands.

to the centre of mass in the lever-balance position—at distances from the centre of mass in inverse proportion to their masses. Their speeds must also be in inverse proportion. This wonderfully simple picture of collisions remains right at the heart of modern physics. It is the way collisions between elementary particles in the Large Hadron Collider in Geneva are described.

The centre-of-mass description pares away everything that is not intrinsic to the collision. As far as it is concerned, all information about whether the balls collide on a boat in Holland or in a sailing ship in the harbour of Marseille is redundant. By a nice twist, other diagrams in Huygens' book on collisions illustrate this rather well but also the issue of the origin of inertial frames that I mentioned in footnote 13. Huygens felt obliged to repeat several times the point that the colliding balls could not possibly 'know' whether the man on the boat or on the bank was causing them to collide. To save on the considerable expense of complete fresh woodcuts (in a nice anticipation of modern data-compression techniques), his subsequent diagrams showed just the interlocking hands of the two men and the suspended balls (Figure 4). His first diagram was already a vignette—it showed nothing of the Dutch landscape. The revised diagram vignettes much more drastically. The two balls are essential, but all that otherwise remains are hands and strings.

But what role, if any, does the rest of the universe play in the collision? About this Huygens says not a word. The earth clearly is involved. The river runs through Holland and the balls must be suspended since otherwise they would fall to the ground. But these are clearly incidentals.

We get a better idea of Huygens' standpoint from the opening proposition of *De Motu Corporum*: "When once a body has been set in motion, it will, if nothing opposes it, continue that motion with the same speed in a straight line." In a less precise form, this principle had already been stated by Descartes. It became Newton's first law of motion. But there's a catch. How can we be sure anything is moving in a straight line with constant speed? This is what Huygens says: “The motion of bodies and their speeds, uniform or nonuniform, must be understood as relative to other bodies that are regarded as being at rest even if these together with the others partake in a further common motion.”

But this is the start of an infinite regress. How are we to confirm that the further common motion is uniform? We must invoke further ‘other bodies’ and so *ad infinitum.* Can we ever stop? I’ll come back to this issue later in the book and propose a terminus, or rather a kind of closing of the circle. I think this is important, being convinced that the explanation of the arrows of time that we observe locally must be sought in the overall behaviour of the universe. Even today, despite the development of sophisticated models of the universe, much discussion of time’s arrows is restricted to what are called local freely falling frames, which are the form taken by inertial frames in Einstein’s general theory of relativity.

I’ll round off this chapter by noting that the principle of energy conservation for purely mechanical processes, including the distinction between potential and kinetic energy, became properly established in the 1840s, especially through an influential paper in 1847 by Hermann von Helmholtz. I have already given the expression for kinetic energy. It will help to give now the expression for potential energy in the case of Newton’s law of universal gravitation. It’s going to be important in the second half of the book. If N bodies of mass $m_i, i = 1, \ldots, N$, that gravitate among themselves can each be regarded as mass points with distances r_{ij} between each pair of them, their Newtonian gravitational potential energy is the negative of the quantity obtained by adding togething all terms of the form

$$\frac{m_i m_j}{r_{12}} \tag{6}$$

for different unequal values of i and j. There are $N(N-1)/2$ such terms and the total Newtonian potential energy is expressed compactly in the form

$$V_{\text{New}} = -\sum_{i<j} \frac{m_i m_j}{r_{ij}}. \tag{7}$$

In Joule’s ecstatic article in the Manchester Courier, you may have wondered what he meant by ‘attraction through space’. In fact, he was intuitively anticipating forces that can be derived from an expression like (6) by differentiation with respect to the inter-particle separation r_{12}. In the case of gravity, this gives Newton’s famous force inversely proportional to the square of the distance between the particles though Joule had in mind analogous electric or magnetic forces.

It may be helpful to note that in the 1840s the modern distinction between force and energy had not yet been established through clear definition of the terms force, kinetic energy and potential energy. We are about to meet the man whom, before any others, we have to thank for the clarification.

5 The First Law of Thermodynamics

In the deeply religious Britain of the mid 19th century, not yet disturbed by Darwin's book of 1859, Joule's vision, formalized in 1850 as the first law of thermodynamics, made a huge and comforting impact. Two men were mainly responsible for this part of the story.

The first was the Scots–Irish William Thomson (1824-1907), who was born in Northern Ireland. He was a remarkably prolific mathematical physicist and engineer, publishing substantial papers on a wide variety of topics at the rate of about one a fortnight from the early 1840s. Both he and his elder brother James, who became a distinguished engineer, were coached and strongly supported by their father, who moved as a professor to Glasgow when William was eight and took his family to Paris to learn French and a year later to Germany. On this trip they were forbidden to read anything but German. However, the 15-year old William had smuggled into his luggage and read in secret the renowned *Théorie Analytique de la Chaleur* of Joseph Fourier (1768-1830) published in 1822.

William's first scientific paper, published anonymously in his first year as a student in Cambridge, defended, against criticism of it by a professor in Edinburgh, Fourier's hugely important method of representing almost any function by the trigonometric sine and cosine functions. The professor had the grace to admit his error. William's deep understanding of Fourier's heat-transfer equation already persuaded him by 1844 that the world must have come into existence through a creation event. This was at least a partial stimulus to a most important paper that he published in 1852. William's father sent him again to Paris when he was 18 to work in the laboratory of the great experimentalist Dragault. While there he had discussions on almost equal terms with leading French mathematicians, including Joseph Liouville (1809-1882). A theorem that Liouville proved will play a central role once we get a bit further in the story.

William Thomson was an important telegraph engineer and inventor, which brought him fame, wealth and honour. Queen Victoria knighted him in 1866 on account of his work on the transatlantic telegraph project; in 1892, in recognition of his achievements in thermodynamics (and opposition to Irish Home Rule!), he became the first British scientist to be enobled, taking the name Baron Kelvin of Largs, Kelvin from the river in Glasgow near his laboratory, Largs from the small town near Glasgow where he built a handsome rsidence.[18] The absolute scale of temperature is measured in kelvins to reflect his accurate determination of its value: -273.15C. Despite many offers, he refused to leave Glasgow, remaining Professor of Natural Philosophy for over 50 years, until his retirement. Towards the end of his life, he made the statement: "There is nothing new to be discovered in physics now, All that remains is more and more precise measurement." This was unlucky to say the least: the revolutionary discoveries of quantum mechanics and relativity (as well as the electron and radioactivity) were made not long after the claim. He was buried in Westminster Abbey next to the grave of Isaac Newton. But despite this honour and all his brilliance he barely features in modern physics apart from his formulation of the laws of thermodynamics and the definition of the scale of absolute temperature, to which we will soon come. He was the last of the great exponents of classical physics and struggled for decades to solve problems that could only yield to the wonders of relativity and quantum mechanics.

[18] I strongly recommend the collection of essays *Kelvin: Life, Labours, Legacy*.

It's time to describe the work for which he does deserve great credit. In 1849 he published a paper that drew wide attention to Carnot's work. It begins with the observation that the presence of heat can be recognized in every natural object and that "there is scarcely an operation in nature which is not more or less affected by its all-pervading influence". It will therefore be desirable, in laying the foundation of a physical theory of heat, to discover or

> imagine phenomena free from all such complication ... in which the relation between cause and effect, traced through the medium of certain simple operations, may be clearly appreciated. Thus it is that Carnot, in accordance with the strictest principles of philosophy, enters upon the investigation of the theory of the motive power of heat.

In many ways, the most interesting thing about Thomson's paper, which there is no need to discuss in detail, is his attitude to caloric, about which he is ambivalent. He mentions "the extremely important discoveries of Mr Joule of Manchester" but concludes that in the present state of science

> the fundamental axiom adopted by Carnot [that heat is an indestructible substance] may be considered as still the most probable basis for an investigation of the motive power of heat; although this, and with it every other branch of the theory of heat may ultimately require to be reconstructed upon another foundation when our experimental data are more complete.

It seems clear that Thomson at this stage clung to the notion of caloric because the full implications of Joule's work had escaped him. He recognized that work could be transformed into heat but had doubts about the transformation of heat into work. In a paper of 1852 to be discussed later, Thomson refers to a letter of 1847 from Joule that provided "an instance of the conversion of heat into the mechanical force of a current" which showed that he had been mistaken to believe that no evidence could be adduced to show "that heat is ever put out of existence". Five years later he was fully persuaded it could and repeated the striking phrase about heat being "put out of existence".

In his 1849 paper, Thomson highlighted Carnot's achievement in a definition in which, as a hyphenated adjective, *thermodynamics* is first named:

> *A perfect thermo-dynamic engine is such that, whatever amount of mechanical effect it can derive from a certain thermal agency, if an equal amount be spent in working it backwards, an equal reverse thermal effect will be produced.*

The influence of Carnot's 'caloric' way of thinking, as expressed in his waterfall analogy, can be seen in Thomson's comment: "At the conclusion of this cycle of operations the total thermal agency has been the *letting down* of H units of heat from the body A, at the temperature S, to B, at the lower temperature T." Carnot's analogy was so plausible—a fixed amount of water does a definite amount of work when passing through a watermill from a higher to a lower level. Some ideas can be very hard to shake off. That was beyond the 25-year old Thomson in 1849 but not, as we will see in a moment, an almost equally young man.

In his discussion of the possible replacement of Carnot's fundamental axiom, Thomson had said this could only be done "when our experimental data are more complete". It seemed any reconstruction was relatively far off. In fact, within a year Rudolf Clausius (1822-1888) brought down the tottering edifice of caloric by a clear argument using a few judiciously chosen words. He had read, and acknowledged graciously, Thomson's paper. Together with Clapeyron's paper, it was Clausius' sole knowledge of Carnot's book, a copy of which he had been unable to acquire.[19]

First, a few biographical details. Clausius was born in Prussia in Köslin (now Koszalin in Poland). His father was a Protestant pastor and school inspector; in his papers and a well-known photograph, Rudolf comes across as a rather stern person. He graduated from the University of Berlin in 1844, where he studied mathematics and physics with some distinguished teachers. Prussia was making great efforts in education. During 1847, he got his doctorate from the University of Halle on optical effects in the earth's atmosphere. He then became professor of physics at the Royal Artillery and Engineering School in Berlin and Privatdozent at the Berlin University. In 1855 he became professor at the Swiss Federal Institute of Technology in Zürich (ETH Zürich), where he stayed until 1867. During that year, he moved to Würzburg and two years later, to Bonn. In 1870 he organized an ambulance corps in the Franco–Prussian War and was wounded in battle, leaving him with a lasting disability. He was awarded the Iron Cross for his services. His wife Adelheid died in childbirth in 1875, leaving him to raise their six children. He continued to teach, but had less time for research. He remarried in 1886 and had another child.

His famous paper "On the Moving Force of Heat and the Laws of Heat which may be Deduced Therefrom" appeared in German in 1850 and was soon translated into English, appearing in the *Philosophical Magazine*. This and regular translation of subsequent papers ensured that his work soon became widely known. They express the key points clearly and crisply; one senses a sharp mind at work. Over 15 years Clausius patiently worked his way through to what is perhaps the most subtle concept in physics.

In laying the new foundation that Thomson thought was still some way off, Clausius examines Carnot's assertion that in a functioning steam engine it is the *transmission* of heat "from a warm body to a cold one" which corresponds to the work produced. He quotes Carnot's assertion "*no heat is lost* in the process" and says he is not sure there is sufficient experimental support for the claim. He argues that

> although no such loss may have been directly proved, still other facts render it exceedingly probable that a loss occurs. If we assume that heat, like matter, cannot be lessened in quantity, we must also assume that it cannot be increased; but it is almost impossible to explain the ascension of temperature brought about by friction otherwise than by assuming an actual increase of heat. The careful experiments of Joule ... establish almost to a certainty, not only the possibility of increasing the quantity of heat, but also the fact that the newly-produced heat is proportional to the work expended in its production.

Clausius also notes that

[19] Besides being very good at introducing names for key concepts such as kinetic energy, Thomson deserves credit for the way he highlighted the work of not only Carnot but also George Green. I must leave the reader to check out Green's remarkable life and achievements in Wikipedia.

> many facts have lately transpired which tend to overthrow the hypothesis that heat is itself a body, and to prove that it consists in a motion of the ultimate particles of bodies. If this be so, the general principles of mechanics may be applied to heat; this motion may be converted into work, the loss of *vis viva* in each particular case being proportional to the quantity of work produced.

He says there is an urgent need to establish whether "the production of work is not only due to an alteration in the *distribution* of heat, but to an actual *consumption* thereof." He praises Thomson's paper but disagrees with the worry that abandoning the principle of indestructible heat would lead immediately to "innumerable other difficulties." Clausius counters "we ought not to suffer ourselves to be daunted by these difficulties, but that, on the contrary, we must look steadfastly into this theory." He continues:

> On a nearer view of the case, we find that the new theory is opposed, not to the real fundamental principle of Carnot, but to the addition "no heat is lost;" for it is quite possible that in the production of work both may take place at the same time; a certain portion of heat may be consumed, and a futher 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.

Thus, according to Clausius the real fundamental principle of Carnot is this: *whenever work is produced by heat and a permanent alteration of the body in action does not at the same time take place, a certain quantity of heat passes from a warm body to a cold one.*[20]

Turning to the task of setting up a new theory of heat without the addition "no heat is lost", he says he will "merely lay down one maxim" founded on that assumption:

> *In all cases where work is produced, a quantity of heat proportional to the work done is consumed; and conversely, by the expenditure of a like quantity of heat, the same amount of heat may be produced.*

Every now and then in the history of science one comes across a few words that, in the context in which they arise, open up quite new vistas. For brevity, nothing can rival Darwin's 'by means of natural selection', but the import of Clausius' words here, clearly justified and prompted by Joule's work, is still great. They do not use those visionary phrases like "the wonderful arrangements of creation" but in sober words give the first truly clear expression of the first law of thermodynamics. It's essential content is the precise equivalence of heat and work. Of course, Clausius could never have formulated it if Joule had not performed his famous experiment. The *joule* is a unit of energy for good reason.

Thomson missed his chance of priority because he baulked at the idea of heat being "put out of existence". Within a year Clausius became caloric's executioner. Discussing the conversion of water into steam and the heat involved, he said scientists distinguish

[20]The exclusion of a permanent alteration of the body (the working medium) rules out the possibility that I mooted at the end of chapter 2 of halting the Carnot cycle, when the working medium is in a different state, at the half-way stage. That would allow complete conversion of heat into work. To be clear I'm not proposing humans can build perpetual-motion machines. What I will do is argue that the universe is not a steam engine and therefore need not be governed by the same laws.

> the *sensible* heat and the *latent* heat. Only the former of these, however, must be regarded as present in the produced steam; the latter is, not only as its name imports, hidden from our perceptions, but has actually *no existence*; during the alteration it has been *converted into work*.

Clausius's *no existence* is the counterpart of Thomson's *put out of existence* in his letter to Joule (but not in his 1949 paper). The difference is that Clausius accepted without reservation what Thomson had put off for another day. Clausius was wasting no time. As he destroyed one concept, he was also working his way towards a new one. This was a much more sophisticated concept than an indestructible substance.

6 The Second Law of Thermodynamics

Clausius's 1850 paper is justly considered his greatest. Besides the succinct formulation of the first law of thermodynamics and significant steps toward the discovery of entropy, it contained the first formulation of the second law. This came about through Clausius's reexamination of the theorem which Carnot thought he had proven—that all reversible heat-engines operating between the same two temperatures have the same efficiency whatever working medium they employ. If not, one could build a perpetual-motion machine. Clausius examined the claim in the context of his newly minted but not yet named first law of thermodynamics. He found the result right but the proof wrong. A tiny modification put it right. His argument begins like Carnot's.

Suppose two substances, one of which produces more work by transmitting a certain amount of heat from a warm heat reservoir A to a colder one B. Use the two substances alternately. The first could do a certain amount of work and the second consume it by reversing the process. At the end, both A and B would be back in their original state (at their original temperatures); further, the work expended and the work produced would exactly balance, so that also, in agreement with Clausius's formulation of the first law, the quantity of heat would have been neither increased or decreased. But the *distribution* of the heat would have been changed—more heat would have been brought from B to A than from A to B. The net outcome would be a transmission of heat from B to A. Clausius's killer conclusion was that

> by repeating both of these alternating processes, without expenditure of force or other alteration whatever, any quantity of heat might be transmitted from a *cold* body to a *warm* one; and this contradicts 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.

The last sentence is nothing more and nothing less than the first enunciation of the second law of thermodynamics. It's a fact all too familiar if you try to keep a house warm in winter without a fire, but had any scientist before Clausius recognized its significance? Carnot was close with the condition for maximising work done: heat must never flow from a hot to a colder body without doing work by the working medium's expansion. But, blinded by caloric's allure, he just missed laying the final foundation stone of thermodynamics. It is through insights like Clausius's, so simple when seen for what they are, that science makes spellbinding advances.

Whereas Carnot invoked the impossibility of perpetual motion to prove all reversible heat engines have the same efficiency, Clausius put thermodynamics on deep and sure foundations by showing the kind of perpetual motion Carnot considered was simply impossible because heat never flows spontaneously from a colder to a hotter body. Ever since the work of Carnot, Clausius and Thomson, thermodynamics has been based on two axioms: energy conservation, which rules out perpetual-motion machines 'of the first kind'—you cannot get something for nothing—and simple statements about the behaviour of heat which rule out perpetual-motion machines 'of the second kind'—even if you have something (energy in the form of heat) you cannot use all of it.

There will be a bit more to say about Clausius's 1850 paper in connection with his discovery of entropy, but it is now time to consider the first of a series of papers that Thomson, reacting to Clausius's paper,started to publish in 1852, the first based on a lecture given in 1851. The title of the series, *The dynamical theory of heat*, and its opening sentence reveal the impact of Clausius's paper. All caloric equivocation has vanished. Humphrey Davy had caused two pieces of ice to melt by rubbing them together so "caloric does not exist" and "the dynamical theory of heat [is] thus established". He mentions the recent conclusions of Rankine[21] and Clausius, describes the work of Mayer and Joule and says the whole theory of the motive power of heat is founded on two propositions. These coincide essentially with what Clausius had already initiated and, it must be said, do somewhat give the impression of a Brit trying to catch up with a German:[22]

> Prop. I. (Joule). When equal quantities of mechanical effect are produced by any means whatever from purely thermal sources, or lost in purely thermal effects, equal quantities of heat are put out of existence or are generated.
>
> Prop. II. (Carnot and Clausius). If an engine be such that, when it is worked backwards, the physical and mechanical agencies are all reversed, it produces as much mechanical effect as can be produced by any thermo-dynamic engine, with the same temperatures of source and refrigerator, from a given amount of heat.

Proposition I is essentially Clausius' first maxim and a formulation of the first law of thermodynamics; since Clausius openly credited the content to Joule, Thomson's designation is clearly justified. However, I think one could argue that its formal elevation to the status of the first law of the new science of thermodynamics was Clausius's service. The same, with important input by Thomson, can be said of the second law. In the years following his 1850 paper, Clausius increasingly referred to the *erster Hauptsatz und zweiter Hauptsatz der Thermodynamik*, which can be translated somewhat awkwardly as 'the first principal law and second principal law of thermodynamics'. Not surprisingly they are simply called the first and second laws in English, but they do have a grander ring in German and helped the second law achieve a special—indeed almost mystical—status.

The coining of words is important; as I already said, Thomson seems to have been the first to use the expression 'thermo-dynamic' and in footnote 22 I mention his expression 'internal energy'. His crediting to Carnot and Clausius of Proposition II above, which is not to be confused with the second law, also stuck. It became known as the Carnot–Clausius principle and, as we will see, played a central role in furious debates in the final decades of

[21] William Rankine (1820-1872), an engineer (in Glasgow like Thomson), played a lesser role in the creation of thermodynamics.

[22] In one respect, he succeeded in that very well by naming and emphasizing the importance of the internal energy of any medium used in a heat engine. Together with the work done by the medium, the internal energy is a key term in the equation that expresses the first law. More generally it is above all due to Thomson and, to some extent, his friend and collaborator Peter Guthrie Tate (1831-1901), as the joint authors of their *Treatise on Natural Philosophy*, that energy, in its two forms kinetic and potential, came very prominently to centre stage, displacing from there Newton's concept of force. But ironically they argued strongly and incorrectly that Newton himself was well aware of the existence and significance of energy and had merely failed to name it. Because of this Thomson's achievement was, and to this day still is, underrated.

the 19th century about the status of thermodynamics as opposed to statistical mechanics, to which we will come shortly.

The immediate continuation of Thomson's paper introduces a rather curious feature in the history of the second law. He writes:

> The demonstration of the second proposition is founded on the following 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 the surrounding objects.*[23]

In fact, there are several curious features. The first, which we have here, is that, in contrast to Clausius's improvement of Carnot's perpetual-motion argument, which Clausius did not formulate as a formal axiom but merely invoked as a fact of nature (heat does not flow spontaneously from a cold to a warm body), Thomson's is the first statement of the second law formal terms. We'll come to the next curiosity in a moment.

I won't give Thomson's proof of the Carnot–Clausius principle since, despite starting from a seemingly different standpoint, it is very like Clausius's proof, and both proofs, in their turn, are closely related to Carnot's proof but modified to take into account the now fully recognized non-existence of caloric. This means, said Thomson, that Carnot was wrong in his assumption that, in a complete cycle of operations, "the medium parts with exactly the same quantity of heat as it receives." For this reason

> Carnot's original demonstration utterly fails, but we cannot infer that the proposition [itself] is false. [Its truth] appeared to me, indeed, so probable, that I took it in connection with Joule's principle ... as the foundation of an investigation of the motive power of heat ... It was not until the commencement of the present year that I found the demonstration given above, by which the truth of the proposition is established upon [the axiom above] which I think will be generally admitted.

Thomson continued:

> It is with no wish to claim priority that I make these statements, as the merit of first establishing the proposition upon correct principles is entirely due to Clausius, who published his demonstration of it in the month of May last year [1850], in the second part of his paper on the motive power of heat. I may be allowed to add, that I have given the demonstration exactly as it occurred to me before I knew that Clausius had either enunciated or demonstrated the proposition. The following is the axiom on which Clausius' demonstration is founded: *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.* It is easily shown, that, although this and the axiom I have used are different in form, either is a consequence of the other.[24]

[23] In a footnote, in an illustration often repeated, Thomson says: "If this axiom be denied for all temperatures, it would have to be admitted that a self-acting machine might be set to work and produce mechanical effect by cooling the sea or earth, with no limit but the total loss of heat from the earth and sea, or, in reality, from the whole material world."

[24] The reader can find the proof of the equivalence in Fermi's book (see footnote 5).

The second curious thing is this. Clausius, despite his knack in the formulation of axioms and propositions, had not, besides not formulating them as an axiom, used the words Thomson put in his mouth. Let me quote again what Clausius actually said:

> Hence by repeating both of these alternating processes, without expenditure of force or other alteration whatever, any quantity of heat might be transmitted from a *cold* body to a *warm* one; and this contradicts 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.

The oddity of this story continues, for in 1854, in his next foundational paper on thermodynamics, Clausius, without referring to his 1850 paper and his just quoted words in it and without any mention of Thomson's 1852 formulation of 'Clausius' axiom', gives a crisp formulation of the second law: *Heat can never pass from a colder to a warmer body without some other change, connected therewith, occurring at the same time.* Critically important here is the subsidiary condition "without some other change, connected therewith, occurring at the same time". This condition is already effectively present in Clausius' 1850 words "without expenditure of force or other alteration whatever" but is neither explicit nor, I think, implicit in Thomson's 1852 formulation of his own axiom (though through the caveat "unaided by any external agency" it is in the one he attributes to Clausius!)

In his 1936 lectures, Fermi emphasized the importance of an addition like Clausius's by formulating 'the postulate of Lord Kelvin' as follows: *A transformation whose only final result is to transform into work heat extracted from a source which is at the same temperature throughout is impossible.* To this, Fermi adds the illuminating footnote:

> An essential part of Lord Kelvin's postulate is that the transformation of the heat into work be the *only* final result of the process. Indeed, it is not impossible to transform into work heat taken from a source all at one temperature provided some other change in the state of the system is present at the end of the process.

This is essentially the comment that I made at the end of chapter 2, namely that heat can be fully transformed into work if the Carnot cycle is stopped at the half-way stage c in Fig. 1. As I noted then, this may be relevant when we come to consider the universe and whether its evolution can be likened in any way to a Carnot cycle.

The reader may well ask where Fermi's '*only* final result' is to be found in Kelvin's words. They aren't. I suspect they are Fermi's; they are certainly sharp and may have been prompted by the widely quoted proposal of Max Planck in 1897: "It is impossible to construct a periodically functioning machine that does nothing except raise a weight and cool a reservoir." Planck asserts that there are numerous different possible formulations of the second law and that all the more or less satisfactory ones, which include those of Clausius and Thomson, are effectively equivalent. Of his own, he says it is mainly chosen for its use of expressions familiar to engineers. A formulation like Planck's or Fermi's is often called the Kelvin–Planck formulation. I don't think there is any reason to doubt Thomson's statement that he came to his formulation before learning of Clausius' work. But what is not in doubt is that Clausius was there first and moreover, already in his 1850 words, got the full formulation right. Thomson did not. This, like the numerous formulations that one can find of the second law, illustrates the delicate nature of its content. It's a tricky matter.

I've already said that the universe is not a steam engine. Except for the heat engines constructed by humans, I think it is questionable whether any naturally occurring ones exist anywhere in the cosmos. This just underlines the point that we may need to think about it differently. This is where Thomson may help us more than Clausius, as the next chapter will explain.

7 The Dissipation of Mechanical Energy

Thomson may have been slow off the mark in realizing caloric's days were numbered, but he did recognize and point out clearly before anyone else something that soon came to be seen as an alarming, if not to say nightmarish, implication of the second law. In 1852, he published his paper "On a Universal Tendency in Nature to the Dissipation of Mechanical Energy." Carnot's single observation, so close to a formulation of the second law, on how one must maximise work done by heat prompted Thomson to note "the remarkable consequences which follow from Carnot's proposition, that there is an absolute waste of mechanical energy available to man when heat is allowed to pass from one body to another at a lower temperature, by any means not fulfilling his criterion of a 'perfect thermo-dynamic engine'." Thomson concluded: "As it is most certain that Creative Power alone can either call into existence or annihilate mechanical energy, the 'waste' referred to cannot be annihilation, but must be some transformation of energy."

To clarify what he means by examples, Thomson introduces the concept of *stores* of mechanical energy and divides them into two classes—*statical* and *dynamical*. The former are weights at a height, electrified bodies, a quantity of fuel; the latter are masses of matter in motion, a volume of space through which undulations of light or radiant heat are passing, a body having thermal motions among its particles (that is, not infinitely cold). As examples of the transformation of energy, he mentioned "When heat is created by any irreversible process (such as friction), there is *dissipation* of mechanical energy, and a full *restoration* of it to its primitive condition is impossible" and "When heat is diffused by *conduction*, there is dissipation of mechanical energy, and perfect *restoration* is impossible." He gave similar examples of irreversible transformation for radiant heat or light and concluded:

> 1. There is at present in the material world a universal tendency to the dissipation of mechanical energy.
>
> 2. Any *restoration* of mechanical energy, with more than an equivalent of dissipation, is impossible in inanimate material processes.
>
> 3. Within a finite period of time past the earth must have been, and within a finite period of time to come the earth must again be, unfit for the habitation of man as at present constituted, unless operations have been, or are to be constituted, which are impossible under the laws to which the known operations going on at present in the material world are subject.

The reference to 'inanimate material processes' in the second of these may reflect a residual religious hope that animate processes might be different.

I'm not sure how Thomson came to the conclusions in his third statement; he gave no detailed arguments for them. He probably had in mind, at least in their earliest rudimentary form, his ideas about the origin of the sun and earth that he developed more fully in the next decade or so. According to them, both bodies had been formed in the not too distant past; for this reason alone the earth could not have been fit for the habitation of man—it would not even have existed. Since in accordance with the not yet formally stated (but with its implications well understood) first law of thermodynamics the sun and earth each had only finite stores of mechanical energy, these must be dissipated within a finite time in the future.

By 1854, when he addressed the British Association, Thomson's thoughts had crystallized in certain key respects. He was of the firm opinion that the ultimate source of all mechanical energy in the solar system was gravitational potential energy of a primordial nebula that, by our epoch, had been largely transformed into hot solid or liquid bodies. He believed that the known mechanical laws could be used with confidence to predict in broad outline the future of the solar system, which was one of progressive cooling through dissipation leading to the end of the world as a habitation for man. The same laws could be used to trace things backwards to an epoch in which all bodies must have been indefinitely remote from each other. However, "such conclusions are subject to limitations, as we do not know at what moment a creation of matter or energy may have given a beginning, beyond which mechanical speculations can not lead us". Thomson insisted such divine intervention could not be ruled out because "all fossil organic remains, are organized forms of matter to which science can point no antecedent except the Will of a Creator, a truth amply confirmed by the evidence of geological history."

This last comment makes it worth saying something about how, in the second half of the 19th century, Thomson, almost single-handedly, transformed ideas about geology. This was a discipline very largely developed in Britain. For this discussion, I draw on Joe D. Burchfield's *Lord Kelvin and the Age of the Earth* (see also the essays in the collection mentioned in footnote 18). In the mid 19th century, the dominant theory in geology was *uniformitarianism*. It had been initially proposed by James Hutton (1726-1797). He had argued that lawful physical processes like those currently operating had been doing so over immense stretches of time and were capable of creating the observed geological record. The 'uniform' in the later coining uniformitarianism drew attention to the assumed uniformity of the relevant physical processes. Hutton postulated a cyclic progression of changes so ancient as to obscure any "vestige of a beginning" and to hold out "no prospect of an end". This did not necessarily imply, as was often assumed, an eternal, essentially unchanging world, but simply a stretch of time so great that its actual length was of no practical significance. Hutton's ideas were developed much further by Charles Lyell (1797-1875), who between 1830 and 1833 published his multi-volume *Principles of Geology*.

Both of the newly discovered laws of thermodynamics enabled Thomson to mount a sustained attack on uniformitarianism that lasted more or less to the end of his life. He thereby intoduced into geology physical principles it had hitherto been lacking. The first law guaranteed that individual bodies only had a definite finite store of energy, while the second ensured that it must inevitably be dissipated. This raised the important question of the length of time dissipation would take and Thomson's interest in the age of the earth. He also challenged Hutton's cyclic progression of changes, which suggested the need for a perpetual motion machine of the second kind and could be ruled out by the second law.

A substantial digression into the impact of thermodynamics on geology in the second half of the 19th century is not warranted but, since it is very much part of the story within a story, I think it is worth quoting the opening of a popular article Thomson published in 1862 in *Macmillan's Magazine*. It opens with the confidence of a man who, like Joule before him, has made a great discovery in science but who then, as a religious person, finds himself forced to equivocate:

> The second great law of thermodynamics involves a certain principle of *irreversible action in Nature.* It is thus shown that, although mechanical energy is *indestructible*, there is a universal tendency to its dissipation, which produces gradual augmentation and diffusion of heat, cessation of motion, and exhaustion of potential energy through the material universe. The result would inevitably be a state of universal rest and death, if the universe were finite and left to obey existing laws. But it is impossible to conceive a limit to the extent of matter in the universe; and therefore science points rather to an endless progress, through an endless space, of action involving the transformation of potential energy into palpable motion and thence into heat, than to a single finite mechanism, running down like a clock, and stopping for ever. It is also impossible to conceive either the beginning or the continuance of life, without an overruling creative power; and, therefore, no conclusions of dynamical science regarding the future condition of the earth can be held to give dispiriting views as to the destiny of the race of intelligent beings by which it is at present inhabited.

Thomson was surely struggling with the possible implications of Darwin's *On the Origin of Species by Natural Selection*, which had been published in 1859. As envisaged by Darwin, evolution must have taken place over a great stretch of time and, rashly as he latter admitted bitterly to friends, he had included an estimate for the age of the Weald, a geological feature near his home. For this he had assumed the sea would eat into and denude chalk cliffs at the rate of one inch per century. This led Thomson, on the basis of his considerations about the physical nature of the sun, to ask sceptically in his 1862 article "What then are we to think of such geological estimates as 300,000,000 years for the 'denudation of the Weald'?" Under the impact of this attack, Darwin discretely omitted from later editions the estimate of the Weald's age, replacing it with bland hand-waving arguments.

For several decades a battle raged between, on the one hand, the geologists and supporters of evolution and, on the other, the physicists. It pitted two 'bulldogs' against each other: Thomas Huxley on Darwin's side and Peter Guthrie Tait, who argued very dogmatically for the line taken by Thomson. Of course, all Thomson's arguments about the nature and age of the sun and earth were hopelessly premature. Antoine Becquerel's discovery of radioactivity in March 1896 proved that the earth had a powerful, completely unexpected internal source of heat and provided radioactive dating as a new method for estimating the age of the earth. After refinement over a few decades, it led to an estimate for the age of the earth about 15 times greater than Darwin's for the age of the Weald. Moreover, in the 1930s, nuclear fusion, not gravitational potential energy, was shown to be the true source of the sun's energy.

It is indicative of the disconcerting implications of Thomson's 1852 paper for his contemporaries that almost immediately Rankine published a paper with the bizarre proposal that

> in all directions round the visible world, the interstellar medium has bounds beyond which there is empty space. If this conjecture be true, then on reaching those bounds the radiant heat of the world will be totally reflected, and will ultimately be reconstituted into foci.

This was therefore a mechanism to achieve 'anti-dissipation'.

Rankine's hope, rebutted by Clausius in 1863, to salvage something from ultimate dissipation was the first attempt to seek a cosmological mechanism capable of reversing the gloomy implications of the second law. Intriguely, and not widely known, is the fact that by no means everyone reacted like Rankine. In fact, for them the second law had a virtue. If the universe is 'winding down', this suggests it must have been 'wound up' at some time in the past. Religious people, serious scientists among them, interpreted this as evidence for divine creation of the universe—a comforting thought. The historian of science Helge Kragh has written an interesting book, *Entropic Creation*, on the subject. Friedrich Engels, as a materialist, liked the eternity implied by the first law but not the second, with its implication of a creation event. Professional scientists, Thomson among them—in 1844 he had, as I said, already surmised a creation event for the universe—tended to avoid publishing papers on the idea, probably fearing it might harm their reputation, but in religious and cultural circles, especially in Germany, the possibility created much discussion up to and even beyond the end of the 19th century. One could see the scientific grounding through the second law of belief in a creation event as an anticipation of big-bang birth of the universe. At the very least, even though Thomson himself did not want to accept it, the existence of the universal phenomenon to which his 1852 paper drew attention did suggest a universe that evolves and has a history.

However, there is one thing that does need to be said. By no means all of the processes we see around us involve dissipation and decay. A cup may shatter but it was made. We are born before we die. All these processes share the common temporal direction that may be called a background arrow, but they do highlight the fact that in the universe around us we observe both creative and destructive processes. The standard view is that creation is only possible because it is accompanied by a greater amount of destruction. This can certainly be argued if we restrict our attention to what might be called our immediate environment, but in *The Janus Point*, for which (as explained in the preface) the present material originally served as an introduction, I argue things might look different in the context of the entire universe.

One reason I have included this chapter on Thomson's paper, which very soon led to the notion of heat death of the universe, is that, unlike Clausius's notion of entropy, which is quantitative, dissipation as understood by Thomson is qualitative. The difficulty with quantification is ambiguity. Definite numbers can only be introduced if well defined boundaries or limits are at hand. What gave Joule a decisive advantage over the unfortunate Mayer, who merely showed that violent shaking increases the temperature of water, was not only enclosure of water in a box, which both had, but Joule's precise measurement of the work done by descent of weights.

The good thing about Thomson's paper is the ease with which we can recognize dissipation without having to quantify it. Consider the classic example, given so many times, of entropy increase: the cup I just mentioned which falls from a table onto the floor and shatters. Nobody can deny the cup has broken and in that sense dissipation has occurred. But, as we will see in the next chapter, no two scientists set the task of calculating an actual entropy increase associated with the unfortunate accident could possibly come up with the same value without prior agreement between themselves of definite spatial and temporal boundaries within which the entropy increase is to be calculated.

While accepting this point, the reader might still wonder why I included a discussion in this chapter of long abandoned ideas about the nature and history of the earth and sun. My main reason, besides the intrinsic interest of the ideas, is that the 1850s and 1860s were the decades in which certain aspects of thermodynamics came to have the appearance of unshakeable truths. In fact, modern discoveries have only changed the details in Thomson's overall picture of the formation, stabilization and decay of the solar system. Gravitational forces have simply been augmented by further forces operating in accordance with the laws of quantum mechanics. As regards the concerns of this book, the really significant change has been the development of cosmology as a mature science. It bears scant resemblnce to Thomson's "endless progress, through an endless space, of action involving the transformation of potential energy into palpable motion and thence into heat".

The real game changer is Hubble's discovery of the expansion of the universe. I think that should raise doubts about at least some aspects of the thermodynamics that Carnot, Clausius and Thomson discovered. With that in mind, I will say something about a point made in chapter 19 of *The Janus Point* that bears directly on Kelvin's notion of universal dissipation of mechanical energy in an expanding universe. It is not my intention here to repeat its contents, though I will say I address an issue that Kelvin did not when he introduced the notion, which I think is very valuable, of stores of mechanical energy. He told us how they are dissipated and ultimately go out of existence. Did he ever ask how they came into existence? I don't think he did. All we have is that 1862 hand waving in Macmillan's Magazine, in which he speaks of "an endless progress, through an endless space, of action involving the transformation of potential energy into palpable motion and thence into heat"

I took me four years to write *The Janus Point*. Throughout the process I was trying to understand the best way to characterise Kelvin's dissipation and Clausius's growth of entropy, to which we are about to come. These days entropy is almost universally called a measure of order, with entropy growth tantamount to an increase of disorder. However, in a little booklet of Peter Atkins I found the possibly more helpful suggestion that entropy measures the *quality* of heat—the same amount of energy in the form of heat at a higher temperature in a small space can do more work than the same amount of heat energy spread out in a large space. I began to think that Kelvin's energy dissipation might have been better characterised as *spreading*, especially for unconfined systems like stars in an expanding universe. The words to express this only came to me near the very end of writing the book :

> Variety, expressed through shapes and ratios, can increase forever. With your thumb, you can press an ink drop on paper into a smudge; the ink, like Boltzmann's particles in a box, is trapped within the area on which you press your thumb. But nature, the artist, holding a fine pen between her fingers and thumb,[25] can, maintaining contrast and with it variety, draw out the same ink as thinly and as far as her inspiration takes her. She can create the finest picture imaginable for eyes that see contrasts. Those

[25] This was my text. My excellent copy editor Sue Warga felt it better not to give nature and the prominent concept *creation measure* of an earlier chapter a gender. I concurred; the printed text reads "nature, the artist, holding a fine pen between fingers and thumb".

> contrasts are ratios made manifest. Ratios are the foundation of so much that we experience. The real numbers permit unending extension of ratios – one real number divided by another – all the way to infinity. It can be done with density contrasts on the inside of a sphere. In his nutshell, Hamlet can be king, not of bland infinite space, but infinite variety.

Shortly before revising this history of thermodynamics, in an article for the online journal *Nautilus*, I found a way, less poetic but more scientific, to argue that Kelvin's 'dissipation', with its negative connotation,[26] might with advantage have been called 'spreading'. This is what I wrote:

> There is a beautiful effect I often go to watch on afternoon walks near my home. A tree hangs above a brook where the water flows smoothly over a ford. If it has rained, drops of water fall from the tree onto the water, creating circular waves that spread out over the flowing water. You can watch the effect for far too long if you need to get on with work. The waves created by drops that hit the water at different points meet and pass through each other, each emerging intact. If the brook had no banks and the water no viscosity, that would create the condition radiation finds in the vast voids of our expanding universe and the patterns would remain as beautiful for ever. That's the difference between an open and a closed system.
>
> Before the banks have their effect in the brook, mechanical energy is first entirely within each falling drop but is then spread out in the circular waves. Except at or near equilibrium Clausius's entropy is a *quantity* difficult to define and measure; Thomson's dissipation is a *qualitative* effect that is both universal and easy to recognise. He mentioned the heat created by friction as an example of dissipation; the illustration has been endlessly repeated. But Thomson loved the river Kelvin in Glasgow and took his baronial name from it. He must so often have seen water drops falling onto the river. If in the title of his 1852 paper, which was so influential, he had changed 'dissipation' into 'spreading'—it would have been a better characterisation—who can say how that might have changed the interpretation of the second law, especially after Hubble's monumental discovery?
>
> The circle is the most perfect geometrical figure and pi, which relates its circumference to its radius, bids fair to be the most perfect number. "Ah," you say, "a thing of beauty may have been born, but it decays as the waves get shallower and shallower." To which I answer that you forget the lessons of *Gulliver's Travels* and the relativity of size. It is only ratios that have physical meaning. The beauty is in the ratios, and they persist forever even in the expanding universe.

[26] Paul Davies has called Kelvin's landmark paper "one of the gloomiest of all times in science".

8 The Discovery of Entropy

If you walk from any point A on the earth's surface your altitude above sea level when you get to any other point B will be the same whatever route you take. The same is true of the difference of altitudes. In fact, the difference is physically more significant since altitude could just as well be measured relative to Mount Everest's peak as to sea level. Quantities that are route independent are very important in physics. In thermodynamics they are called state functions. We already met them as the pressure, volume and temperature of gases when in equilibrium (p. 000.).

There are other important quantities that, in contrast, are route dependent. To give an example that you can readily understand, suppose that in some country you are always forced to travel by toll roads to get from one point to another and that there are many roads, with different tolls, between any two points. To ease traffic flow, the authorities may even pay you to go by some routes. Then the total cost of the tolls paid during the journey will obviously depend on the route you take. To determine the difference in heights above sea level at the end of a journey, all you need to know are the points of departure and arrival. But if you are having to pay tolls and there are different possible routes, you cannot determine the cost of the journey unless the route is specified.

An indication of the way Clausius recognized the full significance of Joule's work is his observation in his 1850 paper that there exist two different kinds of work that can be performed in thermodynamics system. One is *interior work*, for example, the work needed "to overcome the mutual attraction of the particles" of water. (Joule had spoken of the molecules of a compressed spring being "forced asunder".) Now the same amount of heat will need to be expended on the interior work whenever and however it happens, for that depends solely on the nature of the molecules. It's route independent and analogous to altitude—all you need to know are the intial and final states.

The other is *exterior work* done, for example, by successive expansions and compressions of a gas. If the original condition of the gas is again established at the end of the process there is no guarantee that the entire amount of work produced will equal the amount expended. This would be the case in a Carnot heat engine, Fig. 1, taken along *abc* and then simply back to *a* along *cba*. Then no net work would be done. But with the route along *cda* back to *a* the maximum possible work is done. Thus, exterior work is a route-dependent quantity. One has to know the succession of states through which the system is carried by the application of heat or pressure from without.

Clausius's insight about the difference between the two kinds of work overturned the notion of indestructible caloric. For if caloric did exist, the amount of it in the working medium would be fixed given the temperature and volume. Like the height above sea level, it would be a state function. In his 1850 paper Clausius banished a state function; four years later he introduced a replacement, entropy. It is a much more sophisticated concept than caloric and intimately related to the measurement of temperature.

Galileo used the expansion of gas to measure temperature. Relatively soon it was discovered that, if confined and kept at a constant pressure, the volumes of many gases all increase in the same proportion as the temperature is increased. These are the so-called ideal gases. With the freezing and boiling points of water used, as in the Celsius scale, as reference points, such gases reliably measure concordant temperatures as long as they are far from their liquefaction temperature. However, once that is approached the gases behave differently and give discordant temperatures. Nothing singles out the readings given by one gas rather than another. The very concept of temperature is threatened.

Moreover, by the early 19th century ideal-gas thermometers had provided strong evidence for a limiting value of possible temperatures. They showed that $PV = t + a$, where P is the constant pressure of the employed gas, V is its volume, t is the temperature and a is a constant found, to good accuracy, to be the same for all gases. Clearly, if one could cool the gases to $t = -a$ something special would happen: $PV = 0$. Since the pressure P is fixed, the volume of the gas, if ideal, would have to become zero. All tested gases pointed to a limit of possible temperatures, an *absolute zero*. In the Celsius scale, Thomson found it to be $-273.15°$. In his honour, temperatures measured from it are called kelvins.

There still remained the problem of defining an unambiguous notion of temperature and measuring it. Thomson's solution to this problem simultaneously solved another. Carnot had shown that the efficiency of a perfect heat-engine depends only on the two temperatures between which it operates but he had not been able to find its value. This depends on the amount of heat taken from the furnace and the amount that must be wasted when deposited in the condenser. The problem is to pass from the efficiency defined in terms of them to the temperatures at which the transfers take place. Thomson's answer almost looks like cheating. He did not use temperatures to determine amounts of heat but rather amounts of heat to define temperatures.

I'm not going to go through here the subtle arguments that, through the work of Thomson and Clausius, led to the simultaneous solution of these two problems and also the definition of entropy. You will find them in Fermi's sharp and succinct account, which employs Thomson's form of the second law (which, as is standard now, he calls Kelvin's). Here I will simply outline the conditions involved and give the key results. However, for anyone with physics at high-school level among their accomplishments, it's well worth trying Fermi. The story is one wonder after another.

The key axiom is the second law, to which, in either of its forms, Carnot himself came so close. Everything else comes straight from his book: the notions of equilibrium states and reversible heat-engines. As Fermi shows, using the second law twice, the mere ability to order temperatures as hotter or colder makes it possible, through purely mechanical measurements (using those 'gifts of nature'), to fix their difference in a purely thermodynamic absolute scale of temperature. It is defined independently of ideal gases but agrees with their readings when they are far from liquefaction. It is also noteworthy that the boiling and freezing points of water are used to define the Celsius scale of temperature but only one of them to define the absolute scale; the second point is the absolute zero, which is universal in being the same whatever the substance.

Bearing in mind that amounts of heat and work are, as Joule showed, interchangeable, let's now consider the efficiency of an idealized Carnot heat engine. In its first stage, let the working medium take up the heat Q_2 from the furnace and the heat wasted in the condenser be Q_1. Then the amount of work done is proportional to $Q_2 - Q_1$. The expenditure of heat is Q_2, so the efficiency η is

$$\eta = \frac{Q_2 - Q_1}{Q_2}. \tag{8}$$

The denouement, obtained by Thomson's subtle arguments and presented by Fermi, is that one can define consistently the ratio of the temperatures T_2 and T_1 of the furnace and condenser in terms of Q_2 and Q_1 as follows:

$$\frac{T_2}{T_1} = \frac{Q_2}{Q_1}. \tag{9}$$

This definition of temperature agrees with the readings of ideal-gas thermometers and gives the absolute thermodynamic temperature scale.

With amounts of heat expressed by means of (9) in terms of absolute temperatures, the efficiency becomes

$$\eta = \frac{T_2 - T_1}{T_2}. \tag{10}$$

This shows that steam-engines, operating between 100°C and the typical ambient temperature say 20°C, have an efficiency at best of the order

$$\eta = \frac{373 - 293}{373} = \frac{80}{373}, \tag{11}$$

which is barely more than 20%. Newcomen's original steam engine had an efficiency of only 2% (but still managed to create England's "colossal power").

This may all seem to be far removed from the profound difference between past and future, but Carnot's remark, picked up by Thomson, "that there is an absolute waste of mechanical energy available to man when heat is allowed to pass from one body to another at a lower temperature" already indicates the way things will develop. Clausius took the next decisive step. In 1854 he published a mathematical expression for what he called 'the transformational content' of a body. It was entropy in all but name.

He considered a system that, like a Carnot heat engine, can undergo cyclic transformations from and back to a given initial state. However, instead of being brought into contact with just two heat reservoirs there can be many having the absolute temperatures $T_1, T_2, \ldots, T_N$. At them the system takes up or deposits the amounts of heat $Q_1, Q_2, \ldots, Q_{\mathrm{N}}$, these being counted positive or negative accordingly. Moreover, Carnot's strict restriction to reversibility is relaxed. Both reversible and irreversible heat transfers, with finite temperature differences, are allowed. In the latter case, the temperatures that appear in the calculations are always those of the heat reservoirs. Finally, the system needs at least two 'boxes' containing a medium of some kind that can take up and give up heat which can be transferred between them and into them individually from the reservoirs.

Under these conditions, Clausius proved that, for a cyclic process in which all heat reservoirs are visited, the sum

$$\frac{Q_1}{T_1} + \frac{Q_2}{T_2} + \ldots \frac{Q_N}{T_N} \tag{12}$$

is negative unless all heat transfers are reversible, in which case the sum is exactly zero. The proof is one of the wonders I mentioned. It involves an additional heat reservoir at an arbitrary temperature T_0 and as many Carnot heat engines as the number, N, of reservoirs that the system 'visits'. They are constructed precisely to enable them to 'ferry' as much heat between the T_0 reservoir and each of the N reservoirs to ensure that they are all returned to their original state. So too is the system.

The only outcome of the whole cyclic process is that heat has been taken from the source and work may have been done by the heat engines. As the calculations show, that will be the case if the sum (12) is positive. But that would violate the second law (here used, in Kelvin's form, for the third time in Fermi's account) because work would have been done by extracting heat from a source, the T_0 reservoir, at a uniform temperature. Therefore, the sum (12) can only be zero or negative. It is easy to show that the value zero is only possible if all heat transfers are reversible. If any are irreversible, the sum will be negative. This is when heat is allowed to flow from a hotter to a colder medium without doing any work. Carnot was so close to the whole story.

In arriving at the concept of entropy and its implications, Clausius initially considered transformations from one equilibrium state of a system to another solely by reversible transformations. The key thing that led to his success was *to take no account* of any work that might be done by or on the system. He just considered the amounts of heat transferred into or out of its medium and, critically, *the absolute temperatures* at which this happens. Suppose the amount of heat Q flows reversibly into the medium at the absolute temperature T. Then Clausius called

$$\frac{Q}{T} \tag{13}$$

the transformational content. This makes sense—the medium has been changed and one might like to quantify the change. The subtlety is that it is not the obvious quantity Q, the amount of heat, that appears in the definition but Q divided by the absolute temperature. This makes (13) not simply an amount of heat but something that also reflects its nature and measures, as I learned from Peter Atkins, its quality, or rather lack of quality—the magnitude of (13) is greater the lower the temperature, and with low temperature one can do less than with high temperature.

As I mentioned in the previous chapter, increase of entropy is very often characterized as an increase in disorder. Although somewhat anthropocentric in origin, decrease in quality is a useful complement. Another, which is particularly helpful and closely related to Thomson's dissipation of mechanical energy, is decrease in concentration. An example often given for this concerns photons (the 'particles of light' whose existence Einstein was the first to recognize). The earth is bathed in a steady stream of high-energy photons that come from the sun and are absorbed during the day by plants, by soil and, often carcinogenically, by sun-bathing humans. During the night approximately twenty times as many low-energy photons are radiated from the earth into space. The concentration of the associated energy is reduced by a factor ten.

From now on I will not use 'transformational content' for (13) and instead Clausius's later coining entropy. Very often he would consider infinitesimal additions of heat and express the corresponding infinitesimal increase of entropy by the equation

$$\mathrm{d}S = \frac{\mathrm{d}Q}{T}. \tag{14}$$

This is one of the great equations of physics. I have seen it suggested that Clausius chose the letter S for the new concept of entropy to honour Carnot, taking the first letter of Sadi. It's a nice thought and surely appropriate. The notation has stuck.

Equation (14) casts a very interesting light on what happens in an idealized heat engine. Carnot believed the amount of caloric would merely pass through the working medium and remain unchanged—it would be conserved. But in fact what remains unchanged is the amount of entropy in the world. In stage 1 of the cycle the amount of entropy Q_1/T_1 flows into the working medium from the furnace, assumed to be a heat reservoir of infinite capacity; in stages 2 and 4 there is no heat exchange and therefore no change in entropy: in stage 3 the entropy amount Q_3/T_3 is transferred from the working medium to the refrigerator. Because some of the heat energy Q_1 was used to do work in stages 1 and 2 (and was thereby "put out of existence"), the quantity Q_3 transferred to the refrigerator is less than Q_1. However, T_3 is less than T_1 by precisely the amount that ensures $Q_1/T_1 = Q_3/T_3$ and thus equality of the entropy that enters and leaves the working medium. The entropy which the furnace sheds is exactly taken up by the refrigerator.

There is no change in the entropy of the universe. If we accept that entropy increase is synonymous with increase of disorder—whether this is always so is something I have already questioned and discuss further in *The Janus Point*—then, although not all the energy extracted from the furnace can be exploited, the work done in a Carnot cycle does not lead to any increase of disorder. You will recall that, at the end of chapter 1, I noted that, considered objectively and without anthropocentric distortion, a Carnot cycle stopped at the half-way has done work without any wastage. And a complete cycle changes the shape of the universe; we now see it does that without increasing disorder.

We have now very nearly reached the goal which Clausius had had in mind since his paper of 1850: to find a quantity that characterizes the change in some system independent of the actual succession of states through which the system had been carried. The difficulty, any 'exterior work' that might or might not be done by or on the system, is eliminated in equation (14). It makes no reference to any such work. The 'interior work' that heat addition might do by breaking up molecules or simply making them move faster was no concern to Clausius.

The final step to establishing the existence of a new 'route-independent' state function is to suppose that the considered system is carried by heat transfer from an initial equilibrium state A to another equilibrium state B by infinitesimal reversible steps. Then the change in entropy, ΔS, is

$$\Delta S = \int_A^B \frac{\mathrm{d}Q}{T}. \tag{15}$$

The integral sign $\int$ means simpy that one is adding up all the individual infinitesimal contributions $\mathrm{d}Q/T$. The right hand side of equation (15) is the infinitesimal form of (12). Clausius's triumph was his proof that whatever route one took, always in reversible steps between between A and B, the value found for ΔS will be the same. One can also imagine going round in a closed loop of such reversible steps. Then one will find that $\Delta S = 0$.

One can now choose any equilibrium state of the system as reference and give it a nominal value S_0 of the entropy. This is equivalent to measuring altitude from, say, Everest's peak. Then the entropy of any other state is defined relative to it. It is also easy to show that if the heat transfers in (15) are irreversible, the corresponding value of (15) will be greater than for a reversible process. Clausius's entropy is indeed like altitude. If you go for a hike in the mountains along any route and come back to the point from which you set out, you have obviously come back to the same altitude.

Now we come to the entropic equivalent of Thomson's universal tendency in nature to the dissipation of mechanical energy, which, as we know, can be stored in heat. Let the two equilibrium states A and B of the sytem have entropies $S(A)$ and $S(B)$ and suppose any transformation, reversible or irreversible, between A and B. It is easy to show that

$$\int_A^B \frac{\mathrm{d}Q}{T} \tag{16}$$

is less than or equal to $S(B) - S(A)$, with equality holding only when the transformation is reversible. Now comes an argument that, at least as presented by Fermi, is rather curious. He says that if we consider a completely isolated system it clearly cannot take up or give out any heat, so that all the dQs in (16) are zero. This in turn means that $S(B)$ is greater than or equal to $S(A)$. The conclusion is that *for any transformation occurring in an isolated system the entropy of the final state can never be less than that of the initial state.* If the entropy does not stay the same, it must increase. I say the argument is rather curious since all the calculations up to this critical step have been made under the assumption that something does actually happen—heat is exchanged. A conclusion, fatal as we will see in its implications for the universe, seems to be based on *nothing happening*!

In fact, Fermi's two subsequent examples of transformations in an isolated system are so persuasive one cannot doubt the italicized conclusion even if the argument leading to it is strange. One involves the generation of heat by friction, while the other, which fits the situation rather better, relies on the possibility that the system is composite in nature, taking the form of a collection of 'boxes' that are both thermally insulated from each other as well as from the outside world. If two adjacent boxes have different temperatures and the insulation between them is removed, it is clear that heat will flow from the hotter to the colder box. This is a good example of a transformation that does increase the entropy of the complete system.

Fermi also draws another important conclusion: when an isolated system is in the state of maximum entropy (consistent with its energy), it cannot undergo any further transformation, since that would decrease its entropy. Thus, *the state of maximum entropy is the most stable state for an isolated system.* It is in the state of heat death. The extent to which such an argument can be applied to the universe is a major topic in the second half of *The Janus Point*.

For what comes shortly, what Fermi says about the possibility of defining entropy for nonequilibrium states or inhomogeneous systems is important. First, he notes that in the definition of entropy it is critical that the initial and final states are equilibrium states. This is because the passage between them by reversible transformations is necessarily through a sequence of equilibrium states. By continuity, the initial and final states must also be equilibrium states. However, Fermi then comments that in many cases it is possible to define the entropy even for nonequilibrium states. Consider, for example,

> a system composed of several homogeneous parts at different temperatures and pres-

> sures. Let each part, however, have a uniform temperature and pressure. If the different parts are in direct contact with each other, the system will evidently not be in equilibrium, since heat will flow from the hotter to the colder parts, and the differences of pressure will give rise to motion. If, however, we enclose each part in a thermally insulating rigid container, our system will be in equilibrium, and we shall be able to determine its entropy.

For comparison, when Clausius considered the same issue in 1862 he said that if the body is not of uniform temperature throughout then the expression for the entropy

> must not be referred to the entire body, but only to a portion whose temperature may be considered as the same throughout; so that if the temperature of the body varies continuously, the number of parts must be assumed as infinite. In integrating, the expressions which apply to the separate parts may be united again to a single expression for the whole body.

As Clausius was breaking new ground, it's not surprising that he did not anticipate the 'thermally insulating rigid container' that Fermi requires for each homogeneous part. To a significant extent, rigorous thermodynmics treats highly idealized situations. This is seldom a problem. In a laboratory, it is possible to construct containers that, for the states of matter being studied, are to an excellent approximation thermally insulating and rigid. But we do not see that sort of thing around us in the universe at large. And we do not see it in Clausius's words either. This must raise doubts about that part of his arguments.

In fact, there's more to this than Fermi's requirement that there be thermally *insulating* and *rigid* containers. When he emphasises the significance of a state of maximal entropy, he talks about *isolated* systems. Similarly, when discussing entropy and the arrows of time, many authors say their results apply to 'isolated' systems when, strictly, they should say 'insulated' or 'confined'. The distinction is important because conclusions that certainly hold to an excellent approximation under laboratory conditions are then applied to the universe, which is said to be the isolated system *par excellence*. But, it surely is not confined within a rigid insulating container.

Let us continue. Clausius was very good at making important points clearly. As I already mentioned, he was also not averse to making sure their deep significance would be noted. Like Thomson, he knew the role well chosen words can play and had his eye on posterity. In fact, it was only in his 1865 paper that Clausius made the shift from the 'transformational content' that he had hitherto used. He said it could still be used but

> I hold it to be better to borrow terms for important magnitudes from the ancient languages, so that they may be adopted unchanged in all modern languages, I propose to call the magnitude S the *entropy* of the body, from the Greek word $\tau\rho o\pi\eta$, *transformation.* I have intentionally formed the word *entropy* so as to be as similar as possible to the word *energy*; for the two magnitudes to be denoted by these words are so nearly allied in their physical significances, that a certain similarity in designation appears to be desirable.[27]

[27] In ancient Greek, *entropy* would mean 'within transformation' and *energy* 'within work'.

The passage just quoted prepares the way for the dramatic end of the paper. Clausius says he wants to allude, at least briefly, to a subject of which "even a brief statement may not be without interest, inasmuch as it will help to show the general importance of the magnitudes which I have introduced". The second law, he says, asserts that all transformations in nature may take place in one direction spontaneously but not in the other. He says this "leads to a conclusion to which W. Thomson first drew attention" and cites the 1852 paper discussed in the previous chapter. Clausius ends his final significant paper on the second law with these words:

> If for the entire universe we conceive the same magnitude to be determined, consistently and with due regard to all circumstances, which for a single body I have called *entropy*, and if at the same time we introduce the other and simpler conception of energy, we may express in the following manner the fundamental laws of the universe which correspond to the two fundamental theorems of the mechanical theory of heat.
>
> 1. *The energy of the universe is constant.*
> 2. *The entropy of the universe tends to a maximum.*

Since maximal entropy corresponds to thermal equilibrium and heat death of everything, it's hardly surprising that the final words, expressing the second fundamental law of the universe and coming 13 years after Kelvin's widely noted paper, created such an impression. Early civilizations had the concept of inescapable fate; it reappeared in the early modern age in Calvin's doctrine of predestination. However, in both these cases, the fate need not be dreadful. Even if they had no influence on the matter, Calvinists could still hope they would find themselves after death among the blessed in heaven rather than among the damned in hell. In contrast, Clausius's bleak statement offered no hope for the universe. It was widely discussed in educated society and contributed to the discussion about entropic creation that I mentioned earlier. The scientific impact was considerable. When Willard Gibbs [28] published in 1876 the first part of his pioneering 300-page article "On the equilibrium of heterogeneous substances" he placed at the head of it as motto Clausius's 'two fundamental laws of the universe'.

The inevitability of entropy increase of the universe has been widely accepted since Clausius's ominous pronouncement. The ineluctable end in heat death seems to rest on remarkably secure arguments. I have already quoted Einstein on his confidence in the durability of the thermodynamic laws. In Gifford Lectures in 1927, in which he coined the expression 'the arrow of time', Arthur Eddington said "The law that entropy always increases, holds, I think, the supreme position among the laws of Nature ... if your theory is found to be against the second law of thermodynamics I can give you no hope; there is nothing for it but to collapse in deepest humiliation."

I don't want to give you the impression that I aim to overthrow all of thermodynamics and statistical physics; far from it. I do feel confident that there is something irreversible

[28]Gibbs's contribution to thermodynamics and statistical mechanics will be important a little later in the book. His 1876 paper greatly extended the scope of thermodynamics by making it possible to treat substances with different chemical properties. In 1873, his paper "A Method of Geometrical Representation of the Thermodynamic Properties of Substances by Means of Surfaces" entranced the great James Clerk Maxwell. Maxwell enters our story in the next chapter.

about the behaviour of the universe; I see no guarantee that in the far future it will be hospitable for humans. But it might still be as beautiful as the brook I like to visit to watch the falling raindrops and what the create. The otherwise circumspect Clausius may have got carried away when, perhaps encouraged by Thomson's paper, he turned his thoughts to the universe. Is it easy, sensible or even possible to define an entropy of the universe? When he defined the entropy of an inhomogeneous body, he simply said "the number of parts must be assumed as infinite. In integrating, the expressions which apply to the separate parts may be united again to a single expression for the whole body."

In Carnot's work, it is critical that the working medium is throughout confined in a box. Clausius made truly great discoveries by studying such a medium. Only if confined can it be carried through a sequence of equilibrium states. Clausius's definition of the entropy of an inhomogeous system is much less rigorous than Fermi's. How can Clausius be sure an entropy can be defined at all for the universe? It is manifestly very inhomogeneous. Even with the telescopes available in the 19th century, it did not look as if it is compartmentalized into strictly homogeneous regions separated by rigid heat-insulating walls.[29]

Simple-minded application of thermodynamic laws to the universe might be dangerous. What was it Einstein said about the durability of thermodynamics? Its laws will continue to hold "within the framework of applicability of its basic concepts". Does that framework include the universe?

A final comment about the direction of time, especially as applied to an isolated system like the one considered by Fermi in the final step of his discussion. It's a universe unto itself. Why should the state with two boxes with different temperatures be supposed earlier than the one in which they have equal temperatures? Neither Clausius nor Fermi are with us so we cannot ask them. Of course, if Fermi's system is in a laboratory, the direction of processes within it can be defined relative to clocks in the laboratory, to the observed motion of the sun and moon or, indeed, to the background arrow that we see all around us in the mulitude of unidirectional processes in the universe at large. We are inside our universe and can take our direction of time from it. But the universe itself in not inside another.

[29]In fact, most scientists throughout the 19th century had only the vaguest notions about the nature of the universe. There was little interaction between the thermodynamicists and the astronomers, who were at least learning a lot more about the stars if not yet about other galaxies. It is not even clear whether by 'universe' Thomson and Clausius meant all the known and as yet undiscovered celestial bodies or just the solar system. In fact, in his pronouncement on the fate of the universe, Clausius actually used the word *Welt*, which can be translated as either universe or world, the latter clearly suggesting something more than just the earth but perhaps no more than the solar system.

9 Statistical Mechanics

Atomic ideas arose in antiquity, but the motions attributed to atoms were never described in precise mathematical terms. Even after Newton had formulated laws of motion for 'bodies' the efforts to put atomic ideas on a sound footing were fitful. In 1738 the Swiss Daniel Bernoulli explained how the pressure exerted by a gas could be explained by impacts of its molecules on the walls of its container. His results also indicated that the temperature of a gas could be a measure of the *vis viva*, i.e., kinetic energy, of the molecules of which it is composed.

It was only in the 1850s that these early ideas were seriously revived and taken significantly further, initially by Clausius with a paper in 1857 bearing the title "The nature of the motion which we call heat".[30] Although he had effective precursors in the first half of the 19th century, in particular the Britons John Herapath, John Waterston and Joule, Clausius was the effective founder of statistical mechanics, the attempt to explain macroscopic phenomena by microscopic models. Its flowering that then began continues unabated to this day. I'm only going to describe the aspect of the subject that is needed for this book. However, it is worth noting here that Clausius made a clean distinction between, on the one hand, the general axioms on which he and Kelvin had founded what is now called phenomenological thermodynamics and, on the other, statistical mechanics. This was characteristic of Clausius. He rode two horses at once with some skill; as we will see, that created difficulties later in the 19th century for the ultimate hero of statistical mechanics, Ludwig Boltzmann.

It's fortunate for me that the key insight which led to the microscopic interpretation of entropy was developed using the simplest conceivable models of ideal gases in a box. The simplicity of the macroscopic laws such gases were found to obey meant that it was relatively easy to construct models that explained the observed phenomena. Daniel Bernoulli had already made progress in that direction. Clausius, James Clerk Maxwell, whose work I will soon describe, and Boltzmann took things much further. The simplest models of ideal gases had two more or less equivalent forms: either tiny hard balls that bounced off each other elastically in accordance with the laws that Huygens had found (chapter 4) or else point particles that interacted through short-range forces that also ensured conservation of energy and momentum so that outside the interaction range the effect would be the same as an elastic collision of the hard balls. Critically, the balls or particles were always assumed to bounce off the box walls elastically.

Two thirds of Clausius's 1857 paper consists of a general introduction in which he proposes an overall qualitative picture of how he imagines the three states of matter, above all gases, to be like. It is based fair and square on the idea, by then well supported by experiment, that the different chemical elements consist of atoms that have different masses and that, under suitable conditions, the atoms of different elements can combine into molecules to form compounds. Reading Clausius, you soon realize that he followed the literature closely, both in chemistry as well as in the study of gases and liquids. He certainly had an eye for experimental results that could inform theory. One also gets the sense that Clausius had a pretty good feel for when a theoretical idea could actually be developed to the stage at

[30]Brush adopted this as the title of his two-volume historical study of the kinetic theory of gases, which I recommend strongly.

which verifiable predictions could be made. The opening pages of his 1857 paper mark an impressive transformation of earlier inchoate atomic ideas into a clear picture of the three states of matter, above all gases. He had an almost unerring ability to make his conceptual ideas just precise enough to serve as the basis for precise theoretical calculations without being overburdened with too much detail. Since they will be sufficient for my purposes, I will only describe results that relate to ideal gases.

Beginning with gases, he says he shares with August Krönig, whose 1856 paper had prompted Clausius to publish ideas he had been developing already before 1850, the view that the molecules of gases "move with constant velocity in straight lines until they strike against other molecules, or against some surface which to them is impermeable". He agrees that this translational motion can explain the pressure of gases (as, in fact, Bernoulli had already shown 120 years earlier).

However, in a decisive advance, Clausius argues that "this is not the only motion present". The point is that, since molecules, being composed of atoms, are conceived to be extended, they must be capable of rotational motion which can be increased or decreased whenever two molecules collide. There can surely also be vibrational motion corresponding to motion of the atoms within a molecule relative to each other. Although the concept was not coined until later by Maxwell, we can usefully introduce here the general notion of mechanical *degrees of freedom*. They come in two kinds, one related to position, the other to translation from place to place as expressed by velocity or, more precisely, momentum (velocity times the mass of the considered body). If we consider a point particle, its position will be defined by three coordinates. These are positional degrees of freedom. There are also three momentum degrees of freedom associated with motion along the three coordinate directions. However, if one has a molecule made up of two or more atoms, there are more degrees of freedom. For example, a diatomic molecule will be like a dumbell that can rotate and vibrate, so there are further possible motions.[31] In the final third of his paper, Clausius translates these qualitative ideas into concrete calculations for gases and obtains important numerial results, which I shall shortly describe.

However, we can already present the picture at which he arrives. First, the pressure and absolute temperature of an ideal gas will be proportional to the average translational *vis viva* (kinetic energy) of its molecules. Moreover, the frequent collisions between molecules will ensure that the average *vis viva* of the rotational and vibrational motions will, once an equilibrium state has been achieved, be equal to the average *vis viva* of the translational motion.[32] Clausius also argues that in a mixtures of gases in equilibrium all molecules will

[31]Until quantum mechanics was discovered, physicists had great difficulty in understanding how the different kinds of degrees of freedom manifested themselves in quantities that can be measured macroscopically. Thus was done through the capacity of gases to take up heat, expressed as the rise in their temperature under conditions of either constant pressure or constant volume. In the first case, the gas would expand and do work. This would mean its temperature would rise more slowly in the first case than the second. The difference is expressed in terms of the ratio of what are called specific heats at constant pressure and volume. The ratio was found to be 5/3 for monatomic gases and 7/5 for diatomic gases. This result could not be explained unless the only degrees of freedom possessing kinetic energy were the translational and rotational ones. Physicists had to assume that for some reason, presumably rigidity in the binding together of the atoms, vibrational motion was not excited.

[32]Clausius emphasizes more than once that the molecules will have a range of velocities and that it is

on average have the same *vis viva*. This result is referred to as the *equipartition of energy*. It means that the heavier molecules will on average have lower speeds than the lighter ones. In the gaseous state, one can expect the average distance between molecules to be many times greater than the diameters of the molecules.

The reason why the very simple ideal-gas model sufficed to yield the good conceptual understanding of entropy that I discuss in this chapter and later in the book is that the phenomena related to entropy concern the collective properties of systems which possess *many* degrees of freedom. As far as the concept of entropy is concerned, the actual nature of the degrees of freedom plays no role; what counts is the bare existence of many degrees of freedom subject to dynamical laws of a certain kind. Moreover, the striking universality in the behaviour of different substances used as media in Carnot-type heat-engines and in the experiments of Joule on the mechanical equivalence of heat is explained by fact that all degrees of freedom are subject to mechanical laws and accordingly possess kinetic and potential energy, the transformations between which are subject to the fundamental law of energy conservation.

Clausius was certainly right, when justifying his coining of entropy, to make it have "a certain similarity in designation" to energy because the two concepts are "so nearly allied in their physical significances". By the time his claim was made (eight years after his 1857 paper), the new science of statistical mechanics had already notched up several triumphs that strongly supported the ancient atomic hypothesis. There was now a clear microscopic picture that could *explain* observed macroscopic changes. At the microscopic level there were precise mechanical laws, above all the conservation of energy, expressed in simple mathematics that matched the quantitative effects observed macroscopically. There was very encouraging evidence that the dream of the pre-Socratic philosophers—to *understand* nature—was being realized. And the overarching concepts involved in this process, which is still ongoing, are energy and entropy. They reflect respectively law and number. About law and energy enough has been said; about number and entropy, we will see that number comes in critically in two different ways. First, in a box of gas there are *a finite number* of discrete things—atoms or molecules— and, second, there are different ways in which they can be arranged.

To give you some idea of the triumphs of statistical mechanics, I will sketch the arguments that Clausius gave to explain pressure and its relation to kinetic energy. It's one of the simplest successes of statistical mechanics and will give you an idea of the power that the simplest of ideas possess. At the end of his *Theory of Heat*, Maxwell gives a beautifully lucid account of several more statistical-mechanical triumphs. I will include two of them in this chapter.

In a gas that is confined, its molecules can be expected to be moving randomly in all spatial directions. The gas exerts pressure on the wall of its container because the molecules are constantly colliding with the wall and bouncing off it. The things that determine the pressure are the number of molecules that strike the wall in unit time; the momentum of the molecules; and the angle with which they approach the wall. This last requires a calculation to determine the effective average of all possible angles, which can be made

---

only the average *vis viva* that will be proportional to the temperature and pressure of a gas in equilibrium. However, he made no attempt to establish what should be the distribution about the average. As we will see, it was Maxwell, in his most important contribution to statistical mechanics, who took up that challenge.

under the assumption that they are randomly distributed. The result obtained shows that the pressure is what one would expect if exactly one third of the molecules collide head on with the wall while the remaining two thirds move parallel to the wall and therefore do not strike it and make no contribution to the pressure. Since space has three dimensions, this is what one would expect intuitively.

The argument which shows that the pressure depends on the average energy $mv^2/2$ and not, as one might at first think, on the momentum mv, is both neat and nice. It is true that the impulse given to the wall by each molecule depends on its momentum, but one must also remember that the total impulse depends on how many molecules hit the wall in unit time. This depends on their velocity, which means that mv must be multiplied by v, becoming mv^2.

Now it had long been known, in accordance with the ideal gas law $PV = RT$, that, at constant volume, the pressure P of a gas is proportional to its absolute temperature T. It follows from this that $mv^2/2$ is proportional to T: since T, along with P and V, can be measured, this means, as we shall see, that from measurements of these quantities the value of v^2 can be determined. This how it is done.

Two key quantities in the calculation are the number n of molecules in the confined quantity of gas and the mass of each molecule. By the mid 1850s, the discoveries in chemistry had made it abundantly clear that all the molecules of a particular gas must be identical. In particular, they must have the same mass m. The total mass of n identical such molecules is therefore nm, and its weight is $W = nmg$ where g is the force of gravity. It then follows that, if all the the molecules have the same velocity v

$$v^2 = \frac{3gPV}{W}. \tag{17}$$

All the quantities on the right-hand side of this equation can be measured, from which Clausisus deduced the following average speeds of three sample gases at the temperature of melting ice:

for oxygen :	461 metres/sec
for nitrogen :	492 metres/sec
for hydrogen :	1844 metres/sec

Clausius emphasized that these are average velocities; some of the molecules may have velocities markedly different from the average. Although Clausius knew that Joule had published a paper which developed ideas similar to his own, he did not know its contents; in fact, Joule had obtained a broadly similar velocity for hydrogen. It may be noted that in air, which is a mixture of oxygen and nitrogen, the speed of sound v_S is not hugely different from the above values. In dry air at 20°C, it is 343 metres/sec.

Note also the critical role that the weight W plays in the above estimates of the translational speeds. First, it supplies experimental input without which nothing could be predicted; second, it will play a decisive role a little later in our story by helping to put the atomic hypothesis on a firm footing. The point is that, if the hypothesis is to be taken seriously, it will require reliable determination of things like the mass (which is proportional to the weight) and size of the putative atoms and molecules. When that has been determined, we can start to ask a question that has an amazing answer: "How many atoms are there in my little finger?" We will then also see how 'gifts of nature'—things in the macroscopic world with which we can measure and objects that we can count—enable us to peep through the window of theory into the microscopic world and count things there.

Let us continue. With the first part of his paper, Clausius had not obtained essentially new results though nobody hitherto had derived them so clearly. In the remainder of his calculations he broke new ground. This was based upon his earlier qualitative arguments that one must expect the total energy of the molecules in a gas to be divided between translational motion and motion associated with rotation and vibration of each individual molecule. The question then arises naturally of how the total energy is distributed between the three different forms: translational, rotational, and vibrational.

Three things made it possible to get a handle on this issue. First, various people, above all Joule, had measured the mechanical equivalent of heat. This could be used in both directions. If you knew by how much the temperature of a gas is raised by the addition of a certain amount of heat, you could work out by how much the average speed of the translational motion of its molecules is increased. As anticipated in footnote 31, a second input can be gained from the different amounts by which the temperature of a gas rises when heat is put into it in one of two possible ways: with the volume kept constant and with the pressure kept constant. The former, the specific heat at constant volume, is always greater than the latter, the specific heat at constant pressure, because some of the heat applied is expended on mechanical work if the volume is allowed to change. The third and final source of experimental information comes from the ideal gas law $PV = RT$. Using the three available relations, Clausius could determine how much of the heat put into a gas had been converted into the heat of translational motion of its moleclues and how much must have gone into motions of the constituents.

The results, obtained for two broad classes of gases, were illuminating. First, for simple gases that behave in a regular way and also for gases that arise through chemical reactions in which there is no reduction of volume he found that the fraction of heat that goes into translational motion is 0.6315. For the gases for which the chemical reactions that lead to a reduction of volume the fraction is smaller with, moreover, the fraction being smaller the greater the reduction. His conclusion, for any gas, was that

> the *vis viva* of the translatory motion does not alone represent the whole quantity of heat in the gas, and that the difference is greater the greater the number of atoms of which the several molecules of the combination consist. We must conclude, therefore, that besides the translatory motion of the molecules as such, the constituents of these molecules perform other motions, whose *vis viva* also forms a part of the contained quantity of heat.

Out of remarkably simple calculations, Clausius has already obtained a surprisingly clear picture of what a gas must be like. It consists of molecules moving in space between collisions, the molecules themselves being composed of a complex of atoms that can move relative to each other in vibratory motion as well as rotating as a whole. Plato's cave metaphor in which he pictured mortals chained within a cave and only able to see shadows cast by beings that pass in front of the cave opening behind them turns out to have been too pessimistic.

The shadows that Clausius and others sought to understand were not cast from behind but came from within containers and were cast as phenomenological macroscopic effects such as temperature and pressure. Measurement and theory proved Plato wrong. We can make great progress even if it never leads to a definitive understanding of the universe. But it is a book that, as Galileo confidently predicted, we can begin to read.

Very soon after Clausius had published his paper, the Dutch meteorologist Christophe Buys Ballot (1817-1890) pointed out what seemed to be a severe difficulty.[33] He said that if the molecules in gases move in straight lines at great speeds adjacent volumes of gases should mix rapidly but "how then does it happen that tobacco-smoke, in rooms, remains so long extended in immoveable layers." Further, "if sulphuretted hydrogen or chlorine be evolved in one corner of a room, entire minutes elapse before they are smelt in another corner, although the particles [at the speeds estimated by Clausius] must have had to traverse the room hundreds of times in a second."

In response, Clausius granted that he had not specified sufficient details of his model. He suggested that molecules should have a certain effective diameter and be separated from each other by a definite average distance. They would then travel a certain average distance, their *mean free path*, before having a collision that would deflect them significantly from their rectilinear motion. Because of such collisions, molecules would undergo many deflections in passing from one side of a room to another. For the same molecular speed, the time taken to do that would be much greater than if there were no collisions. If, for example, the molecules are separated on average by 8 times their diameters they will travel 62 times the average separation before undergoing a strong collision. On this basis, Clausius could argue that his model showed one could satisfy three stringent conditions at once: explanation of the slow diffusion effects to which Buys Ballot had drawn attention; fulfilment of the ideal gas laws; and the deviations from those laws under high pressures and densities that had been found recently in extremely precise measurements by Henri Regnault (in whose laboratory in Paris the youthful William Thomson had gained that valuable 'work experience'). It was a triumph. The notion of the mean free path entered the conceptual arsenal of theoreticians.

Clausius's two papers, of 1857 and 1858, provided not only explanations of physical effects but also gave theoreticians something definite on which their imaginations could work. With his laws of motion and theory of universal gravitation, Newton had opened a wide window onto the cosmos; Clausius did the same for the window into the microscopic world. Two theoreticians took a really good look through it: in the remainder of this chapter I'll describe what Maxwell saw; in the next, the triumphs and trials of Boltzmann.

I won't say much about Maxell's contributions to statistical mechanics. They were numerous and very important for the development of the subject but concern details that mostly are not directly related to the main concern of this book. One however is vitally important. I will mention it shortly but only say what it is in the next chapter, when we

[33] Buys Ballot made what must rank as one of the most delightful and simultaneously hilarious experiments in physics. It was to test the effect predicted in 1842 by Christian Doppler. In the modern age, we are all familiar with the way the pitch of police and ambulance sirens drops markedly as the emergency vehicles pass us. Of course, no such vehicles existed in the 1840s. Buys Ballot arranged for a band of brass musicians to stand on an open carriage of a steam train on the Utrecht–Amsterdam line. People on the station platform through which the train passed clearly heard that the pitch of the music was higher as the train approached and lower as it receded. If a reconstruction of the actual experiment is not on YouTube, it should be.

come to consider what Boltzmann did with Maxwell's contribution.

Maxwell published his first paper on the kinetic theory of gases in 1860. It opens with the comment that so many properties of matter, especially when in the gaseous state, "can be deduced from the hypothesis that their minute parts are in rapid motion ... that the precise nature of this motion becomes a subject of rational curiosity". Maxwell's style contrasts strikingly with that of Clausius; he follows the tradition of Newton in the *Principia* of proving rigorously a string of propositions from which profound conclusions are then drawn in relatively few words. He notes that two essentially equivalent models can be adopted for the 'minute parts' of matter: they can be represented either as centres of force or as hard elastic bodies, it being evident, he says, that the same results will be obtained. Maxwell opts for the second model and considers two possibilities: the elastic bodies are exactly spherical or have a bounding surface that is not spherical. Besides reproducing results that had already been obtained, Maxwell's paper contained a result of fundamental conceptual significance as well as a prediction that was so counterintuitive Maxwell himself initially thought it disproved the kinetic theory of gases. However, in a famous experiment with his wife, he later confirmed the prediction. This boosted confidence in the theory, which by 1865 gained widespread acceptance in its broad outlines. Attention then turned to details into which we need not go.

To begin with this, it concerned the viscosity of gases. This is illustrated by a pendulum. If it is set swinging in air, its initial amplitude slowly decreases until it eventually comes to rest. This happens partly because of friction at its point of suspension and partly because the pendulum bob has to 'push aside' the air it encounters during its swing. This effect of the air's viscosity is analogous to the effort anyone must exert when walking through water up to their shoulders in a swimming pool. Like Maxwell, everyone assumed that the same pendulum in a gas like air would come to rest more rapidly the denser the air. To his great surprise, Maxwell found from his calculations that changing the density had very little effect. However, with his wife, he performed an experiment that confirmed the prediction.[34] This boosted confidence in the theory, which by 1865 gained widespread acceptance in its broad outlines.

Besides the remarkable vicosity result, Maxwell confirmed what Clausius had argued, namely that the mean energies of the two different kinds of degrees of freedom in his model—translational and rotational—should be equalized even if they are not so at some stage. Comparing his results for perfectly spherical hard elastic spheres with the results for nonspherical bodies, Maxwell concluded from the measurements of the specific heats of gases that his calculations seemed to be "decisive against the unqualified acceptance of the hypothesis that gases are such systems of hard elastic spheres".

Step by step, clever combination of theory and experiment was leading to an increasingly clear picture of atomic phenomena. Indeed, a striking by product of Maxwell's viscosity calculations was his ability to estimate the value of Clausius's mean free path. Taking for air the mean velocity of its molecules to be 1505 feet per second, he found that the mean

[34] It eventually transpired that an equivalent experiment had been made nearly two centuries earlier by Boyle using Robert Hooke's vacuum pump. Boyle had found, to his surprise too, that a pendulum in a chamber from which most of the air had been evacuated would swing at almost the same frequency as in one filled with air and that the time taken for it to come to rest was much the same in the two cases.

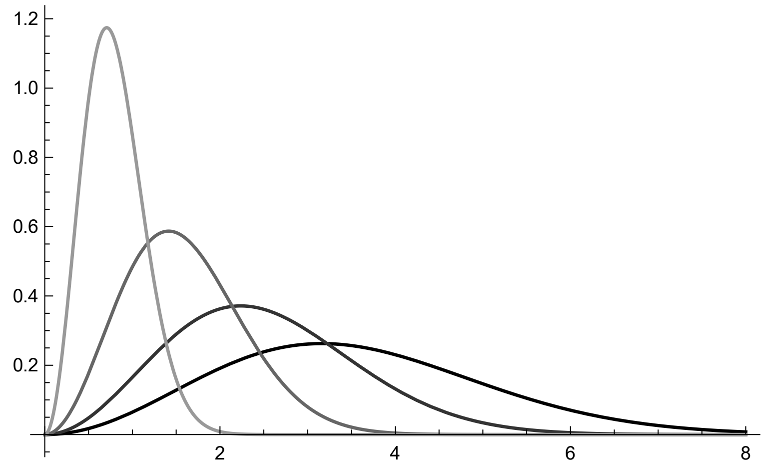


Figure 5: Maxwell velocity distributions.

distance travelled by a molecule between collisions is 1/447000th of an inch, each molecule therefore making 8 077 200 000 collisions per second. The extraordinarily large numbers involved in the atomic hypothesis were beginning to emerge. However, this did not yet yield an estimate of the diameter of an individual molecule—or the number of atoms in your little finger. Some further idea would be needed to get at that.

I will conclude this chapter with the major conceptual advance that, along with the viscosity result, Maxwell made in his first paper on statistical mechanics. Clausius had more than once emphasized that the molecules of a gas could not be expected to all have the same velocity; there must be some distribution of the velocities. However, he had not attempted to find the distribution. Maxwell attacked the problem head on and in barely two pages found the correct answer. He later admitted, in 1866, that his derivation was 'precarious' and provided a revised proof. The important thing is that the result has stood the test of time. Generalized by Boltzmann to include potential as well as kinetic energy, it provided the firm basis for all calculations for ideal gases in thermal equilibrium prior to the discovery of quantum mechanics, in which it still plays an important role. It is known as the Maxwell–Boltzmann distribution.[35]

Maxwell's key idea was that for molecules in one dimension their velocity distribution could be (half) the Gaussian distribution, widely known as the bell curve. In three dimensions, you need to calculate the velocities that follow from this assumption, taking into account the numbers of molecules in 'velocity shells' of radii from zero to the maximum possible value (when a single molecule has all the kinetic energy since it alone is moving). Examples for different values of the energy for the same number of particles are shown in Fig. 5. The horizontal axis is the energy of the particles, the vertical axis gives the number of particles that have that energy. The smooth curves shown in the figure are the approximations that hold when the number of particles is very large. I need to point out that the curves in Fig. 5 represent two quite distinct things, one of which is more fundamental than

[35]Like Maxwell's judgement on his first derivation, my summary of the argument leading to it is 'precarious'. I hope to improve it if and when this text is published as a book. Once again I have to say "Caveat lector".

the other.

Before I discuss them I need to point out the importance of units in the measurement of physical quantities. One of the great services of the French Revolution was to introduce the standard metric system for measuring times, lengths and masses. The second was defined as the fraction $1/24 \times 60 \times 60$ of the earth's rotation period, the metre as a certain fraction of the earth's circumference and weights through the platinum–iridium kilogram held in Paris. In modern metrology, these units and others used to measure things like electric charge have been replaced by definitions in terms of more fundamental physical quantities. The most basic change is that the second is defined in terms of one of the frequencies of the radiation emitted by the caesium atom. This brings with it great advantages: first, quantum mechanics ensures that all caesium atoms are identical, which means that the standard second can be reproduced anywhere; second, the accuracy that can be achieved is remarkable. Time can be measured more accurately than any other physical quantity. At the time of writing, the final step to putting metrology on the most secure basis possible is a corresponding redefinition of the kilogram. Its hitherto existing definition means that the mass of the earth, indeed the universe, increases every two years when the platimum–iridium standard is cleaned and a few atoms are unavoidably removed.

In the light of these comments, now consider a single collection of N particles in a box. If they have an equilibrium distribution then it can be represented by any one of infinitely many curves of the form shown in Fig. 5. You can pass from one to the other simply by changing the units, for example, by replacing centimeters by inches. That clearly cannot affect the physics in the box. The situation is quite different if there are two boxes in which the temperatures are different. That's an absolute fact. Whatever units are chosen, there will be two distinct curves. If the units are changed, the curves will be shifted up and down together in figures like Fig. 5. This is related to the fact that only ratios, of temperatures in this case, are physically meaningful.

The more fundamental thing about the curves is related to the energy ratios of the particles within one and the same box. In whatever units have been chosen, let the average energy of the particles be E. Then one can ask questions like this: what fraction of the particles have energies between E and $2E$ or between $E/2$ and E. These fractions, being ratios, have values that are completely independent of the units that are used. If one determines the values of such fractions for any of the curves in Fig. 5, they always come out the same. All of the curves in the figure are said to be thermal, and they are identical as regards the fractions just described. Any temperatures associated with such curves are always relative to some arbitrary choice of units or to some other system that is in thermal equilibrium. One can say by how much one system is hotter than another, but without such a comparison or the existence of external units temperature by itself has no meaning. It relies on something extrinsic to the system considered. In contrast, the notion of a thermal state is well defined in intrinsic terms. This cannot fail to be critical if we are attempting to consider the whole universe. For it, everything must be defined intrinsically.

There's something more to say about the connection between Maxwell's result and the absolute thermodynamic scale of temperature. Through the latter, temperature is defined in terms of mechanical work quantified by means of macroscopic objects. Maxwell's distribution defines thermal equilibrium mechanically in terms of the kinetic energies of microscopic particles. Galileo's 'book' (the universe) contains macroscopic words—bodies and their motions. Galileo himself, Kepler, Huygens and Newton were the first acute readers of the sentences composed of these words. Using these sentences as guides, the first practitioners of statistical mechanics, above all Clausius and Maxwell, were able to read messages from another world.

10 Boltzmann and the Second Law. I

It was in 1872 that Ludwig Boltzmann published one of the most important papers in the development of statistical mechanics. Despite its significance, it soon led to controversy that continued to the end of Boltzmann's life and beyond his sad suicide in 1906. Indeed, arguments continue to this day.

The task that Boltzmann set himself is easy to understand. It concerned the reconciliation of the typical macroscopic observations made in thermodynamics, above all those of equilibration, with the atomic–molecular hypothesis. Since equilibration will be at the heart of our discussion pretty well throughout the remainder of this book, let's be quite clear what it means. You can easily see an example in a bath tub containing still water. If you disturb the water with a few violent movements of your hands, the waves you create soon subside. The state of equilibrium is reached. In thermodynamics, two adjacent thermally insulated boxes containing gases at two different temperatures provide a typical example. If an opening is made between the boxes, the gases mix and soon the temperature will have been equalized between them. Many different kinds of equilibration can occur. I do need to keep on making the point about systems in a box: equilibration can only take place if the system under consideration is confined in some way or another - like water in a bath tub.

I have suggested several times how the historical development of thermodynamics may have hindered resolution of the problem of the origin of time's arrows. It seems to me that a major factor was the concern with showing how a state that initially is not in equilibrium does tend to equilibrium. What was seldom considered was how, all around us, we find systems that are not in equilibrium. How did they get there?

It was only quite late in his studies, in the 1890s, that Boltzmann was forced to address this question. We will come to that. In 1872, his concern, precisely stated, was this: if the atomic-molecular hypothesis is correct, the question then arises of how equilibration takes place, specifically, what are the molecules of a gas doing as they pass from a macroscopically non-equilibrium state to equilibrium? It had been clear to Clausius and Maxwell that molecules with above average energies would, colliding with those with below-average energies, transfer some of their energy to slower molecules. Maxwell made some beautifully simple calculations of actual collisions in which energy transfer takes place. They showed how one could expect an average equilibrium distribution of the energies to be established and maintained. By plausible arguments Maxwell had found what this distribution should be. Striving for rigour, Boltzmann was more ambitious. He aimed to show that Maxwell had found the only possible equilibrium distribution—there could not be any others.

The method that Boltzmann developed relied critically on the as yet unknown but surely colossal number of molecules in any cubic centimeter of air at atmospheric pressure and room temperature. Just as Maxwell had done, this meant one could speak in terms of a well-defined density of molecules both in ordinary space and in the abstract space of their possible velocities or kinetic energies. Both would be described by distribution functions. Boltzmann made some simplifying assumptions about what he called the initial state. He assumed that, as a result of collisions, any initial state would very soon settle down into one in which the spatial density of the particles could be assumed uniform. He also assumed that the kinetic-energy distribution would soon become the same at each spatial point.

The task was then to show that whatever form the energy distribution function might have at some initial time it would necessarily evolve eventually to the Maxwell distribution and thereafter remain unchanged. Of course, this applied to the average macroscopic distribution function. At the microscopic level, small fluctuations could always happen. Boltzmann was able to find an equation that should govern the equilibration process and would indeed, whatever the initial distribution, lead to the Maxwell distribution. The equation soon became known as the Boltzmann equation and proved its utility and viability in many calculations involving countless different processes. It is one of the most important equations in physics. For example, cast in a form that takes into account quantum effects, it plays a hugely important role in the description of processes that take place in the early universe though, rather remarkably, it is used in that context to describe how the medium that permeates the universe *comes out of equilibirum.*

Apart from this case, which plays a central role in *The Janus Point*, the important thing about the Boltzmann equation is that, by its very construction, it has an inbuilt tendency to reduce deviations from the Maxwell distribution. Unlike the time-reversal symmetric equations taken to describe the collisions of microscopic particles, Boltzmann's equation is time-asymmetric.[36] It was some time before the implications of this mismatch sank in. When they did, they very soon prompted Boltzmann to his greatest contribution to physics. That will be the topic of chapter 12; in the remainder of this chapter I will say a bit more about his 1872 paper and the remarkable consequences of a device he introduced in it to simplify and clarify his equations.

Boltzmann's equation was not his only innovation. In one of his most important contributions, he introduced a quantity which later came to be called the H function. He proved a result about the behaviour of this function that became famous and is known as the H theorem. In essence, the H function is a measure of the deviation of the current energy distribution from the Maxwell distibution for the same total energy of the system. To get a picture of this in concrete terms suppose the curve of the current distribution superimposed on the Maxwell distribution. The two curves will not overlap—there will in general be closed areas between the two curves. The value of the H function is a measure of the sum of these areas. It will tend to zero as the current distribution gets closer and closer to the Maxwell distribution. Boltzmann interpreted such an effect as a microscopic mechanical explanation of the way the entropy of an isolated system tends to a maximum. The actual relationship between the H function and entropy will be an important part of chapter 12.

For the moment, it is sufficient to say that the calculations which Boltzmann made using the equation he had derived confirmed his anticipation. He was able to show that his H function did have an inbuilt tendency to decrease. Boltzmann was naturally delighted about this for two reasons. First, growth of entropy was profoundly significant both in fundamental natural science as well as for technical issues; an understanding of it at a deep level must mark a major advance (Boltzmann published his H-theorem paper seven years

[36]There was nothing wrong with the theory of individual collisions. The problem lay in an implicit assumption, made by both Maxwell and Boltzmann, that concerned the delicate issue of the probability of collisions. It was this that transformed the time-symmetric theory of the collisions into the time-asymmetric theory that Boltzmann inadvertently developed. It is nevertheless remarkable how valuable Boltzmann's theory, when applied appropriately, has proved to be.

after Clausius's epoch-making paper in which he coined the word entropy). Second, and not unrelated, Boltzmann was strongly attracted to the atomic hypothesis and welcomed what he regarded as evidence in support of it.

Let me now turn to the device Boltzmann introduced as to way to think about his distribution function. It turned out to be very valuable and has come to be called *coarse graining*. Instead of assuming that each particle could have any value of the energy, Boltzmann assumed that they could only have energies that are multiples of some very small quantity ϵ. Thus, the possible values of the energy for any individual particle would be

$$0, \epsilon, 2\epsilon, 3\epsilon, \dots\ p\epsilon, \tag{18}$$

where p (a very large integer) and ϵ are chosen such that $p\epsilon$ is equal to the total energy E of the system of N particles. If one particle should happen to have all the energy, then of course all the remaining $N-1$ particles could have no energy at all. In the general case, the energy would be distributed between many particles. Then there could be n_0 with no energy, n_1 with the energy ϵ, n_2 with 2ϵ, etc, up to n_p with energy E. Of course, not all of these numbers could have arbitrary values. They must satisfy two conditions. First, the total number of particles must be N; thus, the numbers n_1, n_2, etc must add up to N. Second, the total energy must have the fixed value E.

Choosing values for E, p, ϵ and all the n_1 one can obtain a value of the distribution function at some time. One can represent this pictorially by means of $p+1$ 'bins' arranged in a line into which one can place as many particles as have the energies $0, \epsilon, 2\epsilon \dots p\epsilon$, the first bin giving the number that have no energy at all. The heights of the bins correspond to the number of particles with corresponding energies. The differences of the heights from the Maxwell distribution are a measure of the value of Boltzmann's H function.

Now collisions of the particles in an actual gas will give rise to changes in the energies of the particles. In the coarse-grained picture, they will lead to a redistribution of the particles between the energy bins. Unaware of the implicit assumption he was making, Boltzmann derived a 'coarse-grained' form of his equation and, using it, showed how an initial non-equilibrium distribution would tend to the Maxwell distribution.

At this point, I want to quote a few truly remarkable sentences in which Boltzmann justified his coarse-graining method. I came across them when reading Boltzmann's 1872 paper. I always like to go back to originals. They are often clearer than second-hand accounts and give one insights into the great personalities who created the edifice of modern science. Having explained that, purely for mathematical convenience and clarity, he would assume the particles could only have the discrete energies (18), this is what Boltzmann said about the conditions he would impose:

> No molecule may have an intermediate or a greater kinetic energy. When two molecules collide, they can change their kinetic energies in many different ways. However, after the collision the kinetic energy of each molecule must always be a multiple of ϵ. I certainly do not need to remark that for the moment we are not concerned with a real physical problem. It would be difficult to imagine an apparatus that could regulate the collisions of two bodies in such a way that their kinetic energies after a collision are always multiples of ϵ. That is not the question here. In any case we are free to study the mathematical consequences of this assumption, which is nothing more than

> an artifice to help us calculate physical processes. For at the end we shall make ϵ infinitely small and $p\epsilon$ infinitely large, so that the series of kinetic energies given in (18) will become a continuous one, and our mathematical fiction will reduce to the physical problem treated earlier.

Boltzmann wrote these words nearly 30 years before the discovery of quantum mechanics, which shattered the belief which had held sway at least since Leibniz had said nature never makes jumps (*natura non facit saltus*). There was clearly not the remotest thought in Boltzmann's mind that jumps actually occur in nature. Nevertheless, the quoted passage is ironic to say the least, especially the difficulty of imagining an 'apparatus' that could regulate the required organizational feats; for that is exactly what quantum physicists had to do—and it took a quarter of a century of the most intense theoretical and experimental work by many brilliant physicists to do it. To boot, to this day nobody knows what the hugely impressive structure and successes of quantum mechanics really mean and whether and how jumps actually happen.

In fact, the above passage and its artifice are even more ironic in that they played a crucial role in the discovery of quantum mechanics. It's a case of irony piled on irony. As we will see, Boltzmann's belief in the atomic–molecular hypothesis came under particularly intense criticism in Germany and Austria in the 1890s. This was due to the strong influence of the positivistic philosophy of science that had developed and which argued that the goal of science should not be to propound theories that explain phenomena but merely to describe phenomena and identify universal patterns in them. In particular, several influential figures argued that thermodynamics should be concerned solely with general principles of the kind espoused in particular by Clausius in his early papers before he turned to statistical atomic–molecular explanations of them.

One such person was the German Max Planck (1858-1947); he published an important textbook *Thermodynamik* in 1897 in which, while granting the successes of the atomic–molecular theory, argued the case for developing thermodynamics on the secure foundations provided by the repeated failure of the attempts to build perpetual-motion mations of either the first or second kind. In 1899 he turned his attention to the outstanding problem of understanding the nature of the thermal equilibrium of radiation in furnaces. For technical reasons it was called the problem of black-body radiation. Remarkably, the best available theory at that time suggested the energy in such an equilibrium state should be infinite, in complete contradiction to the fact that a loaf of bread could be baked in an oven heated by a few logs of wood.

Planck, who does not seem to have been aware or concerned about this problem, was well aware of Boltzmann's statistical ideas and his trick to implement them by introducing discreteness artificially. Thinking the statistical approach was not the way to go, Planck first attempted to solve the problem without using statistics and failed. Seeing no alternative, he turned to the statistical ideas of Boltzmann and, literally in desperation, took his artifice for reality. He introduced discreteness as a fundamental feature of nature. Suddenly he obtained a formula that reproduced the observations of black-body radiation perfectly. His paper, published rather appropriately in 1900, ushered in the quantum age and won him the Nobel Prize for physics in 1918. Sadly, it seems Boltzmann did not learn about or recognize what Planck had done and hanged himself in 1906 unaware of the momentous consequence of his device of 1872, not that anyone else apart from the youthful Einstein had come to sense the implications of Planck's paper.

11 Maxwell's Demon and Thomson on Time

Nowhere in his 1867 paper does Maxwell comment on the mismatch between time-symmetric microscopic collisions and macroscopic equilibration. However, a letter he wrote in 1868 to his friend Peter Guthrie Tait[37] shows he understood the problem very well. He first stated it publicly in his beautiful book *Theory of Heat* published in 1871. Near the end, he says

> it is impossible in a system enclosed in an envelope which permits neither change of volume nor passage of heat, and in which both the temperature and the pressure are everywhere the same, to produce any inequality of temperature or pressure without expenditure of work. This is the second law of thermodynamics, and it is undoubtedly true as long as we can deal with bodies only in mass, and have no power of perceiving or handling the separate molecules of which they are made up.

He continues

> But if we conceive a being whose faculties are so sharpened that he can follow every molecule in its course, such a being ... would be able to do what is at present impossible to us. For we have seen that the molecules in a vessel full of air at uniform temperature are moving with velocities by no means uniform, though the mean velocity of any great number of them, arbitrarily selected, is almost exactly uniform. Now let us suppose that such a vessel is divided into two portions, A and B, by a division in which there is a small hole, and that a being, who can see the individual molecules, opens and closes this hole, so as to allow only the swifter molecules to pass from A to B, and only the slower ones to pass from B to A. He will thus, without expenditure of work, raise the temperature of B and lower that of A, in contradiction to the second law of thermodynamics.

Thus, unlike the law of gravity, the second law does not have a fundamental status. It is statistical in origin, relying for its validity on the presumed huge number of molecules in any finite measureable amount of gas.

In 1874, in a paper I will discuss in this chapter, Thomson called Maxwell's hypothetical being a *Maxwell demon*. The catchy name has helped to ensure that the issue it raises still generates much interest, but I won't discuss it because it does not directly address our central concern: why is it that processes which unambiguously single out a common direction of time are so ubiquitous in nature? Note that, as formulated (in Thomson's form) by Maxwell, the second law refers to what one cannot do with an equilibrium state. How the non-equilibrium states that we find all around us might have arisen is not addressed.

We now come to the reaction to Boltzmann's 1872 paper. The first was probably Thomson's 1874 paper just mentioned; I say probably because it does not mention Boltzmann though Thomson, Maxwell and Tait were well aware of his work. Thomson's paper begins with a graphic image:

[37] It is striking that thermodynamics as a theoretical discipline and its interpretation in terms of statistical mechanics was very largely the work of the Scots-Irish Thomson, the Scots William Rankine and Maxwell, the German Clausius and the Austrian Ludwig Boltzmann. As is still often the case, the continentals invariably referred to the Scots as English, while the Scots called Boltzmann a German.

> The essence of Joule's discovery is the subjection of physical phenomena to dynamical law. If, then, the motion of every particle of matter in the universe were precisely reversed at any instant, the course of nature would be simply reversed for ever after.

I don't think the implications of time-reversal symmetry had been expressed so clearly before. Note, however, that Thomson's "for ever after" seems to presuppose Newtonian time that continues to flow in its wonted forward direction. This conjures up an image of the divinity watching an autonomous river of time continuing to roll forward as all the motions in the universe are instantaneously reversed. In fact, we can imagine ourselves having the divinity's experience when we watch a film run backwards. Our thoughts run forwards, like the river of time, while the diver comes backwards out of the swimming pool.

But, leaving aside these manifestly atypical and fully explicable experiences, we, as beings within the universe, can only register, entirely within our consciousness, the background arrow of the innumerable unidirectional phenomena that surround us. Relative to that arrow we readily recognize reversal of any localized motion, as when the film of a diver is reversed but not the forward train of our thoughts or anything else we can see, say an usher showing somebody to their seat, as we sit in the cinema.

I only come to this at the end of *The Janus Point*, but I want to mention here *psychophysical parallelism*. The notion had precursors (going back at least to Galileo), but it came to especial prominence in the *Principles of Psychology* published in 1890 by William James (older brother of the novelist Henry James). According to it, all perceptions in our consciousness have counterparts in processes in the brain. Putting the idea as simply as I can, a divinity should be able to 'see' how the background arrow in the material universe at large induces, through standard physical processes, an arrow in the physical processes in our brain, which then somehow, through the mystery of consciousness, then becomes our personal psychological arrow.

If we accept this picture, the question then arises of how we are to think about the background arrow. Where precisely, in the material world, is it to be found and how is it to be recognized? What form does it take? Before I attempt to answer this question, I'll continue with what Thomson, after spelling out the purely physical implications of precisely reversing the motion of every particle of matter in the universe, said about the consequences for us:

> And if also the materialistic hypothesis of life were true, living creatures would grow backwards, with conscious knowledge of the future but no memory of the past, and would again become unborn.

Driven I imagine by his religious conviction, Thomson then says "But the real phenomena of life infinitely transcend human science; and speculation regarding consequences of the imagined reversal is utterly unprofitable."

But Thomson had just said what they are! Nevertheless, I struggle to understand what he meant. I think his "grow backwards" and "conscious knowledge of the future but no memeory of the past" must refer in his mind to some absolute river of time as the ultimate arbiter. However, the critical question is this: what would be the actual experience of the living creatures?

The question boils down to this: what form does the relevant information take? Suppose we insist it takes concrete form. Since neuroscience is still in a state of relative infancy, let's look to geology for help. This is the science in which the notion of deep time first took root. Geologists came to the conclusion that the earth must have a past that extended over an immense duration into the past by looking at the rich, effectively *static* structure it exhibited at their time. In the book by Burchfield on which I drew on in chapter 7 I came across a wonderful comment of Jean-Baptiste Lamarck (1744-1829), the well-known precursor of Darwin as a theorist of evolution. For the same reason as Darwin, this directed him to an interest in geology and led him to assert that *the earth's surface is its own historian.* In other words, if the earth were aware of its own instantaneous structure at any instant, it would have memories of its past and anticipate a future.

This suggests to me, as I argued in my *The End of Time*, that a sense of forward progression in time (and even of motion), might be stored in information about positions alone. If this is the case, Thomson's 'living creatures' would be oblivious to the reversal of all motions that he conjectured. Experience, with all its rich awareness of both being and becoming, would have its psychophysical parallel in static structures, *time capsules* as I define them in chapter 1 of *The Janus Point* (and earlier in *The End of Time*).

Elsewhere in the paper Thomson considers a situation, worth looking at here, in which motions are reversed in only a restricted part of the universe. For this he calls up in imagination a whole army of Maxwell demons and, so that they can precisely reverse particle velocities, arms them with molecular cricket bats!

He first considers a situation in which a very large number of particles initially have the same temperature on two sides of a vessel but then, through the action of demons, come to have unequal temperatures. The demons then cease their intervention, so that the average result of the free motions and collisions of the particles "must be to equalize the distribution of the energy among them in the gross". Then, at a later time when the temperatures in the two halves of the vessel have been effectively equalized on average,

> let the motion of every particle become instantaneously reversed. Each molecule will retrace its former path, and at the end of a second interval of time, equal to the former, every molecule will be in the same position, and moving with the same velocity, as at the beginning; so that the given intial unequal distribution of temperature will again be found, with the only difference that each particle is moving in the direction reverse to that of its initial motion. *This difference will not prevent an instantaneous subsequent commencement of equalization, which, with entirely different paths for the individual molecules, will go on in the average according to the same law as that which took place immediately after the system was first left to itself.*

The italics here are mine. In the context of his example of particles confined to a vessel, the argument is in fact very simple and probabilistic in nature. Collisions among a very large number of particles in a perfectly rigid vessel will almost certainly very soon equalize temperatures on average and this will, with almost equal certainty, happen for the two overall possible directions of the particle motions.

However, in a clear anticipation of a situation that confronted Boltzmann with severe difficulties two decades later, Thomson notes that in the vessel he considers spontaneous disequilibration "certainly will happen in the course of some very long time", this being longer the greater the number of particles in the vessel. So far as I know, this is the first recognition of the possible occurrence of spontaneous disequilibration within the study of statistical mechanics. There is also an interesting passage near the end of the paper where Thomson comments that if the rigid isolation of the vessel is removed and the number of molecules in the universe is infinite then disequilibration will never occur. He was clearly pleased with his argument because he said of it "This one instance suffices to explain the philosophy of the foundation on which the theory of the dissipation of energy rests."

Note that in drawing his conclusion Thomson relied on two things: increase without limit in the number of particles capable of dynamical interaction and an implicit assumption that the conditions in the universe outside the vessel are essentially the same as in the vessel. Challenging this assumption is the point of departure of my proposed explanation of time's arrows in *The Janus Point*. But we can already see a problem. Quite apart from the fact that he cannot have had any concept of the universe significantly clearer in 1874 than his one of 1862, Thomson's theory and his conclusion drawn from it would seem, at least at first glance, to imply that thermal equilibrium must be maintained everywhere and at all times in the universe. How did it come about that in Thomson's time, as now, the earth was manifestly suitable for habitation? There's a gap at the heart of his theory.

12 Boltzmann and the Second Law. II

It's time to move on. If Thomson did not mention Boltzmann's H theorem of 1872, Boltzmann failed to mention Thomson's 1874 paper in two papers in 1877. The second of these was particularly important. It did not solve the mystery of entropy increase, but it did lead to a beautiful microscopic concept of entropy. Applied in a slightly modified way, it plays a central role in *The Janus Point*.

Boltzmann was prompted to his entropy concept—his most important contribution to physics—in response to a comment by his friend and colleague Josef Loschmidt, who in 1869 had already proposed a Maxwell-type demon which he hoped would overcome the disturbing implications of the second law. In 1877 he published a paper on the thermal equilibrium of a system of bodies subject to gravitational forces[38] and gave some examples of special initial conditions for which Boltzmann's H theorem certainly will not hold. On this basis, he claimed that his considerations

> also destroy the terroristic nimbus of the second law, which makes it appear to be a destructive principle for the totality of life in the entire universe. At the same time this would open the consoling perspective that mankind is not entirely dependent on the intervention of coal or of the sun as regards the conversion of heat into work, but that there will be a never ending supply of convertible heat at our disposal for all time.

Loschmidt's 'terroristic nimbus' (which, in the translation I have quoted, should perhaps have been 'terrifying nimbus') gives some idea of the impact that the second law had made on many people. However, his hopes for the salvation of mankind were far too optimistic. His examples were, to say the least, artificial and required what is now called fine tuning—purely theoretically they are conceivable but in practice not capable of realization. An example he gave was of perfectly round elastic bodies on a perfectly straight line with velocities exactly along the line. Such a state could be maintained but only if set up with perfect accuracy. And what existence of any interest could be realized with such a scheme?

Much more valuably and apparently unaware of Thomson's 1874 paper, Loschmidt made a comment that stimulated Boltzmann in much the same way that Buys Ballot's intervention had enabled Clausius to greatly strengthen the case for atomism. Loschmidt simply remarked "Evidently quite generally in any system the entire course of events becomes retrograde if at a certain moment the velocities of all its elements are reversed." This is, of course, exactly what Thomson had said in 1874 and, as we will see, led him to insights that anticipated important developments in the 1890s. In Boltzmann's case it led to his greatest contribution to physics and to an equation that some regard as on a par with Einstein's $E = mc^2$.

Loschmidt's comment came to be known as the reversibility paradox. Given any possible state of a system of particles, defined by their masses, positions and velocities, there will

[38]Loschmidt's gravitational forces introduced a quite new element into statistical mechanics. They play an important role in *The Janus Point*. Loschmidt was also the first person to estimate the size of molecules. He did this by combining Maxwell's estimates quoted earlier for gases with the assumption that in the liquid state the molecules are closely packed together. This enabled him to estimate the number of atoms in a given mass of substance. In English literature this is known as Avogadro's number, but, understandably, the Germans call it the Loschmidt *Zahl* (number).

always be a state paired with it in which the positions are the same but all the velocities are exactly reversed. This manifest symmetry might lead one to expect a decrease as as often as an increase of entropy. One never does. That's the paradox.

Boltzmann responded promptly, noting that to apply the laws of mechanics, it is not sufficient to know how the forces act. The initial positions and velocities must also be known. Without them it would be impossible to prove that entropy must always increase. The reversibility already noted by Thomson and now by Loschmidt makes that clear: if entropy increases for a given set of initial conditions, reversal of all velocities at some later time will result in a decrease of entropy. Boltzmann makes no mention of the difficulty this poses for his H theorem, but brings in a quite new issue: the *probability* of possible initial states.

He does this for an example that complements significantly his considerations in 1872 and illustrates the generality of the entropy notion. For the H theorem he had assumed a system of particles that already had a uniform spatial density and a distribution of velocities the same in all directions. He had only considered how the velocity distribution would change in time and tend to Maxwell's. He now considered "a large but not infinite number of absolutely elastic spheres that move in a closed container whose walls are completely rigid and likewise absolutely elastic." No external forces act on the particles.

He then supposes that at some initial time the distribution of the spheres is not uniform. Then without knowing the initial conditions one cannot say the spheres will become more uniformly distributed. For if they do, one could take the state at which they have become much more uniform and simply reverse the velocities at that instant. Then "the spheres would sort themselves out" in such a way that they became non-uniform.

Thus a proof that after a certain time "the spheres must necessarily be mixed uniformly" whatever the initial state cannot be given. He says that this is a consequence of probablility theory "for any non-uniform distribution, no matter how improbable it may be, is still not absolutely impossible." This is a fact about probabilities. For example in a game of Lotto the chance of turning up the seemingly improbable sequence 1,2,3,4,5 is just as likely as any other sequence of the five numbers. In a critical passage, he says:

> It is only because there are many more uniform distributions than non-uniform ones that the distribution of states will become uniform in the course of time. One therefore cannot prove that, whatever may be the positions and velocities of the spheres at the beginning, the distribution must become uniform after a long time; rather one can only prove that infinitely many more initial states will lead to a uniform one after a definite length of time than to a non-uniform one.

The key comment here is "there are many more uniform distributions than non-uniform ones". There was surely little difference between Thomson's and Boltzmann's intuitions as to why non-uniform states will evolve to uniform states. The decisive difference is that Boltzmann goes beyond intuition and expresses things quantitatively. Indeed, soon after the quoted passage he actually speaks of numbers of states and makes one of the most important observations in the history of physics: "One could even calculate, from the relative numbers of the different state distributions, their probabilities, which might lead to an interesting method for the calculation of thermal equilibrium." He wasted no time in doing just that by counting states in the very next issue of the journal. The method by which he did it is so simple.

We must remember that any collection of particles has two distinct kinds of distribution: one of positions in space, the other of kinetic energies. They need to be handled in different ways. In 1877, Boltzmann's main interest was the kinetic energies. The prize he was after was the Maxwell distribution and why it must be established eventually if the energy distribution is initially non-Maxwellian. Because quantities that vary continuously are difficult to handle in probabilistic terms, I defer that to the next chapter and here employ Boltzmann's trick and consider only discrete quantities. Let's begin with spatial distributions.

Number a chequerboard's squares from 1 to 64. Number, say, 48 chequers from 1 to 48 and think about all the ways they can be distributed on the board. Any number can be put any one square. The most non-uniform way to do that is to put all the chequers on just one square. There are 64 ways to do that, one for each square. However, if we are only interested in how uniform our distribution 'looks', all of the possibilities look the same whatever the square on which the pile is placed and there is only one maximally non-uniform distribution. In quantum mechanics, situations like this second one come into play, but I will stick with the first. It corresponds to what is called Boltzmann statistics and is valid when quantum effects can be ignored. Of course, quantum mechanics also introduces discrete effects, but here, as in Boltzmann's original studies, discretization, which reduces the problem to counting permutations, is used as a trick to clarify an underlying idea and permit some simple calculations that are later to be made more rigorous.

Next we can put all the chequers on just two squares. This hugely increases the possibilities. First, there are $64 \times 63 = 4032$ possible choices of the two squares. Call them A and B. Among the 48 chequers, we can initially choose one and put it on A and all the rest on B. There are 48 ways to choose a single chequer. But we can also put two on A and 46 on B. There are $(48 \times 47)/2 = 1128$ ways of doing that; you divide by 2 because the order in which you put the two chequers on A leads to the same distribution. To get the total number of possibilites, you must multiply the 1128 by the 4032 ways of choosing the two squares. By the time you get down to equal division, 24 on A and 24 on B, the number of permutations is vast, astronomical. And that's with 64 chequers. With Avogadro's number 6.022×10^{23} the tally is barely conceivable.

At this point a different kind of chequerboard that is a better model for the universe should it be, as is not impossible, closed up spatially on itself in three dimensions like the earth is in two. To model that 'chequerboard' could be the surface of a sphere divided into areas, say four equal segments by two lines of longitude from one pole to the other and $90°$ apart at the poles. These four segments could in turn each be divided in two along the equator, giving a 'chequerboard' with eight 'squares'. The same game of permutations can be played it. An advantage is that this board has no 'special squares', unlike a real chequerboard with its four distinguished squares in each corner and more along the edges.

The next thing to discuss is the way Boltzmann actually counted states and the number of ways in which they can be realized. For this, we need the modern names *macrostate* and *microstate* for his two fundamental concepts.

Suppose, for any finite number N of chequer squares, labelled $i = 1, 2, \dots N$, we have a finite number P of chequers distributed on them. For some distribution, let the number of chequers on square i be n_i. The numbers n_i define the macrostate of the chequers. Now, supposing they all have their own individual nature, let's give the P chequers the names $a, b, c, \ldots.$ Each different way in which they can be laid on the N squares to give the same numbers n_i of the macrostate just considered defines one of its possible microstates. In general there will be many different microstates that give the same macrostate. (Boltzmann called the number of microstates corresponding to a given macrostate its *complexion*). I will shortly give, for the case of energy distributions, the macrostates and numbers of their microstates that Boltzmann actually gave in 1877 for the example of seven particles.

Now because we have a finite number of chequers and a finite number of squares, the total number of distinct macrostates is finite. The same is true of the microstates. Let the total number of microstates be m_{tot}; it is much greater than the number M of macrostates. The final step to Boltzmann's 1877 notion of entropy brings in probability. For each microstate, let us put lottery balls in a bag bearing a name that identifies it together with the macrostate to which it belongs. The chance of drawing a particular microstate is $1/m_{\text{tot}}$. If the number of microstates that belong to macrostate M_a is m_a, the chance that a microstate drawn at random belongs to macrostate M_a is m_a/m_{tot}, i.e., it is the number of microstates that belong to M_a divided by the total number of microstates. If one makes the assumption that all microstates are equally probable, m_a/m_{tot} is the probability of the corresponding macrostate. I will discuss later the conditions under which it is reasonable to assume that all microstates are equally probable. The issue has generated much debate.

Boltzmann's insight in 1877 was that it is the ratio m_a/m_{tot} that determines the entropy of the considered macrostate. In fact, subsequent development showed that it is very sensible to define the entropy of a macrostate as the logarithm of the number of microstates that it contains. This has to do with a property of probabilities of independent events: if the probability of event a is p_a and of the independent event b is p_b, then the probability that both a and b will occur is the product $p_a p_b$. Now the logarithm of a product is the sum of the logarithms of the individual terms in the product, $\log p_a p_b = \log p_a + \log p_b$, and in phenomenological thermodynamics it is found experimentally that the entropy of independent systems that are then combined is the *sum* of their individual entropies. This matches beautifully the probabilistic count-of-microstates nature of entropy that Boltzmann discovered. It is for this reason that the famous expression for entropy,

$$S = k \log W, \tag{19}$$

where k is the dimensionful Boltzmann's constant and W stands for *Wahrscheinlichkeit* (probability), is engraved on Boltzmann's tombstone in Vienna. Ironically, the equation was never written down by Boltzmann; it is due to Max Planck.

Before we move on, I'd like to recall what I said about Clausius being more than justified when he coined the word 'entropy' to make it have "a certain similarity in designation" to energy because the two concepts are "so nearly allied in their physical significances". One of Richard Feynman's best known quotations is this:

> If, in some cataclysm, all of scientific knowledge were to be destroyed, and only one sentence passed on to the next generation of creatures, what statement would contain the most information in the fewest words? I believe it is the atomic hypothesis that

> all things are made of atoms—little particles that move around in perpetual motion, attracting each other when they are a little distance apart, but repelling upon being squeezed into one another. In that one sentence, you will see, there is an enormous amount of information about the world, if just a little imagination and thinking are applied.

Ultimately, the notion of entropy can be defined because the atomic nature of matter, Feynman's great truth condensed into a single sentence, introduces discreteness and the possibility of counting, while conservation of energy reflects the lawful manner in which discrete things can be rearranged.

Returning to Boltzmann's key idea, it should be said that there is a certain 'fuzziness' about the notion of entropy, both at the macroscopic and microscopic level. This relates to the definition of macrostates, which can be made more or less restrictive. This need not concern us here; I will say something about it later in the book. The question of how continuous distributions are to be treated will be discussed in the next chapter.

Now we come to the important difference between entropy calculations for spatial and energy distributions. For the former, the chequers and squares are all on an equal footing. A checker looks the same whatever square you place it on. The same is true for the position of an atom in a box. However, it's quite diffferent when possible kinetic energies are considered. Here it is in principle possible that one single atom has all the kinetic energy. It then belongs to just one 'square' that has an energy value equal to the total energy available. No kinetic energy remains for any of the other atoms. They must all sit on the 'zero-energy' square. This skews the distribution.

Boltzmann gave a remarkable example of this in his 1877 paper. He considered a system of just seven particles and supposed the total energy to be divided up into seven quanta distributed among the particles. In this case, the checker board is replaced by a row of eight 'bins' labled $0, 1, 2, \ldots 7$. All particles that have no kinetic energy go into bin 0; those with one quantum of kinetic energy (there can be up to seven of them) go into bin 1; particles with two quanta go into bin 2 (there can be at most three of them with one particle in bin 1) and so on. Finally, if one particle has all the energy it goes in bin 7 and the remaining six go in bin 0.

In Boltzmann's example with seven particles, there are 15 different macrostates and 1716 microstates. One macrostate has just one microstate—the one in which all particles have one quantum of energy. Another macrostate has seven microstates; it is the one in which just one particle, for which seven choices can be made, has all the energy and all the remainder have none. One macrostate has 420 microstates. In it, three particles have no energy quanta, two have one quantum, one has two quanta and one has three quanta. Two macrostates have 210 microstates each. Two have 140; five have 105; and three have 42. Thus, just three of the macrostates have 840 out of the total of 1716 microstates: that is only very slightly less than half the total. Boltzmann's seven-particle macrostate with 420 microstates already approximates the Maxwell distribution surprisingly well. Only a modest increase in the number of particles would be sufficient to get impressively good agreement. Note also the point I made earlier: all Maxwell distributions are identical if one only considers the relative distribution of energies between the particles. In his seven-particle model, the numbers that define the macrostates correspond only to fractions of the kinetic energy, not the total amount of it.

At this point I need to clarify the conditions under which the kind of combinatorial arguments that Boltzmann employed are valid. There's no problem with actual numbers, for example, the 48 chequers and 64 squares in my example or the seven particles in Boltzmann's. Any finite positive integers for the numbers can be chosen and 'the game' played with them. In contrast, what is essential is the situation in which those numbers can be chosen. In the statistical mechanics developed from the 1850s, the number of chequers corresponds to the numbers of atoms or molecules assumed to be present in a box. When defining their models, all practitioners of the discipline assumed a large but finite number of particles. The subsequent discovery of quantum mechanics showed that this was not problematic; what then was important was not so much the number of particles but the number of quantum states. Somewhat more problematic for the pre-quantum era is Boltzmann's sub-division of the main 'container' into coarse-graining boxes. For a given number of particles, this will always lead to ambiguity in any quantity used to define the uniformity of the particle distribution. In modern treatments this is usually overcome by a device called 'going to the thermodynamic limit'. In the model with chequers, which is all I need to make the conceptual issues clear, the issue does not arise.

The real issue is this: is the model with a container likely to be a good model of the universe? Before I address this question, let me review what were the main aims behind the development of statistical mechanics. They began with attempts to explain by means of the atomic hypothesis the observed macroscopic properties of ideal gases confined to some kind of box and in the state of equilibrium. This project made significant progress in the two decades up to Boltzmann's H theorem paper of 1872. This modified the original programme, giving it a new goal that was both clear and simple in its aspiration: to show why a system initially out of equilibrium would, under the assumption the particles of the system are governed by simple mechanical laws, tend to the equilibrium state. This process is called equilibration. We observe many examples of it all around us. Maxwell gave the nice example of a cold silver teaspoon put into a cup of hot tea: the handle soon becomes warm, indeed hot. The spoon and the tea come into equilibrium. It may be noted that if the tea is not drunk, it and the spoon will in time come to have the same temperature as the air in the room. This process of equilibration in a larger background is still effectively taking place in a box, though it is one that is not perfectly insulated.

But now suppose that, in otherwise empty space, we have gas particles in a box that are not in equilibrium. Let the walls of the box be removed. The particles will simply fly out into empty space. They will never equilibrate and instead become dispersed through the void of empty space. The question now arises of what remains of the insights and results of the statistical mechanics that Clausius, Maxwell and Boltzmann developed.

I have not made anything remotely like an exhaustive search through the statistical-mechanics literature, but I have looked at much of it devoted to its application to the issues around the arrows of time. I have not seen one single reference to what might be expected if one cannot rely on a box to confine the gas. There is, however, what seems to me a significant remark in the 'bible' of statistical mechanics. This is the *Elementary Principles in Statistical Mechanics, developed with especial reference to the rational foundation of thermodynamics* published in 1902 right at the end of his life by J Willard Gibbs, whose important earlier work has already been mentioned. In his book, he treats dynamical systems with great generality but says he does need to introduce two restrictions, namely that the system cannot be allowed to spread out in an infinite space or have velocities that can grow faster than a certain rate. The reason for this is that he wants to introduce probabilities, and these would lose meaning if he did not impose the restrictions. In this connection, he notes that in thermodynamics a system that can disperse through infinite space cannot attain an equilibrium state. It is clear that as regards the spatial restriction Gibbs is effectively introducing a conceptual box. There is also a 'box' for the velocities.

Finally, let's consider how non-uniform distributions can become uniform, doing this for chequers, rather than the particles Thomson and Boltzmann considered; this will show how general are the situations that came into clear focus with the discovery of thermodynamics and its explanation by statistical mechanics. Let us suppose the following procedure, based on Boltzmann's 1872 model of the effect of collsions, for changing any given chequer distribution. Let there be two bags, one containing red balls numbered 1 to 48 and another white balls numbered 1 to 64. Initially, let there be a strikingly non-uniform distribution of chequers. In the first step of the procedure, one ball is drawn at random from each bag. Wherever it it, the chequer with the number on the red ball is moved to the chequer square with the number on the white ball. Of course, there is a tiny chance it will already be there, in which case nothing is done. This procedure continues step by step forever.

Now since chance determines what happens in each step and, initially, there are many more uniform distributions than the non-uniform one from which we begin, I think you will agree that, with very high probability, the chequer distribution will tend to become more uniform with each step. However, there will always be a small chance that it becomes less so, generally for only a few steps. After a large number of steps, the distribution will almost certainly have become significantly more uniform though there is always the tiniest of chances that the reverse process will set in and the distribution becomes much less uniform. You will recall Boltzmann's comment that in Lotto the probability of drawing the sequence 1,2,3,4,5 from a bag of the first five integer numbers is just as likely as any other sequence, for example, the less 'special looking' 3,1,5,2,4. The point is that all sequences are equally likely. To even begin to comprehend the meaning of entropy in the real world you have to think of the first one hundred integers in a bag and imagine the chance of drawing them out in the sequence $1, 2, 3, 4, \ldots, 97, 98, 99, 100$. It might happen, just as you might toss a fair coin and get heads one hundred times in a row (that's much more probable than getting the first one hundred integers in the correct order).

Once the distribution has got close to the greatest uniformity that is possible, it is almost certain to stay very near to it for a very long time. However, there will always be small fluctuations away from the maximum and, if you wait for long enough, there will eventually come a time at which the distribution becomes markedly non-uniform. If it does get there, it is almost certain to begin to get more uniform again very soon.

Now suppose you know that such a lottery has been going on for aeons of time and you are shown just one distribution, the current one, and you find it is very non-uniform. You know the rules of the game and therefore, without being shown the immediately preceding and following distributions, you can bet with high confidence that in both temporal directions the uniformity will be greater, if not immediately then at least after relatively few steps. If the distribution is already far from uniformity, the chance that it will get even more non-uniform is very small.

Although Thomson never came remotely near giving an argument along these lines in his 1874 paper, his argument that I italicized on p. 72 has the same underlying basis though applied in a continuous rather than discrete form. Boltzmann too arrived at essentially the same conclusion in his first paper of 1877, as is clear from this passage:

> I will mention here a peculiar consequence of Loschmidt's theorem, namely that when we follow the state of the world into the infinitely distant past, we are actually just as correct in taking it to be very probable that we would reach a state in which all temperature differences have disappeared, as we would in following the state of the world into the distant future. This would be similar to the following case: if we know that in a gas at a certain time there is a non-uniform distribution of states, and that the gas has been in the same container without external disturbance for a very long time, then we must conclude that much earlier the distribution of states was uniform and that the rare case occurred that it gradually became non-uniform.

Note that, unlike Thomson (who only considered molecules in a box), Boltzmann is here considering two quite different scenarios: one in which the evolution of the whole universe in two directions of time is considered and another in which an experimentalist is examining gas in a container and will be aware of the background arrow of time. This makes it possible to establish an unambiguous 'before–after' ordering of the states of the gas. But in the case of the whole universe, what criterion is being used to define the direction of time? It was nearly twenty years before Boltzmann was forced to confront this question directly. We will come to that in the next chapter.

13 Boltzmann's Tussle with Zermelo

We now come to the final part in the saga of Boltzmann's work on the second law. It set's the scene for the new interpretation of time's arrows in *The Janus Point*.

First some background. In the decade and a half that followed Boltzmann's 1877 insight, majority opinion among German physicists turned, rather surprisingly, against atomism and the statistical-mechanical interpretation of thermodynamics. That this happened despite the remarkable successes that had been achieved had two main reasons. First, the atomists encountered increasingly great difficulties in arriving at precise models of atoms that matched observations. This applied especially to the failure to explain the ratio of specific heats: energy that theory strongly suggested should be present in internal vibrational degrees of freedom obstinately refused to manifest itself. Equally seriously, the spectral lines of the radiation emitted and absorbed by atoms and molecules were proving extremely difficult to interpret theoretically. What nobody realized was that models based on classical mechanics had no chance of resolving these problems. That had to wait until the full discovery of quantum mechanics in the mid 1920s by Heisenberg and Schrödinger. Thus, failures fully understandable with hindsight were fostering a certain distrust in atomism.

Perhaps more important was the philosophical movement known as positivism. Several prominent thinkers on the continent, especially in Germany and Austria, favoured it. The best known among them was Ernst Mach (1837-1916), best known for his criticism of Newton's concepts of absolute space and time and his discovery of shock waves; te Mach numbers are named for him. Mach had a strong distrust of theory because his study of the history of science had shown how many seemingly plausible theoretical models proved under testing to be inadequate to describe phenomena. Caloric is a good example. Like other positivists, Mach argued that the only proper task of science was to identify and describe phenomena that one could repeatedly observe and describe by means of measurement. However, it was in fact at just about the time Mach finished his university studies that the discipline of theoretical physics really started to come into its own. The great thing about theory is that, from an initial 'bold hypothesis' (to use Karl Popper's expression), one can deduce consequences by strict logical argument and predict hitherto unexpected effects. A classic example was Maxwell's discovery, much to his surprise, that the viscosity of a gas should not depend on its density. In turn this helped in the first determinations of the size of atoms and molecules the results of which were later confirmed by many different independent methods.

Distrust in atomism and adherence to positivism peaked in the 1890s; an important factor was the difficulties encountered by mechanical models in explaining irreversible phenomena, above all the growth of entropy. Although he was careful not prescribe precise models of atoms, Boltzmann was for decades a leading advocate for atomism and was challenged for this reason. Sadly for Boltzmann (and bad for Mach's posthumous reputation) there was a famous occasion when Boltzmann gave a lecture about atoms at the Imperial Academy of Science in Vienna in 1897 after which Mach said loudly from among the listeners "I don't believe that atoms exist."

The opponents of atomism argued that countless experiments had shown that, under appropriate exclusion of external effects, entropy never decreased in a closed system and in general increased. This, it was argued, should simply be accepted as a fundamental law of nature. Attempts to explain away the conflict between the reversibility of mechanical laws and the irreversible growth of entropy violated sound principles of science, which should concern itself solely with the identification and clear economic description of universal phenomena.

It was Boltzmann's misfortune, towards the end of his life, to be a lonely defender in the German-speaking world of the virtues and successes of the rival approach to positivism that is today called theoretical physics. He was a classic example of a prophet unrecognized in his own country. He was taken much more seriously in Britain and in 1894 was invited to attend a meeting of the British Association in Oxford at which there was a serious discussion of the famous H-theorem and its apparent conflict with time-reversal symmetry as raised by Loschmidt. This meeting stimulated Boltzmann to write a letter to *Nature* (in remarkably good English) in which he addressed various questions in the philosophy of science and concluded with a passage that sets the scene perfectly for Boltzmann's last stand defending the virtues of atomism and the mechanical explanation of entropy.

He notes that Mr Culverwell, in a letter to *Nature*, had said the H theorem could not be true, for "if it were true every atom of the universe would have the same average *vis viva*, and all energy would be dissipated." Boltzmann countered that "this argument only tends to confirm my theorem, which requires only that in the course of time the universe must tend" to such a state. Thus there is nothing wrong with the H-theorem; the great mystery is "why this state is not yet reached". He then continued:

> I will conclude this paper with an idea of my old assistant, Dr Schuetz.
>
> We assume that the whole universe is, and rests for ever, in thermal equilibrium. The probability that one (only one) part of the universe is in a certain state, is the smaller the further this state is from thermal equilibrium; but this probability is greater, the greater is the universe itself. If we assume the universe great enough, we can make the probability of one relatively small part being in any given state (however far from the state of thermal equilibrium), as great as we please. ... Assuming the universe great enough, the probability that such a small part of it as our world should be in its present state, is no longer small.
>
> If this assumption were correct, our world would return more and more to thermal equilibrium; but because the whole universe is so great, it might be probable that at some future time some other world might deviate as far from thermal equilibrium as our world does at present. Then the afore-mentioned H-curve would form a representation of what takes place in the universe. The summits of the curve would represent the worlds where visible motion and life exist.

In this passage the summits of the H-curve are, in accordance with the definition of the H-function, places where the entropy is low and the state is ordered. Second, in later amplifications of the passage we will see that Boltzmann added some refinements and these have had the consequence that Dr Schuetz's role in the genesis of the original idea has not received much notice. That seems a bit unfair since the proposal is an early (perhaps even the first) and particularly clear example of what has become known as the *anthropic principle*. This has been much discussed in recent decades for several reasons, the main one being that the known laws of physics seem to be remarkably fine-tuned for the existence of intelligent life or, at the least, the conditions we find around us—visible motion and life, in

Boltzmann's striking expression.[39] The anthropic principle says we should not be surprised to find around us conditions favourable for our existence: if they were not so, we would not be here. Many physicists dislike the anthropic principle; they would much prefer a theory of everything. Even without that ambition, it is possible to raise a serious doubt about Dr Schuetz's idea and Boltzmann's endorsement of it.

This relates to the notion of time capsules mentioned in the discussion of Thomson's paper of 1874. The observable universe that surrounds us not only exhibits 'visible motion'—concentrations of matter that move relative to each other—but much, much more: truly remarkable evidence of a past in everything we examine. What is more, one could imagine that fluctuations out of thermal equilibrium could contain hints of different pasts, but the evidence points overwhelmingly to a unique past. It is found in structures that we can very persuasively interpret as *mutually consistent records* of that past. There are many ways in which a distribution of material particles can be far from thermal equilibrium but surely only a tiny fraction will have a 'time-capsule' structure. Thus, to explain adequately the universe we now observe we need to invoke a Schuetz-type argument twice: once to get to all the localized states in the universe that are strongly out of thermal equilibrium and then once more to get to the states that are both out of thermal equilibrium and are time capsules. I'm not persuaded that mere probability arguments are sufficient to do that; at the least, I have not seen the issue addressed except in a paper by the historian and philosopher of science John Norton that I cite in *The Janus Point*,in chapter 5, which I show how the Schuetz–Boltzmann fluctuation idea had to be abandoned when it was realised that a single brain should fluctuate into existence believing itself to be in a universe than that a huge region that looks like a universe should do so. That same chapter highlights how long the conceptual ideas validly developed to study systems confined in a box have lived on to the present day.

The seeds of the problems Boltzmann faced in his last stand in defence of atomism were actually sown before the 1894 meeting in Oxford. The great French mathematician Henri Poincaré (1854-1912) planted them in a very important (and long, 270 pages) paper that he wrote in 1890. The paper is famous for more than one reason but mainly as the first discovery of what later became known in dynamical theory as chaos. The name is a bit unfortunate; what it refers to is the extreme sensitivity of some dynamical systems to slight differences in the initial conditions. Tiny changes can later lead to huge differences. The effect is often illustrated by the effect of a butterfly flapping its wings in the Amazon delta on the weather in Europe weeks later. In fact, the phenomenon Poincaré found passed more or less unnoted until Edward Lorenz rediscovered it in the context of weather forcasting in 1963. At the time of writing, there's a beautiful illustration of the effect online to be found by consulting Wikipedia.

The 1890 paper, "Sur le probleme des trois corps et les équations de dynamique," was a hurried revision and correction that Poincaré had made when, to his horror, he discovered an error in a paper for which he had been awarded a prize by the King of Sweden. The paper addressed the long-standing problem of the solar system's stability. Might one day a planet be ejected and fly off into interstellar space? There is no need to go into that question here. All that I need to tell you about is the part of his paper in which Poincaré proved what

[39] I like Thomson's 'palpable' (from Latin's *palpare* to 'feel, touch gently') even better.

has become known as his *recurrence theorem*. It and the foundation of its proof known as *Liouville's theorem*[40] are so important for later I will need to spell them out as clearly as I can. Luckily, the ideas are not too difficult to grasp.

Let's start with Liouville's theorem,[41] which applies in particular to the important class of dynamical theories that are called Hamiltonian, after the Anglo-Irish mathematician William Rowan Hamilton (1805-1865) who developed them in brilliant fashion in the 1830s. One can see such theories as a formalization of Newton's laws for systems not subject to friction that especially highlights their deterministic nature. Hamiltonian theory can be developed for any kind of dynamical degrees of freedom, but I shall illustrate it for point particles since the most basic issues around entropy and irreversibility can be illustrated with them.

The key concept is *phase space*. Suppose we have N particles in (three-dimensional) space. At any instant, $3N$ coordinates define the positions of the particles and $3N$ their velocities. There are however rather remarkable mathematical reasons discovered by Hamilton—we needn't go into them you will probably be relieved to know—why it is more convenient to use the particle momenta rather than the velocities. In the case of point particles, these are simply the velocities of each particle multiplied by its mass. Just as two coordinates, say longitude and latitude, can define a point on the two-dimensional surface of the earth, the complete set of $3N + 3N = 6N$ positions and momenta defines a point in phase space.

Any such point defines a possible initial state of the dynamical system. Once it is specified, Hamilton's equations, just like the Newtonian equations that they generalize, determine the evolution uniquely in both time directions. They transport the phase point along curves in phase space. It is a beautifully important fact about these curves that they never cross; they either go on forever or make a closed loop.

We now come to Liouville's theorem. The game Poohsticks in one of the Winnie-the-Pooh books by A. A. Milne illustrates its content. Each player drops a stick on the upstream side of a bridge. The stream carries the sticks along with its flow; the player whose stick first appears on the other side is the winner. The sticks in this game are like the dynamical points which Hamilton's equations carry through phase space. A continuously connected set of points is like a patch of oil carried along by the stream.

Using Hamilton's equations, mathematicians can determine how such a patch behaves in phase space by determining the curve followed in it by each phase point. What they find is one of those beauties that, every now and then, mathematics throws up: the area of the patch remains constant. It may be deformed and stretched out in all possible ways, but the area does not change by a jot. Actually, I should not say area but volume since phase space has $6N$ dimensions; it is only for the simplest two-dimensional phase spaces that an area remains constant. It's worth looking online for the illustration I found after a bit of searching. You probably know how to do that better than me. If in doubt, I always start with Wikipedia as the first port of call.

Poincaré's recurrence theorem relies on Liouville's theorem and, critically, holds provided the total volume of the considered Hamiltonian system's phase space is finite; the technical

[40] Joseph Liouville (1809-1882) was, as I mentioned, one of the mathematicians with whom Thomson had discussions in Paris.

[41] It acquired great significance as the dynamical foundation of statistical mechanics in a book Willard Gibbs published in 1902 shortly before his death.

expression is *has a bounded Liouville measure*. This and the notion of phase space are essential for everything that follows. Now suppose any region however small—the tiniest patch in my previous account of Liouville's theorem—and consider what happens to points within it under the Hamiltonian flow. You can imagine throwing a stick into some small region of a stream and following where the flow takes the stick. In an ordinary stream, the stick would eventually be carried out to sea. But Hamiltonian flow in a region of bounded Liouville measure is not like that. It's more like the flow in a Jacuzzi in which a cork will be carried around all over the place as long as the bath is operated. Of course, the laws of nature hold eternally, so we must consider a Jacuzzi that runs for ever and a cork condemned, like the Flying Dutchman, to be carried around by the flow in a never ending voyage. Note also that at any given point the flow is always the same; it's a very special Jacuzzi.

What Poincaré was able to show by combining Liouville's theorem, the special nature of the flow and the restriction on the phase space to have a bounded Liouville measure was this. For any region in phase space, no matter how tiny, infinitely many flow lines pass through it infinitely often. In the Jacuzzi analogy, a pointlike cork placed initially at any point within an arbitrarily small region of the bath is bound to return infinitely often to that same small region and pass through it. It is not bound to return to exactly the same point, but because the region can be assumed to be arbitrarily small it must return infinitely often arbitrarily close to the point. Liouville's theorem plays a critical role in the proof because the volume of a 'patch' of initial solutions carried along by the dynamical flow cannot avoid coming back to where it once was, basically because in a bounded region 'it runs out of places it can visit'.[42] There is some analogy, for which I am indebted to Jim Hurley, with a girl in sandals stepping around a sand pit and leaving foot prints. However hard she tries to avoid stepping on an already existing footprint, sooner or later she cannot step on any undisturbed sand.

If we remember that in statistical mechanics the motion of the cork stands for what arbitrarily many mass points are doing in ordinary three-dimensional space, we see that Poincaré's result is extraordinarily powerful. A more refined illustration is needed. Imagine a swarm of bees that fly forever in a windowless room. At an initial instant, note where each bee is *and* the direction in which it is flying. Then the recurrence theorem states that the swarm is bound to return infinitely often to a state in which each bee is arbitrarily close to its original position in the room and is flying with direction and speed arbitrarily close to what it had at the initial instant. This gives an inkling of how rigid and mighty are laws of nature. Physicists see exquisite beauty in them; others find them as horrifying as Calvin's doctrine of predestination. It should be said that immense stretches of time—far, far greater than the age of the universe—must pass before recurrence occurs; their length increases rapidly with the number of particles.

Note also that the recurrence theorem does not say that each of the overwhelming majority of the solutions of the system will get arbitrarily close to *any* (or virtually all) points of its phase space. A system for which this is true is said to be *ergodic*. Remarkable though it is, the recurrence theorem only says the system will return infinitely often infinitely close to a point in its phase space through which it has already passed.

Poincaré's proof of his recurrence theorem occupied only a few pages of his monumental

[42] Poincaré showed that in fact there are some exceptional solutions that return to the original region only a finite number of times, but, being exceptional, these can be ignored.

270 page paper of 1890. In it, he does not connect the theorem with thermodynamics, but, like Mach, Poincaré was deeply sceptical about atomism and mechanical explanation of the second law. In 1893, in his brief paper "Mechanism and experience", he used his recurrence theorem to mount an attack.

The opening sentence sets the tone: "Everyone knows the mechanistic conception of nature which has seduced so many good men, and the different forms in which it has been dressed." But how, he asks, can one reconcile the fact that in accordance with the mechanistic hypothesis "all phenomena must be *reversible*" whereas experience provides a number (he could have said an innumerable number) of examples of irreversible phenomena.

After dismissing an attempt by Helmholtz to resolve the problem he says "The English have proposed a completely different hypothesis." He illustrates it by a comparison: "If one had a hectolitre of wheat and a grain of barley, it would be easy to hide this grain in the middle of the wheat; but it would be almost impossible to find it again, so that the phenomenon appears to be in a sense irreversible." At this point, he refers to Maxwell's demon and says "For such a demon, if one believes the mechanists, there would be no difficulty in making heat pass from a cold to a hot body." He calls the development of this idea "the most serious attempt to reconcile mechanism and experience". He continues, with obvious reference to his recurrence theorem,

> But all the difficulties have not been overcome.
>
> A theorem, easy to prove, tells us that a bounded world, governed only by the laws of mechanics, will always pass through a state very close to its initial state. On the other hand, according to accepted experimental laws (if one attributes absolute validity to them, and if one is willing to press their consequences to the extreme), the universe tends towards a certain final state, from which it will never depart. In this final state, which will be a kind of death, all bodies will be at rest at the same temperature.
>
> I do not know if ... the English kinetic theories can extricate themselves from this contradiction. The world, according to them, tends at first toward a state where it remains for a long time without apparent change; and this is consistent with experience; but it does not remain that way forever, if the theorem cited above is not violated; it merely stays there for an enormously long time, a time which is longer the more numerous are the molecules. This state will not be the final death of the universe, but a sort of slumber, from which it will awake after millions of millions of centuries.
>
> According to this theory, to see heat pass from a cold body to a hot one, it will not be necessary to have the acute vision, the intelligence, and the dexterity of Maxwell's demon; it will suffice to have a little patience.
>
> One would like to be able to stop at this point and hope that one day the telescope will show us a world in the process of waking up, where the laws of thermodynamics are reversed.

.

After this beautiful discussion, which I find strengthens rather than weakens the 'English' position, Poincaré claims that "unfortunately, other contradictions arise". However, in the hundred or so words that remain of his paper, the objections he raises are by no means substantiated and Poincaré's hope to demolish the mechanistic philosophy remains little more than an aspiration; he puts no flesh on his arguments. Curiously, there is no mention of Boltzmann in the paper. Was it perhaps some antipathy to the German-speaking world or just a desire, realized through the ironic tone ('a little patience'), to make a bit of fun at the expense of the English?

Liouville's theorem and the recurrence theorem now take centre stage. As you follow the drama unfold, please don't forget that Liouville's theorem holds whether or not there is a bound on the phase space measure, but the recurrence theorem fails without one.

Although Poincaré did not mention Boltzmann, and Boltzmann does not seems to have been aware of Poincaré's papers of 1889, 1890 and 1893, the young German mathematician Ernst Zermelo (1871-1953), who was an assistant to Max Planck in Berlin, literally pounced on them. As you know, Planck was initially sceptical if not hostile to the mechanistic attempts to explain the second law; his famous textbook *Thermodynamik* of 1897 favoured the purely phenomenonogical approach to the subject. Zermelo's intervention, surely encouraged by Planck, had the great virtue of drawing out Boltzmann to make what were in effect his last comments on the second law and its statistical interpretation. Zermelo, for his part, went on to become a famous logician; he found a way to resolve some of the paradoxes in set theory and in it proved an important theorem (every set can be well ordered). In 1935, he resigned from his post in Freiburg on the edge of the Black Forest in protest at Hitler's regime but was restored to an honorary status after the war.

Zermelo's first paper on thermodynamics appeared in 1896 and began with a proof of the recurrence theorem, which he showed must hold "provided that the coordinates and velocities cannot increase to infinity". Zermelo's following attack on atomism in fact does little more than raise the problems to which Poincaré had drawn attention, moreover in Zermelo's case with rather more words and no ironic wit. Specifically, Zermelo supposes "a gas enclosed in a rigid container with elastic sides that are inpenetrable to heat". In such a situation, initial states "instead of undergoing irreversible changes, will come back periodically to their initial states as closely as one likes". Since the mechanistic theory requires quantities like the temperature and entropy to be determined by the instantaneous state, irreversible changes of them cannot be expected to occur. Unlike Poincaré, Zermelo does not grant that an apparent heath death might last for an enormously long time before being woken from slumber though that is a clear implication of the recurrence theorem.

Unlike Poincaré's papers, Zermelo's did rouse the bear; Boltzmann's response appeared within a few months in the same volume of the *Annalen der Physik.* He began by saying that, like Clausius, Maxwell and others, he has "often emphasized as clearly as possible" that Maxwell's velocity distribution law

> is by no means a theorem of ordinary mechanics which can be proved from the equations of motion alone; on the contrary it can only be proved that it has a very high probability ... At the same time I have also emphasized that the second law of thermodynamics is from the molecular viewpoint merely a statistical law. Zermelo's paper shows that my writings have been misunderstood; nevertheless, it pleases me for it seems to be the first indication that these writings have been paid any attention in Germany.

Boltzmann immediately grants that Poincaré's theorem "is clearly correct" but that Zermelo's application of it "to the theory of heat is not". I won't go into the details of

Boltzmann's response but its gist is simple and convincing. Expressed in terms of entropy,[43] it is that in a container at rest holding a great number of molecules their state will for vast stretches of time be very close to the most probable one of maximal entropy. There will always be tiny, essentially unobservable deviations from that state but just occasionally there will be large deviations. The time between such fluctuations will be many, many times greater than their duration. It is true that the entropy of the system can decrease and that therefore the molecular theory of heat cannot derive the second law of thermodynamics as an absolute law that never fails. Indeed, for a small number of molecules (which had never yet been observed) Boltzmann grants violations of the second law must be expected to occur frequently.

However, the real question is whether "the mechanical viewpoint led to some consequence that was in contradiction to experience". This would be the case if a system were observed to pass from a low to a high entropy state and then return to the original low-entropy state *in an observable length* of time. If not, we would never actually see any such violation of the second law. Boltzmann was able to give a simple example involving "a trillion tiny spheres" which showed that the recurrence time must be "comfortingly large". Indeed, the length of the time "makes any attempt to observe it ridiculous". However, Zermelo, surely egged on by Planck, was a true gadfly, and that was not the end of the argument. Before we come to Zermelo's second attack, two further passages from Boltzmann's first response will be important later.

Towards the end of his paper, Boltzmann briefly turns his attention from the molecules in a box and says:

> Naturally, we cannot expect from natural science an answer to the question—how does it happen that at present the bodies surrounding us are in a very improbable state—any more than we can expect from it an answer to the question why phenomena exist at all and unfold in accordance with certain given laws.

Although, like Boltzmann, I suspect that any answer to his second question, if it exists at all, is probably far beyond our present ken, I do think there may well be a satisfactory answer to his first question; among other things it explains why we have those gifts of nature (footnote 2) all around us and why I have written this book—and indeed am able to revise it now. In fact, I think that in the paper we are discussing, Boltzmann already gives some hints about where the answer to his first question, and with it the resolution of the problem of time's arrows and the second law, is to be sought: through consideration of the universe in its totality and not just things we can observe around us. It is not that he says much that is at all explicit. However, having noted that the Poincaré theorem suggests "the entire universe must return to its original state after a sufficiently long time" he asks:

> How shall we decide, when we leave the domain of the observable, whether the age of the universe, or the number of centres of force which it contains is infinite? Moreover, in this case the assumption that the space available for motion, and the total energy, are finite, is questionable.

[43] Boltzmann actually discussed the equivalent behaviour of his H-function.

Brief as these comments are, I argue in that they are critical. This applies especially to the second caveat: if the space available for motion is not finite, the universe does not have a phase space of bounded Liouville measure and the recurrence theorem fails. The system is not in a box. The implications are spelled out in *The Janus Point*.

Here we continue with the 'battle' over the status of the second law fought out in the pages of the *Annalen der Physik*. Zermelo's response to Boltzmann's response came soon. He grants he had not been fully familiar with "Herr Boltzmann's investigations of gas theory" and agreed that one could choose either the Carnot–Clausius principle, in accordance with which entropy never decreases, or the fundamental modification entailed in the mechanical viewpoint. He then continued with a passage that perfectly captures the sceptical positivistic attitude to atomic theory then prevalent among many Continental scientists:

> As for me (and I am not alone in this opinion), I believe that a single principle summarizing an abundance of established experimental facts is far more reliable than a mathematical theorem, which by its nature represents only a theory which can never be directly verified; I prefer to give up the theorem rather than the principle, if the two are inconsistent.

A little later, Zermelo raises, rather effectively, a new issue. He notes that the discussion concerns

> the entropy of *any arbitrary* system free of external influences. How does it happen, then, that in such a system there always occurs only an *increase* of entropy and *equalization* of temperature and concentration differences, but never the reverse? ... It seems to me that probability theory cannot help here, since every increase corresponds to a later decrease, and both must be equally probable or at least have probabilities of the same order of magnitude.

This kind of argument leads Zermelo to conclude that the initial state one would find on examining a system of molecules subject to Poincaré recurrence can just as well lie in a time interval in which the entropy is decreasing as one in which it is increasing. How then does it come that one only always observes an increase of entropy? He argues that "as long as one cannot make comprehensible the *physical origin* of the initial state, one must merely assume what one wants to prove". A little later he says "It is clear *a priori* that the probability concept has nothing to do with time and therefore cannot be used to deduce any conclusions about the *direction* of irreversible processes."

Predictably, Boltzmann's response to Zermelo's response followed in the next volume of the *Annalen der Physik*. I will come to the subtle argument by which Boltzmann countered Zermelo's entropy-decrease-as-likely-increase argument in a moment. But I first want to give what can be seen as the two options that Boltzmann saw at the end of his life as the best ultimate mechanical explanation of the second law of thermodynamics. To this day, almost all such explanations that have been advanced boil down to one or other of his two proposals.

Both come right at the start of the paper, where he suggests that

> the universe—or at least a very large part of it which surrounds us—started from a very improbable state, and is still in an improbable state. Hence, if one takes a smaller

system of bodies in the state in which he actually finds them, and suddenly isolates this system from the rest of the world, then the system will initially be in an improbable state, and as long as the system remains isolated it will always proceed toward more probable states.

Of his two proposals, the first—that the universe as a whole rather than only "a very large part of it which surrounds us"—has by far been taken more seriously for several decades.

Boltzmann's second proposal, later in the same paper, develops the suggestion he had made in *Nature* in 1895 and attributed there to Dr Schuetz. What he now said in 1897 was that in a sufficiently large universe which is in thermal equilibrium as a whole and therefore dead, there must be

> here and there relatively small regions of the size of our galaxy (which we call worlds), which during the relatively short time of eons deviate significantly from thermal equilibrium. Among these worlds the state probability increases as often as it decreases. For the universe as a whole the two directions of time are indistinguishable, just as in space there is no up or down. However, just as at a certain place on the earth's surface we can call "down" the direction toward the centre of the earth, so a living being that finds itself in such a world at a certain period of time can define the direction of time as going from the less probable to more probable states (the former will be the "past" and the latter the "future") and by virtue of this definition he will find that this small region, isolated from the rest of the universe, is "initially" always in an improbable state. This viewpoint seems to me to be the only way in which one can understand the validity of the second law and the heat death of each individual world without invoking an unidirectional change of the entire universe from a definite initial state to a final state.[44]

It was through passages like these that Boltzmann persuaded scientists that the experienced direction of time is aligned with the direction of entropy increase. The 1895 letter to *Nature* did not include the argument that the less probable state will be taken to be the "past". This is one of the ideas for which Boltzmann is famous; it seems have been his own addition to Schuetz's original idea. I think it is clear that neither Thomson in 1874 nor Boltzmann in 1877 had been capable of shedding the instinctive feeling that, whatever may be happening in the world, time flows forward inexorably. Also, I am not aware that Boltzmann ever made it fully explicit that intelligent beings could exist on both sides of a single localized entropy dip and would therefore live in a spatially and temporally part of the universe with *bidirectional arrows of time*.

However, near the end of the section *Application to the universe* in the second part (published in 1898) of his *Lectures on Gas Theory*, having repeated some of the comments made in the exchange with Zermelo, he does say that if unidirectional change of the entire universe from a definite initial to a final state does not occur the situation will be as follows:

> In the entire universe, the aggregate of all individual worlds, there will however in fact occur processes going in the opposite direction. But the beings who observe such

[44] The 'only way' in this last sentence seems to me to imply that Boltzmann's preference for explanation of the entropic arrow was through fluctuations that, without violating time-reversal symmetry, would create transiently existing worlds of low-entropy dips.

> processes will simply reckon time as proceeding from the less probable to the more probable states, and it will never be discovered whether they reckon time differently from us, since they are separated from us by eons of time and spatial distances $10^{10^{10}}$ times the distance to Sirius—and moreover their language has no relation to ours.

This is very close to saying explicitly that there will be bidirectional arrows associated with a single entropy dip.

If the part of Boltzmann's response considered above, which effectively killed Newton's notion of absolute time that flows uniformly "without relation to anything external", has stood the test of time rather well, I am not able to say the same about his answer to what had seemed a killer argument on Zermelo's part: why don't we see entropy decrease around us as often as we see it increase? In fact, Boltzmann makes no direct attempt to counter the argument. He merely says that large deviations from maximal entropy are enormously more rare than small ones. What he might have said but didn't is that if you do manage to catch a state with low entropy it will very soon, if not immediately, return to one of maximal entropy even if initially the entropy does decrease. This is because the probability of departures from maximal entropy get vastly more improbable the greater the departure. The system is therefore likely to find its way back toward equilibrium rapidly. This kind of argument was given later by other physicists and was in fact already anticipated by Boltzmann in 1877.

However, my belief is that the recurrence theorem has led people, beginning with Zermelo and Boltzmann, to think about the problem in quite the wrong way (as I already started to suggest in chapter 1. We need to consider *the conditions under which the problem is formulated*. The situation the founders of statistical mechanics asked us to envisage is extremely artificial but fine for their needs. It did sterling service. The setting is a sealed box that contains a gas consisting of a huge number of molecules under conditions for which the recurrence theorem holds. If we were a god-like observer of the gas with faculties as sharp as Maxwell's demon, we would surely see the gas's entropy rise and fall, mostly with the tiniest of dips but also very, very rarely with large ones. Just as Zermelo argued, the time taken to go into a dip would on average be equal to the time taken to emerge from it. However, as we are not gods and cannot see through the box, we would actually have to make measurements on it. Even if we could do this without disturbance, the chances of finding a significant departure from equilibrium are essentially zero. At least in Boltzmann's time, tests of the molecular theory of gases could never have been made in that way.

By and large, the tests that could be made, both under equilibrium and non-equilibiurm conditions (the latter involving viscous flow, for example), gave rather good support for the molecular hypothesis. In particular, Boltzmann's definition of entropy in terms of a count of microstates gave a beautiful mechanical explanation of Clausius's phenomenological $\mathrm{d}S = \mathrm{d}Q/T$ definition of entropy increments and explained why it is entropy and not heat that is conserved in Carnot's idealized steam engine. The problem lay elsewhere. Boltzmann had identified it when he said science could not explain why "at present the bodies surrounding us are in a very improbable state". Zermelo had echoed this when he said no progress could be made until one can "make comprehensible the *phyiscal origin* of the initial state".

One of the gifts of nature to which I drew attention earlier more than once is the profusion of bodies around us that enable us to make measurements and do experiments. They are all "in a very improbable state". It is their existence which makes it possible to set up systems enclosed in a box in highly non-equilibrium states that then equilibrate and are never observed to disequilibrate. Boltzmann was entirely justified in saying that the mechanical theory of heat in conjunction with probablility arguments provided a fully adequate explanation of what happens under the conditions of such experiments. The problem is not that we never see an equilibrated system disequilibrate; it is the lack of equilibrum all around us that is the problem. You will recall Carnot's comment "The phenomenon of the production of motion by heat has not been considered from a sufficiently general point of view." The problem we face now is not the nature of heat; that has long been clarified. What we need is a point of view that enables us to understand where all the bodies in highly improbable states all around us come from.

It has long been suspected that this will require us to shift our view from the laboratory to the universe. As we have just seen, Boltzmann's writings already contain hints in that direction, but there was hardly any way in which useful ideas about the universe could have been obtained in the 19th century; the necessary theory and above all observations were completely lacking. What is more, no suggestions that go significantly beyond the two that Boltzmann made to Zermelo have been made since their debate. One thing that I do find particularly surprising about the studies of time's arrows, in all the years that have elapsed since their tussle, is the failure to consider the role played throughout the development of thermodynamics and statistical mechanics by the gas container. That work all goes back, very understandably, to Carnot's idealized heat-engine and its conceptualized steam box. The box provided the perfect framework for understanding the properties and nature of heat and the associated studies led to much, much more. But when it came to the ultimate explanation of time's arrows, too many theorists were looking in the wrong direction—into the box and not out into the universe. Once you do that, you have to ask: is the universe in a box?

This is not an idle question. Ultimately it is a box, either physical or conceptual, of some kind that creates conditions under which the Poincaré recurrence theorem holds. It is also the box that, at least in macroscopic situations on human time scales, ensures the ubiquitous irreversible occurrence of equilibration, the mechanical explanation of which led Boltzmann to his greatest achievements. One of the surprising things about the post-Boltzmann study of time's arrows has been the apparently unconscious transfer of arguments appropriate under conditions when the recurrence theorem will hold to circumstances for which that seems very questionable. In this respect, Richard Feynman's discussion of the entropic arrow of time in the first volume of his justly famous *Lectures on Physics* is illuminating.

His comments on the issue of why entropy always increases come immediately after his discussion, which closely follows Boltzmann's count-of-microstates idea, of how one can define the entropy of gas confined in a box. He then considers the problem of the origin of the entropic arrow and says that

> one possible explanation of the high degree of order in the present-day world is that it is just a question of luck. Perhaps our universe happened to have had a fluctuation of some kind in the past, in which things got somewhat separated, and now they are running back together again. This kind of theory is not unsymmetrical, because we can ask what the separated gas looks like either a little in the future or a little in the

> past. In either case, we see a grey smear at the interface, because the molecules are mixing again. No matter which way we run time, the gas mixes. So this theory would say the irreversibility is just one of the accidents of life.

However, he argues that this is not a satisfactory explanation:

> Suppose we do not look at the whole box at once, but only at a piece of the box. Then, at a certain moment, suppose we discover a certain amount of order ... What should we deduce about the condition in places where we have not yet looked? If we really believe that the order arose from complete disorder by a fluctuation, we must surely take the most likely fluctuation which could produce it, and the most likely condition is not that the rest of it has also become disentangled! Therefore, from the hypothesis that the world is a fluctuation, all of the predictions are that if we look at a part of the world we have never seen before, we will find it mixed up, and not like the piece we just looked at. If our order were due to a fluctuation, we would not expect order anywhere but where we have just noticed it. [But] astronomers, for example, have only looked at some of the stars. Every day they turn their telescopes to other stars, and the new stars are doing the same thing as the other stars. We therefore conclude that the universe is not a fluctuation, and that the order is a memory of conditions when things started.

This observation leads Feynman to say that the order is not due to a fluctuation but to a much higher ordering at 'the beginning of time'; he concludes

> This is not to say that we understand the logic of it. For some reason, the universe at one time had a very low entropy for its energy content, and since then the entropy has increased. So that is the way toward the future. That is the origin of all irreversibility ... [This] cannot be completely understood until the mystery of the beginnings of the history of the universe are reduced still further from speculation to scientific understanding.

The volume of Feyman's *Lectures* from which I have quoted these words, all of which Boltzmann could have said (and in some cases did), was published in 1963. That was nearly six decades after Boltzmann's death and over three decades after the expansion of the universe had been firmly established. But despite the wide gap of time and the monumental discovery of the universe's expansion, the conceptual framework of box and recurrence fluctuations is not questioned but merely found wanting. Is it not time to ask whether the universe is in a box?

A *bon mot* often helps. In Boltzmann's interpretaion of entropy, each microstate is a point of phase space. One can also define a *microhistory.* It's just the complete history traced out from a given microstate. For a system in a Poincare-recurrence 'box', i.e., both positions and momenta restricted, the microhistories go on forever and at no stage do anything remotely interesting. Even the rare larger fluctuations lack structure. In a review of Samuel Beckett's two-act *Waiting for Godot*, the Irish theatre critic Vivian Mercier said it is "a play in which nothing happens, twice." How well that comma is placed! In conventional statistical mechanics, microhistories are stories in which nothing happens, not just twice but infinitely many times.