# THERMODYNAMICS AND RELATIVITY*

BY R. C. TOLMAN

1. *Introduction.* We have met to do honor to the memory of Josiah Willard Gibbs. By the labors of this master, the classical principles of thermodynamics were given their most complete and comprehensive expression. As the subject for the tenth memorial lecture, it seems appropriate to discuss the extensions to these classical principles which have since been made necessary by Einstein's discovery of the special and general theories of relativity.

The need for an extension of thermodynamics to relativity arises in two ways. In the first place, the classical thermodynamics was—perhaps unintentionally but nevertheless actually—only developed for systems which were tacitly assumed to be at rest with respect to the observer, and further investigation is necessary for the treatment of thermodynamic systems which are moving relative to the spatial coordinates in use. This further investigation must be carried out with the help of those principles for the inter-comparison of measurements—made by observers in uniform relative motion to each other—which form the subject matter of the special theory of relativity.

In the second place, the older thermodynamics tacitly assumed that the behavior of thermodynamic systems could be described with the help of ideas as to the nature of space and time which we now know to be approximately valid only for a limited range of space-time and in the absence of strong gravitational fields. The considerations of the classical thermodynamics were thus actually limited to the treatment of small enough systems and weak enough gravitational fields so that the deviations from this kind of space-time could be neglected, and the Newtonian theory of gravitation could be applied as a close enough approximation. In order, however, to investigate the thermodynamic behavior of large portions of the universe as we may

* The Tenth Josiah Willard Gibbs Lecture, delivered at Atlantic City, December 29, 1932, under the auspices of the American Mathematical Society, at a joint meeting of the Society with the American Physical Society, and Section A of the American Association for the Advancement of Science.

wish to do in connection with cosmological problems, and in order to obtain even in the case of small systems more precise expressions for the thermodynamic effects of gravity, it becomes necessary to extend thermodynamics to general relativity, and to make use of the more valid ideas as to the nature of space and time and the more precise theory of gravitation which Einstein has now provided.

2. *The Character and Validity of Thermodynamics and Relativity.* In carrying out these proposed extensions of thermodynamics to relativity, it proves possible to combine the known principles of thermodynamics with those of special and general relativity in a very natural manner with only small and apparently rational additions in the way of new hypothesis. Hence the character and validity of the system of relativistic thermodynamics that we obtain is largely dependent on the character and validity of the two component sciences.

In character, the classical thermodynamics may be regarded as a macroscopic, phenomenological science, which has no actual need for that interesting kind of support that can be furnished by the microscopic atomic considerations of statistical mechanics, but which attempts to treat the gross behavior of matter with the help of those generalized descriptions of the results of numerous gross experiments on the mechanical equivalent of heat and on the efficiency of heat engines, which we call the first and second laws of thermodynamics.

As to the validity of thermodynamics, we have feelings of great confidence, on account of the extensive experimental verification which exists, not only directly for the two laws themselves, but for an extraordinary number of consequences which have been drawn from them—often by elaborate but logical trains of deductive reasoning. Further additions to the principles of thermodynamics may be found, such as the newer so-called third law of Nernst and Planck, but we cannot escape the conviction that, so long as the human mind retains its present ideas of rationality, these additions are likely to prove—as in the case mentioned—supplementary rather than destructive.

Relativity, a science which changes as it does our very ideas as to the nature of space and time, has much more fundamental and far-reaching implications than thermodynamics and cannot be so easily characterized. There are, however, certain similarities between the two sciences which may be emphasized.

In the first place, at least in its present stage of development, relativity must also be regarded as a macroscopic theory dealing with ideas as to the nature of space and time which have been directly derived from macroscopic experiences. Indeed, in view of Heisenberg's uncertainty principle and the great difficulties which have been encountered in all attempts to construct a satisfactory relativistic quantum mechanics, we may even doubt whether these ideas as to space and time are really suitable for microscopic considerations. This, however, offers no difficulties if we are to combine with another macroscopic science such as thermodynamics.

In the second place, although we are often inclined to be specially impressed by the wonderful conceptual content of the theory of relativity, we may here emphasize its not negligible character as a phenomenological or descriptive science.

Thus the first of the two postulates of the special theory of relativity may be regarded as a generalized description of many failures to detect the absolute velocity of the earth's motion. And the second postulate may be regarded as a mere empirical statement of that constancy in the velocity of light, which is specially clearly demonstrated in the case of distant double stars by the lack of any effect from the changing motions of the members of the doublet on the time needed for their light to reach the earth.

Turning, moreover, to the two postulates necessary for the general theory of relativity, the principle of equivalence may be regarded not unfairly as a reasonably generalized description of Galileo's discovery that all bodies fall at the same rate. The principle of covariance, however, is on a somewhat different footing, since as first pointed out by Kretschmann—given sufficient mathematical ingenuity—any physical law whatever could undoubtedly be expressed in covariant language the same for all coordinate systems, so that the principle of covariance can imply no necessary physical consequences. Nevertheless, as emphasized by Einstein, the actual phenomena of physics must themselves be independent of the choice of coordinate system, since this is a conceptual introduction on the part of the scientist which may be made in any way that may suit his convenience or please his fancy. Hence the actual employment of invariant forms of expression in searching for the appropriate axioms of physics is desirable in order to avoid the introduction of unsuspected assumptions which might otherwise be insinuated by the use of special coordinates. We can then also, somewhat facetiously, emphasize the phenomenological character of the principle of covariance, by regarding it as a generalized description of the familiar phenomenon, that the purely conceptual activities of man—in inventing imaginary coordinate systems—are likely in first approximation to have no immediate effect on the laws of physics.

As to the validity of the theory of relativity, we have to rely on three different kinds of evidence.

In the first place, we may put its agreement with a great range of diverse facts from different branches of science, which as isolated phenomena can often be attractively explained in terms of pre-relativistic notions, but which as a whole have only been successfully correlated with the help of the theory of relativity.

In the second place, we must put those special observations which distinguish as uniquely as may be between the predictions of relativity and those which would result from other points of view. Here we have in the case of the first postulate of special relativity the demonstration of the Lorentz contraction by the Michelson-Morley experiment and all its now numerous repetitions, if we may include the extensive work of Professor Miller as demonstrating this contraction at least as the primary effect. And we also have the remarkable demonstration of Einstein's time dilation by the beautiful experiments of Kennedy and Thorndike. In the case of the second postulate of the special theory, we have as most important the precise analysis of double star orbits by de Sitter. And turning to the general theory of relativity, we have the entirely satisfactory results of the three crucial tests provided by the rotation of the perihelion of Mercury, the bending of light in passing the sun, and the shift in the wave-length of light originating on the surface of the sun and on that of the companion to Sirius.

Finally, as a third kind of evidence for judging the validity of relativity, we must not neglect the bearings of that wonderful internal coherence of the theory, with its simple foundation and elaborate but logical superstructure, which so well attests the genius of Einstein. Although such qualities can of themselves provide no guarantee as to correspondence with external phenomena, we can, nevertheless, regard them as indicating that such correspondence—when found for our present limited range of observation—is likely to persist over a much wider range of possible experience.

Like all parts of science, the theory of relativity will presumably be subject to future modifications and additions, such, for example, as might be provided by a successful unified field theory. Nevertheless, just as the Einstein theory has retained the Newtonian theory of gravitation as an exceedingly satisfactory first approximation, we may expect at least for a long time that such changes will here—as well as in the case of thermodynamics—be supplementary rather than destructive.

The character of the two sciences of thermodynamics and relativity, which we are going to combine, is thus sufficiently similar so that we may have no hesitations on that score, and may expect the resulting relativistic thermodynamics to be itself a macroscopic theory suitable for use in the description of the gross phenomena of the external world. And the validity of the two component sciences is sufficiently established so that for the present we may concentrate attention, as we shall in what follows, on the rationality of that small amount of additional hypothesis which we must introduce to effect the combination.

3. *The Extension of Thermodynamics to Special Relativity.* We are now ready to consider the actual procedure adopted in the extension of thermodynamics, first to special relativity and then to general relativity. The extension to special relativity, so as to obtain a suitable thermodynamic theory for moving systems, was made by Planck* and by Einstein,† as early as the year 1907, in that brilliant period of development which was initiated by Einstein's publication of the elements of special relativity only two years previous.

a. *Special Relativity and the First Law of Thermodynamics.* In order to appreciate the nature of this extension, let us begin by

* Planck, Berlin Berichte, 1907, p. 542; Annalen der Physik, vol. 26 (1908), p. 1.

† Einstein, Jahrbuch der Radioaktivität und Electronik, vol. 4 (1907), p. 411.

seeing what happens to the first law of thermodynamics when the extension is made.

In the classical thermodynamics for systems at rest with respect to the observer, we have found it important to distinguish two ways in which there can be an interchange of energy between a system and its surroundings, namely, through the flow of heat into the system from its surroundings and through the performance of work by the system on its surroundings. Making use of this distinction, and making use of the principle of the conservation of energy, which requires that any alteration in energy content can only result from interchange with the surroundings, we then write the first law of thermodynamics in the form given by equation (1)

$$\Delta E = Q - W, \tag{1}$$

where $\Delta E$ is the increase in the energy of the system which accompanies the influx in heat $Q$ and the performance of work $W$ against external forces.

In *form* this equation can be taken over without modification into the thermodynamics of moving systems, in the first place, since the special theory of relativity has done nothing to upset the principle of the conservation of energy, and in the second place, since we shall still wish to distinguish between the energy transfer $W$ corresponding to work done against macroscopic external forces and the other modes of transfer which we call the flow of heat $Q$.

In the application of this equation to moving instead of stationary systems, however, an important difference—which would not have been suspected in prerelativistic days—now arises on account of the relations between mass, energy, and momentum made clear by Einstein's work. To illustrate this difference, let us consider—as we usually do in thermodynamics—only very simple systems consisting of a given amount of thermodynamic fluid or working substance which exerts a pressure on its surroundings.

If such a system is *at rest*, the only way it can do work on its surroundings is by a change in volume under this pressure, and the application of the first law equation (1) then gives us simply

$$dE_0 = dQ_0 - p_0 dv_0, \tag{2}$$

where the subscript ($_0$) has been added to indicate that the quantities involved are all referred to coordinates in which the system is at rest.

If such a system is *in motion*, however, its momentum will in general change with its energy content even though we hold the velocity constant, owing to the special relativity relation which associates mass with energy. Hence in applying the first law equation to moving systems, even in the simple case of constant velocity, we shall have to include, in addition to the work done against external pressure, the work done against the external force involved in the change in momentum. We must then write in general

$$dE = dQ - p\,dv + \bar{u}\cdot d\bar{G}, \tag{3}$$

where the last term is the scalar product of the velocity of the system $\bar{u}$ and its change in momentum $d\bar{G}$. Moreover, in making use of this equation, we must employ the special relativity relation connecting the momentum of the system with its energy flow

$$\bar{G} = \frac{E + pv}{c^2}\bar{u}, \tag{4}$$

where $c$ is the velocity of light, and the term $E\bar{u}/c^2$ gives the momentum due to the transport of the energy of the system as a whole, and the term $p\,v\,\bar{u}/c^2$ corresponds to the additional flow of energy resulting from the work done on the moving volume by the action of the external pressure.

With the help of these two expressions for the first law (3), and for the momentum of a moving system (4), we can then obtain transformation equations which will give us expressions for all the quantities involved in the application of the first law to moving systems, in terms of the analogous quantities as they would be measured by a local observer moving with the system. In accordance with the known equations for force, and the Lorentz contraction for moving volumes, we are already provided by the special theory of relativity with the simple transformations for pressure and volume

$$p = p_0, \qquad v = v_0(1 - u^2/c^2)^{1/2}. \tag{5}$$

Furthermore, considering first an adiabatic acceleration in which the velocity of our system is changed without flow of heat or change in internal condition as measured by a local observer, and then considering more general processes in which flow of heat is permitted, we readily obtain as the transformation equations for energy and heat the two expressions

$$(6) \qquad E = \frac{E_0 + p_0 v_0 \dfrac{u^2}{c^2}}{(1 - u^2/c^2)^{1/2}}, \quad dQ = dQ_0(1 - u^2/c^2)^{1/2}.$$

This gives all the apparatus necessary for the application of the first law of thermodynamics to moving systems. It is to be specially noted that so far no new assumptions, beyond those already present in the mechanics of special relativity, have been introduced into our system of thermodynamics, except, if you wish, our procedure in still giving the name heat to that part of the energy transfer which does not take place through the work done against macroscopic external forces.

b. *Special Relativity and the Second Law of Thermodynamics.* Let us now turn to the more characteristically thermodynamic considerations involved in the application of the second law of thermodynamics, and examine the fate of this principle when the extension to special relativity is made.

In the classical thermodynamics the full content of the second law could be conveniently condensed into the very simple expression

$$(7) \qquad \Delta S \geqq \int \frac{dQ}{T},$$

where the left-hand side gives the increase in the entropy content $S$ when a system changes from one state to another, and the right-hand side is to be obtained by dividing each element of heat $dQ$ absorbed by its temperature $T$, and summing up for the whole process by which the system changes from its initial to its final state.

The sign of equality ($=$) in this expression applies to reversible processes which take place with that highest possible efficiency, which would just be sufficient to permit a return *both* of the system and its surroundings to their original state. And the sign of inequality (>) applies to those less efficient, actual processes which we ordinarily encounter in nature. With the help of the relation of equality we can then calculate the entropy of any system by considering an ideal reversible process by which it could be brought from its standard state to the state under consideration. And with the help of the two relations of equality and inequality, we codify all that extraordinary range of information as to the equilibrium and efficiency of physical-chemical processes which is subservient to the second law.

In making the extension to special relativity, it was found possible to take over this expression for the second law of thermodynamics as a postulate without any change at all in form. And this was evidently a rational thing to try to do since it preserves the constancy of entropy for purely mechanical processes, makes the increase of entropy for reversible thermal processes dependent on the transfer of energy in forms other than work, and retains with the help of the sign of inequality those opportunities for irreversibility and spontaneous increase in entropy which lie at the heart of thermodynamic considerations.

In applying this expression to moving systems we must of course substitute values for entropy, heat, and temperature which are appropriate for a moving system, and hence we shall desire transformation equations which will permit us to calculate these quantities in terms of the analogous quantities which would be directly measured by a local observer travelling with the system in question.

In the case of heat, we are already provided by the application of the first law with the transformation equation

$$dQ = dQ_0(1 - u^2/c^2)^{1/2}. \tag{8}$$

In the case of entropy, we are then directly led by the postulate itself to the conclusion that the entropy of a system must be an invariant for the Lorentz transformation

$$S = S_0 \tag{9}$$

owing to the possibility of changing the velocity of a system by a quasi-static reversible adiabatic acceleration, which leaves the internal state and proper entropy $S_0$ unaltered on account of the quasi-static character of the acceleration, and leaves the entropy $S$ unaltered on account of the reversible and adiabatic character of the acceleration. This invariance of entropy is, moreover, in evident agreement with the statistical mechanical interpretation which relates the entropy of a system to the probability of its state, a quantity which could hardly be a function of the velocity with which the observer happens to be moving past the system.

Finally in the case of temperature, by combining the requirements of the postulate itself with the two transformation equations already obtained, it is evident that we are now necessarily led to the relation

$$T = T_0(1 - u^2/c^2)^{1/2} \tag{10}$$

in order that the postulated law (7) may apply to the description of a given change in state both from the point of view of a local observer moving with the system and from the point of view of other observers with respect to which the system is in motion.

c. *Discussion of the Extension to Special Relativity*. This completes all that is necessary for the extension of thermodynamics to special relativity.

It will be seen that the additions in the way of new hypotheses, beyond what is already contained in the special theory of relativity and in the classical thermodynamics, have really been very small and apparently rational. Indeed, it seems fair to say that these additions consist solely in the assumption that the second law of thermodynamics, as expressed in the usual well known form given by (7), will not break down when we turn to the consideration of moving systems, and that the quantity $dQ$ occurring in this expression must still be interpreted as that part of the energy transfer which cannot be considered as work done against macroscopic external forces.

It should also be noted that the results which are given by this extended theory are entirely coherent with the accepted body of theoretical physics. For example, the application of this theory to determine the dynamical properties of a moving enclosure filled with black-body radiation leads to the same results as were originally obtained by Mosengeil* from strictly electromagnetic considerations. And the transformation equation given for heat which we have regarded as derived from an application of the mechanics of special relativity to the behavior of a portion of fluid, agrees with that which can be derived from electromagnetic considerations for the Joule heating effect in a moving electrical conductor. Most important of all, however, it should be noted that the extension has been so devised that any predictions which we make with its help as to the behavior of a given system moving with a *constant* velocity $u$ will completely agree with those which would be made with the help of the classical thermodynamics by a local observer who moves along with the system in question.

It is important to emphasize these qualities of rationality and coherence, since our judgement as to the validity of this extension of thermodynamics must be largely based thereon. Any direct test of the extension would for the present be out of the question, since all the various thermodynamic quantities for moving systems were found to differ from the analogous ones for stationary systems only by terms of the order of $u^2/c^2$ or higher, and we could only expect differences of this practically undetectable order for any thermodynamic theory of moving systems that might be proposed.

The usefulness of the extension consists partly in the ease with which we can now treat problems by simple thermodynamic methods which would otherwise involve complicated kinetic theory or electromagnetic considerations, as in the case of the moving enclosure filled with radiation. The usefulness of the extension depends mainly, however, on the increased insight which we now have into the nature of thermodynamics and thermodynamic quantities. Thus the invariance to the Lorentz transformation for entropy and for the ratio of heat to temperature provided by the special theory of relativity prove essential for the further extension of thermodynamics to general relativity to which we must now turn.

4. *The Extension of Thermodynamics to General Relativity.* In

* Mosengeil, Annalen der Physik, vol. 22 (1907), p. 867. The results of Mosengeil were employed by Planck in his method of obtaining the extension of thermodynamics to special relativity.

the general theory of relativity, the space-time continuum in which physical events take place is regarded as characterized by the formula for interval

$$(11)\qquad \begin{aligned} ds^2 &= g_{11}dx_1^2 + 2g_{12}dx_1dx_2 + \cdots + g_{44}dx_4^2 \\ &= g_{\mu\nu}dx^\mu dx^\nu, \qquad (g_{\mu\nu} = g_{\nu\mu}), \end{aligned}$$

where $x_1$, $x_2$ and $x_3$ are the three spatial coordinates that are being used, $x_4$ is the temporal coordinate, and the $g_{\mu\nu}$ are the ten gravitational potentials. The dependence of these gravitational potentials on the distribution of matter and energy is given by Einstein's ten field equations

$$(12)\qquad -8\pi T^{\mu\nu} = R^{\mu\nu} - \tfrac{1}{2}Rg^{\mu\nu} + \Lambda g^{\mu\nu},$$

where $T^{\mu\nu}$ is the energy-momentum tensor, $R^{\mu\nu}$ and $R$ are obtained from the Riemann-Christoffel tensor by contraction, and $\Lambda$, the so-called cosmological constant, is a quantity which is observationally known in any case to be exceedingly small when expressed in reciprocal square centimeters, and may well be zero. Finally, the motion of free particles and light rays in this space-time continuum is determined by the equation

$$(13)\qquad \delta \int ds = 0,$$

with $ds$ greater than zero for material particles and equal to zero for light rays.

The results predicted by these fundamental equations of general relativity are in satisfactory agreement with all the facts that are now at our disposal, and, in particular, agree with the astronomical observations which have furnished the three crucial tests of relativity.

In order to include thermodynamics within this framework, we must now enquire into the analogues in general relativity of the ordinary first and second laws of thermodynamics.

a. *The Analogue of the First Law in General Relativity.* In the case of the first law the procedure to be adopted is clear. In the classical thermodynamics the first law was an expression of the principle of the conservation of energy as applied to small stationary systems in the absence of a gravitational field, and in relativistic thermodynamics we must evidently use as the analogue of the first law the more general energy-momentum principle provided by relativistic mechanics.*

This principle can be expressed by the very simple *tensor* equation

$$(14)\qquad (T^{\mu\nu})_\nu = 0,$$

and may be regarded as an immediate result of Einstein's field equations (12), since the tensor divergence of the expression there given for the energy-momentum tensor $T^{\mu\nu}$ can be shown to be necessarily identically equal to zero. For purposes of computation it is often more convenient to rewrite this equation in the *tensor density* form

$$(15)\qquad \frac{\partial \mathfrak{T}_\mu^\nu}{\partial x^\nu} - \frac{1}{2}\mathfrak{T}^{\alpha\beta}\frac{\partial g_{\alpha\beta}}{\partial x^\mu} = 0.$$

And to obtain an insight into the nature of the principle, it is sometimes useful to rewrite it in the form of an ordinary divergence as expressed by the non-tensor yet nevertheless *covariant* equation

$$(16)\qquad \frac{\partial}{\partial x^\nu}(\mathfrak{T}_\mu^\nu + \mathfrak{t}_\mu^\nu) = 0,$$

where the pseudo-tensor density of potential energy and momentum $\mathfrak{t}_\mu^\nu$ is defined for *all systems* of coordinates in such a way that we can substitute $\partial \mathfrak{t}_\mu^\nu/\partial x^\nu$ for the second term of (15).

To remind us of the physical significance of these familiar equations of relativistic mechanics, it will be recalled that the equations reduce in the absence of a gravitational field to the ordinary principles of special relativity for the conservation of the energy and momentum directly associated with matter and radiation. In general, however, in the presence of gravitational fields, it will be evident from the third form (16) in which the equations have been written, that they will lead to conservation laws only when we include—along with the energy and momentum directly associated with matter and radiation—the potential energy and momentum of the gravitational field, which

---

* Tolman, Proceedings of the National Academy, vol. 14 (1928), p. 268; Physical Review, vol. 35 (1930), p. 875.

corresponds to the presence of the pseudo-tensor density $\mathfrak{t}_\mu{}^\nu$ in the equation in this form (16).

This general result proves to be of great importance for relativistic thermodynamics by permitting—even in the case of isolated systems—an increase in the energy directly associated with matter and radiation at the expense of the potential energy that we assign to the gravitational field. For example, if we consider a system composed of a perfect fluid having the proper macroscopic density of energy $\rho_{00}$ and proper pressure $p_0$ as measured by a local observer at the point of interest, and having no flow of heat, it is known that we can write

$$T^{\mu\nu} = (\rho_{00} + p_0)\frac{dx^\mu}{ds}\frac{dx^\nu}{ds} - g^{\mu\nu}p_0 \tag{17}$$

as an expression for the energy-momentum tensor. And if we substitute this expression into the above equations of relativistic mechanics, we can obtain for any infinitesimal element of the fluid of proper volume $\delta V_0$ the relation

$$\frac{d}{dt_0}(\rho_{00}\delta V_0) + p_0\frac{d}{dt_0}(\delta V_0) = 0. \tag{18}$$

From one point of view there is nothing surprising about this result since it merely states that a local observer who examines the behavior of an element of the fluid small enough so that the gravitational *curvature* of space-time can be neglected will find the rate of change in energy content related in the expected way to the work done against the external pressure. From another point of view, however, as this same equation can be applied to each one of all the elements into which the total fluid of the system can be divided, the result may seem somewhat surprising, since it leads to the possibility of systems in which the proper energy of every element of the fluid may be simultaneously decreasing or increasing, according as the system is expanding or contracting. Moreover, since it is this proper energy immediately associated with matter and radiation which determines the possibilities for entropy increase, we shall later find in relativistic thermodynamics an escape from certain restrictions imposed in the classical thermodynamics by the usual form of the principle of the conservation of energy.

Just as in the previous case of special relativity, we note that the extension of thermodynamics to general relativity involves, so far as the first law is concerned, no new hypothetical material beyond that already contained in relativistic mechanics. And we may now turn to the relativistic analogue of the second law of thermodynamics.

b. *The Analogue of the Second Law in General Relativity.* To guide us in obtaining a suitable postulate to serve as the relativistic second law of thermodynamics, we must make use of the two fundamental ideas of general relativity which are expressed by the principles of covariance and equivalence. In accordance with the principle of covariance, our postulate must be expressed in covariant form the same for all coordinate systems, to avoid the danger of being influenced in its selection by a spurious simplicity when referred to some particular system of coordinates. And in accordance with the principle of equivalence, our postulate must reduce to the thermodynamic requirements of special relativity, when applied to an infinitesimal element of fluid, using natural coordinates for the point of interest.

These two principles have been sufficient to lead with considerable confidence to the expression*

$$\frac{\partial}{\partial x^{\mu}}\left(\phi_0 \frac{dx^{\mu}}{ds}(-\ g)^{1/2}\right)\delta x_1\,\delta x_2\,\delta x_3\,\delta x_4 \geqq \frac{\delta Q_0}{T_0} \tag{19}$$

as the appropriate postulate to take as the relativistic analogue of the ordinary second law of thermodynamics. The quantity $\phi_0$ in this expression is the proper entropy density of the thermodynamic fluid under consideration as measured at the point of interest by a local observer; the quantities $dx^{\mu}/ds$ are the components of the macroscopic *velocity* of the fluid at that point; and the other quantities on the left-hand side of the expression have their usual significance. The significance of the right-hand side of the expression is more difficult to grasp, and will be specially treated in a forthcoming article by Robertson and myself.† The quantity $\delta Q_0$ may be taken as the heat, measured by a local observer at rest in the fluid at the point of interest, which

* Tolman, Proceedings of the National Academy, vol. 14 (1928), pp. 268, 701; Physical Review, vol. 35 (1930), p. 896.

† Tolman and Robertson, submitted to the Physical Review.

flows into an element of the fluid having the instantaneous proper volume $\delta V_0$ during the proper time $\delta t_0$, where these quantities are so chosen as to make

$$\delta V_0\, \delta t_0 = \sqrt{-g}\, \delta x_1\, \delta x_2\, \delta x_3\, \delta x_4, \tag{20}$$

and the quantity $T_0$ is taken as the temperature ascribed to this heat by the local observer.

The two signs of equality $(=)$ and inequality $(>)$ in the expression refer respectively to the two cases of reversible and irreversible processes, and in applying the principle to irreversible processes we are to regard an increment in coordinate time $\delta x_4$ as positive when taken in the direction to correspond to a positive increment in proper time $\delta t_0$ as measured in the ordinary manner by a local observer.

To show the agreement of this postulated expression for the relativistic second law with the principle of covariance, we have merely to note that it is a tensor equation of rank zero—both sides being scalar invariants—and hence is true in all coordinate systems if true in one. To show its agreement with the principle of equivalence we must see what it reduces to in natural coordinates for the point of interest. Introducing such coordinates $x, y, z, t$, *and making use of the transformation equations for entropy, heat, and temperature provided by the special theory of relativity*, we find, however, that our principle then reduces to

$$\left[\operatorname{div}(\phi \bar{u}) + \frac{\partial \phi}{\partial t}\right] \delta x\, \delta y\, \delta z\, \delta t \geqq \frac{\delta Q}{T}, \tag{21}$$

where $\phi$, $\bar{u}$, $\delta Q$ and $T$ are now the quantities referred to our present coordinate system which we ordinarily designate as entropy density, velocity, heat absorbed, and temperature. And we see that this result does relate the change in the entropy of the element of fluid, instantaneously contained in the coordinate range $\delta x\, \delta y\, \delta z$, to the absorbed heat and temperature in the way

$$\frac{d}{dt}(\phi\, \delta x\, \delta y\, \delta z)\delta t \geqq \frac{\delta Q}{T}, \tag{22}$$

which is required by the second law of thermodynamics in special relativity.

At the present stage of observational knowledge, our belief in the validity of the proposed postulate is primarily based on this agreement with the two principles of covariance and equivalence. In addition, however, it may be emphasized that the principle has been chosen so as to be simply the immediate covariant re-expression of the special relativity form of the second law; and past experience has shown, notably for example in the cases of the fundamental formulas for space-time interval and geodesic trajectory, that these simplest possible covariant generalizations, when feasible, are likely to be correct. Furthermore, it may be remarked that the conclusions which have so far been drawn from this extension of thermodynamics to general relativity appear—at least after due reflection—to be reasonable and illuminating.

It must be emphasized, nevertheless, that these qualities are not sufficient to prove the validity of the postulate, since other covariant expressions might be found which would also reduce to the special relativity law in natural coordinates. Hence the postulate must be regarded as a real generalization with a range of validity to be finally determined only by the correspondence between observation and prediction.

5. *Consequences of Relativistic Thermodynamics.* To complete our discussion, we must now consider the possible consequences of relativistic thermodynamics. The technical modifications in thermodynamic theory needed to secure its extension to relativity may have seemed too trivial and obvious to warrant the expectation that these consequences could be very novel or interesting. Nevertheless, the actual effect of replacing classical ideas as to the nature of space and time by relativistic ideas is so fundamental as to lead to important differences between the results of classical and relativistic thermodynamics. Three examples may now be given to illustrate both the essential novelty and the inherent rationality of the consequences of relativistic thermodynamics.

a. *Temperature Gradient in Gravitational Field.* In the classical thermodynamics we have become accustomed to the conclusion that a system which is in thermal equilibrium will necessarily have uniform temperature throughout. As a result of relativistic thermodynamics, however, it is found that this conclusion must be modified in the presence of appreciable gravitational fields. Thus if we consider a spherical distribution of material held together in static equilibrium by its own gravitational forces, corresponding to the line element

$$(23) \qquad ds^2 = -e^{\mu}(dr^2 + r^2 d\theta^2 + r^2 \sin^2 \theta d\phi^2) + e^{\nu} dt^2,$$

where $\mu$ and $\nu$ are functions of $r$ alone, it can be shown* that the condition for thermal equilibrium is given by

$$(24) \qquad \frac{d \log T_0}{dr} = -\frac{1}{2}\frac{d\nu}{dr},$$

where $T_0$, the proper temperature as measured by a local observer, decreases as we go outward instead of remaining constant. And for the more general static case of a system, not necessarily having spherical symmetry, but corresponding to the more general line element

$$(25) \qquad ds^2 = g_{ij}\, dx^i\, dx^j + g_{44} dt^2, \qquad (i, j = 1, 2, 3),$$

where $g_{ij}$ and $g_{44}$ are independent of $t$, it has been shown by Ehrenfest and myself† that the condition for thermal equilibrium would be given by a constant value throughout the system for the product

$$(26) \qquad T_0(g_{44})^{1/2} = \text{const.}$$

rather than for the proper temperature $T_0$ itself.

From an observational point of view, this new conclusion may not be very important, since the change in temperature with position would only be

$$(27) \qquad \frac{d \log T}{dr} = 10^{-17}\ \text{cm}^{-1}$$

in a gravitational field having the intensity of that at the earth's surface. From a theoretical point of view it is of interest, however, in leading to a modification of one of the most cherished results of the classical thermodynamics.

It should be emphasized, nevertheless, that this modification is entirely reasonable from the point of view of relativity. In accordance with that theory all forms of energy have mass and weight, and it is hence indeed not surprising that a temperature

* Tolman, Physical Review, vol. 35 (1930), p. 904.

† Tolman and Ehrenfest, Physical Review, vol. 36 (1930), p. 1791.

gradient is necessary to prevent the flow of heat from regions of higher to those of lower gravitational potential when thermal equilibrium has been reached. Furthermore, since the pressure of black body radiation in equilibrium with a thermal system would evidently have to increase as we go to lower gravitational levels, in order to support the weight of radiation above, we can also see from purely mechanical considerations that temperature must increase with depth.

b. *Possibility for Reversible Processes at a Finite Rate.* To turn to a second example, we have long been accustomed to believe as a result of the classical theory that *reversible* thermodynamic processes taking place at a *finite rate* could never occur in nature, since a finite rate of change would make it impossible to achieve that maximum efficiency which would permit a restoration both of the system and its surroundings to their original states.

To illustrate this we may consider a fluid contained in a cylinder with non-conducting walls and provided with a piston as shown in Fig. 1. The *reversible* expansion of this fluid with the piston moving out at a *finite rate* would evidently be impossible, in the first place since there would be friction between the piston and the walls, and in the second place since the fluid in flowing in behind the moving piston would not be able to maintain as high a pressure and hence do as much external work as at an infinitesimal rate of expansion. It has hence been concluded in the past that the reversible expansion of a fluid at a finite rate would under all circumstances be impossible.

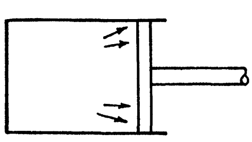

Fig. 1
Classical Cylinder

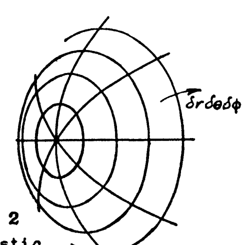


Fig. 2
Relativistic
Non-static Universe

In relativistic thermodynamics, nevertheless, this situation is altered in an important manner by new possibilities for changes to take place in the proper volume of an element of fluid, because of changes in the gravitational potentials which were neglected in the older theory. To illustrate this, Fig. 2 attempts to give a symbolic two-dimensional representation of the space-like coordinates, corresponding to the line element

$$(28)\quad ds^2 = -\frac{e^{g(t)}}{\left[1+\frac{r^2}{4R^2}\right]^2}(dr^2 + r^2 d\theta^2 + r^2\sin^2\theta d\phi^2) + dt^2$$

which characterizes the homogeneous distribution of fluid that we usually call a non-static model of the universe. These coordinates are so selected that any element of fluid lying in a given coordinate range $\delta r\ \delta\theta\ \delta\phi$ will remain permanently therein. The proper volume of such an element of the fluid as measured by a local observer will then be

$$(29)\qquad \delta V_0 = \frac{e^{(3/2)g(t)}}{\left[1+\frac{r^2}{4R^2}\right]^3} r^2 \sin\theta\ \delta r\ \delta\theta\ \delta\phi,$$

and this will change with the time $t$ at a finite rate on account of the occurrence of the function $g(t)$ in the gravitational potentials and hence also in the expression for proper volume.

These changes in proper volume take place, however, in the first place quite obviously without any friction of moving parts, and in the second place with a perfect balance between internal and external pressures owing to the uniformity of conditions throughout the fluid. Thus the two sources of irreversibility in the previous classical illustration are completely eliminated. Furthermore, there can be no irreversible heat flow in the model under consideration, again owing to the uniform conditions throughout the fluid, and the possibility of irreversible processes within each element of the fluid can be eliminated by the choice of a sufficiently simple substance. As a consequence it has proved possible to construct conceptual models of this kind* which can expand or contract at a finite rate, and nevertheless satisfy the relativistic condition for reversibility by maintaining the sign of equality instead of inequality in the simple expression

* Tolman, Physical Review, vol. 37 (1931), p. 1639; ibid., vol. 38 (1931), p. 707.

$$\frac{d}{dt}(\phi_0 \delta V_0) \geqq 0, \tag{30}$$

which results from the application of the general form of the relativistic second law (19) to the present case.

This possibility for reversible thermodynamic processes taking place at a finite rate seems very strange from the point of view of the classical thermodynamics. Nevertheless, from the relativistic point of view the result seems both rational and, indeed, perhaps inevitable, when we recall that the principles of *relativistic mechanics* alone, without bothering about thermodynamics, have been found sufficient to show* the possibility of constructing models of the kind in question, which—when filled with such simple fluids as a perfect monatomic gas or black-body radiation—would expand at a finite rate to an upper limit and then precisely retrace their steps to their original density, pressure, and temperature.

c. *Possibility for Irreversible Processes without Reaching a Final State of Maximum Entropy*. As a final example of the differences between classical and relativistic thermodynamics, let us now turn to the classical conclusion that the ultimate result of irreversible processes taking place in a system having no interaction with its surroundings would necessarily be a state of maximum entropy where further change would be impossible.

To illustrate the classical reasons for belief in this principle, let us consider a simple homogeneous system consisting of a gas of uniform pressure, temperature, and composition throughout. The state of such a system can be specified by its energy $E$, volume $v$, and the number of mols $N_1$, $N_2$, etc., of the different chemical substances which it contains, and in accordance with a fundamental expression, provided by the work of Gibbs† himself, its entropy can then be determined with the help of the general equation

$$dS = \frac{dE}{T} + \frac{p}{T}dv + \left(\frac{\partial S}{\partial N_1}\right)dN_1 + \left(\frac{\partial S}{\partial N_2}\right)dN_2 + \cdots . \tag{31}$$

* Einstein, Berlin Berichte (1931), p. 235; Tolman, Physical Review, vol. 38 (1931), p. 1758.

† Allowing for the difference in notation and form, equation (31) is equivalent to the fundamental equation (12) given in the *Collected Works of Gibbs*, Longmans, Green and Company, 1928, vol. I, p. 63.

In applying this expression, however, to the special case of a system having no interaction with its surroundings, it was necessary in the classical thermodynamics to take the energy change $dE$ equal to zero to agree with the classical principle of energy conservation. It was also necessary, for a system having no interaction with its surroundings, to take either the volume-change $dv$ equal to zero or the pressure $p$ itself equal to zero to prevent the performance of work on the surroundings. In both cases the only chance for entropy increase and hence for change would then lie in the readjustment of chemical composition. At constant energy and volume, moreover, this would have to cease when the entropy reached the maximum possible value compatible with the fixed value of energy and volume; while at constant energy and zero pressure the gas would be infinitely dilute and a final state of maximum entropy would be reached when all the molecules of gas had disassociated into their atoms.

In the relativistic treatment of this problem, nevertheless, the application of relativistic mechanics alone is sufficient to show a very different state of affairs. In the relativistic treatment, the analogue of the preceding homogeneous isolated system will evidently be a non-static model of the universe, since the line element of these models corresponds to a completely homogeneous distribution of fluid with no interaction with anything outside the system itself. By the application of relativistic mechanics, however, it has been shown by the work of Ward and myself,* that there is a great class of such models, obtained by taking the cosmological constant $\Lambda$ equal to zero or less, which could apparently undergo a continued succession of expansions and contractions without ever coming to rest. To be sure, the equations that we now have available for our highly idealized models are only sufficient to describe the expansion of the models to their upper limit and return, and not sufficient to describe their passage through the exceptional point at their lower limit of volume. Nevertheless, on physical grounds we must be inclined to assume that contraction to the lower limit of volume would be followed by renewed expansion.

It is to be specially noted now, however, that our conclusion

* Tolman, Physical Review, vol. 30 (1932), p. 320; Tolman and Ward, ibid., vol. 39 (1932), p. 835.

as to the possibility of such behavior is based on mechanical equations alone and is quite independent of assumptions as to the reversibility or irreversibility of any thermodynamic processes that may take place in the fluid, although in the case of irreversible processes it can be shown that successive expansions have a tendency to go to larger and larger volumes before returning. The possibility thus provided for an unending succession of irreversible expansions and contractions seems very strange from the point of view of classical thermodynamics, but can nevertheless be shown to be reasonable from the point of view of relativistic thermodynamics.

The application of the relativistic second law to these non-static models leads to the general result

$$\frac{d}{dt_0}(\phi_0 \delta V_0) \geqq 0, \tag{32}$$

which states that the proper entropy of each element of the fluid as measured by a local observer can only remain constant or increase with local time. With a fluid of composition simple enough to eliminate the possibility for changes other than in density, we obtain the conditions necessary for constant entropy, and are led to the reversible expansions and contractions at a finite rate previously considered. With a slightly more complicated fluid, however, such as a gas which tends to disassociate as the density is lowered, we can obtain the irreversible increases in entropy in which we are now interested, since the composition of the fluid will then lag behind the changes in volume, and the processes of disassociation and recombination will always take place in the direction of an equilibrium which has not been attained.

From a classical point of view, such a continuous irreversible increase in entropy in an isolated system which undergoes a never-ending succession of expansions and recontractions to an earlier volume might seem impossible, since the energy of the isolated system would have to remain constant, and with a given value of energy and volume there would be a definite, maximum possible value of entropy.

From the relativistic point of view, nevertheless, it is easy to see that this continued increase in entropy could be made possible by the failure in relativistic mechanics of the ordinary principle of energy conservation. Indeed the application of relativistic mechanics to the models in question leads to the result

$$\frac{d}{dt_0}(\rho_{00}\delta V_0) + p_0\frac{d}{dt_0}(\delta V_0) = 0, \tag{33}$$

which shows that the proper energy $(\rho_{00}\delta V_0)$ of each element of fluid in the model will decrease as the model expands and increase as it contracts, and since the general effect of irreversibility would be to give higher pressures during compression than during the preceding expansion, we are thus provided with the needed mechanism for a gradual increase in proper energy and hence also in the proper entropy of the elements of fluid.

The difference between the classical and relativistic points of view can also be illustrated in a somewhat different manner, if we return to the classical cylinder and relativistic model of the universe given by Figures 1 and 2, and consider a continued succession of expansions and contractions, in both cases providing for irreversibility by again introducing for example a gas which tends to disassociate when the pressure is lowered.

The classical man who is engaged in moving the piston of his cylinder successively in and out, reports that the entropy inside the cylinder is all the time getting greater and greater, because the gas cannot disassociate fast enough on the way out nor recombine fast enough on the way in to maintain equilibrium. He also reports, however, that the energy inside his cylinder—which is now no longer an isolated system—is also getting greater and greater, since the pressure has a tendency to be too high to correspond to equilibrium on compression and too low on expansion, so that he does more work on the way in than he gets back on the way out. He then finally remarks that this continued increase in energy indeed provides in the case of this non-isolated system an opportunity for the entropy to go on increasing forever, but that he personally is getting very tired of the silly experiment, and is going to have to stop:—not now because the irreversible processes will ever lead to an unsurpassable maximum of entropy but because he actually won't have the necessary energy to push in the piston one time more.

How different the remarks of the relativistic man as he sits calmly by and—in his mind's eye—watches his conceptual model expand and contract. "Yes," he says, "I notice that the proper entropy of each element of fluid is all the time increasing owing to successive processes of disassociation and recombination under conditions that do not correspond to equilibrium." "I do not worry, however, since I also notice that the proper energy of each element of fluid is also increased after a succession of expansions and contractions sufficiently to allow for this increase in entropy, and I know that the principles of relativistic mechanics not only permit such increases in proper energy, but indicate—without reference to reversibility or irreversibility—the continued succession of expansions and contractions that I observe."

This completes the discussion of examples illustrating the kind of conclusions that we may expect from the extension of thermodynamics to general relativity.

6. *Conclusion.* Much remains to be done in the further application and development of relativistic thermodynamics. The application to systems in which heat flow is taking place from one portion of the fluid to another has not yet been undertaken and might lead to interesting results. It could be of special importance when we have made more progress in the treatment of non-homogeneous models of the universe which may well be necessary for a better understanding of the actual universe. The development of relativistic thermodynamics to include an appropriate treatment of fluctuations might also be attempted with the help of the statistical methods introduced by Gibbs himself. The effect of fluctuations on the behavior of cosmological models might be specially important at certain stages of their history.

In trying to estimate the significance of the applications of relativistic thermodynamics that have already been made, we must not misjudge the nature of the two applications to cosmological models that were described in the foregoing. It should be emphasized that the homogeneous cosmological models which we now consider are not only very highly simplified and idealized, but at best are constructed to agree throughout their entire extent with that small sample of the actual universe which lies within the range of some $10^8$ light years. Furthermore, it must be remembered that among the different possible kinds of homogeneous cosmological model, there is a class which would expand never to return as well as the class that could undergo an unending succession of expansions and contractions; and we do not now have sufficient data so that we could assign the actual universe to either class. Hence we must be very careful in extrapolating to the actual universe any conclusions that we may draw as to the behavior of our conceptual models.

The dangers of long-range scientific extrapolation were specially emphasized to you by Professor Bridgman in his beautiful Gibbs memorial lecture of last year, and these dangers, as we see, are specially present in cosmological speculations, based on the observation of a small fragment of the total universe for an inappreciable time span. Nevertheless, it certainly seems significant that conceptual models of the universe can be constructed, which are permitted by the principles of relativistic thermodynamics to exhibit behavior in serious conflict with the classical conclusions, that reversible processes could never take place at a finite rate, and that the end result of irreversible processes would necessarily be a stationary condition of maximum entropy. Hence at the very least, it would now seem desirable to extrapolate to a cautious position, in which we no longer dogmatically assert that the principles of thermodynamics necessarily require a universe created at a finite time in the past and fated for stagnation and death in the future. Indeed, the chief duty and glory of theoretical science is certainly not merely to describe in complicated language those facts that are already known but to extrapolate—as cautiously and wisely as may be—into regions yet unexplored but pregnant with human interest.

California Institute of Technology