On the Principle of Carathéodory

H. A. Buchdahl

Additional information on Am. J. Phys.

On the Principle of Carathéodory

H. A. Buchdahl

University of Tasmania, Hobart, Tasmania

I

IN most treatments of thermodynamics the Second Law is stated in one or other of the original forms resulting from Clausius and Thomson. However, Carathéodory, in his axiomatic development of thermodynamics, replaced the traditional statements of the Second Law by what has become known as the *Principle of Carathéodory*. A more widespread knowledge of the methods of Carathéodory[1] seems desirable, not least because of their great didactic value. Experience shows that they can be understood by undergraduates in the second or third year of a course in physics. These methods involve, however, one mathematical theorem (the *Theorem of Carathéodory*), the usual proofs[2] of which are often so unpalatable to the physicist, that the theorem itself may form a serious obstacle to a proper understanding of the whole treatment. We shall therefore give in a subsequent paper an alternative proof of this theorem for the case of three variables, which may be found more attractive than the proof usually offered in text books. It is desirable in the development of any physical theory that there should be a clear-cut division between empirical content and mathematical method. In the case we are considering, once Carthéodory's Theorem is understood as a theorem in pure mathematics, the existence of a certain single-valued function of the variables of state is at once seen to be an immediate consequence of the generalized empirical knowledge which is contained in the Second Law (in the form of Carathéodory's Principle). In the usual treatments, the existence of this function is generally proved with the aid of abstract engines and cycles—a method which may leave some students without much appreciation of what has been proved, and without too clear an understanding of the phenomenological meaning of entropy.

II. The Principle of Carathéodory

Keeping the preceding remarks in mind, and in order to emphasize the close analogy which exists between the 'physical argument' and the 'mathematical argument' of this treatment we propose to deal in this paper with some general considerations concerning the Principle of Carathéodory. The latter may be stated as follows:

In the neighborhood of any arbitrary initial state J_0 of a physical system there exist neighboring states J which are not accessible from J_0 along adiabatic paths.

This principle thus takes as a starting point the empirical recognition that if two states, J_0 and J, of a given adiabatically enclosed[3] thermodynamic system be prescribed, and granted (i) that the transition from J_0 to J is mechanically possible, and (ii) that such a transition would not violate the demands which the First Law of Thermodynamics already imposes upon it; then the transition from J_0 to J may nevertheless be impossible, while at the same time the reverse transition is possible. We then say that the *thermodynamic weight* of J_0 exceeds that of J.

Let us consider an elementary example. If $J_0 \rightarrow J$ stands for the phrase 'the transition of the system from the state J_0 to the state J,' let $J_0(h_0, t_0)$ be the state of the Joule paddle-wheel apparatus, the contents of the calorimeter being at temperature t_0, and the mass m at height h_0. Let $J(h, t)$ be a second state of the system, where $t < t_0$ and $h > h_0$, in such a way that the energy difference of the contents of the calorimeter, to which corresponds the temperature difference $t_0 - t$, is just accounted for by the potential energy difference of the mass $(= mg(h - h_0))$ in the two states of the system respectively. Then we know empirically that $J_0(h_0, t_0) \rightarrow J(h, t)$ is impossible, notwithstanding the fact that neither the First Law, nor the laws of mechanics would

---

[1] Carathéodory, *Math. Annalen.* **67**, 355 (1909) and *Sitz. d. Preu. Akad. d. Wiss.*, p. 39 (1925).

[2] Born, *Physikalische Ztschr.* **22**, 251 (1921) and Chandrasekhar, *Stellar Structure* (Univ. Chicago Press, 1939), Ch. I, p. 21 and *Carathéodory*, (1909), *loc. cit.*, p. 369.

---

[3] A system is said to be adiabatically enclosed if a state of equilibrium of the system can be disturbed only by mechanical means, as by shaking, stirring, or the passage of electric currents.

be violated in this transition; $J(h, t) \rightarrow J_0(h_0, t_0)$ is, however, possible.

III. Properties of Systems and Single-Valued Functions

In order to simplify the following considerations, which are of a fairly general kind, let a system K consist of a gas within an envelope,[4] the volume occupied by the gas being v, at a pressure p. We regard the mechanical variables p and v as the independent *variables of state*, i.e., the quantities p and v define the state of the system, and are within certain limits variable at will. In a manner which we need not consider here, the conditions for thermal equilibrium lead us to associate with given values of p and v a number t, such that two such systems K and K' can be in thermal equilibrium if, and only if, the corresponding numbers t and t' are equal.[5] That is, empirical knowledge concerning the thermal equilibrium of physical systems leads to the definition of a single-valued function $t(p, v)$ of the variables of state, which expresses a new property of the system, *viz.*, the property of being, or of not being in thermal equilibrium with another system when the two are brought into non-adiabatic contact. Any other such definitive property,[6] often expressed in the form of a natural law, may similarly be expected to lead to the attachment of a certain number to every given state, i.e., to the definition of a new single-valued function of the variables of state associated with the system, which expresses this property. The First Law of Thermodynamics is an excellent example; it generalizes the result of a very great number of experiments in the statement that *the mechanical work W done by a system in any adiabatic transition between two given states depends upon these states alone, not on the manner of transition*. The definition of a new single-valued function of the variables of state, the energy U of the system, is an immediate consequence of this statement. The term 'quantity of heat' (Q) then appears merely as an abbreviation for the difference between the actual work done in a given non-adiabatic transition and the change in the value of the energy function which occurs in it. Thus if U_0, U are the values of the energy in the initial and final states respectively, then

$$Q = (U - U_0) + W. \tag{1}$$

IV. The Existence of the Entropy Function

After the preliminary observations of the preceding section we return to the consideration of Carathéodory's Principle. As we have seen, the latter expresses a definitive property of the system, *viz.*, the property that, when adiabatically isolated, the possibility or impossibility of $J_0 \rightarrow J$ depends upon J_0 and J alone, subject to certain other well-defined conditions being already satisfied. Accordingly we may expect the principle to lead to a new single-valued function of the variables of state S,[7] such that S is a measure of the thermodynamic weight of the state J. We call S the entropy of the system. It follows at once, the sign of S being suitably chosen, that $J_0 \rightarrow J$ is possible if $S \geqslant S_0$, and impossible if $S < S_0$; for the condition of accessibility cannot be expressed in any essentially different way in terms of a pair of numbers which must enter into the relations quasi-symmetrically. Moreover, let $S > S_0$; then $J_0 \rightarrow J$ is possible. But having effected $J_0 \rightarrow J$, $J \rightarrow J_0$ is now impossible, for now $S(final) < S(initial)$. That is, $J_0 \rightarrow J$ is irreversible. Clearly $J_0 \rightarrow J$ is reversible only if $S_0 = S$. The last results may be summed up as follows:

A transition of an adiabatically enclosed system is impossible, possible reversibly, or possible irreversibly according as the entropy of the initial state is greater than, equal to, or less than that of the final state.

This at once gives rise to the corollary that in any adiabatic transition of a system the entropy can never decrease. This is the so-called Principle of Increase of Entropy, which shows that unlike mass, energy, charge etc. entropy obeys a one-sided conservation law.

We shall not pursue the physical consequences

---

[4] The envelope is not to be regarded as forming a part of the system.

[5] The apparent existence of more than one value of t for given p and v, e.g., water near 4°C, would show the incorrectness of the assumption that p, v constituted a sufficient set of independent variables of state.

[6] A property of this type evidently cannot depend on the previous history of the system.

[7] It is conceivable that it could define more than one new function determining the mutual accessibility of different states, but it is difficult to see how this could come about on the basis of a law of the type under consideration. However, it appears that ultimately we must rely upon the confirmation obtained from a mathematical treatment of the problem.

of the Principle of Carathéodory beyond this point; for the elucidation of the phenomenological meaning of entropy as 'transition potential' has been dealt with at sufficient length for our purpose.

V. Plausibility Arguments Based on Carathéodory's Principle

Finally we examine briefly how the considerations above indicate to us how to begin the mathematical formulation of the consequences of Carathéodory's Principle. To do this it is sufficient to consider a system L with three independent variables of state (such as the two aforementioned systems K and K' in thermal equilibrium), which we take to be v, v' and the common temperature t. Now Carathéodory's Principle speaks of arbitrary adiabatic transitions. It applies therefore *a fortiori* to quasi-static[8] adiabatic transitions. During an infinitesimal part of it the work done by L is $pdv+p'dv'$; and since the transition is adiabatic this work must, in virtue of the definition of energy, be equal to the change dU in the energy $U(v, v', t)$ of L, i.e.,

$$(\partial U/\partial v+p)dv+(\partial U/\partial v' \\ +p')dv'+(\partial U/\partial t)dt=0. \quad (2)$$

Thus the quasi-static adiabatic transitions of L are subject to a condition of the form

$$dQ \equiv P(x, y, z)dx+Q(x, y, z)dy \\ +R(x, y, z)dz=0, \quad (3)$$

where P, Q, R are certain functions[9] of the inde-

[8] A transition of a system L is said to be quasi-static if, in the course of it, L passes through a continuous series of states of equilibrium. This is equivalent to a reversible transition, which necessarily proceeds at an infinitesimal rate.

[9] Equation (3) preserves its form under any substitution of independent variables. Thus, if the set x, y, z be given as functions of the new set x', y', z', then the equation becomes $P'dx'+Q'dy'+R'dz'=0$, where $P'(x', y', z')=P\partial x/\partial x' +Q\partial y/\partial x'+R\partial z/\partial x'$, etc. Note that the Q here has nothing to do with the symbol for quantity of heat.

pendent variables x, y, z. Interpreting the latter as right angled Cartesian coordinates of a 'picture space' A, every state of equilibrium of L is represented by a point in A. By definition a quasi-static transition must therefore be represented by a continuous curve C in A. If the transition is also adiabatic, C is restricted by Eq. (3). In other words: The 'quasi-static adiabatics' of L are the solution curves of the differential Eq. (3).

But in Sec. IV we tentatively accepted the existence of a certain function S, and we saw that it remains constant in a quasi-static adiabatic transition. That is, as a consequence of the Second Law there exists a function S such that the equation

$$Pdx+Qdy+Rdz=0 \quad (4)$$

implies

$$dS=0, \quad (5)$$

and which has the properties described in Sec. IV. Hence if we are on the right track we may expect that there exists another function $\omega(x, y, z)$, such that[10]

$$Pdx+Qdy+Rdz \equiv \omega dS. \quad (6)$$

We have arrived at the tentative Eqs. (5) and (6) by means of a direct physical 'plausibility argument' based on Carathéodory's Principle. These equations may now be put on a rigorous basis through an application of the Theorem of Carathéodory, the statement, and a new proof of which will form the substance of another paper,[11] in accordance with our original intention of delimiting the mathematical core of the consequences of the Second Law.

[10] For quite unrestricted P, Q, R this is, in general, impossible; in fact, the 'condition of integrability,' $P(\partial Q/\partial z-\partial R/\partial y)+Q(\partial R/\partial x-\partial P/\partial z)+R(\partial P/\partial y-\partial Q/\partial x)$ $=0$, must be satisfied. See Forsyth, *Differential equations* (Macmillan, 3rd ed. 1903), pp. 282–284.

[11] Buchdahl, see article in this issue.

I should like to mention one case where the artist—quite unknowingly—has been able to provide valuable data for science. The clay from which the Greek potter made his beautiful vases more than 2000 years ago always contained some magnetic oxide of iron. At a certain stage of the cooling, after firing, the iron particles are very susceptible to the action of magnetic forces, and orient themselves in the direction of the earth's magnetic field. The direction of this magnetization was fixed permanently when the vase cooled and since we know the vase must have been always in a vertical position during the firing, the scientific man can find the direction of this magnetization and thus fix the inclination or 'dip' of the earth's magnetic field at the time and for the place where the vase was made. By this curious observation we have been able to extend our knowledge of the secular variations in the earth's magnetic field to a remote epoch more than 2000 years before the importance of such measurements was recognized.—E. RUTHERFORD (1932).