

VI. *The Law of Partition of Kinetic Energy.*
By Lord RAYLEIGH, F.R.S.*

THE law of equal partition, enunciated first by Waterston for the case of point molecules of varying mass, and the associated Boltzmann-Maxwell doctrine respecting steady distributions have been the subject of much difference of opinion. Indeed, it would hardly be too much to say that no two writers are fully agreed. The discussion has turned mainly upon Maxwell's paper of 1879 †, to which objections ‡ have been taken by Lord Kelvin and Prof. Bryan, and in a minor degree by Prof. Boltzmann and myself. Lord Kelvin's objections are the most fundamental. He writes § : "But, conceding Maxwell's fundamental assumption, I do not see in the mathematical workings of his paper any proof of his conclusion 'that the average kinetic energy corresponding to any one of the variables is the same for every one of the variables of the system.' Indeed, as a general proposition its meaning is not explained, and it seems to me inexplicable. The reduction of the kinetic energy to a sum of squares leaves the several parts of the whole with no correspondence to any defined or definable set of independent variables."

In a short note ¶ written soon afterwards I pointed out some considerations which appeared to me to justify Maxwell's argument, and I suggested the substitution of Hamilton's *principal* function for the one employed by Maxwell ¶. The views that I then expressed still commend themselves to me; and I think that it may be worth while to develop them a little further, and to illustrate Maxwell's argument by applying it to a particular case where the simplicity of the circumstances and the familiarity of the notation may help to fix our ideas.

But in the mean time it may be well to consider Lord Kelvin's "Decisive Test-case disproving the Maxwell-Boltzmann Doctrine regarding Distribution of Kinetic Energy"**, which appeared shortly after the publication of my note. The following is the substance of the argument :—

"Let the system consist of three bodies, A, B, C, all movable only in one straight line, KHI:

* Communicated by the Author.

† 'Collected Scientific Papers,' vol. ii. p. 713.

‡ I am speaking here of objections to the dynamical and statistical reasoning of the paper. Difficulties in the way of reconciling the results with a kinetic theory of matter are another question.

§ Proc. Roy. Soc. vol. l. p. 85 (1891).

¶ Phil. Mag. Apr. 1892, p. 356.

¶ See also Dr. Watson's 'Kinetic Theory of Gases,' 2nd edit. 1893.

** Phil. Mag. May 1892, p. 466.

"B being a simple vibrator controlled by a spring so stiff that when, at any time, it has very nearly the whole energy of the system, its extreme excursions on each side of its position of equilibrium are small:

"C and A, equal masses:

"C, unacted upon by force except when it strikes L, a fixed barrier, and when it strikes or is struck by B:

"A, unacted on by force except when it strikes or is struck by B, and when it is at less than a certain distance, HK, from a fixed repellent barrier, K, repelling with a force, F, varying according to any law, or constant, when A is between K and H, but becoming infinitely great when (if at any time) A reaches K, and goes infinitesimally beyond it.

"Suppose now A, B, C to be all moving to and fro. The collisions between B and the equal bodies A and C on its two sides must equalize, and keep equal, the average kinetic energy of A, immediately before and after these collisions, to the average kinetic energy of C. Hence, when the times of A being in the space between H and K are included in the average, the average of the *sum of the potential and kinetic energies of A* is equal to the average kinetic energy of C. But the potential energy of A at every point in the space HK is positive, because, according to our supposition, the velocity of A is diminished during every time of its motion from H towards K, and increased to the same value again during motion from K to H. Hence, the average kinetic energy of A is less than the average kinetic energy of C!"

The apparent disproof of the law of partition of energy in this simple problem seems to have shaken the faith even of such experts as Dr. Watson and Mr. Burbury*. M. Poincaré, however, considering a special case of Lord Kelvin's problem †, arrives at a conclusion in harmony with Maxwell's law. Prof. Bryan ‡ considers that the test-case "shows the impossibility of drawing general conclusions as to the distribution of energy in a *single* system from the possible law of permanent distribution in a large number of systems." It is indeed true that Maxwell's theorem relates in the first instance to a large number of systems; but, as I shall show more fully later, the extension to the time-average for a single system requires *only* the application of Maxwell's assumption that all *phases*, i. e. all states, defined both in respect to configuration and *velocity*, which are consistent with the energy condition

* Nature, vol. xlvi. p. 100 (1892).

† *Revue générale des Sciences*, July 1894.

‡ "Report on Thermodynamics," Part II. § 26. Brit. Ass. Rep. 1894.

lie on the same path, *i. e.* are attained by the system in its free motion sooner or later. This fundamental assumption, though certainly untrue in special cases, would appear to apply in Lord Kelvin's problem; and, if so, Maxwell's argument requires the equality of kinetic energies for A and C in the time-averages of a *single* system.

In view of this contradiction we may infer that there must be a weak place in one or other argument; and I think I can show that Lord Kelvin's conclusion above that the average of the sum of the potential and kinetic energies of A is equal to the average kinetic energy of C, is not generally true. In order to see this let us suppose the repulsive force F to be limited to a very thin stratum at H, so that A after penetrating this stratum is subject to no further force until it reaches the barrier K; and let us compare two cases, the whole energy being the same in both.

In case (i.) F is so powerful that with whatever velocity (within the possible limits) A can approach, it is reflected at H, which then behaves like a fixed barrier. In case (ii.) F is still powerful enough to produce this result, except when A approaches it with a kinetic energy nearly equal to the whole energy of the system. A then penetrates beyond H, moving slowly from H to K and back again from K to H, thus remaining for a relatively long time beyond H. Lord Kelvin's statement requires that the average total energy of A should be the same in the two cases; but this it cannot be. For during the occasional penetrations beyond H in case (ii.) A has nearly the whole energy of the system; and its enjoyment of this is *prolonged* by the penetration. Hence in case (ii.) A has a higher average total energy than in case (i.); and a margin is provided which may allow the average *kinetic* energies to be equal. I believe that the consideration here advanced goes to the root of the matter, and shows why it is that the possession of potential energy may involve no deduction from the full share of kinetic energy.

Lord Kelvin's "decisive test-case" is entirely covered by Maxwell's reasoning—a reasoning in my view substantially correct. It would be possible, therefore, to take this case as a typical example in illustration of the general argument; but I prefer for this purpose, as somewhat simpler, another test-case, also proposed by Lord Kelvin. This is simply that of a particle moving in two dimensions; and it may be symbolized by the motion of the ball upon a billiard-table. If there is to be potential energy, the table may be supposed to be out of level. The reconsideration of this problem may perhaps be thought superfluous, seeing that it has been ably treated

already by Prof. Boltzmann*. But his method, though (I believe) quite satisfactory, is somewhat special. My object is rather to follow closely the steps of the general theory. If objections are taken to the argument of the particular case, they should be easy to specify. If, on the other hand, the argument of the particular case is admitted, the issue is much narrowed. I shall have occasion myself to make some comments relating to one point in the general theory not raised by the particular case.

In the general theory the coordinates † of the system at time t are denoted by q_1, q_2, \dots, q_n , and the momenta by p_1, p_2, \dots, p_n . At an earlier time t' the coordinates and momenta of the same motion are represented by corresponding letters accented, and the first step is the establishment of the theorem usually, if somewhat enigmatically, expressed

$$dq'_1 dq'_2 \dots dq'_n dp'_1 dp'_2 \dots dp'_n = dq_1 dq_2 \dots dq_n dp_1 dp_2 \dots dp_n. \quad (1)$$

In the present case q_1, q_2 are the ordinary Cartesian coordinates (x, y) of the particle; and if we identify the mass with unity, p_1, p_2 are simply the corresponding velocity-components (u, v) ; so that (1) becomes

$$dx' dy' du' dv' = dx dy du dv. \quad (2)$$

For the sake of completeness I will now establish (2) *de novo*.

In a possible motion the particle passes from the phase (x', y', u', v') at time t' to the phase (x, y, u, v) at time t . In the following discussion t' and t are absolutely fixed times, but the other quantities are regarded as susceptible of variation. These variations are of course not independent. The whole motion is determined if either the four accented, or the four unaccented, symbols be given. Either set may therefore be regarded as definite functions of the other set. Or again, the four coordinates x', y', x, y may be regarded as independent variables, of which u', v', u, v are then functions.

The relations which we require are readily obtained by means of Hamilton's principal function S , where

$$S = \int_{t'}^t (T - V) dt. \quad (3)$$

In this V denotes the potential energy in any position, and T

* Phil. Mag. vol. xxxv. p. 156 (1893).

† Generalized coordinates appear to have been first applied to these problems by Boltzmann.

is the kinetic energy, so that

$$T = \frac{1}{2} u^2 + \frac{1}{2} v^2 = \frac{1}{2} \dot{x}^2 + \frac{1}{2} \dot{y}^2. \quad \dots \quad (4)^*$$

S may here be regarded as a function of the initial and final coordinates; and we proceed to form the expression for δS in terms of $\delta x'$, $\delta y'$, δx , δy . By (3)

$$\delta S = \int_{t'}^t (\delta T - \delta V) dt, \quad \dots \quad (5)$$

and

$$\begin{aligned} \int \delta T dt &= \int (\dot{x} \delta \dot{x} + \dot{y} \delta \dot{y}) dt \\ &= \int \left(\dot{x} \frac{d \delta x}{dt} + \dot{y} \frac{d \delta y}{dt} \right) dt \\ &= \left[\dot{x} \delta x + \dot{y} \delta y \right]_{t'}^t - \int (\dot{x} \delta x + \dot{y} \delta y) dt; \end{aligned}$$

so that

$$\delta S = \left[\dot{x} \delta x + \dot{y} \delta y \right]_{t'}^t - \int_{t'}^t (\dot{x} \delta x + \dot{y} \delta y + \delta V) dt.$$

By the general equation of dynamics the term under the integral sign vanishes throughout, and thus finally

$$\delta S = u \delta x + v \delta y - u' \delta x' - v' \delta y'. \quad \dots \quad (6)$$

In the general theory the corresponding equation is

$$\delta S = \sum p \delta q - \sum p' \delta q'. \quad \dots \quad (7)$$

Equation (6) is equivalent to

$$\left. \begin{aligned} u' &= -dS/dx', & u &= dS/dx, \\ v' &= -dS/dy', & v &= dS/dy. \end{aligned} \right\} \quad \dots \quad (8)$$

It is important to appreciate clearly the meaning of these equations. S is in general a function of x , y , x' , y' ; and (*e. g.*) the second equation signifies that u is equal to the rate at which S varies with x , when y , x' , y' are kept constant, and so in the other cases.

We have now to consider, not merely a single particle, but an immense number of similar particles, moving independently of one another under the same law (V), and distributed at time t over all possible phases (x , y , u , v). The

* As is not unusual in the integral calculus, we employ the same symbols x , &c. to denote the current and the final values of the variables. If desired, the final values may be temporarily distinguished as x'' , &c.

most general expression for the law of distribution is

$$f(x, y, u, v) dx dy du dv, \dots \dots \dots (9)$$

signifying that the number of particles to be found at time t within a prescribed range of phase is to be obtained by integrating (9) over the range in question. But such a distribution would in general be *unsteady*. If it obtained at time t , it would be departed from at time t' , and *vice versa*, owing to the natural motions of the particles. The question before us is to ascertain what distributions are steady, *i. e.* are maintained unaltered notwithstanding the motions.

It will be seen that it is the spontaneous passage of a particle from one phase to another that limits the generality of the function f . If there be no possibility of passage, say, from the phase (x', y', u', v') to the phase (x, y, u, v) , or, as it may be expressed, if these phases do not lie upon the same *path*, then there is no relation imposed upon the corresponding values of f . An example, given by Prof. Bryan (*l. c.* § 17), well illustrates this point. Suppose that $V=0$, so that every particle pursues a straight course with uniform velocity. The phases (x', y', u', v') and (x, y, u, v) can lie upon the same path only if $u'=u, v'=v$. Accordingly f remains arbitrary so far as regards u and v . For instance, a distribution

$$f(u, v) dx dy du dv \dots \dots \dots (10)$$

is permanent whatever may be the form of f , understood to be independent of x and y . In this case the distribution is *uniform* in space, but uniformity is not indispensable. Suppose, for example, that *all* the particles move parallel to x , so that f vanishes unless $v=0$. The general form (9) now reduces to

$$f(x, y, u) dx dy du; \dots \dots \dots (11)$$

and permanency requires that the distribution be uniform along any line for which y is constant. Accordingly, f must be independent of x , so that permanent distributions are of the form

$$f(y, u) dx dy du, \dots \dots \dots (12)$$

in which f is an arbitrary function of y and u . If either y or u be varied, we are dealing with a different path (in the sense here involved), and there is no connexion between the corresponding values of f . But if while y and u remain constant, x be varied, the value of f must remain unchanged, for the different values of x relate to the same path.

Before taking up the general question in two dimensions, it may be well to consider the relatively simple case of motion in one dimension, which, however, is not so simple but that

it will introduce us to some of the points of difficulty. The particles are supposed to move independently upon one straight line, and the phase of any one of them is determined by the coordinate x and the velocity u . At time t' the phase of a particle will be denoted by (x', u') , and at time t the phase of the same particle will be (x, u) , where u will in general differ from u' , since we no longer suppose that V is constant, but rather that it is variable in a known manner, *i. e.* is a known function of x . The number of particles which at time t lie within the limits of phase represented by $dx du$ is $f(x, u) dx du$, and the question is whether this distribution is steady, and in particular whether it was the same at time t' . In order to find the distribution at time t' , we regard x, u as known functions of x', u' , and transform the multiple differential. The result of this transformation is best seen by comparison with intermediate transformations in which $dx du$ and $dx' du'$ are compared with $dx dx'$. We have

$$dx du = dx dx' \times \frac{du}{dx'}, \dots \dots \dots (13)$$

$$dx' du' = dx dx' \times \frac{du'}{dx} \dots \dots \dots (14)$$

In du/dx' of (13) x is to be kept constant, and in du'/dx of (14) x' is to be kept constant. If we disregard algebraic sign, both are by (8) equal to $d^2S/dx dx'$, and are therefore equal to one another. Hence we may write

$$dx du = dx' du' ; \dots \dots \dots (15)$$

and the transformation is expressed by

$$f(x, u) dx du = f_1(x', u') dx' du', \dots \dots (16)$$

where $f_1(x', u')$ is the result of substituting for x, u in $f(x, u)$ their values in terms of x', u' . The right-hand member of (16) expresses the distribution at time t' corresponding to the distribution at time t expressed by the left-hand member, as determined by the laws of motion between the two phases. If the distribution is to be steady, $f_1(x', u')$ must be identical with $f(x, u)$; in other words $f(x, u)$ must be such a function of (x, u) that it remains unchanged when (x, u) refers to various phases of the motion of the same particle. Now, if E denote the total energy, so that

$$E = \frac{1}{2}u^2 + V, \dots \dots \dots (17)$$

then E remains constant during the motion; and thus, if for the moment we suppose f expressed in terms of E and x , we

see that x cannot enter, or that f is a function of E only. The only permanent distributions accordingly are those included under the form

$$f(E) dx du, \dots \dots \dots (18)$$

where E is given by (17), and f is an arbitrary function.

It is especially to be noticed that the limitation to the form (18) holds only for phases lying upon the same path. If two phases have different energies, they do not lie upon the same path, but in this case the independence of the distributions in the two phases is already guaranteed by the form of (18). The question is whether all phases of given energy lie upon the same path. It is easy to invent cases for which the answer will be in the negative. Suppose, for example, that there are two centres of force O, O' on the line of motion which attract with a force at first proportional to distance but vanishing when the distance exceeds a certain value less than the interval OO' . A particle may then vibrate with the same (small) energy either round O or round O' ; but the phases of the two motions do not lie upon the same path. Consequently f is not limited by the condition of steadiness to be the same in the two groups of phases. In all cases steadiness is ensured by the form (18); and if all phases of equal energy lie upon the same path, this form is necessary as well as sufficient.

All the essential difficulties of the theory appear to be raised by the particular case just discussed, and the reader to whom the subject is new is recommended to give it his careful attention.

In the more general problem of motion in two dimensions the discussion follows a parallel course. In order to find the distribution at time t' corresponding to (9) at time t , we have to transform the multiple differential, regarding x, y, u, v as known functions of x', y', u', v' . Here again we take the initial and final coordinates x, y, x', y' as an intermediate set of variables. Thus

$$dx' dy' du' dv' = dx' dy' dx dy \times \begin{vmatrix} \frac{du'}{dx'} & \frac{dv'}{dx'} \\ \frac{du'}{dy'} & \frac{dv'}{dy'} \end{vmatrix}, \dots \dots (19)$$

$$dx dy du dv = dx dy dx' dy' \times \begin{vmatrix} \frac{du}{dx'} & \frac{dv}{dx'} \\ \frac{du}{dy'} & \frac{dv}{dy'} \end{vmatrix} \dots \dots (20)$$

In the determinants of (19), (20) the motion is regarded as a function of x, y, x', y' , and the three quantities which do not appear in the denominator of any differential coefficient are to be considered constant. This was also the understanding in equations (8), from which we infer that the two determinants are equal, being each equivalent to

$$\begin{vmatrix} \frac{d^2S}{dx dx'} & \frac{d^2S}{dx dy'} \\ \frac{d^2S}{dx' dy} & \frac{d^2S}{dy dy'} \end{vmatrix} \dots \dots \dots (21)$$

Hence we may write

$$dx dy du dv = dx' dy' du' dv', \dots \dots (22)$$

an equation analogous to (15). By the same reasoning as was employed for motion in one dimension it follows that, if the distribution is to be steady, $f(x, y, u, v)$ in (9) must remain constant for all phases which lie upon the same path. A distribution represented by

$$f(\mathbf{E}) dx dy du dv, \dots \dots \dots (23)$$

where

$$\mathbf{E} = \frac{1}{2}u^2 + \frac{1}{2}v^2 + \mathbf{V}, \dots \dots \dots (24)$$

will satisfy the conditions of steadiness whatever be the form of f ; but this form is only *necessary* under the restriction known as Maxwell's assumption or postulate, viz. that all phases of equal energy lie upon the same path.

It is easy to give examples in which Maxwell's assumption is violated, and in which accordingly steady distributions are not limited to (23). Thus, if no force act parallel to y , so that \mathbf{V} reduces to a function of x only, the component velocity v remains constant for each particle, and no phases for which v differs lie upon the same path. A distribution

$$f(\mathbf{E}, v) dx dy du dv \dots \dots \dots (25)$$

is then steady, whatever function f may be of \mathbf{E} and v .

That under the distribution (23) the kinetic energy is equally divided between the component velocities u and v is evident from symmetry. It is to be observed that the law of equal partition applies not merely upon the whole, but for every element of area $dx dy$, and for every value of the total energy, and at every moment of time. When x and y are prescribed as well as \mathbf{E} , the value of the resultant velocity itself is determined by (24).

Another feature worthy of attention is the spacial distribution; and it happens that this is peculiar in the present problem. To investigate it we must integrate (23) with respect to u and v , x and y being constant. Since x and y are constant, V is constant; so that, if we suppose E to lie within narrow limits E and $E + dE$, the resultant velocity U will lie between limits given by

$$U dU = dE. \quad (26)$$

If we transform from u, v to U, θ , where

$$u = U \cos \theta, \quad v = U \sin \theta, \quad . . . (27)$$

$du dv$ becomes $U dU d\theta$; so that on integration with respect to θ we have, with use of (26),

$$2\pi F(E) dE . dx dy. \quad (28)$$

The spacial distribution is therefore *uniform*.

In order to show the special character of the last result, it may be well to refer briefly to the corresponding problem in three dimensions, where the coordinates of a particle are x, y, z and the component velocities are u, v, w . The steady distribution corresponding to (23) is

$$f(E) dx dy dz du dv dw, \quad (29)$$

in which

$$E = \frac{1}{2} U^2 + V = \frac{1}{2} u^2 + \frac{1}{2} v^2 + \frac{1}{2} w^2 + V. \quad . . . (30)$$

Here equation (26) still holds good, and the transformation of $du dv dw$ is, as is well known, $4\pi U^2 dU$. Accordingly (29) becomes

$$4\pi F(E) dE . (2E - 2V)^{\frac{3}{2}} dx dy, \quad . . . (31)$$

no longer uniform in space, since V is a function of x, y .

In (31) the density of distribution decreases as V increases. For the corresponding problem in *one* dimension (18) gives

$$F(E) dE . (2E - 2V)^{-\frac{1}{2}} dx, \quad (32)$$

so that in this case the density *increases* with increasing V .

The uniform distribution of the two-dimensional problem is thus peculiar. Although an immediate consequence of Maxwell's equation (41), see (41) below, I failed to remark it in the note before referred to, where I wrote as if a uniform distribution in the billiard-table example required that $V = 0$. In order to guard against a misunderstanding it may be well to say that the uniform distribution does not necessarily extend over the whole plane. Wherever $(E - V)$ falls below zero there is of course no distribution.

We have thus investigated for a particle in two dimensions the law of steady distribution, and the equal partition of energy which is its necessary consequence. And we see that "the only assumption necessary to the direct proof is that the system, if left to itself in its actual state of motion, will, sooner or later, pass through every phase which is consistent with the equation of energy" (Maxwell). It will be observed that so far nothing whatever has been said as to time-averages for a single particle. The law of equal partition, as hitherto stated, relates to a large number of particles and to a single moment of time.

The extension to time-averages, the aspect under which Lord Kelvin has always considered the problem, is important, the more as some authors appear to doubt the possibility of such extension. Thus Prof. Bryan (Report, § 11, 1894), speaking of Maxwell's assumption, writes:—"To discover, if possible, a general class of dynamical systems satisfying the assumption would form an interesting subject for future investigation. It is, however, doubtful how far Maxwell's law would be applicable to the *time-averages* of the energies in any such system. We shall see, in what follows, that the law of permanent distribution of a very large number of systems is in many cases not unique. Where there is more than one possible distribution it would be difficult to draw any inference with regard to the average distribution (taken with respect to the time) for one system."

The extension to time-averages appears to me to require nothing more than Maxwell's assumption, without which the law of distribution itself is only an artificial arrangement, sufficient indeed but not necessary for steadiness. We shall still speak of the particle moving in two dimensions, though the argument is general. It has been shown that at any moment the u -energy and the v -energy of the group of particles is the same; and it is evident that the equality subsists if we integrate over any period of time. But if this period be sufficiently prolonged, and if Maxwell's *assumption be applicable*, it makes no difference whether we contemplate the whole group of particles or limit ourselves to a single member of it. It follows that for a single particle the time-averages of u^2 and v^2 are equal, provided the averages be taken over a sufficient length of time.

On the other hand, if in any case Maxwell's assumption be untrue, not only is the special distribution unnecessary for steadiness, but even if it be artificially arranged, the law of equal time-averages does not follow as a consequence.

Having now considered the special problem at full—I hope

it may not be thought at undue—length, I pass on to some remarks on the general investigation. This proceeds upon precisely parallel lines, and the additional difficulties are merely those entailed by the use of generalized coordinates. Thus (1) follows from (7) by substantially the same process (given in my former note) that (22) follows from (6). Again, if E denote the total energy of a system, the distribution

$$f(E)dq_1 \dots dq_n dp_1 \dots dp_n, \dots (33)$$

where f is an arbitrary function, satisfies the condition of permanency; and, if Maxwell's assumption be applicable, it is the *only* form of distribution that can be permanent.

As I hinted before, some of the difficulties that have been felt upon this subject may be met by a fuller recognition of the invariant character of the expressions. This point has been ably developed by Prof. Bryan, who has given (*loc. cit.* § 14) a formal verification that (33) is unaltered by a change of coordinates. If we follow attentively the process by which (1) is established, we see that in (3) there is no assumption that the system of coordinates is the same at times t' and t , and that accordingly we are not tied to one system in (33). Indeed, so far as I can see, there would be no meaning in the assertion that the system of generalized coordinates employed for two different configurations was the same*.

We come now to the deduction from (33) of Maxwell's law of partition of energy. On this Prof. Bryan (*loc. cit.* § 20, remarks:—"Objections have been raised to this step in Maxwell's work by myself ('Report on Thermodynamics,' Part I. § 44) on the ground that the kinetic energy cannot in general be expressed as the sum of squares of *generalized momenta* corresponding to generalized coordinates of the system, and by Lord Kelvin (*Nature*, Aug. 13, 1891) on the ground that the conclusion to which it leads has no intelligible meaning. Boltzmann (*Phil. Mag.* March 1893) has put the investigation into a slightly modified form which meets the first objection, and which imposes a certain restriction upon the generality of the result. Under this limitation the result is perfectly intelligible, and the second objection is therefore also met." At this point I find myself in disagreement with all the above quoted authorities, and in the position of maintaining the correctness of Maxwell's original deduction.

Prof. Boltzmann considers that "Maxwell committed an

* It would be like saying that two points lie upon the same curve, when the character of the curve is not defined.

error in assuming that by choosing suitable coordinates the expression for the *vis viva* could always be made to contain only the squares of the momenta." This is precisely the objection which I supposed myself to have already answered in 1892. I wrote, "It seems to be overlooked that Maxwell is limiting his attention to systems *in a given configuration*, and that no dynamics is founded upon the reduced expression for T . The reduction can be effected in an infinite number of ways. We may imagine the configuration in question rendered one of stable equilibrium by the introduction of suitable forces proportional to displacements. The principal modes of isochronous vibration thus resulting will serve the required purpose."

It is possible, therefore, so to choose the coordinates that for a given configuration (and for configurations differing infinitely little therefrom) the kinetic energy T , which is always a quadratic function of the velocities, shall reduce to a sum of squares with, if we please, given coefficients. Thus *in the given configuration*

$$T = \frac{1}{2} \dot{q}_1^2 + \frac{1}{2} \dot{q}_2^2 + \dots + \frac{1}{2} \dot{q}_n^2; \dots \dots \dots (34)$$

and, since in general $p = dT/d\dot{q}$,

$$p_1 = \dot{q}_1, \quad p_2 = \dot{q}_2, \quad \&c.,$$

so that

$$T = \frac{1}{2} p_1^2 + \frac{1}{2} p_2^2 + \dots + \frac{1}{2} p_n^2. \dots \dots \dots (35)$$

Whether the coordinates required to effect a similar reduction for other configurations are the same is a question with which we are not concerned.

The mean value of p_r^2 for all the systems in the given configuration is, according to (33),

$$\frac{\int p_r^2 \cdot F\{V + \frac{1}{2} p_1^2 + \dots + \frac{1}{2} p_n^2\} dp_1 \dots dp_n}{\int F\{V + \frac{1}{2} p_1^2 + \dots + \frac{1}{2} p_n^2\} dp_1 \dots dp_n} \dots (36)*$$

The limits for each variable may be supposed to be $+\infty$; but the large values do not really enter if we suppose $F(\bar{E})$ to be finite for moderate, perhaps for nearly definite, values of E only.

It is now evident that the mean value is the same for all the momenta p ; and accordingly that for each the mean value of $\frac{1}{2} p^2$ is $1/n$ of the mean value of T . This result holds good for every moment of time, for every configuration, for every value of E , and for every system of resolution (of which there

* Confer Bryan, *loc. cit.*

are an infinite number) which allows T to be expressed in the form (35).

In the case where the "system" consists of a single particle, (35) is justified by any system of rectangular coordinates; and although we are not bound to use the same system for different positions of the particle, it would conduce to simplicity to do so. If the system be a rigid body, we may measure the velocities of the centre of inertia parallel to three fixed rectangular axes, while the remaining momenta refer to rotations about the principal axes of the body. If Maxwell's assumption hold good, a permanent distribution is such that in one, or in any number of positions, the mean energy of each rotation and of each translation is the same. And under the same restriction a similar assertion may be made respecting the time-averages for a single rigid body.

There is much difficulty in judging of the applicability of Maxwell's assumption. As Maxwell himself showed, it is easy to find cases of exception; but in most of these the conditions strike one as rather special. It must be observed, however, that if we take it quite literally, the assumption is of a severely restrictive character; for it asserts that the system, starting from any phase, will traverse *every* other phase (consistent with the energy condition) *before* returning to the initial phase. As soon as the initial phase is recovered, a cycle is established, and no new phases can be reached, however long the motion may continue.

We return now to the question of the distribution of momenta among the systems which occupy a given configuration, still supposing the coordinates so chosen as to reduce T to a sum of squares (35). It will be convenient to fix our attention upon systems for which E lies within narrow limits, E and $E + dE$. Since E is given, there is a relation between $p_1, p_2, \dots p_n$, and we may suppose p_n expressed in terms of E and the remaining momenta. By (35)

$$p_n dp_n = dT = dE,$$

since the configuration is given, and thus (33) becomes

$$f(E)dE \cdot dq_1 \dots dq_n \cdot p_n^{-1} dp_1 \dots dp_{n-1} \dots \quad (37)$$

For the present purpose the latter factors alone concern us, so that what we have to consider is

$$\frac{dp_1 dp_2 \dots dp_{n-1}}{\sqrt{\{2T - p_1^2 - p_2^2 - \dots - p_{n-1}^2\}}}, \quad \dots \quad (38)$$

in which T , being equal to $E - V$, is given. For the moment we may suppose that $2T$ is unity.

The whole number of systems is to be found by integrating (38), the integral being so taken as to give the variables all values consistent with the condition that $p_1^2 + p_2^2 + \dots + p_{n-1}^2$ is not greater than unity. Now

$$\int \dots \int \frac{dp_1 dp_2 \dots dp_{n-1}}{\sqrt{\{1 - p_1^2 - \dots - p_{n-1}^2\}}} = \frac{\pi^{\frac{1}{2}n - \frac{1}{2}}}{\Gamma(\frac{1}{2}n - \frac{1}{2})} \int_{-1}^{+1} (1 - p_1^2)^{\frac{1}{2}n - \frac{3}{2}} dp_1,$$

and . . . (39)

$$\int_{-1}^{+1} (1 - p_1^2)^{\frac{1}{2}n - \frac{3}{2}} dp_1 = \frac{\Gamma(\frac{1}{2}) \Gamma(\frac{1}{2}n - \frac{1}{2})}{\Gamma(\frac{1}{2}n)}. \quad \dots (40)$$

in which $\Gamma(\frac{1}{2}) = \sqrt{\pi}$. Thus the whole number of systems is

$$\frac{\{\Gamma(\frac{1}{2})\}^n}{\Gamma(\frac{1}{2}n)},$$

or on restoration of $2T$, equal to $2E - 2V$,

$$\frac{\{\Gamma(\frac{1}{2})\}^n}{\Gamma(\frac{1}{2}n)} \{2E - 2V\}^{\frac{1}{2}n - 1}. \quad \dots (41)$$

To this we shall return later; but for the present what we require to ascertain is the distribution of one of the momenta, say p_1 , irrespectively of the values of the remaining momenta. By (39), (40) the number of systems for which p_1 lies between p_1 and $p_1 + dp_1$ in comparison with the whole number of systems is

$$\frac{\Gamma(\frac{1}{2}n)}{\Gamma(\frac{1}{2}) \Gamma(\frac{1}{2}n - \frac{1}{2})} \left\{ 1 - \frac{p_1^2}{2T} \right\}^{\frac{1}{2}n - \frac{3}{2}} \frac{dp_1}{\sqrt{2T}}. \quad \dots (42)$$

This is substantially Maxwell's investigation, and (42) corresponds with his equation (51). As was to be expected, the law of distribution is the same for all the momenta. From the manner of its formation, we note that the integral of (42), taken between the limits $p_1 = \pm \sqrt{2T}$, is equal to unity.

Maxwell next proceeds to the consideration of the special form assumed by (42), when the number n of degrees of freedom is extremely great*. This part of the work seems to be very important; but it has been much neglected, probably because the result was not correctly stated.

Dropping the suffix as unnecessary, we have to consider the form of

$$\left\{ 1 - \frac{p^2}{2T} \right\}^{\frac{1}{2}n - \frac{3}{2}}$$

when n is very great, the mean value of p^2 becoming at the

* The particular cases where $n=2$, or $n=3$, are also worthy of notice.

same time small in comparison with $2T$. If we write

$$T = nK = \frac{1}{2}nP^2, \dots \dots \dots (43)$$

we have

$$\text{Limit} \left\{ 1 - \frac{p^2}{2T} \right\}^{\frac{1}{2}n - \frac{1}{2}} = e^{-p^2/4K} = e^{-p^2/2P^2} \dots \dots (44)$$

The limit of the fraction containing the Γ functions may be obtained by the formula

$$\Gamma(m + 1) = e^{-m}m^m \sqrt{(2m\pi)};$$

and the limiting form of (42) becomes

$$\frac{e^{-p^2/4K}}{\sqrt{(2\pi)}} \frac{dp}{\sqrt{(2K)}}, \text{ or } \frac{e^{-p^2/2P^2}}{\sqrt{(2\pi)}} \frac{dp}{P} \dots \dots (45)$$

It may be observed that the integral of (45) between the limits $\pm\infty$ is unity, and that this fact might have been used to determine the numerical factor.

Maxwell's result is given in terms of a quantity k , analogous to K , and defined by

$$\frac{1}{2}p^2 = k \dots \dots \dots (46)$$

It is

$$\frac{1}{\sqrt{(2\pi)}} \frac{1}{K} e^{-\frac{k}{2K}} dk \dots \dots \dots (47)$$

The corresponding form from (45) is

$$\frac{1}{\sqrt{(2\pi)}} \frac{1}{2\sqrt{(kK)}} e^{-\frac{k}{2K}} dk \dots \dots \dots (48)$$

In like manner if we inquire what proportion of the whole number of systems have momenta lying within the limits denoted by $dp_1 dp_2 \dots dp_r$, where r is a number very small relatively to n , we get

$$\frac{e^{-(p_1^2 + p_2^2 + \dots + p_r^2)/4K}}{\{\sqrt{(2\pi)}\}^r} \frac{dp_1 dp_2 \dots dp_r}{\{\sqrt{(2K)}\}^r}, \dots \dots (49)$$

or, if we prefer it,

$$\frac{e^{-(p_1^2 + p_2^2 + \dots + p_r^2)/2P^2}}{\{\sqrt{(2\pi)}\}^r} \frac{dp_1 dp_2 \dots dp_r}{P^r} \dots \dots (50)$$

These results follow from the general expression (38), in the same way as does (45), by stopping the multiple integration at an earlier stage. The remaining variables range over values which may be considered in each case to be unlimited.

If the integration between $\pm \infty$ be carried out completely, we recover the value unity.

The interest of the case where n is very great lies of course in the application to a gas supposed to consist of an immense number of similar molecules*, or of several sets of similar molecules; and the question arises whether (45) can be applied to deduce the Maxwellian law of distribution of velocities among the molecules of a single system at a given instant of time. A caution may usefully be interposed here as to the sense in which the Maxwellian distribution is to be understood. It would be absurd to attempt to prove that the distribution in a single system is necessarily such and such, for we have already assumed that every phase, including every distribution of velocities, is attainable, and indeed attained if sufficient time be allowed. The most that can be proved is that the distribution will *approximate* to a particular law for the greater part of the time, and that if sensible deviations occur they will be *transitory*.

In applying (45) to a gas it will be convenient to suppose in the first instance that all the molecules are similar. Each molecule has several degrees of freedom, but we may fix our attention upon one of them, say the x -velocity of the centre of inertia, usually denoted by v . In (45) the *whole* system is supposed to occupy a given configuration; and the expression gives us the distribution of velocity at a given time for a single molecule among all the systems. The distribution of velocity is the same for every other molecule, and thus the expression applies to the statistics of all the molecules of all the systems. Does it also apply to the statistics of all the molecules of a single system? In order to make this inference we must assume that the statistics are the same (at the same time) for all the systems, or, what comes to the same thing (if Maxwell's assumption be allowed), that they are the same for the same system at the various times when it passes through a given configuration.

Thus far the argument relates only to a single configuration. If the configuration be changed, there will be in general a change of potential energy and a corresponding change in the kinetic energy to be distributed amongst the degrees of freedom. But in the case of a gas, of which the statistics are assumed to be regular, the potential energy remains approximately constant when exclusion is made of exceptional conditions. The same law of distribution of velocity then applies to every configuration, that is, it may be asserted

* The terms "gas" and "molecule" are introduced for the sake of brevity. The question is still purely dynamical.

without reference to the question of configuration. We thus arrive at the Maxwellian law of velocities in a single gas, as well as the relation between the velocities in a mixture of molecules of different kinds first laid down by Waterston.

The assumptions which we have made as to the practical regularity of statistics are those upon which the usual theory of ideal gases is founded; but the results are far more general. Nothing whatever has been said as to the character of the forces with which the molecules act upon one another, or are acted upon by external agencies. Although for distinctness a gas has been spoken of, the results apply equally to a medium constituted as a liquid or a solid is supposed to be. A kinetic theory of matter, as usually understood, appears to require that in equilibrium the whole kinetic energy shall be equally shared among all the degrees of freedom, and within each degree of freedom be distributed according to the same law. It is included in this statement that temperature is a matter of kinetic energy only, *e. g.* that when a vertical column of gas is in equilibrium, the mean velocity of a molecule is the same at the top as at the bottom of the column.

Reverting to (37), (41), in order to consider the distribution of the systems as dependent upon the coordinates independently of the velocities, we have, omitting unnecessary factors,

$$\{E - V\}^{\frac{1}{2}n-1} dq_1 dq_2 \dots dq_n \dots \dots (51)$$

If $n=2$, *e. g.* in the case already considered of a single particle moving in two dimensions, or of two particles moving in one dimension, or again whatever n may be, provided V vanish, the first factor disappears, so that the distribution is *uniform* with respect to the coordinates $q_1 \dots q_n$. If $n > 2$ and V be finite, the distribution is such as to favour those configurations for which V is least.

“When the number of variables is very great, and when the potential energy of the specified configuration is very small compared with the total energy of the system, we may obtain a useful approximation to the value of $\{E - V\}^{\frac{1}{2}n-1}$ in an exponential form; for if we write (as before) $E = nK$,

$$\{E - V\}^{\frac{1}{2}n-1} = E^{\frac{1}{2}n-1} e^{-V/2K} \dots \dots (52)$$

nearly, provided n is very great and V is small compared with E . The expression is no longer approximate when V is nearly as great as E , and it does not vanish, as it ought to do, when $V = E$.” (Maxwell).

In the case of gas composed of molecules whose mutual influence is limited to a small distance and which are not subject to external forces, the distribution expressed by (51)

is uniform in space except near the boundary. For if q_1 denote the x -coordinate of a particular molecule, and if we effect the integration with respect to all the coordinates of other molecules as well as the other coordinates of the particular molecule, we must arrive at a result independent of x , provided x relate to a point well in the interior. That is to say in the various systems contemplated the particular molecule is uniformly distributed with respect to x . The same is true of y and z , and thus the whole spacial distribution is uniform. If the single system constituting the gas has uniform statistics, it will follow that the distribution in it of molecules similar to the particular molecule is uniform.

The uniformity of the distribution is disturbed if an external force acts. In illustration of this we may consider the case of gravity. From (52) the distribution with respect to the coordinates of the particular molecule will be

$$e^{-gz/2K} dx dy dz,$$

and the same formula gives the density of molecules similar to the particular molecule in a single system.

The main purpose of this paper is now accomplished ; but I will take the opportunity to make a few remarks upon some general aspects of a kinetic theory of matter. Many writers appear to commit themselves to absolute statements, but Kelvin* and Boltzmann and Maxwell fully recognize that conclusions can never be more than *probable*. The second law of thermodynamics itself is in this predicament. Indeed it might seem at first sight as if the case were even worse than this. Mr. Culverwell has emphasized a difficulty, which must have been pretty generally felt, arising out of the reversibility of a dynamical system. If during one motion of a system energy is dissipated, restoration must occur when the motion is reversed. How then is one process more probable than the other? Prof. Boltzmann has replied to this objection, upon the whole I think satisfactorily, in a very interesting letter †. The available (internal) energy of a system tends to zero, or

* Witness the following remarkable passage:—"It is a strange but nevertheless a true conception of the old well-known law of the conduction of heat to say that it is very improbable that in the course of 1000 years one-half the bar of iron shall of itself become warmer by a degree than the other half; and that the probability of this happening before 1,000,000 years pass is 1000 times as great as that it will happen in the course of 1000 years, and that it certainly will happen in the course of some very long time."—('Nature,' vol. ix. p. 443, 1874.)

† 'Nature,' vol. li. p. 413 (1895).

rather to a small value, only because the conditions, or phases as we have called them, corresponding to small values are more probable, *i. e.* more numerous. If there is considerable available energy at any moment, it is because the condition is then exceptional and peculiar. After a short interval of time the condition *may* become more peculiar still, and the available energy *may* increase, but this is improbable. The probability is that the available energy will, if not at once, at any rate after a short interval, decrease owing to the substitution of a more nearly normal state of things.

There is, however, another side to this question, which perhaps has been too much neglected. Small values of the available energy are indeed more probable than large ones, but there is a degree of smallness below which it is *improbable* that the value will lie. If at any time the value lies extremely low, it is an increase and not a decrease which is probable. Maxwell showed long ago how a being capable of dealing with individual molecules would be in a position to circumvent the second law. It is important to notice that for this end it is not necessary to deal with individual molecules. It would suffice to take advantage of local reversals of the second law, which will involve, not very rarely, a *considerable number* of neighbouring molecules. Similar considerations apply to other departures from a normal state of things, such, for example, as unequal mixing of two kinds of molecules, or such a departure from the Waterston relation (of equal mean kinetic energies) as has been investigated by Maxwell and by Tait and Burbury.

The difficulties connected with the application of the law of equal partition of energy to actual gases have long been felt. In the case of argon and helium and mercury vapour the ratio of specific heats (1.67) limits the degrees of freedom of each molecule to the three required for translatory motion. The value (1.4) applicable to the principal diatomic gases gives room for the three kinds of translation and for two kinds of rotation. Nothing is left for rotation round the line joining the atoms, nor for relative motion of the atoms in this line. Even if we regard the atoms as mere points, whose rotation means nothing, there must still exist energy of the last-mentioned kind, and its amount (according to the law) should not be inferior.

We are here brought face to face with a fundamental difficulty, relating not to the theory of gases merely, but rather to general dynamics. In most questions of dynamics a condition whose violation involves a large amount of potential energy may be treated as a *constraint*. It is on this

principle that solids are regarded as rigid, strings as inextensible, and so on. And it is upon the recognition of such constraints that Lagrange's method is founded. But the law of equal partition disregards potential energy. However great may be the energy required to alter the distance of the two atoms in a diatomic molecule, practical rigidity is never secured, and the kinetic energy of the relative motion in the line of junction is the same as if the tie were of the feeblest. The two atoms, however related, remain two atoms, and the degrees of freedom remain six in number.

What would appear to be wanted is some escape from the destructive simplicity of the general conclusion relating to partition of kinetic energy, whereby the energy of motions involving larger amounts of potential energy should be allowed to be diminished in consequence. If the argument, as above set forth after Maxwell, be valid, such escape must involve a repudiation of Maxwell's fundamental postulate as practically applicable to systems with an immense number of degrees of freedom.

VII. *On Swan's Prism Photometer, commonly called Lummer and Brodhun's Photometer.* By Prof. C. G. KNOTT, D.Sc.*.

IN 1849 William Swan, subsequently Professor of Natural Philosophy in the University of St. Andrews, read a paper "On the Gradual production of Luminous Impressions on the Eye and other Phenomena of Vision" before the Royal Society of Edinburgh (see Transactions, vol. xvi.). This paper contains some results of high interest, but I have no recollection of ever having seen it referred to in modern literature on the subject.

On April 4, 1859, Professor Swan gave a second paper on the same subject, much briefer than the first, and entirely occupied with descriptions of greatly improved forms of apparatus (see Transactions, vol. xxii.). Among the forms of apparatus described is his "Prism Photometer." This is simply and solely the form of photometer described in 1889, exactly thirty years later, by Lummer and Brodhun, and named after them in all recent literature (see *Zeitschrift für Instrumentenkunde*, Bd. ix.). I cannot do better than give Swan's description in full, and reproduce his own diagram.

He writes:—"An arrangement which, from an imperfect

* Communicated by the Author: read before the Royal Society of Edinburgh, Dec. 17th, 1899.