Maxwell and the normal distribution: A colored story of probability, independence, and tendency toward equilibrium

Balázs Gyenis

Institute of Philosophy, HAS RCH, 30 Országház str, Budapest, Hungary

Abstract

We investigate Maxwell's attempt to justify the mathematical assumptions behind his 1860 Proposition IV according to which the velocity components of colliding particles follow the normal distribution. Contrary to the commonly held view we find that his molecular collision model plays a crucial role in reaching this conclusion, and that his model assumptions also permit inference to equalization of mean kinetic energies (temperatures), which is what he intended to prove in his discredited and widely ignored Proposition VI. If we take a charitable reading of his own proof of Proposition VI then it was Maxwell, and not Boltzmann, who gave the first proof of a tendency towards equilibrium, a sort of H-theorem. We also call attention to a potential conflation of notions of probabilistic and value independence in relevant prior works of his contemporaries and of his own, and argue that this conflation might have impacted his adoption of the suspect independence assumption of Proposition IV.

Keywords: Maxwell, kinetic theory, statistical mechanics, normal distribution, independence, H-theorem, second law of thermodynamics, condition A

1. Introduction

James Clerk Maxwell's early work on the kinetic theory of gases was a major step-stone in the introduction of probabilistic methods into physics. Proposition IV of Maxwell's (1860) *Illustrations of the Dynamical Theory of Gases*, his first derivation of the velocity distribution law, is frequently cited as "one of the most important passages in physics" (Truesdell, 1975, p. 34) and as such had been thoroughly investigated in the works of S. G. Brush (1958; 1971; 1976; 1983), M. C. Dias (1994), C. W. F. Everitt (1986a), E. Garber (1970), C. C. Gillispie (1963), P. M. Heimann (1970), T. M. Porter (1981), C. Truesdell (1975), J.

Uffink (2007) and other historians of science. Proposition IV shows that, given certain assumptions, the three velocity components of molecules of a box of gas follow the normal distribution and the speed (the magnitude of velocity) follows what nowadays is called the Maxwell-Boltzmann distribution.

This paper provides a new conceptual and historical analysis of Maxwell's derivation. We make four contributions to the literature. First, the paper sheds new light on the logical structure of Proposition IV and gives the first detailed analysis of how Maxwell attempted to physically justify the three main mathematical assumptions on which it rests. This allows us to show, second, that contrary to the common historical wisdom molecular collisions did play an essential role in establishing the conclusion of Proposition IV: one of its three mathematical assumptions requires a prior lemma showing that collisions among particles of the same mass bring about an equilibrium velocity distribution, and we argue that Maxwell indeed made an attempt to lend credence to such an approach to equilibrium on the basis of his Propositions I–III.

To substantiate this claim we take a look at Maxwell's remarks preceding Proposition IV as well as at his attempt to arrive at a proof of tendency towards equilibrium in his discredited and widely ignored Proposition VI. We present a surprisingly simple proof of tendency towards equilibrium that rests on Propositions I–III, both in the relevant special case when all particles have the same mass and in the general case when masses differ. Although this proof of equalization of temperatures is interesting in its own right and is not known by experts, we argue that it is not novel since its general case is nothing but a charitable reconstruction of Maxwell's Proposition VI. Since the charitable reading of Proposition VI provides a proof of tendency towards equilibrium in the general case, and since the special case of this proof is really simple, it is reasonable to assume that Maxwell's brief remarks preceding Proposition IV are at least partially motivated by this special case. Additionally, if the charitable reading is tenable then Maxwell preceded Ludwig Boltzmann by at least 6 years in providing a mechanical argument for tendency towards equilibrium, and this priority of Maxwell should be more widely recognized.

Fourth, the paper contributes to the scholarly discussion of the crucial and arguably unjustified probabilistic independence assumption of Proposition IV. The conflation of different interpretations of probability have already been pointed out in the literature as a potential source of Maxwell's mistake in accepting this independence assumption; besides furthering this analysis we also provide a novel interpretation according to which Maxwell, instead of conflating different interpretations of probability, might have conflated different notions of independence. We distinguish two notions – probabilistic independence and value (or parameter) independence –, argue that their difference was not clear in the relevant works of Clausius and Herschel, and point out that in fact Herschel mistakenly emphasized that it is the satisfaction of value independence that is crucial for the proof that Maxwell allegedly have adapted as his Proposition IV. Thus if Maxwell indeed adapted the proof from Herschel's review article then he also likely to have accepted Herschel's assessment of the assumption on which "the whole force of the proof turns." This in turn could explain why Maxwell

Illustrations locus	Content	Section(s)
Proposition I–III	Analysis of particle collisions; *Condition M* appears as a hidden assumption of Proposition II.	4, 6, 7
'If a great many equal spherical particles were in motion [...]'	Claim that collisions would lead to a stable kinetic energy (temperature) and velocity distribution.	5, 7
Proposition IV	Derivation of velocity distribution.	2, 3, 6, 7
Proposition V	Derivation of relative velocity distribution.	(2)
Proposition VI	Collisions lead to an equalization of temperatures of a mix of gases.	5, 7

Table 1: *Linear structure of the beginning of the* Illustrations of the Dynamical Theory of Gases ((1860)) *with the sections where they are analyzed in this paper.*

was content with the physical justification of the independence assumption since value independence of the velocity components would have appeared immediately clear. Although there seems to be no conclusive evidence deciding which of these two conflations were committed by Maxwell (maybe both), we point out that in his prior and contemporaneous work Maxwell means value independence when he is discussing "independence".

The paper is organized as follows. Section 2 starts with a reconstruction of Maxwell's proof of Proposition IV and identifies the three mathematical assumptions (of existence, independence, and symmetry) that are required for the proof to work. We work backwards from here, first historically, then conceptually. We recall Maxwell's own wording of his proof of Proposition IV and contrast it with a proof of Herschel that has been identified by historians as its likely source. Section 3 also briefly reviews the existing historical literature regarding Maxwell's attempt to justify the mathematical assumptions of Proposition IV. Section 4 reviews the particle collision model of Maxwell's Proposition I–III and identifies a crucial assumption (Condition M) that is needed for Proposition II to work. Section 5 shows how Proposition I–III and Condition M can be used to argue for the equalization of mean kinetic energies and contrasts the argument with Maxwell's Proposition VI. Section 6 distinguishes between two independence notions and calls attention to the conflation of these two notions in the works of Herschel, Clausius, and Maxwell. Section 7 puts the pieces back together by analyzing the extent to which the three aforementioned mathematical assumptions of Proposition IV – that of existence (Section 7.1), independence (Section 7.2), and symmetry (Section 7.3) – could be physically justified by Maxwell's collision model.

2. Proposition IV – Maxwell's derivation

As we know from a 1859 letter to George Gabriel Stokes (Maxwell, 1859, p. 277), Maxwell's interest in the kinetic theory of gases was aroused by the papers of Rudolf Clausius. Clausius was mainly interested in explaining heat in terms of molecular motion, and in his 1857 article he used an elastic sphere model to establish a connection between the average kinetic energy and the pressure of the gas. This treatment met the criticism of a Dutch meteorologist, C. H. D. Buys–Ballot (1858): if, as calculations suggest, the molecules of a gas move at speeds of several hundreds meters per second, odors released in one corner of a room should almost instantly be noticed in the other corner. To answer the objection, Clausius (1859) attempted to show that repeated collisions prevent molecules from traveling for great distances in straight lines. These considerations led him to introduce the notion of the mean free path (for further historical details see Brush et al. (1986a)). Maxwell developed the approach further in his *Illustrations of the Dynamical Theory of Gases* (Maxwell, 1860).

The main goal of Herapath, Waterston, Krönig and Clausius was to establish the kinetic theory as a bridge between thermodynamics and atomic theory. However, the focus of the *Illustrations* is on problems of viscosity, diffusion and heat conduction: Maxwell attempted to treat these as special cases of a general process in which momentum or energy is transported by molecular motion. To investigate these transport properties he relied on results he derived about the velocity distribution of a gas in his Proposition IV.

Proposition IV aims to show that in a box of gas the number of particles whose velocity component in a particular direction lies between v_x and $v_x + dv_x$ is proportional to $e^{-\frac{v_x^2}{a^2}}$ and that the number of particles whose speed lies between v and $v + dv$ is proportional to $v^2 e^{-\frac{v^2}{a^2}}$, where a is a parameter which gets determined later. In essence the proof relies on the fact that the only solution of the functional equation

$$f^2(0) \cdot f(|\vec{v}|) = f(v_x) \cdot f(v_y) \cdot f(v_z) \tag{1}$$

is the Gaussian

$$f(v_x) = C \cdot e^{A \cdot v_x^2}. \tag{2}$$

Equation (1) arises from three assumptions: that

(A1) a stationary velocity distribution exists;

(A2) the components of velocity in an orthogonal coordinate system are independent;

(A3) the velocity distribution only depends on the magnitude of the velocity.

If we denote the stationary velocity distribution of (A1) with $\mathbf{f}(\vec{v})$ and if by "independence" we mean probabilistic independence (treating the velocity components as random variables) then (A2) translates to $\mathbf{f}(\vec{v}) = \mathbf{f}_x(v_x)\mathbf{f}_y(v_y)\mathbf{f}_z(v_z)$ and (A3) translates to rotational symmetry of $\mathbf{f}$, namely that there exists a

function g for which $\mathbf{f}(\vec{v}) = g(|\vec{v}|)$. Noting that $\mathbf{f}_x, \mathbf{f}_y$, and $\mathbf{f}_z$ are distributions it is easy to show that $f \doteq \mathbf{f}_x = \mathbf{f}_y = \mathbf{f}_z$ and that $g(.) = f^2(0) \cdot f(.)$, leading to equation (1) .

Let us quote Maxwell's own proof (with a minor change of notation).

> Prop. IV. To find the average number of particles whose velocities lie between given limits, after a great number of collisions among a great number of equal particles.
>
> Let N be the whole number of particles. Let v_x, v_y, v_z be the components of the velocity of each particle in three rectangular directions, and let the number of particles for which v_x lies between v_x and $v_x + dv_x$, be $Nf(v_x)dv_x$, where $f(v_x)$ is a function of v_x to be determined.
>
> The number of particles for which v_y lies between v_y and $v_y + dv_y$ will be $Nf(v_y)dv_y$; and the number for which v_z lies between v_z and $v_z + dv_z$ will be $Nf(v_z)dv_z$, where f always stands for the same function.
>
> Now the existence of the velocity v_x does not in any way affect that of the velocities v_y or v_z, since these are all at right angles to each other and independent, so that the number of particles whose velocity lies between v_x and $v_x + dv_x$, and also between v_y and $v_y + dv_y$, and also between v_z and $v_z + dv_z$, is
>
> $$Nf(v_x)f(v_y)f(v_z)dv_xdv_ydv_z.$$
>
> If we suppose the N particles to start from the origin at the same instant, then this will be the number in the element of volume $(dv_xdv_ydv_z)$ after unit of time, and the number referred to unit of volume will be
> $$Nf(v_x)f(v_y)f(v_z).$$
>
> But the directions of the coordinates are perfectly arbitrary, and therefore this number must depend on the distance from the origin alone, that is
>
> $$f(v_x)f(v_y)f(v_z) = \phi(v_x^2 + v_y^2 + v_z^2).$$
>
> Solving this functional equation, we find
>
> $$f(v_x) = Ce^{Av_x^2}, \quad \phi(v^2) = C^3 e^{Av^2}.$$

(Maxwell, 1860, pp. 289-290)

By applying the fact that the total number of particles is finite and fixing the constants, $f(v_x)$ turns out to be the normal distribution, $\frac{1}{a\sqrt{\pi}}e^{-\frac{v_x^2}{a^2}}$. Thus,

Maxwell concludes, the number of particles whose velocity in a given direction lies between v_x and $v_x + dv_x$ is

$$dN_{v_x}^{v_x+dv_x} = N\frac{1}{a\sqrt{\pi}}e^{-\frac{v_x^2}{a^2}}dv_x \tag{3}$$

and the number whose speed lies between v and $v + dv$ is

$$dN_v^{v+dv} = N\frac{4}{a^3\sqrt{\pi}}v^2 e^{-\frac{v^2}{a^2}}dv. \tag{4}$$

According to Proposition V we can obtain similar results for the distribution of relative velocities of pairs of randomly chosen particles: the components of relative velocities also follow the normal distribution. Proposition VI intends to show that collisions lead to the equalization of the mean kinetic energies of systems of particles of different masses. From these results, mean values of various functions of velocities used in determining the viscosity and diffusion coefficients could be expressed in terms of a, the standard deviation of the velocities. But it's not just the derivation of mean values that had significance, for many physical properties of gases depend on the actual distribution of molecular velocity. Although the qualitative idea that even at constant temperature different molecules have different velocities was often assumed in verbal formulations – e.g. in Clausius' (1857a, pp. 113-116) explanation of evaporation – equations (3) and (4) are the first quantitative, experimentally verifiable results of the kinetic theory of heat.

3. Proposition IV – historical treatment

On the basis of indirect evidence, several authors (Gillispie, 1963; Brush et al., 1986a) have suggested that Maxwell simply adapted the derivation of Proposition IV from a work in statistics, most probably relying on a July 1850 *Edinburgh Review* article by Sir William Herschel on Quetelet's collection of essays on probability (Herschel, 1850). The review was also reprinted in Herschel's *Essays* in 1857, and we know from an early 1858 letter (Campbell & Garnett, 1882, pp. 301-302) that Maxwell read and "liked" these essays.

In this 57 pages long, typical Victorian-style review, Herschel first gives his own analysis of the subject and praises Quetelet's work on the derivation of the law of least squares. But, to make the subject more palpable to non-mathematical readers, Herschel comes up with a new, "simple, general, and perfectly elementary proof of the principle in question," based on the following three assumptions:

> We set out from three postulates. 1st, that the probability of a compound event, or of the concurrence of two or more independent simple events, is the product of the probabilities of its constituents considered singly; 2nd, that there exists a relation or numerical law of connexion (at present unknown) between the amount of error

committed in any numerical determination and the probability of committing it, such that the greater the error the less its probability, according to some regular LAW of progression, *which must necessarily be general and apply alike to all cases, since the causes of error are supposed alike unknown in all;* and it is on this ignorance, and not upon any peculiarity on cases, that the idea of probability in the abstract is founded; 3dly, that errors are equally probable if equal in numerical amount, whether in excess, or in defect of, or in any way beside the truth. This latter postulate necessitates our assuming the function of probability to be what is called in mathematical language *an even function,* or a function of the square of the error, so as to be alike for positive and negative values; and the postulate itself is nothing more than the expression of our state of *complete* ignorance of the causes of error, and their mode of action. (Herschel, 1850, pp. 19-20)

In the course of the derivation Herschel adds that the independence assumption applies to deviations in a rectangular direction: "the observed oblique deviation is equivalent to the two rectangular ones, supposed concurrent, and which are essentially independent from one another" (Herschel, 1850, p. 20). Together with this later addition Herschel's three main assumptions – existence of a distribution of deviations, probabilistic independence of the deviations in a rectangular direction, and spherical symmetry of the distribution – are similar to those of Maxwell. Herschel's derivation of the functional form, albeit in a verbalized form, proceeds in the same way as Maxwell's, the main difference being that Herschel considers two rather than three dimensions and that he shows no interest in calculating the distribution of the magnitude of the errors.

Even if Maxwell adopted Herschel's derivation it is natural to ask whether Maxwell tried to justify the mathematical assumptions of Proposition IV. According to Maxwell the conclusion of Proposition IV is intended to apply "after a great number of collisions among a great number of equal particles" takes place. Proposition IV of the *Illustrations* is preceded by three other propositions that describe the nature and effects of particle collisions. Yet we find that the historical literature pays very little to no attention to these prior findings of Maxwell; we typically find half-sentence-remarks about the role Propositions I–III played in the argumentation, with as many different half-sentence-remarks as there are historians making them.

The only pertaining remark Porter (1981, pp. 98-99) – who explicitly deals with the history of statistics and probability theory that lies in the background of Proposition IV – makes is that Maxwell's "kinetic gas model [...] was subject to a nearly infinite number of minute causes of deviation, and it displayed perfect statistical regularity," and this somehow "led Maxwell to the insight that the distribution of molecular velocities probably conformed to [the error] curve." It is rather unclear from this remark how Maxwell is supposed to have been led to his insight or even just what the insight of Maxwell was; understood strictly Porter's description is misleading since Maxwell did not make a probabilistic

statement about the conformity of the actual velocity distribution to the normal distribution.

Truesdell (1975, p. 34) mentions that Maxwell's work "begins [...] with consideration of the impact of two spheres" but he doesn't attribute any importance to this part of Maxwell's paper.

Other historians generally assume that Maxwell's analysis of molecular collisions has something to do with underpinning Proposition IV, but we find no clear consensus about the motivations. The longest treatment, referred to by many subsequent historical articles, comes from Brush (1958, p. 243-244) (repeated almost verbatim in Brush (1976, pp. 185-186)) whose length still permits it to be quoted in full:

> [Maxwell] began by pointing out that, if two elastic spheres collide, all directions of rebound are equally likely. He apparently believed that this fact would ensure, not only that all directions of motion were equally probable in the gas, but also that the probability distribution for each component of the velocity was independent of the values of other components.

In other words Brush suggests that Maxwell's analysis of particle collisions was aimed at justifying the probabilistic independence assumption (A2) and the rotational symmetry assumption (A3) but he provides no details about how Maxwell's justification is supposed to have proceeded. It is also not clear what Brush means by "pointing out"; as we will see later in this respect Heimann (1970, p. 194) is slightly more accurate when he says in his half sentence remark that "Maxwell assumed that all motions [of the particles] were equally probable" (although, as we will see, Maxwell's *assumptions* only *implied* that after rebound all *directions* of motion are equally probable).

According to Brush et al. (1986a, p. 7) Maxwell claimed that "when a moving sphere collides with with another fixed sphere randomly placed in its path, the direction in which it rebounds is distributed with equal probability over each element of solid angle" (this is incorrect since Maxwell did not assume any of the particles to be fixed), and they further state – without providing any further details or evidence – that Maxwell used this claim to argue that "the distribution law has an explicit functional form".

Uffink (2007, p. 944) mentions en passant that Maxwell realized that the "velocities [of the particles] will suffer incessant change due to mutual collisions" but he implies that Maxwell's reference to collisions is, with Maxwell's own remark made 20 years later about certain derivations in the works of others, "rather for the sake of enabling the reader to form a mental image of the material system than as a condition for the demonstration" (ibid. p. 948). Uffink also emphasizes that the assumptions of Maxwell's Proposition IV, despite appearances and in contrast to Maxwell's subsequent proof of 1867, "do not refer to collisions at all" (ibid. p. 948; see also his contrast with Maxwell's 1867 proof on ibid. p. 950).

Dias (1994) is the only exception in giving an extensive treatment of Maxwell's collision model but the focus of the paper is on exploring the connections be-

tween Maxwell's Proposition II and later debates of the 1890s on the Stosszahlansatz; Dias does not address what role, if any, Maxwell's collision model played in justifying the assumptions of Proposition IV.

The scant attention paid to the justification question may stem from historians unanimously following Brush's 1958 reconstruction of Proposition IV as a proof resting on *two* assumptions, that of probabilistic independence (A2) and of rotational symmetry (A3), and they are at best only tacit about the need for a third assumption (A1) that a stationary velocity distribution exists (see i.e. Brush (1976, pp. 185-186, pp. 342-343), Heimann (1970, p. 194), Porter (1981, p. 98), Uffink (2007, 945)). They do so despite acknowledging the analogy with Herschel's derivation pointed out by Gillispie, five years after Brush's 1958 reconstruction, in 1963. As we have seen Hershel explicitly emphasized a third assumption, that there must "exists a relation or numerical [...] regular LAW of progression, *which must necessarily be general and apply alike to all cases*" (Herschel, 1850, pp. 19-20; emphasis in original), which adapted to the kinetic theory translates as the assumption of the existence of a stationary velocity distribution, our (A1). By only explicitly acknowledging two assumptions instead of three the natural guess might have been that what Maxwell tried to do with his collision model was to physically justify one or both of these two assumptions. Since without a detailed analysis of Maxwell's collision model only superficial analogies are apparent between Maxwell's prior results and assumptions (A2)–(A3) these prior results of Maxwell were supposedly aimed at justifying, and since Maxwell himself later abandoned the independence assumption (A2) describing it as being "precarious" (Maxwell, 1867, p. 437), it could have been tempting to glance over the opening part of Maxwell's paper as being non-consequential to the argument presented in Proposition IV.

We are going to argue that the most plausible role the collision model played in Maxwell's thinking is in justifying assumption (A1), although our analysis of Maxwell's collision model also sheds light on the extent to which his prior results are capable to support assumptions (A2) and (A3) beyond the superficial analogies. To do this we start with a brief reconstruction of his Propositions I–III.

4. Propositions I–III and Condition M – particle collisions

In the beginning of the first chapter, "On the Motion and Collision of Perfectly Elastic Spheres," Maxwell presents us with three Propositions concerning the interaction of particles. Proposition I and Proposition III are straightforward exercises in Newtonian mechanics that establish the motion of two particles – modeled by perfectly elastic spheres of non-zero radii s, S and of mass m, M – after they collide. It is assumed that the particles do not interact aside such pairwise collisions. In modern notation Maxwell calculates the encounter operator $(\vec{v}, \vec{V}, \rho) \mapsto (\vec{v}', \vec{V}')$ where $\vec{v}, \vec{V}$ are the velocities of the particles before the impact, $\vec{v}', \vec{V}'$ are the velocities after impact, and ρ is a further parameter describing the impact in the following way. Let $\vec{r} = \vec{v} - \vec{v}_{CM}$
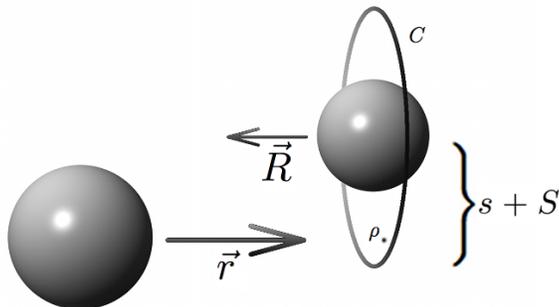
Figure 1: $\vec{r}, \vec{R}$: velocities relative to the center of mass. C: circle of radii $s + S$ centered on the second particle in the plane perpendicular to $\vec{r}$. ρ: point where the line of motion of the first particle intersects circle C.

be the velocity of the first particle relative to the velocity of the center of mass $\vec{v}_{CM} = \frac{1}{m+M}(m\vec{v} + M\vec{V})$ before impact and let C be a circle of radius $s + S$ that lies in the plane perpendicular to $\vec{r}$, C being centered on the center of the second particle. For a collision to take place the line of motion of the first particle needs to intersect C; ρ is the position of this intersection within C. (See Figure 1.)

Due to conservation of momentum the collision does not affect the velocity of the center of mass: $\vec{v}'_{CM} = \vec{v}_{CM}$. Maxwell shows that a perfectly elastic collision alters the direction of the relative velocity but it keeps its magnitude intact:

$$\vec{r}' = r \cdot \vec{u}(\rho), \tag{5}$$

and therefore the velocity after impact can be expressed as

$$\vec{v}' = \vec{v}_{CM} + r \cdot \vec{u}(\rho), \tag{6}$$

where $r = |\vec{r}|$ and $\vec{u}(\rho)$ is a direction of unit length depending on the position ρ within the circle C. (Mutatis mutandis for the $\vec{R} = \vec{V} - \vec{v}_{CM}$ velocity of the second particle relative to the center of mass.)

Maxwell also claims to have proven that whenever a collision takes place all directions of rebound are equally likely in the center of mass frame. What his Proposition II actually shows is that if the assumption that

Condition M the impact parameter ρ is distributed uniformly within the circle C

holds then

Condition A the $\vec{u}(\rho)$ directions of rebound relative to the center of mass are distributed uniformly on the surface of the unit sphere.

10

The name Condition A is due to S. H. Burbury (1894) who invoked it in a debate on the pages of *Nature* with G. H. Bryan about Boltzmann's H-theorem. Condition M is the terminology we introduce here to distinguish Condition A from the assumption Maxwell actually made during the proof of Proposition II. Since in Maxwell's particle model Condition M and Condition A imply each other they tend to be identified (for such an exchange of conditions see i.e. Dias (1994, p. 347)). Maxwell himself neither emphasizes nor attempts to justify Condition M; he merely asserts that "within this circle [C] every position [ρ] is equally probable" (Maxwell, 1860, p. 288), and in the coming years he keeps referring to the conclusion of Proposition II as if it were unconditional (see e.g. Maxwell (1863, p. 339); for a further analysis of Condition A and the *Nature* debate see Dias (1994)).

The conclusion drawn from Propositions I–III is that the $\vec{v}'$ velocity of a particle after impact consists of two components: the first is the $\vec{v}_{CM}$ velocity of the center of mass, which is determined by the $\vec{v}$ and $\vec{V}$ incoming velocities, and the second is the $\vec{r}'$ velocity relative to the center of mass, whose magnitude is determined by the incoming velocities, but whose direction is not determined by the incoming velocities, but may with equal probability be in any direction.

5. Proposition VI – tendency towards equilibrium

After being done with Propositions I–III, but before turning his attention to Proposition IV, Maxwell makes the following remark about the effect of many collisions:

> If a great many equal spherical particles were in motion in a perfectly elastic vessel, collisions would take place among the particles, and their velocities would be altered at every collision; so that after a certain time the *vis viva* will be divided among the particles according to some regular law, the average number of particles whose velocity lies between certain limits being ascertainable, though the velocity of each particle changes at every collision. (Maxwell, 1860, p. 289)

Thus Maxwell claims that after sufficient number of collisions the kinetic energies become distributed along a non-changing distribution. This implies that some sort of process that brings about this distribution of kinetic energies takes place, apparently due to the collisions these particles go through.

On the surface of it Maxwell's remark strikes as an expression of hope rather than an argument. Is there any evidence that Maxwell had good reasons to hold that Propositions I–III are capable of supporting a claim about tendency towards equilibrium? The answer is affirmative, and this becomes clear if we jump ahead to Proposition VI in which Maxwell argues that when two gases mix their temperatures equalizes. This argument is unknown in the historical and in the physics literature and since besides being relevant to the question of justification of an assumption of Proposition IV it is also interesting and historically significant on its own right, we now briefly reconstruct it.
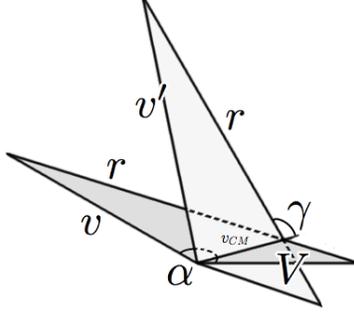
Figure 2: $\vec{v}, \vec{V}$: incoming velocities; $\vec{v}_{CM}$: velocity of center of mass; $\vec{r} = \vec{v} - \vec{v}_{CM}$: velocity of the first particle relative to the center of mass; $\vec{v}'$: velocity of the first particle after rebound; $\alpha = \angle(\vec{v}, \vec{V})$: angle between incoming velocities; $\gamma = \angle(\vec{r}', \vec{v}_{CM})$: angle of rebound.

5.1. The reconstructed argument for tendency towards equilibrium

When two particles, modeled as perfectly elastic spheres of mass m and M, collide the difference between their kinetic energies after the collision can be readily expressed as a function of the difference between their kinetic energies before the collision:

$$\left(\frac{mv'^2}{2} - \frac{MV'^2}{2} \right) = C_0 \cdot \left(\frac{mv^2}{2} - \frac{MV^2}{2} \right) + C_1 \cdot \cos\alpha + C_2 \cdot \cos\gamma, \qquad (7)$$

where α is the angle enclosed by the incoming velocities whose magnitude is v and V and γ is the angle enclosed by the velocity of the center of mass $\vec{v}_{CM}$ with the relative rebound velocity $\vec{r}'$ (see Figure 2; for calculation of the coefficients see the Appendix). Propositions I–III together with Condition M entail that $\vec{r}'$ may point to all directions with equal probability, which means there is an equal chance that the angle of rebound γ falls between $\bar{\gamma}$ and $\bar{\gamma} + d\bar{\gamma}$ or between $180^o - \bar{\gamma}$ and $180^o - \bar{\gamma} - d\bar{\gamma}$.

Consider now collisions between many particles during a time interval Δt, and let us write $(k, l) \in I_{v,V}^{\alpha}$ whenever k is a particle of mass m, l is a particle of mass M, and k and l collide during Δt with respective incoming speeds v, V and angle of income α. We assume that Δt is long enough for $|I_{v,V}^{\alpha}|$ to be large but short enough to allow for at most one collision for each particle. Let us sum up equations (7) for all collision pairs in $I_{v,V}^{\alpha}$ by assuming that Condition M holds for all these particle collisions independently. Since then all directions of rebound occur with equal frequency and since C_2 only depends on m, M, v, V, and α which are now all fixed, the $C_2 \cdot \cos\gamma$ terms approximately cancel and we get

$$\sum_{(k,l) \in I_{v,V}^{\alpha}} \left(\frac{mv_k'^2}{2} - \frac{MV_l'^2}{2} \right) \approx C_0 \cdot |I_{v,V}^{\alpha}| \cdot \left(\frac{mv^2}{2} - \frac{MV^2}{2} \right) + C_1 \cdot |I_{v,V}^{\alpha}| \cdot \cos\alpha. \quad (8)$$

12

Let us further assume that all directions of incoming velocities for the colliding particles occur with approximately equal frequency – an assumption that Maxwell took for granted in his later derivations, i.e. in Maxwell (1867) – and that the number of collisions is independent of the enclosed angle of income (that $|I_{v,V}^{\alpha}|$ is the same for all α, or at least that $|I_{v,V}^{\alpha}| = |I_{v,V}^{-\alpha}|$ for all α given a fixed v and V) and sum up for all possible angles of income. Since C_1 only depends on m, M, v, and V which are now all fixed, the $C_1 \cdot |I_{v,V}^{\alpha}| \cdot \cos \alpha$ terms approximately cancel and we get

$$\sum_{\alpha} \sum_{(k,l) \in I_{v,V}^{\alpha}} \left(\frac{m v_k'^2}{2} - \frac{M V_l'^2}{2} \right) \approx C_0 \cdot \sum_{\alpha} |I_{v,V}^{\alpha}| \cdot \left(\frac{mv^2}{2} - \frac{MV^2}{2} \right). \quad (9)$$

By further summing up for all possible magnitudes of incoming velocities v, V on both sides of the equation we account for all collisions between particles of mass m and M during Δt exactly once, so a division with the total number of such collisions and substituting for C_0 yields

$$\frac{\widehat{mv'^2}}{2} - \frac{\widehat{MV'^2}}{2} \approx \left(\frac{m-M}{m+M} \right)^2 \cdot \left(\frac{\widehat{mv^2}}{2} - \frac{\widehat{MV^2}}{2} \right), \quad (10)$$

meaning that the difference between the mean kinetic energies of the colliding particles decrease after the collisions!

Any change in the difference between the mean kinetic energies (between the temperatures) of *all* m-particles and of *all* M-particles during Δt is due to collisions between particles of different mass (expressed by equation (10)) since the total kinetic energy of particles which do not collide and the total kinetic energy of those particles that collide with another particle of the same mass does not change during Δt. If the particles that undergo collision during Δt are randomly drawn from the set of all particles then the larger the initial temperature difference between m-particles and M-particles the larger the chance that the drawn $I_{v,V}^{\alpha}$ samples of colliding pairs also exhibit a mean kinetic energy difference of the same sign. If the sign is the same then since due to (10) the mean kinetic energy difference between the colliding pairs decreases during Δt it follows that the temperature difference of all m-particles and M-particles also decreases. (If the sign is different then the temperature difference between the gases *increases*, underlining that the present argument for tendency towards equilibrium is statistical.) By repeating the argument for successive Δt time intervals it follows that the difference between temperatures of m-particles and M-particles tends to vanish. When the temperature difference vanishes or becomes small the collisions continue to produce small temperature fluctuations (due to the mean kinetic energy differences that may still occur in the random sample of colliding particle pairs).

Note that the proof becomes extremely simple when we mix gases of different temperature whose particles have the same mass. In this case equation (7) simplifies to

$$\left(\frac{mv'^2}{2} - \frac{mV'^2}{2} \right) = C_2 \cdot \cos \gamma \quad (11)$$

and one only needs to invoke the first cancellation argument provided by Condition M to conclude that the temperature difference decreases as collisions take place. What we thus have is a simple, elegant, and elementary mechanical argument for a statistical tendency towards equilibrium that is based on the assumptions that

(i) particles can be modeled as perfectly elastic spheres,

(ii) the number, size, speed etc. of these particles is such that the time evolution can be split into successive time intervals during which each particle has an approximately equal independent chance to participate in at most one, pairwise collision, and

(iii) in this interval all directions of rebound occur with approximately the same relative frequency (that for any pair of incoming velocity and for the colliding particle pairs that have these incoming velocities the relative frequency of the impact parameter is roughly the same for all positions on the collision circle).

Although we know well from later developments that assumptions (ii) and (iii) are difficult to square with the strict dynamical evolution entailed by (i), its simplicity and its geometrically intuitive probabilistic assumptions could very well allow the argument to fit in an undergraduate course on classical mechanics.

5.2. Proposition VI in the light of the reconstruction

How can it be that this simple argument for tendency towards thermal equilibrium is unknown by historians and physics educators alike? The conclusion that collisions bring about an equalization of temperatures is what Maxwell draws in his Proposition VI after he derives a formula he interprets as (10) by invoking his earlier results about collisions. Prima facie, however, Maxwell's proof seems rather baffling. What he seems to have calculated in his Proposition VI is the difference of kinetic energies of particles of mass m and M after they collide in a specially arranged manner where (using the notation of the previous subsection)

(p1) $\gamma = 90^o$,

(p2) $\alpha = 90^o$,

(p3) $v = \widehat{v}$, $V = \widehat{V}$.

By substituting assumptions (p1)-(p3) to equation (7) one gets

$$\frac{m\widehat{v}'^2}{2} - \frac{M\widehat{V}'^2}{2} = \frac{(m-M)^2}{(m+M)^2}\left(\frac{m\widehat{v}^2}{2} - \frac{M\widehat{V}^2}{2}\right),\tag{12}$$

which is the equation that Maxwell's own derivation would have really established. Maxwell curiously assumes that $\widehat{v}' = \widehat{v'}$ and he also interprets (12) as

if he had just derived (10) and proceeds to draw the conclusion about the long term equalization of mean kinetic energies on the basis of it.

Thus the calculations of Proposition VI make Maxwell's argument seem to be dependent on a particular arrangement of incoming velocities and rebound angle of a pair of particles, and it seems entirely unclear how he thought that an argument about a particular collision of a pair of particles could allow for a valid inference about averages over all collisions, or how the magnitude of velocity after impact that initially equals the mean velocity could be substituted for the mean of velocities after impact, or how the kinetic energy calculated from a mean velocity could be substituted for a mean kinetic energy. On this basis Maxwell's proof, if mentioned at all, is dismissed in the historical literature as a "rather lame argument" and the few historians who do devote more than a footnote to it find it "amazing" that the otherwise brilliant Maxwell or "any of his contemporaries who bothered to examine the argument in detail would have accepted it" (Brush, 1976, p. 344).

In the light of the simple argument for equalization of kinetic energies we outlined in the previous subsection – an argument that rest purely on Maxwell's Propositions I–III and invokes the same sort of calculations that Maxwell performs during his proof of Proposition VI – it seems plausible that Maxwell was being too terse instead of lame. It is clear from the geometry of collisions and from the argument we gave above that for equation (7) averaging over all directions of rebound yields the same result as making assumption (p1), as a result of which the third term cancels. This then justifies focusing on collisions with perpendicular rebound angle as being sufficiently typical for the purposes of the proof. It is also geometrically clear that averaging over all directions of incoming velocities yield the same result as making assumption (p2), as a result of which the second term cancels. This then would justify focusing on collisions with perpendicular income angle as being sufficiently typical for the purposes of the proof.

Propositions I–III motivates assumption (p1) but assumption (p2) is in need of justification. Instead of invoking a further assumption that all directions of incoming velocities are equally likely (like we did above, and as Maxwell did in his later works), Maxwell sets the magnitude of the incoming velocities v, V to equal their mean values given by Proposition IV (assumption (p3)), and he further assumes (p4) that the magnitude of the relative velocity $|\vec{v} - \vec{V}|$ equals its mean value given by Proposition V. Since Proposition V shows that the square of the mean relative velocity equals the sum of squares of the mean velocities, and therefore that the two incoming velocities must be perpendicular, (p3) and (p4) entail (p2). Assumptions (p3) and (p4) are also assumptions about typical velocities, but they are not appropriate for the purposes of the proof. Albeit the proof that proceeds directly along (p3) and (p4) is clearly suspect, (p3) and (p4) might have been intended as mere short-cut for a more detailed argument that rests on an averaging procedure over the incoming velocities, along the lines we sketched above.

Albeit this reconstruction still renders the proof of Proposition VI messy and misleading, it allows us to see the arrangement choices Maxwell makes during his derivation as geometrically intuitive shortcuts to an averaging procedure that rests on an application of the results of his Propositions I–III. If we make the charitable assumption that Maxwell simply didn't properly clean up his proof, as opposed to making assumptions whose fallaciousness should have been obvious, then his reconstructed Proposition VI becomes a convincing demonstration of a tendency towards equilibrium that lives up to the standards of its own time. Although the proof does depend on unjustified probabilistic assumptions such as Condition M, in this regard it does not fare worse than Boltzmann's later attempts (starting with Boltzmann (1868)) to arrive at his celebrated H-theorem, and thus it should reclaim its historical precedence over Boltzmann's demonstrations as being the first good "bad" mechanical proof of tendency towards equilibrium.

The only historical side-remark I'm aware of that also suggests a direct connection between Maxwell's Propositions I–III and Boltzmann's H-theorem is a side remark of Dias (1994, p. 347) who says that

> Maxwell (1860) invoked Proposition II to show how collisions randomized the directions of motion of the molecules, bringing about isotropy of pressure, as demanded by the equation of state for perfect gases. In this use, Proposition II becomes a kind of "proof" of how equilibrium of pressure (transmission of linear momentum) sets in, and thus is, metaphorically speaking, a sort of H-Theorem.

Although the connection between Proposition II and the randomization of direction of motions of molecules is clear, I find no evidence for Dias' claim that Maxwell made any use of Proposition II to argue for the onset of the isotropy of pressure (Dias provides none). From Dias' reference to linear momentum I suspect that his remark was based on a misinterpretation of what Proposition VI aimed to show.

Maxwell never seem to have revisited Proposition VI. Around 1863 he became dissatisfied with Clausius' mean free path method and started to develop a theory valid for any kind of force laws acting between particles, which he came to treat as centers of force rather than colliding spheres with determinate radii. With his new approach serious problems emerge concerning Proposition I and III: as Maxwell's 1867 calculations show, the motion of particles subject to a never-vanishing repulsive force law in general depends on their incoming relative velocity, so the angle of rebound also depends on this incoming relative velocity. Although in his *On the Dynamical Theory of Gases* (1867) he chooses a particular inverse 5th-power force law for which the angle of reflection is independent of the incoming relative velocity, this choice is poorly motivated and gets abandoned later. His writings after 1867 indicate that he grew skeptical of the possibility of justifying the Second Law on a mechanical basis (Uffink, 2007, pp. 951–952).

6. Value vs. probabilistic independence

Before analyzing the extent to which Proposition I–III, Condition M, and the argument for an approach to equilibrium can be used to justify the mathematical assumptions of Proposition IV it is worthwhile to take a look at how Maxwell, Herschel, and Clausius understood what these mathematical assumptions, in particular the independence assumption (A2), amounted to.

Let us start with Herschel's understanding of the independence assumption of his own proof. We note first that there is a small but important difference between Herschel's original 1850 review article and its 1857 reprint, which is the version that Maxwell have likely read. In a footnote added to the 1857 reprint Herschel attempts to clarify what he means by independence of the deviations of errors in different directions and emphasizes the importance of this independence assumption:

> That is, *the increase or diminution [of the deviation] in one or [sic!]*
> *which may take place without increasing or diminishing the other.*
> On this, the whole force of the proof turns. (Herschel, 1857, p. 400),
> (Brush et al., 1986a, p. 11).

Although Maxwell's likely adaptation of Herschel's derivation became common wisdom in the historical literature, so far it has gone unnoticed by historians that in the quoted emphatic footnote Herschel actually mischaracterizes the crucial necessary independence assumption of the derivation. A mathematically correct derivation requires *probabilistic independence* of deviations in rectangular directions; according to Herschel, however, the whole force of the proof turns on the assumption that a deviation in one direction does not impose a restriction on the values a deviation in another direction may take, which is not the same as probabilistic independence of the deviations.

Let us briefly distinguish these notions of independence. Let ξ and υ denote physical quantities (i.e. velocity components) that can take values in the set X and Y respectively. Suppose that we may not know what other quantities serve as inputs that determine the momentary values of ξ and υ, but let us assume that we might know about certain constraints that connect the values ξ and υ can take in the following way. Let $d : X \to \mathcal{P}(Y)$ assign to each value $x \in X$ the set of values $d(x) \subseteq Y$ that variable υ may take if we condition upon ξ taking the value $\xi = x$. If d is single valued for all $x \in X$ then if ξ takes a value x then υ must take the only possible value in $d(x)$: we may thus say that ξ *value determines* υ. If $d(x)$ has more than one element for at least one $x \in X$ then υ *is partially value independent from* ξ: for some value variable ξ may take there are several possibilities that υ may take that are not determined by the value of ξ. If $d(x) = Y$ for all $x \in X$ then υ *is value independent from* ξ: no matter what value variable ξ takes that does not influence the possible values υ may take. Finally when ξ and υ are random variables then υ *is probabilistically independent from* ξ if for all x, y: $Pr(\xi \leq x \wedge \upsilon \leq y) = Pr(\xi \leq x) \cdot Pr(\upsilon \leq y)$, or, alternatively, when $g_{\xi,\upsilon}(x,y) = g_\xi(x)g_\upsilon(y)$ for the respective density functions.

These independence notions are not equivalent with each other. Value independence does not imply probabilistic independence: indeed the domains of two value independent variables do not even need to be endowed with a probability measure. Probabilistic independence also does not imply value independence: a value x of ξ may prevent υ taking a value y without ruining probabilistic independence of ξ and υ if the probability of the occurrence of x and y is zero. Value independence is symmetric, and value independent variables are also partially value independent but the converse is not the case.

It is clear both from Herschel's derivation and from Maxwell's proof of Proposition IV that their "independence" assumption is meant to entail the factorization of the probability distribution function. If by "independence" one means probabilistic independence then factorization does follow and the proof is correct; factorization however does not follow from value independence, which Herschel mistakenly emphasized to be the crucial assumption of his derivation.

Now even if Maxwell adapted the derivation from Herschel's article it is natural to assume that he had reasons to believe that the required mathematical assumptions can be motivated by physical considerations about gases. Without such physical considerations the application of the mere mathematical proposition to gases would remain unjustified. But Maxwell's attempt at physically justifying independence of the velocity components would have clearly depended on his understanding of the sense of independence he had believed to guarantee the success of the proof. We are going to return to this point in Section 7.2.

Let us now turn to Clausius' understanding of the meaning of independence. In his summary of Maxwell's Propositions I–III Clausius (1862, p. 424) claims that the velocity of center of mass $\vec{v}_{CM}$ quantifies the extent to which the velocity after impact is "dependent" upon the velocities before impact while the velocity relative to the center of mass $\vec{r}'$ quantifies the extent to which the velocity after impact is "independent" from the velocities before impact. It is not clear what Clausius meant here by "independence" since what he writes is compatible with several different readings. Given Condition M $\vec{r}'$ may point in any direction no matter what the incoming velocities $\vec{v}$ and $\vec{V}$ were, so it is value independent since setting values for $\vec{v}$ and $\vec{V}$ does not restrict the range of possible values for the direction of $\vec{r}'$. In the same vein $\vec{v}'$ is partially value independent from $\vec{v}$ and $\vec{V}$ since setting values for $\vec{v}$ and $\vec{V}$ does narrow the range of possible values of $\vec{v}'$ but $\vec{v}'$ may still take several values, not just one. Clausius could also have meant that the direction of $\vec{r}'$ is probabilistically independent from $\vec{v}$ and $\vec{V}$: conditioning upon any value for $\vec{v}$ and $\vec{V}$ does not alter the *probability* that $\vec{r}'$ points in a certain direction.

Upon a closer look what Clausius emphasizes in the text and in the accompanying footnote is that the incoming velocities do not determine the direction of the velocities relative to the center of mass after impact and in turn they do not fully determine the velocities after impact either, as these latter may take several (indeed, infinitely many) different possible values for any fixed values of the incoming velocities. These remarks of Clausius are suggestive that he intended to make a point about the partial value independence of the velocities before and after impact, but there is no evidence that Clausius even

distinguished between value independence and probabilistic independence.

Finally let us make a brief remark about Maxwell's own usage of the term "independent". The term appears once more in the context of a generalization of Proposition IV and Proposition V (ibid. p. 315) but the appearance does not shed any more clarity on its intended interpretation than the proof of Proposition IV. In all other contexts in which the term "independent" appears in the *Illustrations* (see p. 288, 298, 300) it is clear that Maxwell understands it as referring to value independence. Taking a systematic look at Maxwell's prior and contemporary work on color vision (see esp. (Maxwell, 1855, pp. 120-121)) I also find that Maxwell systematically uses the term independent in the sense of value independence.[1]

Indeed I find that ambiguous usage of the term "independent" is pervasive in and before the 1860s. That different concepts are named by the same term would not in itself be a problem if users of the term were conscious of their non-interchangeability. Unfortunately it seems that this consciousness was missing in the works of those, such as Clausius and Herschel, that most likely have influenced Maxwell's thinking, and there is no evidence that Maxwell was aware of such distinction in and before his 1860 work on the kinetic theory.

7. Proposition IV – physical justification of its assumptions

7.1. (A1) The existence of a stationary velocity distribution

Let us now return to Maxwell's Proposition IV. The paragraph preceding the statement of Proposition IV, which we have already quoted, claims that after large number of collisions the kinetic energy "will be divided among the particles according to some regular law, the average number of particles whose velocity lies between certain limits being ascertainable" (Maxwell, 1860, p. 289), meaning that this number is not going to be further altered by subsequent collisions, that is, that the velocity distribution becomes stationary.

Indeed without the assumption of stationarity it would be hard to understand what enabled the presumption of the time-independent velocity distribution f right in the beginning of the proof of Proposition IV. But if Maxwell

---

[1]Judging by the number of his publications the topic Maxwell most vehemently pursued in the years prior to and contemporaneous with his early research on the kinetic theory was that of color vision. A previous draft of this paper argued that Maxwell's own experimental work on color vision provided him familiarity with statistical reasoning; in particular Maxwell have encountered a use of distributions which is not rooted in error, ignorance, or uncertainty, and hence understanding probabilities as being rooted in the physical phenomena was not entirely new to him when he started working on the kinetic theory. On the basis of his drawn figures it seems likely that this is also true in particular for the normal distribution. The draft addressed the parallels Maxwell draws between color mixing in his three dimensional color space and addition of vectors in three dimensions in mechanics. Finally it also elaborated on the claim that in these prior works Maxwell understands the term independence as value independence. At the request of a referee these sections got relocated in a separate historical note.

succeeded to show that particle collisions lead to stationarity prior to Proposition IV then it would have been sufficient for him to assume, for the purposes of the proof of Proposition IV, the existence of the right kind of particle collisions. And this is what Maxwell actually says he does, to wit:

> Prop. IV. To find the average number of particles whose velocities lie between given limits, *after a great number of collisions among a great number of equal particles.* (Emphasis added.)

That his analysis of particle collisions was in part intended to establish the existence of a stationary velocity distribution is also clear from Maxwell's later remarks about his 1860 proof:

> The only case in which I have determined the form of this [velocity distribution] function [in the 1860 paper] is that of one or more kinds of *molecules* which have *by their continual encounters brought about a distribution of velocity* such *that* the number of molecules whose velocity lies within given limits *remains constant.* (Maxwell, 1867, p. 437) (Emphasis added.)

In addition we have also seen that the idea that molecular collisions could lead to a stationary velocity distribution was more than baseless wishful thinking. We showed that Propositions I–III are indeed sufficient for a proof of tendency towards equilibrium, which is what Maxwell intended to do with his Proposition VI. In the special case when the masses are equal, which is the relevant case for Proposition IV, we have also seen that the proof of equalization of kinetic energies becomes so simple as to potentially warrant only verbal referencing. The availability of a proof of reaching equalization of temperatures lends credence for the availability of a proof of reaching a stationary velocity distributions.

Thus we can conclude that Maxwell's argument consists of two steps: a 'lemma' showing that particle collisions lead to stationarity, and a 'theorem' that makes use of the lemma. It is then of course no surprise that reference to particle collisions only appears in the statement of the theorem and is absent from the proof following the statement, since all the work having to do with particle collisions was previously delegated to the lemma; the proof of the theorem itself only makes use of the assumption of stationarity that follows from the lemma and its assumption about collisions. Establishing the conclusion of course requires both establishing the lemma and establishing the theorem, and thus Maxwell's 1860 argument does crucially depend on particle collisions, contrary to what historians of science (i.e. Brush et al. (1986a, p. 8), Uffink (2007, p. 948)), who only focus on the proof that follows the statement of Proposition IV, claim.

The extent to which Maxwell succeeded in establishing the lemma is, of course, another question. Establishing that a certain tendency towards thermal equilibrium does follow from the model assumptions is clearly significant, but it is unclear to me how equalization of kinetic energies would strictly imply

the onset of a stationary velocity distribution. If we conceptually divide the particles of the gas into two large groups – say, red and blue – then Maxwell's results imply that after collisions continue for a long time the mean velocity of all red particles will equal the mean velocity of all blue particles. This puts a constraint on how the velocity distribution of all particles may change after collisions. Since the division of particles to red and blue is arbitrary we end up with many such similar constraints, and jointly all these constraints *may* be sufficient to force the velocity distribution to become stationary. So much is clear, but whether this is so requires an explicit argument rather than a hopeful gesture.

Although this attempt stays close to the line of reasoning Maxwell himself utilizes in the end of his proof of Proposition VI he does not elaborate on the details of such or similar musings. He apparently thought that the evidence provided by Propositions I–III on the nature of particle collisions and his scattered remarks preceding the statement of Proposition IV are sufficient to establish that collisions lead to a stationary velocity distribution straightforwardly. Some followers of Maxwell shared this view, although it progressively became more clear that a proof of it, if possible at all, may not be easily befitted to a margin. Yet as late as 1890 we find proofs to the effect that Propositions I–III and Condition A implies stationarity:

> [Proposition] 6. (a) Every distribution of velocities among the molecules which satisfies the condition that for given [velocity of the center of mass of a pair of colliding particles] V all directions of [the velocity of rebound relative to the center of mass] R are equally probable is [...] stationary. (Burbury, 1890, p. 299)

One of the main difficulties with these proofs is that they sneak in a probabilistic independence assumption in one form or another without appropriate justification.

7.2. (A2) Probabilistic independence of velocity components

The most crucial assumption of the proof of Proposition IV is the factorization of the velocity distribution to its marginals, which is a way to state the probabilistic independence of the velocity components. Let us re-quote, in its entirety, what Maxwell had to say about establishing this latter:

> [...] the existence of the velocity v_x does not in any way affect that of the velocities v_y or v_z, since these are all at right angles to each other and independent[.]

The obvious reading of this claim is that setting any value for the velocity component v_x does not affect what values v_y and v_z could possibly take; in other words, that v_x is value independent from v_y and v_z. This obvious reading renders the claim obviously true and in no need of further justification. As we have seen before, Herschel emphasizes that value independence is the crucial assumption that is needed to justify the factorization of the distribution to its
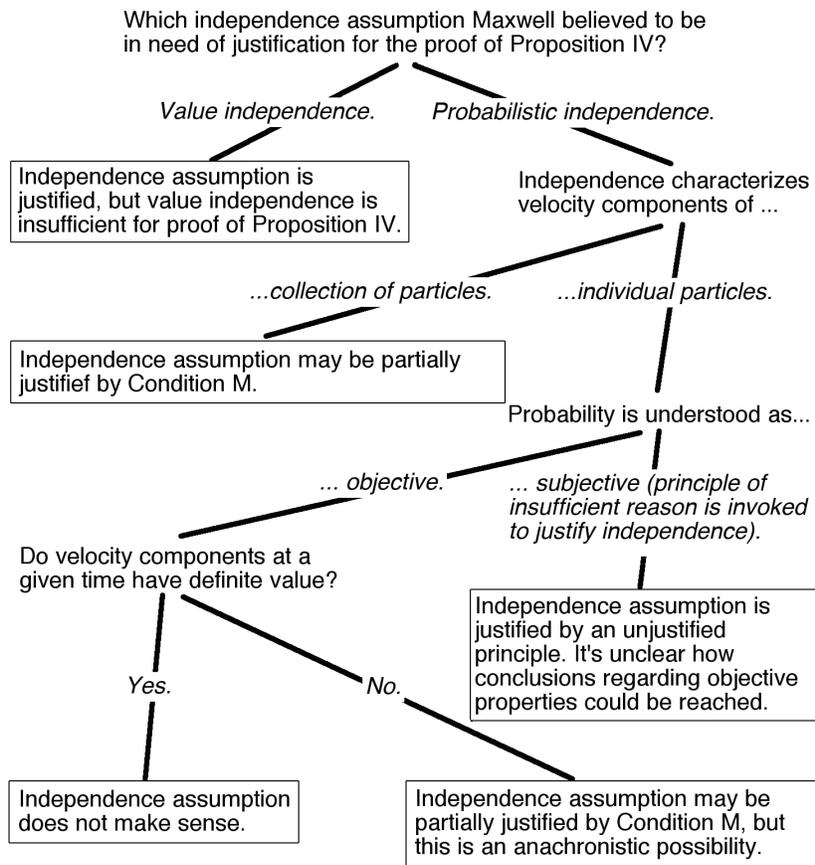
Which independence assumption Maxwell believed to be
in need of justification for the proof of Proposition IV?

Value independence. *Probabilistic independence.*

Independence assumption is
justified, but value independence is
insufficient for proof of Proposition IV.

Independence characterizes
velocity components of ...

...collection of particles. *...individual particles.*

Independence assumption may be partially
justifief by Condition M.

Probability is understood as...

... objective. *... subjective (principle of
insufficient reason is invoked
to justify independence).*

Do velocity components at a
given time have definite value?

Independence assumption is
justified by an unjustified
principle. It's unclear how
conclusions regarding objective
properties could be reached.

Yes. *No.*

Independence assumption
does not make sense.

Independence assumption may be
partially justified by Condition M, but
this is an anachronistic possibility.

Figure 3: *Various ways of understanding the term "independent" in the proof of Proposition
IV.*

marginals. Given that Maxwell himself consistently used the term "independent" in the sense of value independence in his other works prior to 1860 it is then possible that he simply took what Herschel said for granted, and only bothered to point out that value independence of the velocity components in different directions clearly holds. Value independence of course does not imply probabilistic independence, and hence the assumption that would have been truly required for the mathematical proof remained unjustified.

According to another reading Maxwell was merely sloppy stating his assumption in the quoted claim, and what he really meant was that setting any value for the velocity component v_x does not affect the *probability* of v_y and v_z taking their values. This reading seemingly finds support in what he says in his 1867 recollection that the 1860 proof was

> [...] founded on the assumption that the probability of a molecule having a velocity resolved parallel to x lying between given limits is not in any way affected by the knowledge that the molecule has a given velocity resolved parallel to y. (Maxwell, 1867, p. 437)

Aside from the curious allusion to 'knowledge' this recollection correctly states that the proof of Proposition IV requires the assumption of probabilistic independence of v_x and v_y. What the recollection does not make clear, however, is whether Maxwell also intended to rely on the assumption of probabilistic independence (as opposed to the assumption of value independence) *back in 1860*, and so this later reflection of Maxwell in itself can not be taken as evidence for an awareness of a distinction between probabilistic independence and value independence at the time when he produced his first proof. All we know is that in 1867 Maxwell found this reformulated independence requirement sufficiently "precarious" to warrant the construction of a different proof.

In the rest of this section we follow the latter reading, namely that Maxwell intended to have assumed a sort of probabilistic independence. It is still not entirely clear what could have been the intended carrier of the assumption of probabilistic independence according to Maxwell's 1867 recollection (see Figure 3 for some possibilities). Although Maxwell seemingly talks about the velocity components of an individual particle, the velocity components of individual particles have a definite value at any given time, rendering any talk about a particle having a velocity component with some probability obscure (unless we either interpret probability as ignorance – the above mentioned appearance of the term 'knowledge' may indicate that – or unless with a great dose of anachronism we give up definiteness of velocity values along the lines of an objective interpretation of the collision probabilities). It is more in line with the frequentist allures of the 1860 proof to take the probabilistic independence assumption to be a statement about a collection of particles, namely that the distribution of velocities along the y axis for all particles is the same as the distribution of velocities along the y axis for the particles whose velocity component along the x axis is set to some value.

Does the *Illustrations* provide a justification for the assumption of probabilistic independence of velocity components of a collection of particles, or does

at least have the resources to do so? With our reconstruction of Propositions I–III at hand the obvious guess is to trace back the assumption of probabilistic independence to Condition M. Recall that Condition M states that whenever two particles collide the impact parameter ρ is distributed uniformly within a circle C. The 'whenever' part of this assumption presumably states that ρ is distributed uniformly regardless what are the incoming velocities of the pair of colliding particles or what is the configuration of other particles that are not participating in the collision. In a precise formulation this would presumably translate to conditional probabilistic independence of ρ from the incoming velocities of the pair and from other physical quantities characterizing the remaining particles. If so then Condition M already furnishes a sort of probabilistic independence assumption regarding the velocities of a particle pair, and it is not inconceivable that some clever aggregation procedure would then lend credence to the assumption of probabilistic independence of velocity components of a collection of particles.

Maxwell himself does not hint at such argument in the *Illustrations*, but associating Condition M (and its consequence, Condition A) with a sense of probabilistic independence that is required for a derivation of the velocity distribution was so deep that even in 1894, during an ignited debate on the pages of *Nature* about the validity of Boltzmann's H-theorem, Condition A and the Stosszahlansatz (the crucial independence assumption about factorization of the velocity probability distribution that appears in Maxwell's 1867 proof) gets identified and interchanged without justification or warning (for a detailed account see Dias (1994)). If the non-interchangeability of these assumptions did not occur to prominent physicists who had the advantage of many years of developments that followed Maxwell's 1860 proof, it seems reasonable to assume that it might have not occurred to Maxwell either. Hence Maxwell might have thought that Condition M, whose geometrical meaning is at least apparent, and which (as Gyenis (2005) ventures) may be loosely motivated on the basis of Clausius' mean free path approach, justifies the probabilistic independence of v_x and v_y.

It would be interesting to see, but I do not know of a formally satisfying investigation of the relationship of Condition M and the assumption of probabilistic independence of velocity components of a collection of particles. Clearly, any such investigation would require that first someone says exactly what Condition M states.

7.3. (A3) Rotational symmetry of the velocity distribution

Maxwell's justification of the assumption of rotational symmetry of the velocity distribution is physically intuitive: nothing in the argument he gave for the factorization of the velocity distribution depends on the particular choice of the oblique coordinate system, and the physical behavior of the colliding particles should clearly not depend on how we choose to describe their behavior. Requiring that the factorization holds for other oblique coordinate systems entails that the distribution only depends from the distance from the origin. This

result also follows when we only resort to non-dependence from choice of orthogonal coordinate systems, in which case it can be conveniently mathematically expressed as

$$f(v_x)f(v_y)f(v_z) = \phi(v_x^2 + v_y^2 + v_z^2),$$

which is what Maxwell writes. (Note that the orthogonality assumption is not essential, only convenient. With any oblique coordinate system the functional equation (1) still has a unique solution, a generalized normal distribution.)

To sum up, among the three mathematical assumptions of Maxwell's Proposition IV (A1) stationarity was intended to be justified on the basis of prior assumptions about particle collisions, (A2) probabilistic independence of velocity components was either misunderstood as value independence or was intended to be justified on the basis of prior assumptions about particle collisions, and (A3) rotational symmetry was supposed to follow from the non-dependence of the argument for the assumption of probabilistic independence of velocity components from the choice of coordinate system. Maxwell's analysis and assumptions regarding particle collisions hence played a crucial motivational role for the mathematical assumptions of Proposition IV.

Finally note that if the number of particles is finite then the actual distribution of velocities at any time must be discrete and have a bounded support. However the only measurable solution of the functional equation (1), the Gaussian (2), is a continuous function with infinite support. To reconcile this tension one would need to give up at least one of the assumptions: one could insist that stationarity, rotational symmetry, or probabilistic independence only holds approximately, or that Maxwell's velocity distribution f merely provides an approximate description. This is one of a number of issues relating to Proposition IV that we leave here undiscussed as they have been thoroughly addressed in the literature; for an overview see Uffink (2007, pp. 945–946).

8. Conclusions

The mathematical essence of the proof of Proposition IV is that a velocity distribution $\mathbf{f}(\vec{v})$ follows the Maxwell velocity distribution law if it satisfies the mathematical assumptions of probabilistic independence $\mathbf{f}(\vec{v}) = \mathbf{f}_x(v_x)\mathbf{f}_y(v_y)\mathbf{f}_z(v_z)$ and rotational symmetry $\mathbf{f}(\vec{v}) = f(|\vec{v}|)$. The mathematical essence of a proof may however mislead about what in it needs to be endowed with physical meaning and justification. We argued that emphasizing the importance of the implicit assumption of the existence of a stationary velocity distribution $\mathbf{f}(\vec{v})$ make the need for an argument for an approach to equilibrium clear and this in turn may shed light on the role Maxwell's widely ignored molecular collision model played in his *Illustrations of the Dynamical Theory of Gases*. In particular we showed that for perfectly elastic particles Maxwell's Proposition I–III – under the assumptions that their time evolution satisfies the constraint that it can be split into successive time intervals during which each particle has an approximately equal independent chance to participate in at most one, pairwise collision, and that in this interval the impact parameter is approximately equally distributed

on the collision circle (Condition M) – yields a good "bad" mechanical proof of approach to equilibrium temperature, a precursor to Boltzmann's H-theorem, which also lends partial credence to the availability of a proof of reaching a stationary velocity distribution. We argued that Condition M of Proposition II furnishes a sort of probabilistic independence assumption regarding velocities that could have motivated the assumption of probabilistic independence, but we also suggested another potential explanation for the apparent lack of effort on the part of Maxwell to justify probabilistic independence, namely that Maxwell might have conflated the notions of probabilistic independence and value independence. We argued that distinction of meaning of these two notions was not yet available in the 1850's and we presented circumstantial evidence for these claims on the basis of relevant prior works of Clausius, Herschel, and Maxwell.

Section 7 explored the ways how Propositions I–III, Condition M, and the approach to equilibrium furnished by a reconstructed version of Proposition VI could have played a role in Maxwell's physical justification of the three mathematical assumptions of Proposition IV; we have seen that there are multiple interpretative possibilities. In closing let me choose the interpretation which seems to me most plausible. On a physical basis Maxwell knew that after a sufficient amount of time a box of gas settles into a thermal equilibrium, and he made the important realization that supplementing a standard Newtonian collision model with the seemingly innocuous Condition M allows the proof of such an approach to thermal equilibrium if the particles are modeled as perfectly elastic spheres. Maxwell implicitly assumed that the onset of this equilibrium also entails that the velocity distribution function becomes stationary. Thinking that the existence of a stationary velocity distribution can be proved on a physical basis he looked around for mathematical proofs that could be used to characterize properties of such stationary distributions, and he recalled Herschel's characterization of the distribution of errors from reading his *Essays* as such. Unfortunately Herschel mischaracterized the crucial independence assumption of his own proof, and Maxwell implemented Herschel's proof together with this mischaracterization to the context of the kinetic theory. In turn the physical justification of the so-mischaracterized independence assumption (and as a consequence also the assumption of rotational symmetry) seemed straightforward. Maxwell later realized that this independence assumption is "precarious" and since other considerations also led him to abandon Clausius' mean free path approach that lies behind the collision model, he also abandoned the attempt to establish stationarity of the velocity distribution via Condition M. It took several decades until the scientific community came to a clear grasp of the distinctions between various interpretations of probability and various notions of independence, and thus Maxwell's early mistake in his first paper on the kinetic theory was by no means an obvious one.

Acknowledgement

Appendix

Let $\vec{v}$ and $\vec{V}$ be the incoming velocities and $\vec{v}'$ and $\vec{V}'$ be the rebound velocities of two spheres of masses m and M that undergo a perfectly elastic collision. Due to conservation of momentum the velocity of the center of mass is unchanged by the collision:

$$\vec{v}_{CM} \doteq \frac{m\vec{v} + M\vec{V}}{m + M} = \frac{m\vec{v}' + M\vec{V}'}{m + M} \doteq \vec{v}'_{CM}. \tag{13}$$

Let $\vec{r}$, $\vec{R}$ (and $\vec{r}'$, $\vec{R}'$) denote the incoming (and rebound) velocities relative to this center of mass, that is

$$\vec{r} \doteq \vec{v} - \vec{v}_{CM} = \vec{v} - \frac{m\vec{v} + M\vec{V}}{m + M} = \frac{M}{m + M}(\vec{v} - \vec{V}) \tag{14}$$

$$\vec{R} \doteq \vec{V} - \vec{v}_{CM} = \vec{V} - \frac{m\vec{v} + M\vec{V}}{m + M} = \frac{m}{m + M}(\vec{V} - \vec{v}) \tag{15}$$

$$\vec{r}' \doteq \vec{v}' - \vec{v}'_{CM} = \vec{v}' - \vec{v}_{CM} \tag{16}$$

$$\vec{R}' \doteq \vec{V}' - \vec{v}'_{CM} = \vec{V}' - \vec{v}_{CM}, \tag{17}$$

and hence noting that $v \doteq \|\vec{v}\|$, $V \doteq \|\vec{V}\|$, $v' \doteq \|\vec{v}'\|$ etc. we get

$$r^2 = \vec{r} \cdot \vec{r} = \left(\frac{M}{m + M}\right)^2 \left(\vec{v} \cdot \vec{v} + \vec{V} \cdot \vec{V} - 2\vec{v} \cdot \vec{V}\right) \tag{18}$$

$$= \left(\frac{M}{m + M}\right)^2 \left(v^2 + V^2 - 2vV\cos(\alpha)\right), \tag{19}$$

$$R^2 = \vec{R} \cdot \vec{R} = \left(\frac{m}{m + M}\right)^2 \left(\vec{v} \cdot \vec{v} + \vec{V} \cdot \vec{V} - 2\vec{v} \cdot \vec{V}\right) \tag{20}$$

$$= \left(\frac{m}{m + M}\right)^2 \left(v^2 + V^2 - 2vV\cos(\alpha)\right), \tag{21}$$

$$v_{CM}^2 = \vec{v}_{CM} \cdot \vec{v}_{CM} = \frac{m\vec{v} + M\vec{V}}{m + M} \cdot \frac{m\vec{v} + M\vec{V}}{m + M} = \tag{22}$$

$$= \left(\frac{1}{m + M}\right)^2 \left(m^2\vec{v} \cdot \vec{v} + M^2\vec{V} \cdot \vec{V} + 2mM\vec{v} \cdot \vec{V}\right) \tag{23}$$

$$= \left(\frac{1}{m + M}\right)^2 \left(m^2 v^2 + M^2 V^2 + 2mMvV\cos(\alpha)\right) \tag{24}$$

with $\alpha \doteq \angle(\vec{v}, \vec{V})$.

From (16), (17), and from $r' = r$, $R' = R$ (which we get from (5) where we utilized that the collision is perfectly elastic) we get

$$v'^2 = (\vec{v}_{CM} + \vec{r}') \cdot (\vec{v}_{CM} + \vec{r}') = v_{CM}^2 + r'^2 + 2\vec{v}_{CM} \cdot \vec{r}' \tag{25}$$

$$= v_{CM}^2 + r^2 + 2v_{CM}r\cos(\gamma) \tag{26}$$

$$V'^2 = (\vec{v}_{CM} + \vec{R}') \cdot (\vec{v}_{CM} + \vec{R}') = v_{CM}^2 + R'^2 + 2\vec{v}_{CM} \cdot \vec{R}' \tag{27}$$

$$= v_{CM}^2 + R^2 - 2v_{CM}R\cos(\gamma) \tag{28}$$

with $\gamma \doteq \angle(\vec{v}_{CM}, \vec{r}')$ and noting that $\angle(\vec{v}_{CM}, \vec{R}') = 180^o - \angle(\vec{v}_{CM}, \vec{r}')$.

Multiplying (26) with $m/2$, (28) with $M/2$, subtracting them from each other, and making judicious substitutions for r^2 (eq. (19)), R^2 (eq. (21)), and v_{CM}^2 (eq. (24)) yields

$$
\begin{aligned}
\frac{mv'^2}{2} - \frac{MV'^2}{2} \;=\;& \frac{m}{2}\left(v_{CM}^2 + r^2 + 2v_{CM}r\cos(\gamma)\right) \\
& - \frac{M}{2}\left(v_{CM}^2 + R^2 - 2v_{CM}R\cos(\gamma)\right) \\
=\;& \frac{m}{2}\left(\frac{1}{m+M}\right)^2\left(m^2v^2 + M^2V^2 + 2mMvV\cos(\alpha)\right) \\
& + \frac{m}{2}\left(\frac{M}{m+M}\right)^2\left(v^2 + V^2 - 2vV\cos(\alpha)\right) + \frac{m}{2}2v_{CM}r\cos(\gamma) \\
& - \frac{M}{2}\left(\frac{1}{m+M}\right)^2\left(m^2v^2 + M^2V^2 + 2mMvV\cos(\alpha)\right) \\
& - \frac{M}{2}\left(\frac{m}{m+M}\right)^2\left(v^2 + V^2 - 2vV\cos(\alpha)\right) + \frac{M}{2}2v_{CM}R\cos(\gamma) \\
=\;& \frac{1}{2}\left(\frac{1}{m+M}\right)^2\big(m^3v^2 + mM^2V^2 + 2m^2MvV\cos(\alpha) \\
& \qquad\qquad + mM^2v^2 + mM^2V^2 - 2mM^2vV\cos(\alpha) \\
& \qquad\qquad - m^2Mv^2 - M^3V^2 - 2mM^2vV\cos(\alpha) \\
& \qquad\qquad - m^2Mv^2 - m^2MV^2 + 2m^2MvV\cos(\alpha)\big) \\
& + (mr + MR)v_{CM}\cos(\gamma) \\
=\;& + \frac{1}{2}\left(\frac{1}{m+M}\right)^2\big(mv^2(m^2 + M^2 - 2mM) \\
& \qquad\qquad - MV^2(m^2 + M^2 - 2mM)\big) \\
& + 2\left(\frac{1}{m+M}\right)^2(m-M)mMvV\cos(\alpha) \\
& + (mr + MR)v_{CM}\cos(\gamma) \\
=\;& C_0 \cdot \left(\frac{mv^2}{2} - \frac{MV^2}{2}\right) + C_1 \cdot \cos(\alpha) + C_2 \cdot \cos(\gamma)
\end{aligned}
$$

with

$$
\begin{aligned}
C_0 &\doteq \left(\frac{m-M}{m+M}\right)^2, & (29)\\[2mm]
C_1 &\doteq 2\frac{(m-M)}{(m+M)^2}mMvV, & (30)\\[2mm]
C_2 &\doteq (mr + MR)v_{CM}, & (31)
\end{aligned}
$$

which proves equation (7). Clearly, $0 \le C_0 < 1$.

References

Boltzmann, L. (1868). Studien über das Gleichgewicht der lebendigen Kraft zwischen bewegten. *Wiener Berichte*, *58*, 517–560.

Brush, S. G. (1958). The Development of Kinetic Theory of Gases: IV Maxwell. *Annals of Science*, *14*, 243–255.

Brush, S. G. (1971). James Clerk Maxwell and the kinetic theory of gases: A review based on recent historical studies. *American Journal of Physics*, *31*, 631–640.

Brush, S. G. (1976). *The Kind of Motion we call Heat* volume VI of *Studies in Statistical Mechanics*. Amsterdam, New York, Oxford: North-Holland Publishing Company.

Brush, S. G. (1983). *Statistical Physics and the Atomic Theory of Matter From Boyle and Newton to Landau and Onsager*. Princeton Series in Physics. Princeton, New Jersey: Princeton University Pres.

Brush, S. G. (2003). *The Kinetic Theory of Gases: Anthology of Historical Papers with Historical Commentary*. London: Imperial College Press.

Brush, S. G., Everitt, C. W. F., & Garber, E. (1986a). Kinetic theory and the properties of gases: Maxwell's work in its nineteenth-century context. In Brush et al. (1986b).

Brush, S. G., Everitt, C. W. F., & Garber, E. (Eds.) (1986b). *Maxwell on Molecules and Gases*. Cambridge, Massachusetts: The MIT Press.

Burbury, S. H. (1890). On some problems in the kinetic theory of gases. *Philosophical Magazine*, *30*, 298–317.

Burbury, S. H. (1894). Boltzmann's minimum function. *Nature*, *51*, 78–79.

Buys-Ballot, C. H. D. (1858). Ueber die Art von Bewegung, welche wir Wärme und Elektrizität nennen. *Annalen der Physik*, *103*, 240–259.

Campbell, L., & Garnett, W. (1882). *The Life of James Clerk Maxwell*. London: Macmillan and Co.

Clausius, R. (1857a). The nature of motion which we call heat. *Philosophical Magazine*, *14*, 108–127. English translation of his 1857b. See also in Brush (2003), 111–134.

Clausius, R. (1857b). Ueber die Art der Bewegung, welche wir Wärme nennen. *Annalen der Physik*, *100*, 353–380.

Clausius, R. (1858). Ueber die Mittlere Länge der Were, welche bei Molecularbewegung gasförmigen Körper von den einzelnen Molecülen zurückgelegt werden, nebst einigen anderen Bemerkungen über die mechanischen Wärmetheorie. *Annalen der Physik*, *105*, 239–258.

Clausius, R. (1859). On the mean lengths of the paths described by the separate molecules of gaseous bodies. *Philosophical Magazine*, *17*, 81–01. English translation of his 1858. See also in Brush (2003), 135–147.

Clausius, R. (1862). On the conduction of heat by gases. *Philosophical Magazine*, *23*, 416–435, 512–535.

Dias, P. M. C. (1994). "Will someone say exactly what the H-theorem proves?" A study of Burbury's Condition A and Maxwell's Proposition II. *Archive for History of Exact Sciences*, *46*, 341–366.

Garber, E. W. (1970). Clausius and Maxwell's kinetic theory of gases. *Hist. Studies Phys. Sci.*, *2*, 299–319.

Gillispie, C. C. (1963). Intellectual factors in the background of analysis by probabilities. In A. C. Crombie (Ed.), *Scientific Change* (pp. 431–453). New York: Basic Books.

Gyenis, B. (2005). Maxwell and the normal distribution. M.Sc. research paper, Department of History and Philosophy of Science, University of Pittsburgh.

Heimann, P. M. (1970). Molecular forces, statistical representation and maxwell's demon. *Hist. Phil. Sci.*, *1*, 189–211.

Herschel, J. F. W. (1850). Quetelet on probabilities. *Edinburgh Rev.*, *92*, 1–57.

Herschel, J. F. W. (1857). *Essays from the "Edinburgh" and "Quarterly" Reviews*. London: Layton.

Maxwell, J. C. (1855). On the theory of colours in relation to colour-blindness. a letter to dr G. Wilson. In W. D. Niven (Ed.), *The Scientific Papers of James Clerk Maxwell* (pp. 119–125). New York: Dover Publications, Inc. volume I. (1965th ed.). Appeared in the *Transactions of the Royal Society of Arts*, Vol. IV. Part III.

Maxwell, J. C. (1859). Letter from Maxwell to George Gabriel Stokes, May 30, 1859. In Brush et al. (1986b).

Maxwell, J. C. (1860). Illustrations of the dynamical theory of gases. *Philosophical Magazine*, *19, 20*, 19–32, 21–37. Quotations from Brush et al. (1986b), 286–318.

Maxwell, J. C. (1863). On the conduction of heat in gases. In Brush et al. (1986b). Unpublished manuscript in which Maxwell tried to correct some of the omissions and errors of his 1860 paper.

Maxwell, J. C. (1867). On the dynamical theory of gases. *Proceedings of the Royal Society of London*, *15*, 167–171. Also appeared in *Philosophical Magazine*, 32: 129–145 (1866), 35: 129–145, 185–217 (1868). Quotations from Brush et al. (1986b), 420–472.

Porter, T. (1981). A statistical survey of gases: Maxwell's social physics. *Historical Studies in the Physical Sciences*, *12*, 77–116.

Truesdell, C. (1975). Early kinetic theories of gases. *Arch. Hist. Exact Sci.*, *15*, 1–66.

Uffink, J. (2007). Compendium of the foundations of classical statistical physics. In J. E. J. Butterfield (Ed.), *Philosophy of Physics, Part B* Handbook of the Philosophy of Science. The Netherlands: Elsevier.