

Common fallacies in the derivation of Boltzmann's equation

This article has been downloaded from IOPscience. Please scroll down to see the full text article.

1973 J. Phys. A: Math. Nucl. Gen. 6 1800

(<http://iopscience.iop.org/0301-0015/6/11/015>)

View [the table of contents for this issue](#), or go to the [journal homepage](#) for more

Download details:

IP Address: 171.66.16.73

The article was downloaded on 02/06/2010 at 04:41

Please note that [terms and conditions apply](#).

Common fallacies in the derivation of Boltzmann's equation

Erling Dahlberg

Royal Institute of Technology, Department of Plasma Physics, S-100 44 Stockholm 70, Sweden

Received 5 April 1973, in final form 8 June 1973

Abstract. The failure to distinguish between valid proofs and false or inconclusive assertions in standard presentations of Boltzmann's equation is noted. For example, it is false to assert that $|d\gamma| = |d\gamma'|$, where $d\gamma$ is the solid-angle element of the scattering direction, and impossible to define an impact azimuth ϵ so that $|d\epsilon| = |d\epsilon'|$.

When tested by logic, the standard derivations of Boltzmann's collision term in most modern textbooks are seen to be false or at least inconclusive. On the other hand, the proofs given in the early texts are quite satisfactory (eg Boltzmann 1910, 1912, Jeans 1921, Kennard 1938, Chapman and Cowling 1939). If the fallacies become widely known, we may hope that future students will again be offered presentations that are not in principle incomprehensible. A full account of some common malpractices and suggested remedies is available elsewhere (Dahlberg 1972); the aim of the present note is to create a reasonable doubt in the minds of scientists who are already familiar with the Boltzmann equation. Our hope is that this may help to reduce further spreading of the prevalent acceptance of such pseudo-proofs in kinetic theory.

It is sufficient for our purpose to consider the so called heuristic derivation for classical, elastic, short-range encounters between spherically symmetric particles of a single species. We shall use a reference frame with fixed axis directions for specifying particle velocities; further we shall employ primes to denote the post-collision (final) counterpart of an unprimed pre-collision (initial) quantity. Thus c_1 and c'_1 denote the initial and final velocities of particle 1 and analogously c_2 and c'_2 for particle 2. A collision is completely specified by eight scalar quantities, eg (c_1, c_2, γ') or (c'_1, c'_2, γ) , where

$$\gamma' = (c'_2 - c'_1)/|c'_2 - c'_1| \quad (1)$$

$$\gamma = (c_2 - c_1)/|c_2 - c_1| \quad (2)$$

are unit vectors representing the direction of the respective relative velocities.

To obtain the collision term it is necessary to establish that the transformation from (c_1, c_2, γ') to (c'_1, c'_2, γ) has a unit jacobian, ie

$$|dc_1 dc_2 d\gamma'| = |dc'_1 dc'_2 d\gamma|. \quad (3)$$

This result follows from the laws of mechanics; however, for many modern authors it seems to have been a stumbling block. In this connection we should note Boltzmann's strong warning (Boltzmann 1912, § 27) that an equation like (3) "has no meaning whatsoever except when (*hat überhaupt nur eine Bedeutung, wenn*) it refers to" a change of variables in a definite integral over all the variables c_1, c_2, γ' .

The most common error is to base equation (3) on two separate assertions:

$$|dc_1 dc_2| = |dc'_1 dc'_2| \quad (4)$$

and

$$|d\gamma'| = |d\gamma| \quad (5)$$

which are both obviously false. If relations like these were possible we should be able to transform an integral

$$\iiint f dc_1 dc_2 d\gamma'$$

either at constant (c_1, c_2) into

$$\iint dc_1 dc_2 \int f d\gamma$$

which is clearly impossible since γ is completely determined by (c_1, c_2) (cf equation (2)); or at constant γ' into

$$\int d\gamma' \iint f dc'_1 dc'_2$$

which is equally impossible since (c'_1, c'_2) are not sufficient to specify (c_1, c_2) for a given γ' .

Some authors prefer to specify the collision using the 'impact parameter' b and some—usually not well defined—impact azimuth ϵ presumably measured in a frame determined by the initial relative velocity. Although it is true that the impact parameter is the same for a collision and its so called inverse, it is easy to show that *no definition of ϵ can exist for which $|d\epsilon| = |d\epsilon'|$* . Here ϵ' denotes the impact azimuth of the inverse collision. To see this we need only consider a transition from (nearly) glancing collisions for which one must have $d\epsilon \simeq +d\epsilon'$ to (nearly) head-on collisions for which clearly $d\epsilon \simeq -d\epsilon'$.

Yet another group of authors use a different and deceptively simple argument. They first consider the nine-dimensional transformation from (c_1, c_2, g') to (c'_1, c'_2, g) , where $g = g\gamma = c_2 - c_1$ is the relative velocity. Here it is easily established that

$$|dc_1 dc_2 d\gamma'(g')^2 dg'| = |dc'_1 dc'_2 d\gamma g^2 dg| \quad (6)$$

and, further, conservation of energy implies $g' = g$ or

$$(g')^2 dg' = g^2 dg. \quad (7)$$

Thus, it is concluded that equation (3) must follow. Presumably, the authors—and certainly most readers—see nothing more in this than a simple cancellation. However, a full justification (see appendix) requires lengthy mathematics and depends on the special properties of the transformation from (c_1, c_2, g') to (c'_1, c'_2, g) . Since this would presumably require more space and give less physical insight than a correct and more traditional proof, we see no reason for employing this technique in the presentation of the Boltzmann equation.

It may now seem natural to wonder why the above-described pseudo-proofs have come to prevail over the correct approach employed originally. In that approach the particles were mostly thought of as hard spheres or 'billiard balls' and the line of impact or apse-line direction α' was a natural choice for completing the collision specification. There was then no obvious need to *stress* the additional and special advantage of this parameter, namely, that it is conserved (or rather exactly reversed) when we go from a

collision to its inverse, ie

$$|d\alpha| = |d\alpha'|. \tag{8}$$

This special property can be used to show that

$$|(dc'_1 dc'_2)_\alpha| = |(dc_1 dc_2)_\alpha| \tag{9}$$

where the subscript α is meant as an—important but usually omitted—indication that the relation (9) refers to a transformation at constant apse-line direction. One may now conjecture that the originator(s) of false assertions like equations (4) and (5) held the scattering direction to be a more ‘physical’ parameter than the ‘line of impact’. It may then have been tempting to take it for granted that, whatever specifying parameter be used, relations analogous to equations (8) and (9) should obtain. The fact that later authors have so often chosen to accept and employ the false assertions is understandable but scientifically regrettable.

In derivations of the left member of Boltzmann’s equation there sometimes occurs a related obscurity. Such derivations are commonly made by considering—in the absence of collisions—the change (in a time δt) in the number of particles either in a fixed phase-space element $dc_0 dr_0$ or in an element following the phase-space motion from c_0, r_0 to c_t, r_t . Since it is in fact necessary to impose restrictions on the velocity dependence of the forces, the *necessity*—and not a mere statement—of such restrictions must be manifest in the argument if the derivation shall be convincing. This is not the case if it is simply taken as obvious that

$$dc_t dr_t = dc_0 dr_0 \tag{10}$$

because it is in proving this that the need for restrictions comes in. Even less recommendable does it seem to base the argument on a statement that $dc_t = dc_0$ and $dr_t = dr_0$ as two separate assertions. Since $c_t = c_t(c_0, r_0)$ and $r_t = r_t(c_0, r_0)$, it is not even well defined what such a statement should mean. Clearly, it cannot be given the simple interpretation—analogue to equation (9) above—that

$$\begin{aligned} (dc_t)_r &= (dc_0)_r \\ (dr_t)_c &= (dr_0)_c. \end{aligned} \tag{11}$$

Although equations (11) can be shown to be true, they are not sufficient for deriving equation (10) as can be seen from the simple counter example

$$\begin{aligned} c_t &= c_0 + r_0 \\ r_t &= -c_0 + r_0 \end{aligned} \tag{12}$$

for which equations (11) would obtain but which yields $dc_t dr_t = 8 dc_0 dr_0$.

In conclusion, it may also be worth noting that—although this is seldom emphasized—Boltzmann’s equation and its derivation must be modified if the velocities are referred to position-dependent axis directions, eg $c = v_1\hat{r} + v_2\hat{\theta} + v_3\hat{\phi}$ or $c = w_1\hat{R} + w_2\hat{\phi} + w_3\hat{z}$ using spherical or cylindrical coordinates, as might seem natural for some problem geometries. (Alternatively, one could of course employ the corresponding generalized momenta.) Putting $f(t, r, c) = f_{sph}(t, r, v) = f_{cyl}(t, r, w)$ to define the distribution function as a function of time, space and the respective velocity components, it is easily seen

that the collision term is not affected and that we still have

$$d\mathbf{v}_t d\mathbf{r}_t = dv_0 d\mathbf{r}_0 \tag{13}$$

$$d\mathbf{w}_t d\mathbf{r}_t = d\mathbf{w}_0 d\mathbf{r}_0 \tag{14}$$

although, of course, relations analogous to (11) no longer obtain. The expression for the substantial derivative, Df/Dt , ie the left member of Boltzmann's equation, will however be formally different. For the examples above we thus find (on dropping the subscripts sph and cyl and extending the conventional notation, so that eg $\mathbf{G} \cdot (\partial f/\partial \mathbf{v})_{\mathbf{r}} = \Sigma G_i (\partial f_{\text{sph}}/\partial v_i)_{v_j, v_k, \mathbf{r}}$ etc)

$$\frac{Df}{Dt} = \frac{\partial f}{\partial t} + \mathbf{v} \cdot \left(\frac{\partial f}{\partial \mathbf{r}} \right)_{\mathbf{v}} + \mathbf{G} \cdot \left(\frac{\partial f}{\partial \mathbf{v}} \right)_{\mathbf{r}} = \frac{\partial f}{\partial t} + \mathbf{w} \cdot \left(\frac{\partial f}{\partial \mathbf{r}} \right)_{\mathbf{w}} + \mathbf{H} \cdot \left(\frac{\partial f}{\partial \mathbf{w}} \right)_{\mathbf{r}} \tag{15}$$

where

$$\mathbf{G} = \left(\frac{F_r}{m} + \frac{v_2^2 + v_3^2}{r} \right) \hat{\mathbf{r}} + \left(\frac{F_\theta}{m} - \frac{v_1 v_2}{r} + \frac{v_3^2}{r} \cot \theta \right) \hat{\boldsymbol{\theta}} + \left(\frac{F_\phi}{m} - \frac{v_1 v_3}{r} - \frac{v_2 v_3}{r} \cot \theta \right) \hat{\boldsymbol{\phi}} \tag{16}$$

$$\mathbf{H} = \left(\frac{F_R}{m} + \frac{w_2^2}{R} \right) \hat{\mathbf{R}} + \left(\frac{F_\phi}{m} - \frac{w_1 w_2}{R} \right) \hat{\boldsymbol{\phi}} + \frac{F_z}{m} \hat{\mathbf{z}}. \tag{17}$$

Acknowledgment

The author is deeply grateful to Professor T G Cowling of Leeds University for his encouragement and valuable comments.

Appendix. On the handling of jacobians

First, let us recall a well known and often used fact. Consider the transformation A :

$$\mathbf{x} = A\mathbf{u} \tag{A.1}$$

where \mathbf{x} denotes an n -vector etc, and its jacobian

$$\partial \mathbf{x} / \partial \mathbf{u} = J_A \mathbf{u}$$

and assume that A equals its own inverse,

$$A = A^{-1}. \tag{A.2}$$

Then the jacobian of the inverse must also be the same function

$$\partial \mathbf{u} / \partial \mathbf{x} \equiv J_{(A^{-1})}(\mathbf{x}) = J_A(\mathbf{x}). \tag{A.3}$$

Further, one has quite generally

$$\partial \mathbf{x} / \partial \mathbf{u} = 1 / (\partial \mathbf{u} / \partial \mathbf{x}). \tag{A.4}$$

Before drawing any conclusions we should note that equations (A.3) and (A.4) are relations of two different kinds, and that a statement that the two jacobians are equal may be misleading. Equation (A.3) indicates the equality of two functions, or

$$(\partial \mathbf{x} / \partial \mathbf{u})|_{\mathbf{u}=\mathbf{a}} = (\partial \mathbf{u} / \partial \mathbf{x})|_{\mathbf{x}=\mathbf{a}}, \quad \forall \mathbf{a} \tag{A.3a}$$

whereas equation (A.4) represents the equality of the values of two functions for *corresponding* arguments

$$(\partial \mathbf{x} / \partial \mathbf{u})|_{\mathbf{u}=\mathbf{a}} = \{1 / (\partial \mathbf{u} / \partial \mathbf{x})\}|_{\mathbf{x}=A\mathbf{u}=A\mathbf{a}}. \quad (\text{A.4a})$$

However, for the important special case that A is a *linear* transformation, so that $\partial \mathbf{x} / \partial \mathbf{u} = c = \text{constant}$, equation (A.3a) would be valid also for corresponding arguments, and we then obtain

$$\left(\frac{\partial \mathbf{x}}{\partial \mathbf{u}}\right)^2 = \left(\frac{\partial \mathbf{u}}{\partial \mathbf{x}}\right)^2 = c^2 = 1, \quad \text{if } A = A^{-1} \text{ and linear.} \quad (\text{A.5})$$

The importance of the linearity of A should be obvious, the omission of any mention thereof is, however, quite common in similar arguments in the literature.

Second, consider the transformation

$$\begin{pmatrix} \mathbf{x} \\ y \end{pmatrix} = B \begin{pmatrix} \mathbf{u} \\ v \end{pmatrix} = \begin{pmatrix} X(\mathbf{u}, v) \\ Y(\mathbf{u}, v) \end{pmatrix} \quad (\text{A.6})$$

with the inverse

$$\begin{pmatrix} \mathbf{u} \\ v \end{pmatrix} = B^{-1} \begin{pmatrix} \mathbf{x} \\ y \end{pmatrix} = \begin{pmatrix} U(\mathbf{x}, y) \\ V(\mathbf{x}, y) \end{pmatrix} \quad (\text{A.7})$$

Assume also that what we are really interested in is the restriction of B to some hypersurface, which may be given by

$$y = f(v). \quad (\text{A.8})$$

Introducing the symbol C for the restricted transformation, we have

$$\mathbf{x} = C\mathbf{u}, \quad \mathbf{u} = C^{-1}\mathbf{x}. \quad (\text{A.9})$$

(The transformations B and C could, for instance, represent relations between pre- and post-collision variables for a binary encounter together with necessary specifiers, and the restriction (A.8) could correspond to a prescribed energy loss.) It may now happen that a determination of the jacobian, $J_B \equiv \partial(\mathbf{x}, y) / \partial(\mathbf{u}, v)$, appears to be much simpler than a direct calculation of the jacobian $J_C \equiv \partial \mathbf{x} / \partial \mathbf{u}$ of (A.9). Thus we have

$$d\mathbf{x} dy \equiv dx_1 dx_2 \dots dx_n dy = |J_B| d\mathbf{u} dv \quad (\text{A.10})$$

$$dy = f' dv. \quad (\text{A.11})$$

We may now ask whether one can make a 'cancellation' here, or if there are some non-trivial conditions that must be imposed before we can draw the conclusion from (A.10) and (A.11) that

$$d\mathbf{x} = |J_B / f'| d\mathbf{u} \quad (?) \quad (\text{A.12})$$

We recall here Boltzmann's warning, referred to in the text; we may also note the commendable practice of Whittaker and Watson (1927) who write $(d\mathbf{x} dy)$ instead of $d\mathbf{x} dy$ as an indication that the multidimensional differential element cannot always be thought of as a simple product of differentials.

Let us first consider a two-dimensional example:

$$\begin{pmatrix} x \\ y \end{pmatrix} = B \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} X(u, v) \\ Y(u, v) \end{pmatrix}, \quad \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} U(x, y) \\ V(x, y) \end{pmatrix} \quad (\text{A.13})$$

with the restriction

$$y = f(v). \tag{A.14}$$

Thus, we shall carry out a change of variables in the integral

$$I_1 = \int_D F(x, y) dx \tag{A.15}$$

where y , of course, is to be determined by

$$y = f(V(x, y)). \tag{A.14a}$$

(We shall tacitly assume that everything is sufficiently well behaved so that we need not discuss any exceptional cases.) The restriction (A.14) is also seen to imply

$$Y(u, v) - f(v) = 0 \tag{A.14b}$$

so that

$$Y_u du + (Y_v - f') dv = 0$$

where $Y_u \equiv (\partial Y / \partial u)_v$ etc. Consequently

$$\begin{aligned} \frac{dx}{du} &= X_u + X_v \frac{dv}{du} = X_u - \frac{X_v Y_u}{Y_v - f'} = \frac{(X_u Y_v - X_v Y_u - X_u f')}{Y_v - f'} \\ &= \frac{J_B - X_u f'}{Y_v - f'} = \frac{-J_B}{f'} \frac{(1 - f' X_u / J_B)}{(1 - Y_v / f')} = \frac{-J_B}{f'} \frac{(1 - V_y f')}{(1 - Y_v / f')} \end{aligned}$$

where the last equality follows from the well known inversion formula

$$V_y = X_u / (X_u Y_v - X_v Y_u) = X_u / J_B. \tag{A.16}$$

Thus, instead of (12), we obtain

$$(dx) = \left| \frac{J_B}{f'} \right| \left| \frac{1 - V_y f'}{1 - Y_v / f'} \right| (du) \tag{A.17}$$

or more explicitly

$$dx = \left| \frac{\partial(x, y)}{\partial(u, v)} \frac{dv}{dy} \frac{1 - (\partial v / \partial y)_x dy / dv}{1 - (\partial y / \partial v)_u dv / dy} \right| du. \tag{A.17a}$$

We see that the simple cancellation leading to equation (A.12) will produce a correct result *only under very special conditions*.

Next we should try to find an n -dimensional counterpart of the one-dimensional formulae (A.16)–(A.17a). To do this for the transformations (A.6)–(A.9), we shall consider the following integral:

$$I = \int_x \int_y F \delta(y - f(V)) |1 - f'(V) V_y| dy dx. \tag{A.18}$$

Using the properties of the δ function we find on one hand putting $z = y - f(V)$,

$$I \int_x \int_y F \delta(z) |\partial z / \partial y| dy dx = \int_x F(x, y) dx \tag{A.18a}$$

ie the integral corresponding to (A.15), and on the other hand

$$I = \int_{\mathbf{u}} \int_v F(\mathbf{X}, Y) \delta(Y - f(v)) |(1 - f'(v)V_y)J_B| dv d\mathbf{u} = \int_{\mathbf{u}} F(\mathbf{X}, Y) \left| \frac{(1 - V_y f') J_B}{(1 - Y_v/f') f'} \right| d\mathbf{u} \quad (\text{A.18b})$$

where we have again assumed that everything is well behaved in the range considered. Thus we find that the n -dimensional result is completely analogous to the one-dimensional equation (A.17). For reference we give the explicit formulae

$$(d\mathbf{x}) = \left| \frac{J_B (1 - V_y f')}{f' (1 - Y_v/f')} \right| (d\mathbf{u}) \quad (\text{A.19})$$

$$(d\mathbf{x}) = \left| \frac{\partial(\mathbf{x}, y)}{\partial(\mathbf{u}, v)} \frac{dv}{dy} \frac{1 - (\partial v/\partial y)_x dy/dv}{1 - (\partial y/\partial v)_u dv/dy} \right| (d\mathbf{u}). \quad (\text{A.19a})$$

We may also note the n -dimensional analogue of equation (A.16), which may be useful when the inverse (A.7) of the transformation (A.6) is not known explicitly

$$\left(\frac{\partial v}{\partial y} \right)_x = \frac{\partial \mathbf{x} / \partial \mathbf{u}}{\partial(\mathbf{x}, y) / \partial(\mathbf{u}, v)}. \quad (\text{A.20})$$

References

- Boltzmann L 1910 *Vorlesungen über Gastheorie* vol 1, 2nd edn (Leipzig: Barth)
 ——— 1912 *Vorlesungen über Gastheorie* vol 2 2nd edn (Leipzig: Barth)
 Chapman S and Cowling T G 1939 *The Mathematical Theory of Non-Uniform Gases* (Cambridge: Cambridge University Press)
 Dahlberg E 1972 *Royal Institute of Technology, Stockholm, Sweden* Report TRITA-EPP-72-25
 Jeans J H 1921 *The Dynamical Theory of Gases*, 3rd edn (Cambridge: Cambridge University Press)
 Kennard E H 1938 *Kinetic Theory of Gases* (New York: McGraw-Hill)