On the Elementary Explanation of Diffusion Phenomena in Gases

W. H. Furry

Harvard University, Cambridge 38, Massachusetts

The mathematical theory of transport phenomena in gases, including the phenomena of diffusion, was put on a satisfactory foundation by the development of Enskog's method for the solution of Boltzmann's equation. A discussion of the method and of most of the calculations made with it is given in the treatise published several years ago by Chapman and Cowling. Several chapters are devoted to the exposition of the formulas and numerical results obtained from the calculations, and to comparisons with experimental data. The book seems well suited to meet the needs of research workers dealing with transport phenomena.

Despite the satisfactory state of the mathematical theory and of its exposition, there remains a need for simple physical interpretations of these processes. Many, and probably most, scientists are inclined to feel that while the precise mathematical theory of a phenomenon has its place in the final working out of any particular problem, a simple conception of the physical nature of the processes involved is of much greater usefulness in developing new ideas, choosing methods, and designing apparatus. For teaching purposes, elementary explanations are obviously indispensable.

In the case of transport phenomena, the present standard answer to the call for a physical explanation is given by arguments based on the concept of mean free path. As applied to viscosity and thermal conductivity, these arguments succeed in accounting for the main experimental facts in quite a satisfactory fashion. For diffusion phenomena, however, treatments based on the concept of mean free path lead to results that are in distinct qualitative disagreement with experiment, and also, of course, with the results of the detailed mathematical theory. The fact that such arguments do not provide a correct physical approach to diffusion problems has been particularly emphasized by Chapman.

Elementary treatments based on the idea of the mean free path continue, nevertheless, to enjoy widespread use, and in most presentations no indication whatever is given that the results are less satisfactory in the case of diffusion than they are for the other transport phenomena. The need for a better elementary discussion of diffusion problems has in recent years been felt mainly through the desire for a physical explanation of thermal diffusion. Repeated attempts at an explanation in terms of mean free path have been made, the most nearly successful being that of Fürth.

The demand for a simple physical explanation

---

1 D. Enskog, Dissertation (Upsala, 1917).
3 S. Chapman, Phil. Mag. 5, 630 (1928).
4 Prominent exceptions are Knudsen's Kinetic theory of gases, Herzfeld's Kinetische Theorie der Wärme, and Jeans' Dynamical theory of gases. Jeans mentions the specific cases He−A and H−CO to which we refer in Sec. 1.
of thermal diffusion receives a satisfactory answer through a method devised by Frankel and stated by him only very briefly. Examination at somewhat greater length shows that it certainly gives an appropriate physical approach to the problem, and provides results as correct and complete in every respect as could reasonably be expected of an elementary physical discussion. By contrast, Fürth's treatment, despite its skillful presentation and partial success, would seem to be useful mainly in showing the inappropriateness of the concept of mean free path as an approach to this problem.

Frankel's method has been used by Cacciapuoti in a detailed treatment of thermal diffusion for the case of rigid spherical molecules. In the present paper we wish not only to give a more general discussion of the application to thermal diffusion, but also to show that Frankel's approach is the really suitable one for the elementary explanation of ordinary diffusion as well.

The idea used by Frankel is essentially an improved form of a suggestion made long ago by Stefan, who used it to give a detailed and quite satisfactory treatment of diffusion for the case of rigid spherical molecules. This calculation was outstandingly good compared to other approximate work on the problem, but for some reason was rather generally ignored. Stefan's basic physical hypothesis seems to have had the approval of Maxwell, and to have been used by him in obtaining certain results. Maxwell did not, however, give any clear explanation of his way of using the idea. It received mention in the subsequent literature only occasionally, and then usually only as a special assumption which might be taken into account as modifying the results obtained by the mean-free-path approach, not as an independent basis for a direct physical approach to diffusion problems. The revival of this independent method had to wait until the rather recent work of Frankel.

We shall give a clear formulation of the Stefan-Frankel method and apply it to two fundamental problems in connection with which the traditional elementary treatments are decidedly unsatisfactory. In Sec. 1 we shall discuss in greater detail some of the difficulties of the mean-free-path treatment of diffusion. The general formulation of the method of Stefan and Frankel is given in Sec. 2. In Sec. 3 we apply this method to the problem of ordinary diffusion in a binary mixture of gases, with particular regard to the question of the dependence of diffusion coefficient on concentration, which was given altogether incorrectly by the traditional elementary treatment. In Sec. 4 we apply the method to the problem of thermal diffusion.

1. CRITICISM OF THE MEAN-FREE-PATH METHOD

General Nature of the Method

The elementary treatment of transport phenomena based on the concept of mean free path begins by considering all the molecules that cross a certain element of area per unit time. The transport of the quantity in question across this area is then calculated on the assumption that its value for each molecule is determined by the position at which the molecule suffered its last collision, the average coordinates of this position being determined by using the idea of mean free path.

For the cases of transport of momentum (viscosity) and of energy (thermal conductivity) there is no reason, either in principle or in the results obtained, for refusing to accept this as a thoroughly appropriate simple physical approach to these problems. In the case of diffusion (transport of molecular types themselves), however, the method is neither really plausible nor even tolerably successful. We shall return later to the discussion of its lack of plausibility, which provides arguments favoring the point of view of Stefan and Frankel. Its lack of success will be illustrated at once.

Failure to Agree with Experiment

The most striking evidence of the failure of the traditional elementary method to give results
agreeing with experimental facts is obtained by considering the dependence of the diffusion coefficient on concentration. If in a binary mixture only a small fraction of the molecules are of type 1, the result obtained for the diffusion coefficient is

\[ D_{12} = A \beta_2 \lambda_1. \] (1 rare)

Here \( \beta \) is the mean speed and \( \lambda_1 \) the mean free path of molecules of type 1; the constant \( A \) is given the value \( \frac{1}{8} \) by the usual simple argument, but was given the value \( \pi/8 \) by Meyer on the basis of a more pretentious calculation. The occurrence of \( \beta_1 \) and \( \lambda_1 \) only corresponds to a physical picture in which a few molecules of type 1 diffuse through the preponderant type 2 gas. If, on the other hand, molecules of type 2 are rare and type 1 is preponderant, the result is

\[ D_{12} = A \beta_1 \lambda_2. \] (2 rare)

For the general case, an argument given by Meyer and repeated almost universally gives

\[ D_{12} = A (\gamma_1 \beta_1 \lambda_1 + \gamma_2 \beta_2 \lambda_2). \] (3)

Here \( \gamma_1 \) and \( \gamma_2 \) are the fractional numerical concentrations:

\[ \gamma_i = n_i / n = n_i / (n_1 + n_2), \] \( \gamma_1 + \gamma_2 = 1, \) (4)

\( n_i \) being the number of molecules of type \( i \) per unit volume. The dependence of \( D_{12} \) on \( \gamma_1 \) and \( \gamma_2 \) is rather more complicated than is shown explicitly in Eq. (3), because \( \lambda_1 \) and \( \lambda_2 \) themselves depend on concentration.

Without going into the details of the dependence on concentration, we can see at once that for decidedly dissimilar molecules Eqs. (1) to (3) predict that the value of \( D_{12} \) will change rapidly with changes of concentration. This is true especially if the two types of molecules differ greatly in mass. According to the usual ways of calculating mean free path, one finds that:

\[ \langle \lambda_1 \rangle_{1 \text{ rare}} = (m_2 / m_1)^{1/2}, \] \[ \langle \lambda_2 \rangle_{2 \text{ rare}} = (m_2 / m_1)^{1/4}. \] (5)

Since, of course,

\[ \beta_1 / \beta_2 = (m_2 / m_1)^{1/4}, \] (6)

we have, from Eqs. (1) and (2),

\[ (D_{12})_{1 \text{ rare}} = m_2, \] \[ (D_{12})_{2 \text{ rare}} = m_1. \] (7)

Thus, for instance, the value of \( D_{12} \) for the diffusion of a small admixture of hydrogen through carbon dioxide should be 22 times as large as its value for the case of a small admixture of carbon dioxide diffusing through hydrogen. For the case of helium and argon the corresponding ratio should be ten. Such predictions are flatly contradicted by experiment. For example, in both of the rather extreme cases mentioned the total change of \( D_{12} \) with concentration is not more than 15 percent.

This pronounced difficulty with the results of the mean-free-path approach can be alleviated somewhat by resort to considerations based on "persistence of velocity," which effectively make the ratio of mean free paths different from that indicated in Eq. (5). The calculations required, however, are long and complicated, and for cases such as those mentioned there is no way of seeing without considerable labor that the correction would be just enough to make \( D_{12} \) essentially independent of concentration. The partial success of such attempts at "patching up" the results cannot prevent the judgment that the basic approach used is not at all well adapted to the physical nature of the process. It can scarcely be doubted that the physicists of 70 years ago would have discarded the method of Meyer and given

\[ Eq. (6). \] Both Fürth and Rai and Kothari (reference 5) calculate mean free paths incorrectly, taking all molecules except the one in question to be at rest. Perhaps the propensity to this oversimplification comes from the fact that most considerations of mean free path in the last decade have referred to neutron scattering.

\[ Chapman \text{ and Cowling, reference 1, p. 248.} \] The data given do not extend to the limiting concentrations indicated in Eq. (7), but can be extrapolated roughly either graphically or by the use of the results of the Enskog-Chapman theory, which agree well with experiment.

\[ Jeans, reference 11, pp. 313-315, gives a brief indication of one such procedure. The results for the case of rigid spherical molecules are fairly good, but no other cases have been treated. The correction affects only the dependence of \( D_{12} \) on concentration owing to difference in masses of the molecules, and only some very different and probably much more complicated procedure could have any bearing on dependences arising from other causes (see footnote 19). \]
that of Stefan a prominent place in the literature of the subject, if it had not been for the fact that reliable experimental evidence was lacking at that time.

Lack of Physical Plausibility

The simple treatment of the problem of diffusion by means of the concept of mean free path actually lacks plausibility both as to method and as to results. First as to method: In diffusion the transport associated with the passage of a molecule through a given element of area depends only on the fact of its passage and on its permanent property of belonging to one or another species. The question of the locale of its last collision is, accordingly, much less pertinent than it is in the cases of the other transport phenomena, in which the property transported—momentum or energy—is strongly affected by that collision. The previous positions of the particle do, of course, provide statistical evidence as to its probable species, and this is the only evidence available. But instead of yielding a fairly decisive clue, as it does in the problems of viscosity and thermal conductivity, the last collision here serves in large part just to mark the point at which the trail becomes quite hard to retrace. The Meyer method contents itself with the evidence easily obtainable, and gets results that are too inaccurate to be really acceptable. It is physically correct to seek more evidence from considerations of “persistence of velocities,” but the resulting task is tedious, unilluminating and not even completely feasible. We shall see that the Stefan-Frankel method succeeds in obtaining genuinely decisive evidence from consideration of single collisions. This is possible because in this method the effect of the collision on a dynamical quantity—the momentum—is studied.

The implausibility of the result of the simple mean-free-path treatment of diffusion is evident from the form of the result itself, as shown in Eq. (3), without any consideration of the underlying method. We may regard the process of diffusion as one in which each gas is expanding to fill the whole volume uniformly, but the expansion of each gas is hindered by collisions of its molecules with molecules of the other type (or types). Collisions between molecules of the same type affect only the fortunes of individuals, not the flow of the gas. The diffusion coefficient which measures the rapidity of the process, should, accordingly, be capable of being calculated approximately in terms of quantities that are of the general nature of mean free paths but are to be evaluated by taking into account collisions of unlike molecules only. These quasi free paths might be denoted by \( \lambda'_1 \). Since Eq. (3) contains, instead, the ordinary mean free paths \( \lambda_1, \lambda_2 \), which depend on other quantities besides \( \lambda'_1 \), \( \lambda'_2 \) and may differ from the latter by large factors, it is obviously physically incorrect.

It is sometimes suggested that this defect of Eq. (3) be corrected by simply replacing \( \lambda_1, \lambda_2 \) by \( \lambda'_1, \lambda'_2 \). Since

\[
\lambda'_1 = \gamma_2 \lambda_1 \left( \frac{\text{rate}}{\text{rate}} \right), \quad \lambda'_2 = \gamma_1 \lambda_2 \left( \frac{\text{rate}}{\text{rate}} \right),
\]

this gives a result independent of concentration. The obtaining of a qualitatively improved result by such irresponsible doctoring of a wrong formula can only be regarded as fortuitous; the only consistent procedure for improving Eq. (3) would be an elaborate one based on the general idea of “persistence of velocities.” It is unfortunate that this meaningless trick has come to be rather commonly identified as Stefan’s contribution to the theory, while his radically different, simple and physically correct method has been neglected.

2. THE STEFAN-FRANKEL METHOD

Basic Idea of the Method

Stefan’s way of approaching the problem of diffusion is just the opposite of the traditional elementary approach, which uses the concept of mean free path. Instead of starting from an assumed gradient of concentration and calculating from it the flux of each molecular species relative to the motion of the gas as a whole, we begin by assuming the relative flux of the different species and then proceed to calculate the corresponding gradient of concentration. We shall describe the procedure for a mixture of two components; this is the standard case both in the rigorous theory and in the various attempts at elementary explanation.

\[\text{Footnote:}\] The discussion in at least one otherwise excellent elementary textbook offers serious possibilities of confusion by failing to make this point clear; see Millikan, Roller and Watson, *Mechanics, molecular physics, heat, and sound* (Ginn, 1937), pp. 218–219.
In a homogeneous gas mixture of moderate density the number density $n_i$ of each molecular species is related to its partial pressure $p_i$ by the simple equation

$$p_i = n_i k T,$$

which is well known from the most elementary discussions of kinetic theory. Also, if $p$ is total pressure and $n$ is total number density,

$$p = nk T.$$

The partial pressure $p_i$ can then be expressed in terms of the fractional numerical concentration $\gamma_i$, defined in Eq. (4):

$$p_i = \gamma_i n k T = \gamma_i p_i.$$  

If the mixture is not homogenous but has a small gradient of concentration, Eq. (11) will still apply, at least in excellent approximation, to each small part of the volume. The total pressure being supposed constant, we find by differentiating Eq. (11) that

$$\frac{\partial}{\partial x} p_i = n k T \frac{\partial}{\partial x} \gamma_i = p \frac{\partial}{\partial x} \gamma_i.$$  

This relation between gradient of partial pressure and gradient of concentration holds both for constant temperature and for moderate temperature gradients.

The concentration gradient is thus determined if we can calculate the gradient of partial pressure. The latter quantity can readily be computed from suitable assumptions about the motions of the various species of molecule. Indeed, it is evident that the gradient of the partial pressure is equal to the time rate, per unit volume, of transfer of momentum from molecules of other types to molecules of the type in question.

**Transfer of Momentum in Molecular Encounters**

**Encounter between two molecules.**—The transfer of momentum in an encounter between two molecules is most readily calculated by using the reference system of the center of mass. If $V$ is the relative velocity, so that

$$V = c_1 - c_2,$$

then before the encounter the velocity of the first molecule in this system is $m_2 V/(m_1 + m_2)$ and its momentum is

$$m_1 m_2 V/(m_1 + m_2) = \mu V;$$

here $m_1$, $m_2$ are the masses of the molecules, and $\mu$ is the “reduced mass.” The velocity of the second molecule before the encounter is then $-m_1 V/(m_1 + m_2)$, and its momentum is $-\mu V$. In the reference system chosen both paths have the same angle of deflection $\psi$, and the momentums after the encounter are still opposite in direction and each of magnitude $\mu V$. The momentum transferred to the first molecule consists of a component

$$\delta p_{11} = \mu V \cos \psi - \mu V = -\mu V (1 - \cos \psi)$$

in a direction parallel to $V$ and a component

$$\delta p_{12} = \mu V \sin \psi b_1,$$

in a perpendicular direction; here $b_1$ is a unit vector normal to $V$. Since all directions of $b_1$ normal to $V$ are equally probable, the sum of the contributions given by Eq. (16) is zero when we consider the effect of many encounters.

For given relative speed $V$ of two spherically symmetrical molecules of given types the angle of deflection $\psi$ depends only on the impact parameter $b$; this is the distance of closest approach of the centers of mass of the two molecules that would result if they were to pursue their original paths without any deflection. If the molecules are not spherically symmetrical, $\psi$ depends also on their orientations. Since, however, all orientations are equally probable, this merely makes necessary an averaging process; we shall not concern ourselves with this, and shall treat simply the case of spherical molecules.

**Interaction between sets or beams of molecules.**—We now consider a small fraction of the molecules of type 1, having velocities essentially equal to $c_1$ and number density $dm_1$, together with a similar fraction of type 2, with velocity $c_2$ and number density $dm_2$. The number of collisions between members of the two groups per unit volume and unit time with impact parameter between $b$ and $b + db$ is

$$dn_1 dm_2 V \cdot 2\pi b db.$$  

Then, using Eq. (15), we find that the momentum transferred to members of the first group from members of the second is, per unit volume and unit time,

$$-dn_1 dm_2 \mu V \int_{b}^{\infty} 2\pi b db (1 - \cos \psi)$$
The quantity
\[ \sigma(V) = \int_0^\infty 2\pi \delta b (1 - \cos \psi), \]
(19)
can be called the cross section for transfer of momentum. Its special suitability for our purpose comes from the particular factor \(1 - \cos \psi\) in the integrand. In the study of scattering and absorption of fast particles the quantity of this kind most generally used is the "total cross section,"
\[ \sigma_{\text{tot}} = \int_0^{b_{\text{max}}} 2\pi \delta b = \pi b_{\text{max}}^2, \]
(20)
where \(b_{\text{max}}\) is the largest value of the impact parameter that gives a deflection regarded as appreciable. In the Enskog-Chapman theory of transport phenomena the quantity \(\sigma(V)\) defined by Eq. (19) is the cross section most important for diffusion, while for the phenomena of viscosity and thermal conductivity the cross section
\[ \sigma_2(V) = \int_0^\infty 2\pi \delta b (1 - \cos \psi) \]
(21)
plays the main role. The factor \(1 - \cos \psi\) is a maximum for backward scattering; \(1 - \cos \psi\), for scattering at right angles. It is assumed, of course, that the law of force between molecules is such that these factors fall off rapidly for large values of \(b\), so that the integrals in Eqs. (19) and (21) will converge.

In terms of the cross section \(\sigma(V)\) defined by Eq. (19) the momentum transferred to members of the first group of molecules from members of the second group, per unit volume and unit time, can be written
\[ -dn_1 dn_2 V \sigma(V). \]
(22)

**Transfer of momentum between molecular species.** —The number densities \(dn_1\) and \(dn_2\) of our groups of molecules can conveniently be expressed in terms of velocity-distribution functions. The number density of molecules of type 1 having their three components of velocity in the ranges \(u_1\) to \(u_1 + du_1\), \(v_1\) to \(v_1 + dv_1\), \(w_1\) to \(w_1 + dw_1\) is
\[ dn_1 = n_1 f_1(u_1, v_1, w_1) du_1 dv_1 dw_1 = n_1 f_1(c_1) du_1. \]
(23)
Similarly,
\[ dn_2 = n_2 f_2(u_2, v_2, w_2) du_2 dv_2 dw_2 = n_2 f_2(c_2) du_2. \]
(24)
The total densities \(n_1\) and \(n_2\) are, in general, functions of (macroscopic) position. The functions \(f_1\) and \(f_2\) are taken to be normalized so that
\[ \int f_1 du_1 = \int f_2 du_2 = 1, \]
(25)
but the forms of these functions may, of course, depend on position. We need consider only cases in which \(n_1, n_2, f_1, f_2\) are independent of \(y\) and \(z\), and \(f_1, f_2\) depend only on the respective speeds \(c_1, c_2\) and the \(x\)-components \(u_1, u_2\):
\[ n_i = n_i(x); \quad f_i = f_i(c_i, u_i; x). \]
(26)
The concentration gradient and the net flow of each species are then parallel to the \(x\)-axis.

The gradient of the partial pressure \(p_i\), obtained by integrating expression (22), is then directed along the \(x\)-axis, and by Eq. (12) we have
\[ \rho(\partial \gamma_1/\partial x) = -\int d\gamma_1(1) \int d\gamma_2 V_\sigma(V). \]
(27)
Here \(d\gamma_1\) and \(d\gamma_2\) can be expressed in terms of the distribution functions \(f_1, f_2\) as shown in Eqs. (23) and (24). It is convenient sometimes to use the notation of mean values, defined by
\[ \langle F \rangle_\mu = \int f_1 du_1 \int f_2 du_2 F, \]
(28)
where \(F\) is any function of the velocity components, and \(\langle F \rangle_\mu\) is in general a function of position. Then Eq. (27) can be written in the form
\[ \rho(\partial \gamma_1/\partial x) = -n_1 n_2 \mu V_\sigma(V) \langle F \rangle_\mu. \]
(29)

**Relative Flow of Molecular Species.**

To determine the actual forms of the distribution functions \(f_1\) and \(f_2\) it is necessary to resort to the Enskog-Chapman theory. Elementary arguments can be conducted by adopting more or less crude forms for these functions, choosing them to fit assumed values of the relative flow of the two species and of the temperature gradient. The relative flow can be expressed in terms of the mean drift velocities\(^\text{17}\) \(\langle c_1 \rangle_\mu\), \(\langle c_2 \rangle_\mu\), where, by

\(^{17}\) Not to be confused with the mean *speeds* of thermal agitation, \(c\). The vector quantities \(\langle c \rangle_\mu\) have macroscopic significance.
Eqs. (28) and (25),

$$\langle c_2 \rangle_{av} = \int c_2 \, dc_2.$$  

(30)

In the sort of case chosen for consideration, $\langle c_1 \rangle_{av}$ and $\langle c_2 \rangle_{av}$ are, of course, directed parallel to the $x$-axis. It is convenient also to suppose that the mean drift velocity of all the molecules\footnote{This is not, in general, the same as the velocity of the center of mass of all the molecules. The mean drift velocity of all the molecules is the more useful quantity, since, for example, under our assumptions (26), with $p=\text{const.}$, Eq. (31) will hold automatically if the gas mixture is contained in a stationary vessel; the velocity of the center of mass will not in general vanish under these conditions.} is zero:

$$n_1\langle c_1 \rangle_{av} + n_2\langle c_2 \rangle_{av} = n(\gamma_1\langle c_1 \rangle_{av} + \gamma_2\langle c_2 \rangle_{av}) = 0.$$  

(31)

For any chosen form of $f_1$ and $f_2$, the calculation indicated in Eq. (27) or (29) will give the concentration gradient.

**The Diffusion Coefficients**

As will be illustrated in the calculations of the following sections, and as is quite plausible in advance, the concentration gradient always turns out to be a homogeneous linear function of the quantities $\langle c_1 \rangle_{av}$ and $\partial T/\partial x$. The standard way of indicating this fact in writing equation of diffusion is, for cases satisfying Eqs. (26) and (31),

$$\gamma_1\langle c_1 \rangle_{av} = -D_{12}\frac{\partial \gamma_1}{\partial x}+(D_T/T)(\partial T/\partial x).$$  

(32)

Here $D_{12}$ is the diffusion coefficient, and $D_T$ is the coefficient of thermal diffusion. From Eqs. (29) and (32) we have

$$D_{12} = \frac{\gamma_1 p\langle u_1 \rangle_{av}}{n_1 n_2 \mu V_2 V_\gamma(V)_{av}}, \quad \frac{\partial T}{\partial x} = 0$$  

(33)

and

$$k_T = \frac{D_T}{D_{12}} = -\frac{T n_1 n_2 \mu V_2 V_\gamma(V)_{av}}{p(\partial T/\partial x)}$$

$$\langle c_1 \rangle_{av} = 0$$  

(34)

as formulas for calculating the diffusion coefficient and the "thermal diffusion ratio" $k_T$.

**3. SIMPLE METHODS OF ESTIMATING THE ORDINARY DIFFUSION COEFFICIENT**

The formulation of the Stefan-Frankel method which has just been given may not seem very elementary, since it not only involves multiple integrals but refers repeatedly to distribution functions and mean values, concepts that are not very familiar to many students. Technically, it can be called completely elementary, because it does not use the differential-integral equation which is the foundation of the rigorous theory. The rather complicated looking formulation which we have given has its uses in showing the power of the method, when suitably applied, in coming remarkably close to the results of the rigorous theory. It should be emphasized, however, that the Stefan-Frankel method is capable of being used in the simplest conceivable way to give qualitatively good results. This will be our first example.

**Calculation of $D_{12}$ on the Crudest Assumptions**

We shall proceed in a way that is analogous to the calculation of the pressure of a gas given in many first-year college textbooks. The molecules of each of the two species will be taken to be divided into groups moving parallel to the coordinate axes with fixed speeds. For the present purpose it suffices to take only one direction of motion, that parallel to the $x$-axis. Since all of the molecules of each species can act as obstacles to the motion of those of the other species, it is reasonable to assign approximately half of the molecules of each type to each sense of motion. The speeds will be taken to be of the order of magnitude of the speeds of thermal agitation, $c_1$ and $c_2$, but smaller by a factor of the order of 2 to allow for the actual obliqueness of the motions of most of the molecules; we write $s_1$, $s_2$ for such suitably chosen multiples of $c_1$ and $c_2$.

The general idea of a collision cross section and the calculation of the transfer of momentum per unit volume and unit time between two species of molecules can, of course, be explained well enough for the present purpose much more simply than is done in Sec. 2. Without taking the space to do this, we only give the calculation on the basis of Eq. (22).

Writing $j$ for the net flow of molecules of type 1, we have, by Eq. (31),

$$j = n_1\langle c_1 \rangle_{av} = -n_2\langle c_2 \rangle_{av}.$$  

(35)

In the case considered, the vector $j$ lies in the positive $x$-direction. There are various possible choices of the number densities and speeds of the
four sets of molecules which are consistent with Eq. (35). We deal with the two possibilities that most naturally suggest themselves.

**Flow ascribed to differences in number densities.**—If we ascribe the same speed \(s_1\) to all molecules of type 1 and the same speed \(s_2\) to all those of type 2, we can satisfy Eq. (35) by listing the sets as follows:

<table>
<thead>
<tr>
<th>Set No.</th>
<th>Type of molecule</th>
<th>Number density</th>
<th>Velocity</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>(\frac{n_1}{s_1} + \frac{n_2}{s_2})</td>
<td>(s_1)</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>(\frac{n_1}{s_1} - \frac{n_2}{s_2})</td>
<td>(-s_1)</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>(\frac{n_1}{s_1} + \frac{n_2}{s_2})</td>
<td>(-s_2)</td>
</tr>
<tr>
<td>4</td>
<td>2</td>
<td>(\frac{n_1}{s_1} - \frac{n_2}{s_2})</td>
<td>(s_2)</td>
</tr>
</tbody>
</table>

Summing the four contributions of the form (22) for collisions between molecules of sets 1 and 3, 2 and 4, 1 and 4, and 2 and 3, we find that all terms cancel except those that are linear in \(j\), and the result is

\[
\rho \partial \gamma_1 / \partial x = -\frac{1}{2} j \left[ \left( \frac{n_2}{s_1} + \frac{n_1}{s_2} \right) - \frac{n_2}{s_1} \left( \frac{n_1}{s_1} - \frac{n_2}{s_2} \right) + \frac{n_1}{s_2} - \frac{n_2}{s_2} \right] \times \frac{\mu(s_1 + s_2)\sigma(s_1 + s_2) - \frac{n_2}{s_1} \left( \frac{n_1}{s_1} - \frac{n_2}{s_2} \right) \times \mu(s_1 - s_2)\sigma(s_1 - s_2)}{\left( 1 + \gamma_1 + \left( \frac{s_1}{s_2} \right)^2 \gamma_2 \right)}.
\]

(36)

It will be supposed throughout that the molecules of the two species have equal masses or that the first species is the lighter, so that

\[
m_1 < m_2, \quad \varepsilon_1 > \varepsilon_2, \quad s_1 > s_2;
\]

(37)

this justifies our writing \(s_1 = s_2\) for \(s_1 - s_2\) as an argument of \(\sigma\), and makes Eqs. (36), (38) and (40) unsymmetrical in the indices 1 and 2. From Eqs. (32), (35) and (36) we now get

\[
(D_{12})^{-1} = \frac{\rho}{(\gamma_1 / \gamma_2) m_0} \left[ 1 + \gamma_1 + \frac{s_1}{s_2} \gamma_2 \right] \times \mu(s_1 + s_2)\sigma(s_1 + s_2) - \frac{n_2}{s_1} \left( \frac{n_1}{s_1} - \frac{n_2}{s_2} \right) \times \mu(s_1 - s_2)\sigma(s_1 - s_2).
\]

(38)

To simplify this expression further, we may suppose that the masses \(m_1\) and \(m_2\) are not very different, so that \(s_1/s_2\) and \(s_2/s_1\) are approximately 1. Then the result is of the same order of magnitude as the result of the mean-free-path approach, and has the same sort of dependence on temperature and pressure. The same is true of all the results of this section, and we shall not trouble to discuss the rest of our formulas from this point of view.

To study the dependence on concentration, with regard to which the mean-free-path approach fails so badly, we must consider Eq. (38) without supposing the masses nearly equal.\(^\text{19}\) To obtain a simplification when \(s_1/s_2\) is not restricted to values near unity, we consider the velocity dependence of \(\sigma(V)\).

For rigid spherical molecules, \(\sigma\) is a geometric quantity independent of \(V\). On this assumption we obtain from Eq. (38),

\[
(D_{12})^{-1} = \frac{\rho}{(n^2\mu/\rho)^V} \left[ 1 + \gamma_1 + \frac{s_1}{s_2} \gamma_2 \right]^{-1}.
\]

(40)

Thus the dependence of \(D_{12}\) on concentration is given by a factor

\[
\left[ 1 + \gamma_1 + \frac{m_1}{m_2} \gamma_2 \right]^{-1}.
\]

(41)

In the extreme case, for which \(m_1 < m_2\), the coefficient is twice as large for \(\gamma_1 < 1\) as for \(\gamma_2 < 1\). This is a much weaker dependence on concentration than is given by Meyer’s method, although the change is in the same direction.

For Maxwellian molecules, exerting forces varying as the inverse fifth power of the distance, \(\sigma(V)\) is inversely proportional to \(V\), as we shall show in Eq. (47). With \(\sigma^{-1}(V)\) independent of \(V\), Eq. (38) becomes

\[
(D_{12})^{-1} = \left( n^2 \mu / \rho \right)^V \left[ V \sigma^{-1}(V) \right].
\]

(42)

This is exactly independent of concentration.

**Flow ascribed to differences in velocities.**—If we take the two beams of each species to have equal
number densities, we can satisfy Eq. (35) by the following scheme:

<table>
<thead>
<tr>
<th>Set No.</th>
<th>Type of molecule</th>
<th>Number density</th>
<th>Velocity</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>(\frac{1}{2}n_1)</td>
<td>(s_1 + j/n_1)</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>(\frac{1}{2}n_1)</td>
<td>(-s_1 - j/n_1)</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>(\frac{1}{2}n_2)</td>
<td>(-s_2 - j/n_2)</td>
</tr>
<tr>
<td>4</td>
<td>2</td>
<td>(\frac{1}{2}n_2)</td>
<td>(s_2 - j/n_2)</td>
</tr>
</tbody>
</table>

The contribution of the form (22) for collisions of sets 1 and 3 is

\[-(n_1n_2/4)\mu[(s_1 + s_2 + (j/n_1) + (j/n_2))/2] \times \sigma[(s_1 + s_2 + (j/n_1) + (j/n_2))/2].\] (43)

The contribution for sets 2 and 4 is of opposite sign and has the sign of \(j\) changed. The other two contributions, from collisions between 1 and 4, and 2 and 3, are obtained90 from the first two by changing the sign of \(s_2\). Since the mean drift velocities are always minute compared to the mean speeds of thermal agitation, the contributions to \(V\) depending on \(j\) are minute compared to \(s_1\) and \(s_2\), and we can combine the first two contributions to obtain

\[-(n_1n_2/4)\mu[[d/d(V) V^2\sigma(V)] V = s_1 + s_2 \\
-2[(j/n_1) + (j/n_2)]].\] (44)

The sum of the other two contributions to \(p(\partial\sigma/\partial x)\) is obtained from expression (44) by changing the sign of \(s_2\). Then, multiplying the sum of all four contributions by \(-n/pj\) and simplifying algebraically, we obtain the result,

\[(D_{12})^{-1} = (n^2\mu/2p)\left[[(d/dV) V^2\sigma(V)] V = s_1 + s_2 \\
+[(d/dV) V^2\sigma(V)] V = s_1 - s_2\right].\] (45)

This is exactly independent of concentration for any law of force and for all values of \(m_2/m_1\).

The two crude calculations we have made involve two opposite, extreme assumptions about the way the net flow is to be obtained from the four beams of molecules. The second calculation gives a result that is always independent of concentration. The first gives a moderate dependence on concentration for rigid spherical molecules, and none for Maxwellian molecules. The behavior of actual gases is roughly intermediate between that to be expected for rigid spheres and that to be expected for Maxwellian molecules, and is usually much closer to the latter. Thus our very crude considerations based on the Stefan-Frankel approach indicate that \(D_{12}\) is nearly independent of concentration, so that they succeed much better than Meyer’s discussion based on the concept of mean free path.

**Exact Result for Maxwellian Molecules**

For a gas composed of rigid spherical molecules the viscosity would be proportional to the square root of the absolute temperature. Maxwell’s experiments92 indicated that the coefficient of viscosity of air is very nearly proportional to the absolute temperature. Maxwell decided that, since gases ought to behave simply, the proportionality should be exact. This would mean that the mean free path itself should be proportional to the square root of the absolute temperature; in other words, that collision cross sections should be inversely proportional to relative velocity. From this Maxwell concluded that the law of force must be that of the inverse fifth power, and based on this his famous treatment of kinetic theory.93

Maxwell’s discussion of the behavior of collision cross sections for inverse-power forces was based on the actual solution of the dynamical problem, but Frankel94 pointed out that the results could be obtained easily by a dimensional argument. The problem of solving the equations of motion for the force law

\[F = K/V^r\] (46)

and then evaluating an integral such as those in Eqs. (19) and (21) is a well-defined one, and must give \(\sigma(V)\), with dimensions \([LT^{-2}]\), in terms of just three quantities: the force constant \(K\), with dimensions \([ML^{−1}T^{−2}]\); the reduced mass \(\mu\), with dimensions \([M]\); and the relative speed \(V\), with dimensions \([LT^{-1}]\). It follows at once that

\[\sigma(V) = \text{const.} \cdot (K/\mu V^2)^{2/\nu}.\] (47)

Thus

\[\nu \sigma(V) \propto V^{(\nu-6)/(\nu-1)}.\] (48)

For “Maxwellian” molecules, with \(\nu = 5\), \(\nu \sigma(V)\)

90 The signs affixed to these two contributions are determined, as in the other case, by stipulation (37).
92 Reference 9, p. 1.
93 Reference 9, p. 26. Five years later—in 1871 (reference 9, p. 343)—Maxwell presented a number of results based on the hypothesis of rigid spherical molecules, having evidently realized that actual molecules are not strictly “Maxwellian” and that the behavior of most actual gases is roughly intermediate between these two cases.
is independent of \( V \). For "harder" molecules, with \( \nu > 5 \), \( V_\nu(V) \) increases as \( V \) increases. Rigid spherical molecules can be included formally as the case \( \nu \to \infty \), with \( \sigma(V) \) independent of \( V \).

For the Maxwellian case the factor \( V_\nu(V) \) can be removed from the quantity to be averaged in Eq. (29), and we find without approximation, taking account of Eq. (13),

\[
\rho(\partial \gamma_1/\partial x) = -n_1 n_2 \mu [V_\sigma(V)] \times [(c_{12})_u - (c_{21})_u].
\]  

(49)

By Eq. (35) this reduces to

\[
\rho(\partial \gamma_1/\partial x) = -n_1 j \mu [V_\sigma(V)].
\]  

(50)

Multiplying, as usual, by \(-n/pj\), we get

\[
(D_{12})^{-1} = (n_1^2 \mu / \rho) [V_\sigma(V)].
\]  

(51)

This result, which is independent of concentration, was obtained without making any assumption about the distribution functions \( f_1 \) and \( f_2 \), and hence holds exactly. The possibility of treating transport phenomena in the case \( \nu = 5 \) without the necessity of determining distribution functions gave Maxwell's treatment of kinetic theory its great elegance.

Since Eq. (50) was derived without any stipulation other than Eq. (35) about the distribution functions, it is true exactly even for nonuniform temperature of the gas mixture. Equation (50) is just Eq. (32) without the last term. Thus for Maxwellian molecules the phenomenon of thermal diffusion does not appear.

Use of Approximately Correct Distribution Functions

We now return to the general case, in which the evaluation of the right-hand member of Eq. (29) requires the use of definite distribution functions. Instead of the very crude picture of the distribution functions used in the first calculations of this section, we shall now use the best simple assumption available.

For \( j = 0 \), the distribution functions are known exactly, being just Maxwell's distribution law,

\[
f_{i}^{(0)}(c_i) = (\beta_i / \pi)^{1/2} \exp(-\beta_i c_i^2), \quad i = 1, 2
\]  

(52)

with

\[
\beta_i = m_i / 2kT.
\]  

(53)

The distributions for \( j \neq 0 \) differ from these only very slightly, since the mean drift velocities are extremely small compared to the mean speeds of thermal agitation. We shall, accordingly, make simple small changes in Eqs. (52) so as to secure the net flow \( j \) defined by Eq. (35), again taking \( j \) to lie in the positive \( x \)-direction. The resulting distributions will not be the true ones, but will differ from them only slightly. Since even our very crude assumptions gave fairly good results, the present procedure seems well worth trying.

We set

\[
f_1 = (1 + \alpha_1 u_1) f_1^{(0)}.
\]  

(54)

The correction term is an odd function of \( u_1 \), and hence does not affect the normalization. By Eq. (35) we have

\[
j = n_1 \int f_1 u_1 d\omega_1 = a_1 n_1 \int f_1^{(0)} u_1^2 d\omega_1
\]  

\[
= a_1 n_1 (u_1^2)_{av} = \frac{1}{2} a_1 n_1 (c_1^2)_{av} = a_1 n_1 kT / m_1.
\]  

(55)

Substituting the resulting value of \( a_1 \) in Eq. (54), we have

\[
f_1 = [1 + (m_1 j / n_1 kT) u_1] f_1^{(0)}.
\]  

(56)

Similarly,

\[
f_2 = [1 - (m_2 j / n_2 kT) u_2] f_2^{(0)}.
\]  

(57)

Since the total net flow of molecules is zero, it follows from equipartition of energy that there is no net flow of energy, that is, no thermal conduction. Thus the assumption that Eqs. (56) and (57) hold for the element of volume under consideration is consistent with the assumption that there is no temperature gradient.

The numerator of the expression for \( D_{12} \) in Eq. (33) is

\[
\gamma_1 \rho j / n_1 = \rho j / n = jkT,
\]  

(58)

and the denominator is

\[
n_1 n_2 \int f_1 d\omega_1 \int f_2 d\omega_2 (u_1 - u_2) V_\sigma(V).
\]  

(59)

Since the value of \( V \) is left unchanged when the signs of both \( u_1 \) and \( u_2 \) are changed, the only parts of the integrand that give nonvanishing integrals are those linear in \( j \) and quadratic in \( u_1 \) and \( u_2 \). Expression (59) can accordingly be
written

\[
(\mu j/kT) \int f_1(\omega_1) d\omega_1 \int f_2(\omega_2)(u_1 - u_2) \times (n_m u_1 - n_m u_2) V_\sigma(V).
\]

(60)

This expression can be simplified by introducing as variables of integration the velocity \( \mathbf{C} \) of the center of mass and the relative velocity \( \mathbf{V} \),

\[ \mathbf{c}_1 = \mathbf{C} + (m_1/M) \mathbf{V}, \quad \mathbf{c}_2 = \mathbf{C} - (m_1/M) \mathbf{V}, \]

(61)

where \( M \) is an abbreviation for \( m_1 + m_2 \). Then we have

\[ m_c c_1^2 + m_c c_2^2 = MC^2 + \mu V^2. \]

(62)

Also

\[ n_m m_1 u_1 - n_m m_2 u_2 = (n_m m_1 - n_m m_2) C_2 + n_m V_2. \]

(63)

We have

\[ d\omega_1 d\omega_2 = 1 - m_1/M \int dC_1 dV_1 dC_2 dV_2 dV_3, \]

(64)

so that

\[ d\omega_1 d\omega_2 = dC_1 dC_2 dC_3 dV_1 dV_2 dV_3. \]

(65)

If we make the definitions

\[ B = M/2kT, \quad \beta = \mu/2kT, \]

(66)

we see that

\[ \beta_1 \beta_2 = B \beta, \]

(67)

and also, by Eq. (62), that

\[ \beta c_1^2 + \beta c_2^2 = BC^2 + \beta V^2. \]

(68)

Expression (60) then becomes

\[
(\mu j/kT) (\beta/\pi)^3 \\
\times \int \int \int e^{-\beta r_1} dV_1 dV_2 dV_3 (B/\pi)^3 \\
\times \int \int \int e^{-\beta C^2} dC_3 dC_2 dC_1 \\
\cdot V_3 [(n_m m_1 - n_m m_2) C_2 + n_m V_2] V_\sigma(V).
\]

(69)

Here the term in \( C_2 \) is an odd function of \( C_2 \) and contributes nothing to the integral. The other term depends on \( C \) only through the factor \((B/\pi) e^{-\beta r_1}\), which is a normalized function like those defined in Eq. (52), so that its integral over the \( C \)-space gives the factor unity. By symmetry we can replace \( V_3^2 \) by \( V^3 \), so that our expression reduces to

\[
(n \mu^2 j/3kT) (\beta/\pi)^3 \\
\times \int \int \int e^{-\beta r_1} dV_1 dV_2 dV_3 V_3^3 V_\sigma(V).
\]

(70)

We now integrate over the direction of the vector \( \mathbf{V} \):

\[ dV_3 dV_2 dV_1 = V_3^2 dV_3 V_1 \sin \theta_1 d\theta_1 d\phi_1; \]

\[ \int \int \int f(V) dV_3 dV_2 dV_1 = 4\pi \int_0^\infty V_3^2 f(V) dV. \]

(71)

Writing also

\[ \beta^3 V = g, \]

(73)

we find that the formula of Eq. (33), which has (58) for its numerator and (70) for its denominator, has become

\[ D_{12} = 3kT/16\mu n \Omega_{12}^{(1)}(1), \]

(74)

where

\[ \Omega_{12}^{(1)}(\beta) = 2^{-1} \pi^{-1} \int_0^\infty e^{-\beta^2 r^2 + 2\beta V_3 V_0} dV_3, \]

(75)

with \( V \) related to \( g \) by Eqs. (73) and (66). This is the notation used by Chapman and Cowling. \( ^{34} \)

Equation (74) makes \( D_{12} \) exactly independent of concentration. What is more, Eq. (74) is just the formula given by the first approximation of the rigorous Chapman-Enskog theory. The evaluation of the second approximation in some cases, and of higher approximations in very special cases, shows that the first approximation in general gives very nearly the correct result.

For Maxwellian molecules the particular forms postulated for \( f_1 \) and \( f_2 \) do not matter, and Eqs. (42), (45) and (74) all agree with the exact result of Eq. (51). For rigid spherical molecules our crude results given in Eqs. (40) and (45) agree fairly well with the good result of Eq. (74). As already mentioned, the first of these crude results shows a false dependence on concentration,

\footnote{This is the theorem that the total kinetic energy is equal to the kinetic energy of the total mass, moving with the speed of the center of mass, plus the kinetic energy of relative motion.}

\footnote{Reference 2, p. 157.}
though not a very strong one. Besides this, both of the crude results for rigid spherical molecules show an asymmetric dependence on the molecular weights, because the oversimplified velocity distributions allow only one kind of molecule to make “overtaking” collisions. The asymmetry can be expressed by a factor \( [1/(m_1/m_2)]^3 \), which never differs greatly from unity.

For the case of rigid spherical molecules Stefan obtained a result equivalent to Eq. (74) by using essentially the same argument as has been given here. The result given by Maxwell for rigid spherical molecules differs from Stefan’s by the factor \( 2\sqrt{2}/3 \), or by about 6 percent. It is not easy to guess how this difference in numerical factor came about.

The great superiority of Stefan’s approach over Meyer’s as regards physical principles is shown most strikingly by what happens when one postulates that Maxwell’s inverse fifth-power law of force holds between the molecules. When this is done in Meyer’s method, the only effect is to make \( D_{12} \) proportional to \( T^2 \), rather than to \( T^3 \), as it is for rigid spherical molecules; the method remains a crude one, and difficulties with the dependence on concentration remain. On the other hand, when we introduce the inverse fifth-power law into Stefan’s method we note at once that the necessity for guessing at the velocity distributions, which is the one difficulty with the method, is removed, and we are led to claim that the result obtained holds exactly. This claim is borne out by the rigorous Enskog-Chapman theory. The calculations by Stefan’s method for this case are essentially just a simple restatement, in more physical language, of the part of Maxwell’s famous 1866 theory which deals with diffusion.

4. THE ELEMENTARY EXPLANATION OF THERMAL DIFFUSION

General Qualitative Considerations

A main achievement of Frankel’s argument about thermal diffusion is the demonstration that the effect vanishes for Maxwell’s inverse fifth-power law. This has already been brought out in the preceding section. When \( \bar{v}\sigma(V) \) is independent of \( V \) we have, by Eq. (29),

\[
\dot{\rho} (\partial \gamma_1 / \partial x) = -n_1 n_2 (\mu V_r \sigma (V)_{av})
= -n_1 n_2 \mu [\bar{v}\sigma(V)] (\bar{V}_2)_{av}
= -n_1 n_2 \mu [V\sigma(V)] (\bar{\mu})_{av} - (\bar{\mu})_{av}.
\]

(76)

The gradient of concentration accordingly depends only on the relative flow of the two species, and is unaffected by a temperature gradient.

Frankel also remarks that if the intermolecular force depends on distance more or less strongly than by the inverse fifth power, a temperature gradient will produce an effect which will be of one sign if the molecules are “harder” than Maxwellian molecules and of the opposite sign if they are “softer.” The discussion can be confined to the usual case of “hard” molecules, in which the force varies with distance more strongly than by the inverse fifth power. Then \( \bar{v}\sigma(V) \) increases as \( V \) increases, and there is a disproportionately large transfer of momentum in the collisions in which \( V \) is larger.

For the case in which the molecules differ in mass, Frankel argues that, while both kinds of molecule have higher average speeds when coming from the hotter part of the gas than when coming from the colder part, the difference is greater for the lighter molecules than for the heavier ones. Thus more of the collisions with large values of \( V \) will have the lighter molecule coming from the hotter region rather than vice versa, and the net transfer of momentum from heavy to light molecules will be directed toward the hotter part of the gas. This means that the gradient of partial pressure, \( \dot{\rho} (\partial \gamma_1 / \partial x) \), will include a part proportional to \( (\partial T / \partial x) \), with a positive coefficient, besides the term proportional to the relative flow of the two species. Thus—speaking always of the case of “hard” molecules—there will be an effect of thermal diffusion acting to produce an increased concentration of the lighter species of molecules in the hotter region. Since species 1 is taken to be the lighter species, we see that \( D_T \) as defined in Eq. (32) is positive.

It is easy to extend the argument to cover cases in which the two kinds of molecule differ otherwise than in their masses. Suppose that one species has a longer mean free path in the gas mixture than the other species. Then the difference in average speeds of molecules coming from
the hotter and colder regions will be greater for the species having the longer mean free path. Thus, by the same argument that Frankel gives for the effect of a difference of mass, there will be an effect of thermal diffusion acting to increase the concentration in the hotter region of the molecular species which has the longer mean free path.

The obvious reason for a difference of mean free paths is that one species of molecules may be "smaller" than the other, in the sense that the effective range of the force it exerts is less. Then thermal diffusion will act to concentrate the smaller molecules in the hotter region. The coefficient \( D_T \) in Eq. (32) will be positive if species 1 has the smaller molecules.

Another possible source of a difference of mean free paths is that the effective collision cross section for unlike molecules may be different from that for like molecules of either species. If the cross section for unlike molecules is smaller than that for like molecules, the species with the smaller concentration will have the longer mean free path, and thermal diffusion will act to increase the concentration of this species in the hotter region. Since, as shown in Eq. (32), the thermal diffusion coefficient is defined so that it is positive when the effect is to increase the concentration of species 1 in the hotter region, it will in this case be proportional to \( \gamma_2 - \gamma_1 \). If the cross section is greater for unlike molecules than for like, the coefficient will be proportional to \( \gamma_1 - \gamma_2 \).

The first suggestion of the possibility of an effect of this sort seems to have been made by H. M. Mott-Smith, who considered, for the case of isotopic mixtures, the quantum-mechanical effect of the symmetry properties of the wave functions for identical particles. This effect would usually influence thermal diffusion only slightly, and no observations confirming it have been published. In the case of strongly dissimilar molecules the cross section for like molecules of one species may be very different from that for like molecules of the other species, and elementary considerations are not capable of yielding a very good estimate of the part of the thermal diffusion coefficient that depends on the difference of concentration. A discussion based on the exact theory has been given by Chapman.  

An interesting case which has been investigated experimentally is that of neon-ammonia mixtures. If we call the lighter molecules, those of ammonia, species 1, the effect of the difference of masses makes a positive contribution to \( D_T \), but, since the molecules of ammonia are considerably larger than those of neon, the effect of the difference of sizes makes a negative contribution. These two effects largely cancel, but the latter effect is the larger, so that usually \( D_T \) is negative, and the neon is concentrated in the hotter region. When the mixture contains more than 75 percent neon, however, the sign of \( D_T \) changes and it becomes positive. The partial cancellation of the other two effects makes possible the observation of a contribution proportional to \( \gamma_2 - \gamma_1 \).

Frankel remarked that it is clear, if his approach is used, that the thermal diffusion coefficient is proportional to the product of the concentrations. This is obvious, for example, from a glance at Eq. (34). General qualitative considerations about the velocity distributions have now yielded the following conclusions about the thermal diffusion coefficient:

(1) It vanishes for Maxwellian molecules and, ceteris paribus, will have one sign for "harder" molecules, the other for "softer" molecules.

(2) Its main dependence on the concentrations is through a factor \( \gamma_2 \gamma_5 \), which does not affect the sign.

(3) For "hard" molecules, the larger mass of the molecules of species 2 gives rise to a positive contribution to \( D_T \).

(4) A difference in size of the molecules gives rise to a contribution to \( D_T \). For "hard" molecules this contribution is positive if the molecules of species 2 are the larger ones.

(5) A difference in collision cross sections for like and unlike molecules gives rise to a contribution to \( D_T \) proportional to the difference of the concentrations. For "hard" molecules this contribution is proportional to \( \gamma_2 - \gamma_1 \) if the cross section for unlike molecules is the greater.

The first two of these conclusions are the most important, and are valid in general. The

---

25 It should be mentioned that the opposite convention is also often used.
26 Private communication, April 1940.
29 Rai and Kothari (reference 5) obtain conclusions (2), (3), (4) and (5) by algebraic manipulation of a simplified mean-free-path formalism which is in at least one respect incorrect (see footnote 15). They give no physical reasoning to explain these results. Manipulations can accomplish a great deal, but conclusion (1) remains completely inaccessible by the mean-free-path approach.
special contributions described in the other three conclusions are clearly distinguishable only if the properties of the two species of molecules are not very different. For very dissimilar molecules the dependence on concentration becomes considerably more complicated, and cannot be discussed adequately by elementary methods. The situation in this regard is like that met with in connection with the viscosity and thermal conductivity of mixtures of highly dissimilar gases.

The information which we have obtained by discussing the physics of thermal diffusion comprises all the valid information that we could hope to obtain by any calculation conducted on an elementary basis. This fact must be regarded as a distinct virtue of Frankel’s approach. An approximate formula for the thermal diffusion ratio for binary mixtures of not too dissimilar gases can be written down without any further observation except that the ratio \( k_T \) is, by definition—Eq. (34)—a dimensionless quantity. We have then

\[
k_T = D_T / D_{12} \approx \gamma_1 \gamma_2 \left[ (\nu - 5) / (\nu - 1) \right] \times [ \gamma_1 (m_2 - m_1) / (m_2 + m_1) + \gamma_2 (\sigma_{12} - \sigma_{11}) / (\sigma_{12} + \sigma_{11}) + \gamma_3 (\gamma_2 - \gamma_1) [1 - 2 \sigma_{12} / (\sigma_{12} + \sigma_{11})] ].
\]

(77)

Here \( \sigma_{12}, \sigma_{11}, \sigma_{13} \) are collision cross sections, the subscripts indicating the types of molecule involved, and \( C_1, C_2, C_3 \) are numbers of the order of magnitude of unity. The factor involving \( \nu \) and giving expression to the first of our five conclusions is obtained by referring to Eq. (48). The actual dependence on \( \nu \) is, of course, more complicated, but the factor written gives the only sensitive dependence.

The so-called thermal diffusion constant,

\[
\alpha = k_T / \gamma_1 \gamma_2,
\]

(78)

is often used in calculations on separation of isotopes by thermal diffusion. It is not independent of concentration unless the molecules are somewhat similar and the term in \( \gamma_2 - \gamma_1 \) in Eq. (77) is negligible. It may usually be expected to depend to some extent on the temperature, but is independent of pressure.

**Calculation from Assumed Velocity Distributions**

Although all of the reliable information about thermal diffusion that can be obtained by an elementary approach has already been found above without any actual calculations, there is a certain interest in carrying such a calculation through and seeing how expression (34) for \( k_T \) actually reduces to the form of Eq. (77).

In this sort of calculation the idea of mean free path must be used, as in the discussion already given. The Stefan-Frankel approach eliminates the idea of mean free path from the treatment of ordinary diffusion, but it still has to be used in dealing with thermal diffusion, just as for thermal conductivity. The method actually puts the elementary treatment of ordinary diffusion on a much better basis than any of the other transport phenomena, but leaves the elementary treatment of thermal diffusion subject to just the same difficulties and inaccuracies as those of thermal conductivity and viscosity.

The calculations can be based either on a one-dimensional picture of four “beams” of molecules, like our crudest treatments of ordinary diffusion, or on the use of suitably modified Maxwellian distributions of velocities. Either procedure is capable of giving all the information obtainable by elementary means, namely, Eq. (77). Both sorts of calculation have been carried through with care by Cacciapuoti for the case of rigid spherical molecules that differ in mass only.

The crude one-dimensional treatment will be omitted here in favor of a three-dimensional treatment, which is somewhat simpler than Cacciapuoti’s. This treatment will show that Frankel’s approach is capable of giving precisely the factor by which the main dependence on the law of molecular force is determined in the rigorous theory; this contradicts a previous statement by the writer.\(^{20}\) It will also take into account the simultaneous occurrence of both ordinary diffusion and thermal diffusion, and thus lead to the complete equation of diffusion, Eq. (32).

The approximate velocity-distribution functions are taken in the form

\[
f_i = f_i^{(0)}(1 + a_i \omega_i - b_i \omega_i^2), \quad i = 1, 2,
\]

(79)

where the \( f_i^{(0)} \) are the Maxwellian distributions defined by Eq. (52). The net flow of molecules of species 1 in the positive \( x \)-direction is

\[
j = n_1 \int u_1 d\omega_1 = n_1 [a_1 c_T (c_T)^2 + b_1 c_T / c_T']
\]

(80)

Similarly, by Eq. (35), we set

\[
j = -n_2 \int u_2 d\omega_2 = n_2 [-a_3 c_T / c_T' + b_3 c_T (c_T)^2].
\]

(81)

The transfers of kinetic energy in the negative \( x \)-direction

\(^{20}\) R. C. Jones and W. H. Furry, *Rev. Mod. Physics* 18, 156 (1946). This paper was delayed in publication, and the statement in question is one in which the writer concurred in 1940. The more general statement that “the weakness of Frankel’s argument is that it cannot take into consideration the details of the asymmetry of the velocity distribution function itself” remains true; this is, however, also just the weakness of the elementary treatments of thermal conductivity and viscosity.
by molecules of types 1 and 2 are

\[-q_{12} = -n_{1} \int m_{1} \cdot \mathbf{v}_{1} \cdot \mathbf{c}_{1} \, \text{d}x_{1} = -n_{1} m_{1} (b_{1} c_{1})_{12} - \frac{3k_{b}}{2m_{1}} \mathbf{v}_{1} \cdot \mathbf{c}_{1} (\partial T/\partial x),\]

\[-q_{21} = -n_{2} \int m_{2} \cdot \mathbf{v}_{2} \cdot \mathbf{c}_{2} \, \text{d}x_{2} = -n_{2} m_{2} (b_{2} c_{2})_{21} - \frac{3k_{b}}{2m_{2}} \mathbf{v}_{2} \cdot \mathbf{c}_{2} (\partial T/\partial x).\]

From Eqs. (80) to (83) we get

\[-q_{12} = -j \cdot \frac{3k_{b}}{2m_{1}} \mathbf{v}_{1} \cdot \mathbf{c}_{1} (\partial T/\partial x),\]  

\[-q_{21} = -j \cdot \frac{3k_{b}}{2m_{2}} \mathbf{v}_{2} \cdot \mathbf{c}_{2} (\partial T/\partial x).\]

Here the terms proportional to \(j\) represent the kinetic energy transferred by the net flows of molecules of the two species, as the total flow of molecules is taken to be zero these two terms are equal but of opposite signs. The last terms of Eqs. (84) and (85) must be attributed to the thermal conductivity of the gas mixture.

The coefficient of thermal diffusion will turn out to be proportional to the difference of the two terms, one containing the factor \(b_{1}\) and the other \(b_{2}\). Since neither the rigorous theory nor experimental results provide any way of separating the thermal conduction of a gas mixture into parts to be attributed to its constituents, we are obliged to choose values for \(b_{1}\) and \(b_{2}\) by means of crude considerations based on the idea of mean free path. The choice of these expressions, particularly as regards numerical factors, is partly a matter of taste. We set

\[5b_{1} m_{1} c_{1} (kT/m_{1})^{3} = \frac{3k_{b}}{2m_{1}} \frac{3n_{1} m_{1} c_{1} \lambda_{c}}{(\partial T/\partial x)}.\]  

(86)

The choice of factors is based on the approximate relations that hold for each gas taken separately:

- thermal conductivity \(\equiv \frac{3k_{b}}{2m_{1}} \cdot \text{viscosity} \cdot \text{coefficient of thermal conductivity},\)

- \(c_{1} = \frac{3k_{b}}{2m_{1}} \cdot \text{viscosity} \cdot \text{coefficient of thermal conductivity},\)

where \(\lambda_{c}\) is the mean free path. The numerical factor \(\frac{3}{2}\) is perhaps as well justified as any.

From Eqs. (80) to (86) we obtain

\[a_{1} = (m_{1}/n_{1} kT)^{1/2} (5kT/m_{1}) \hat{b}_{1},\]  

(87)

\[a_{2} = (m_{2}/n_{2} kT)^{1/2} (5kT/m_{2}) \hat{b}_{2},\]  

(88)

\[b_{1} = (m_{1} c_{1} \lambda_{c} / 4 \pi k^{2} T^{2}) (\partial T/\partial x),\]  

(89)

\[b_{2} = (m_{2} c_{2} \lambda_{c} / 4 \pi k^{2} T^{2}) (\partial T/\partial x).\]  

(90)

From Eqs. (28) and (29) we have

\[nkT (\partial \gamma_{12} / \partial x) = -n_{1} m_{1} \int \hat{a}_{1} \, \text{d}x \int \hat{b}_{1} \, \text{d}x \int V (V).\]  

(92)

The calculation of the terms in \(a_{1}\) and \(a_{2}\) is the same as that made in the last part of Sec. 3, and the terms in \(b_{1}\) and \(b_{2}\) can readily be evaluated in the same way. On making the transformation described in Eqs. (61) to (68) and performing the integration over \(\text{d}C_{1} \text{d}C_{2} \text{d}C_{3}\), we get from Eq. (92),

\[nkT (\partial \gamma_{12} / \partial x) = -n_{1} m_{1} \beta (\partial / \partial x)^{2} \int \hat{a}_{1} \, \text{d}x \int \hat{b}_{1} \, \text{d}x \int V (V).\]  

(93)

The right-hand member of Eq. (93) can at once be written in terms of the quantities \(\Phi_{12}(r)\) of the Chapman-Cowling notation, defined by Eq. (75). We then have

\[nkT (\partial \gamma_{12} / \partial x) = -n_{1} m_{1} \beta (kT/m_{1}) (\partial V / \partial x),\]  

(94)

When the values of \(a_{1}\), \(a_{2}\), \(b_{1}\), \(b_{2}\) given in Eqs. (87) to (90) are substituted in Eq. (94), this equation becomes

\[nkT (\partial \gamma_{12} / \partial x) = -16 \pi \mu_{1} / 3 \lambda_{12} \beta (1 - 4n_{1} m_{1}^{2} / 3),\]  

(95)

This is the same as the general equation of diffusion, Eq. (32), in virtue of Eq. (35) and the identifications

\[D_{12} = \frac{3k_{b}}{16\pi \mu_{1} \lambda_{12} \beta (1)},\]  

(96)

\[k_{T} (D_{12} / D_{21}) = k_{T} (4n_{1} m_{1} / 3kT) (\gamma_{12} / \gamma_{21}),\]

(97)

The first of these is, of course, the same as Eq. (74).

The resulting expression for \(k_{T}\) is proportional to \(\gamma_{12}\). In the first brackets it contains the same factor as occurs in the result of the rigorous theory to determine the strongest dependence of \(k_{T}\) on the force law. From Eqs. (48), (73) and (75) we can evaluate this factor:

\[2k_{T} (kT/m_{1}) (\partial V / \partial x) = \frac{kT / 2k_{T} \beta (kT/m_{1}) (\partial V / \partial x)}{4 \pi \beta (kT/m_{1}) (\partial V / \partial x)},\]

(98)

In the last form written the insensitive factor involving the gamma-function, which increases by only about 50 percent as \(r\) changes from 5 to \(\infty\), has been replaced by a constant.

The main dependences of \(k_{T}\) on the force law and on concentration are, accordingly, given correctly by the first

\[\text{Chapman and Cowling, reference 2, p. 167; the meaning of abbreviations is explained on p. 164.}\]
factors of Eq. (97). For the dependence on the differences of molecular properties and for any further dependences on concentration we must examine the last factor. This factor, which involves the mean free paths, cannot be regarded as giving reliable information when the molecules are very dissimilar. We shall deal with it only for the case in which the differences between the molecules are small. In this case, if we write for $\bar{v}_1$ and $\bar{v}_2$ the root-mean-square values, in order to have convenient numerical factors, we find

$$(\lambda_1\lambda_2/m_1)-(\lambda_2\lambda_3/m_2) = (3kT)[(\lambda_1/m_1)-(\lambda_2/m_2)]$$

$$\equiv (3kT/m^2)\lambda[3(m_0-m_2)/(m_0+m_2)+\lambda_1-\lambda_2/\lambda]. \quad (99)$$

Here $m$ and $\lambda$ are mean values, from which the values for the two species of molecules differ only slightly. Then $m\equiv2\mu$.

Substitution of Eqs. (98) and (99) in Eq. (97) gives

$$k_T=(D_T/D_0)=(\lambda\sigma_0/2\sqrt{\gamma_0})\gamma_0[3(\mu_0-m_2)/(m_0+m_2)+\lambda_1-\lambda_2/\lambda]. \quad (100)$$

Here $\sigma(V)$ is to be evaluated at the argument shown in Eq. (98), which is a typical value of $V$. Then $\lambda\sigma$ is of the order of magnitude unity, and numerical factors are, of course, not to be taken seriously. When the term $(\lambda_1-\lambda_2)/\lambda$ is expressed approximately in terms of differences of molecular properties, the result takes on precisely the form of Eq. (77).

5. CONCLUSION

The traditional elementary theory based on the idea of mean free path succeeds very well in the treatment of viscosity and thermal conductivity of single gases; the results are really uncertain only as to numerical constant factors. For mixtures of strongly dissimilar gases the dependences of these effects on concentration are not given accurately by the elementary treatment, but there are no very large discrepancies.

In the case of diffusion the mean-free-path approach gives results as good as those for viscosity and thermal conductivity in only one case—that in which the differences between the two molecular species are practically negligible. This is the classical problem of "self-diffusion," and is well realized in most cases in which the differences are isotopic only, not chemical. As we pointed out at some length, the mean-free-path theory encounters serious difficulties in the problem of diffusion of strongly dissimilar gases. The more difficult problem of thermal diffusion presents questions with which the mean-free-path theory is completely unable to deal.

The approach used by Stefan for the problem of ordinary diffusion makes it possible to find by the crudest sort of arguments results that are quite as good as the usual elementary results for viscosity and thermal conductivity. When rather more care is used, Stefan's method gives precisely the result of the first approximation of the rigorous theory, which is very accurate in nearly all cases.

Frankel's treatment of thermal diffusion combines Stefan's procedure with some arguments of the mean-free-path variety. It seems fair to ascribe the remaining inaccuracies of Frankel's theory to the latter ingredient, and credit its general success to the distinctly superior qualities of Stefan's approach. The combination devised by Frankel makes it possible to obtain results which, for mixtures of slightly dissimilar components, are of just the same quality as the traditional elementary theory of viscosity and thermal conductivity of single gases; namely, the only important uncertainties are in the values of numerical constants [see Eq. (77)]. For mixtures of highly dissimilar gases it becomes impossible to check the dependence on concentration in detail, just as in the elementary theory of viscosity and thermal conductivity.

The sages do not consider that making no mistakes is a blessing. They believe, rather, that the great virtue of man lies in his ability to correct his mistakes and continually to make a new man of himself.—Wang Yang-Ming.