2. A THEORETICAL FRAMEWORK FOR MASS TRANSFER

This introductory chapter explains the concepts and equations we will need for understanding the subsequent chapters. All of the issues treated here have already been discussed by others, and the only reason for going over them once more is to make sure that a) everything is neatly and unambiguously defined, and b) the reader is spared the trouble of having to look everything up. Well-informed readers may find this chapter a bit boring, but are nevertheless advised to read it because the perspective and some of the definitions differ from what is customary in literature.

2.1 NOTATIONAL MATTERS

Before we engage on our mission, it is necessary to straighten out some notational matters, most of which concern the kinds of vectors that are used, and the operations applied to them. Basically, two kinds of vectors are used throughout this work:

- **Spatial vectors**, which represent quantities having both a magnitude and a direction. These vectors have, in principle, three elements, one for each co-ordinate in space. Vectors of this type are denoted by a bold-faced symbol. Examples of spatial vectors are the velocity of a body (denoted by \( \mathbf{u} \)) or of a chemical species \( i \) (\( \mathbf{u}_i \)), or the flux of that substance (\( N_i \)).

- **Component vectors**, which contain data related to the components of a mixture. For example, a mixture contains certain masses of its component species, denoted by \( m_i \). We can order these data in a component vector, which has therefore as many elements as there are species, at least, if it is to be complete. As a rule, the length of these vectors is related to the number of components \( n \), but exactly how depends on their purpose. The vectors are represented as underlined symbols (\( \underline{m} \) in this case).

A combined form is also possible. The fluxes of the individual components are ordered in spatial vectors, but these can, in turn, be ordered in component vectors. The result is a component vector with spatial vectors as its elements. The flux component vector is denoted by \( \underline{N} \).

In transfer phenomena gradients of composition-related extensive quantities frequently play an important role, for they often represent the driving forces in the process. These gradients, can also be ordered in a vector of the combined type. As an example, consider the gradients of the concentration \( c \), which are ordered in the vector \( \nabla \underline{c} \). We adopt the following notation convention:

\[
\nabla \underline{c} = \begin{pmatrix} \nabla c_1 \\ \nabla c_2 \\ \vdots \end{pmatrix} \equiv \nabla \underline{c} \tag{1}
\]

For many properties ordered in component vectors, both scalar and vectorial, it is meaningful to define the total value as the sum of all elements. In the above examples, the sum of all component masses is the total mass of the mixture, and the sum of all fluxes is the total flux. For example, consider the concentration vector \( \underline{c} \), then
Chapter 2

\[ c = \sum c \] means \[ c = \sum_{i=1}^{n} c_i \] (2)

where \( c \) stands for the total concentration. It is essential that \( c_i \) is summed over all \( n \) components. In many literature sources the total of an extensive property is denoted with the subscript ‘t’, but this is not strictly necessary, because it is a scalar and therefore cannot be mistaken for the vector (\( c \) versus \( \mathbf{c} \)). If we define a unit vector \( \mathbf{I} \) whose elements are all equal to 1, we could also write \( \mathbf{c} = \mathbf{I} \cdot \mathbf{c} \) instead of \( c = \sum c \). Further, no strict discrimination will be made between column matrices and vectors, and multiplication of a matrix by a vector is considered a legal operation as long as their sizes allow it.

2.2 Defining systems

When speaking of mass transfer, we are confronted with the question: what is transferred and in what does this happen? Clearly, mass is transferred in space, but this is too general a description. We are not concerned with the motion of cars on a highway, nor are we interested in the motions of individual molecules or atoms. Our scales of interest are those of local transport in (chemical) industrial equipment, roughly in the range between a micrometre and a millimetre. The physical surroundings in which the transfer takes place is usually called the medium. The gas, liquid, solid or, if you wish, plasma in which the transfer takes place is often also referred to as system, mixture, solution or fluid. The nature of the medium depends on its components, the thermodynamic conditions and its quantitative composition.

2.2.1 Components

In a mixture, we usually lump all entities with like properties together as a component. Obvious examples are chemical elements and compounds, but a group of similar compounds can also be considered as one component. It is not always simple to discern the components of a mixture. For example, how many components does an aqueous solution of NaCl contain? This depends on the circumstances. In simple bulk diffusion, without external forces, the smallest possible number of components is two: water and NaCl, because the Na\(^+\) and Cl\(^-\) ions cannot move independently to a significant extent without violating the electroneutrality principle. In the presence of an alternating electrical field, however, the two ionic species must be treated as separate components. Electroneutrality, which glued the ions together in the bulk diffusion case, can now be violated on a small scale, but enough to be responsible for the conductivity of the electrolyte. In solutions containing different salts, electroneutrality can be preserved in different ways, so that even for simple bulk diffusion, the ionic species must be defined as independent components.

If microporous structures such as membranes and catalysts are involved, these solid matrices can also be considered as a component, although they themselves are not transferred. More precisely: we are usually interested in transport relative to those structures. Nevertheless, they can be treated as a component because the interaction between them and the other, mobile, components influences the transfer rates. We can even go one step further by defining virtual components: in surface diffusion for example, the free sites at a surface may be thought of as a species with its own diffusional behaviour.
It is hard to give an unambiguous definition of a component. Any physical body or chemical substance that is of influence on a mass transfer process may be called a component. Often, the choice which (combinations of) substances or bodies are called components depends on circumstantial needs and limitations, but we must define how we choose them.

### 2.2.2 Thermodynamic conditions

The condition of a medium can be important. Conditions that we often encounter are:

- **isothermal**, which means that at a given time its temperature is constant throughout its entire extent,
- **isobaric**, in which the pressure is constant,
- **homogeneous**, which means that the total thermodynamic state is spatially constant, and since the state of a thermodynamic phase is determined by its temperature, pressure and composition†, homogeneity implies that these quantities must be constant throughout the mixture. Another implication is that the system consists of one phase only. A homogeneous mixture is not a very interesting one from our point of view, because on our scales of interest nothing will change, and from now on such a medium will be called strictly homogeneous. The term homogeneous will be reserved for single-phase systems both isothermal and isobaric, but possibly with varying composition.

### 2.2.3 Composition

Each component contributes to the extensive properties of a mixture, for example to its mass, volume, number of moles or internal energy. The composition is defined by the extent that the components contribute to such a property (so there are many different ways of defining composition). Mostly relative contributions of the species are used; this makes the composition independent of the total amount of mixture. These fractional contributions of the components to the property $Q$ can be ordered in composition vectors, which we shall denote as $\mathbf{x}^Q$. Compositions of mixtures are most conveniently described by the mass, mole or volume fractions of the chemical species they contain. It is customary to denote the mass, mole and volume fraction vectors by $\mathbf{\omega}$, $\mathbf{x}$ and $\mathbf{\varepsilon}$ respectively.

At this point it is useful to pay some attention to the ‘problem’ that there are different ways in which a composition-dependent quantity may be represented. If, in an $n$-component mixture, the mole fractions are chosen as the composition variables, they together form the $n$-vector $\mathbf{x}$. The sum of any complete set of fractions equals unity: $\sum \mathbf{x} = 1$. Therefore, $\mathbf{x}$ is a dependent vector with one redundant element. If the $k$th element is discarded, an independent $(n-1)$-vector $\mathbf{y}$ remains:

$$ y_i = \begin{cases} x_i & \text{if } 1 \leq i < k \\ x_{i+1} & \text{if } k \leq i \leq n-1 \end{cases} \quad (3) $$

Component $k$ is called the dependent component, while the others are said to be independent. We see that the most obvious choice for $k$ is $n$, because if $k \neq n$, the element numbers of $\mathbf{y}$ no longer correspond to the substance numbers as they were chosen originally. This

† This is true only in the absence of other possible driving forces such as electric and magnetic fields and interfacial tension differences. In this thesis these effects are assumed non-existent.
correspondence is lost for elements with numbers \( k \) and higher. This brings about undesired and unnecessary component bookkeeping. Further, we can always arrange the components in such a way that the dependent one bears the number \( n \). This does not imply that there is nothing to be said for \( k \) being a free number. In that case, we do not have to fiddle with the component order if we want to change the dependent component. Anyway, the choice for \( k = n \) is so common that it is often stated only implicitly and therefore we will also adopt it. The notations \( x \) and \( y \) for the two different composition vectors will be used from now on.

The \( x, y \)-ambiguity has some consequences for the derivatives of composition dependent functions. Any composition-dependent function, \( Q \) say, can take two forms, one operating on \( x \), \( Q(x) \), and another operating on \( y \), \( Q(y) \). Theoretically derived functions usually operate on \( x \), while empirical functions mostly operate on \( y \). The partial derivatives of these two functions are related by:

\[
\frac{\partial Q(y)}{\partial y_i} = \frac{\partial Q(x)}{\partial x_i} - \frac{\partial Q(x)}{\partial x_n}, \quad i \in \{1, \ldots, n-1\}
\]

It is assumed here that \( y \) is truly independent. The left derivative is called a constrained derivative, while those on the right are called unconstrained. As an example, consider the equation defining the activity coefficient \( \gamma_i \) of component \( i \) in a mixture:

\[
\ln \gamma_i = \left( \frac{\partial \mu_j g^E}{\partial n_j} \right)_{T,p} = \ln \left( \frac{g^E}{RT} \right) = \ln \left( \frac{g^E}{RT} \right) = \sum_{j=1}^{n} \frac{\partial g^E}{\partial x_j} (\delta_{ij} - x_j) = g^E + \sum_{j=1}^{n-1} \frac{\partial g^E}{\partial y_j} (\delta_{ij} - y_j)
\]

The last step is not obvious, but follows from substitution of eq. (4). Again, this is nothing new, but it is important to what is coming, and besides, in literature it is not always clearly indicated which composition vector or which derivative is used.

### 2.2.4 Frames of reference

Consider a mixture of \( n \) components, in which we want to observe the flows or fluxes of a certain extensive quantity \( Q \), for example mass, number of moles, volume or internal energy. We want to say something about flows, and therefore about velocities. Just as the concept of position, velocity has no independent meaning as such; it is a relative property. Therefore, a frame of reference must be fixed, which we may choose arbitrarily. But first, some things must be set straight about our conception of mixtures.

Usually, we consider a mixture as a continuum: its properties are continuous and more or less well-behaved functions in the space it occupies. However, at very small (read: molecular) scales this picture is not true: the smooth property functions break down into useless noise and lose their meaning. The continuum model is only applicable as long as the scales we work on do not enter this molecular domain.

Now the difference between intensive and extensive properties becomes important. We cannot speak of the magnitude of an intensive quantity of a certain amount of mixture without referring to its average. For example, the temperature of an amount of water is its average.
temperature. This is not so for extensive quantities: an amount of mixture represents a certain amount of the quantity, no matter how this may be distributed over the mixture. Because an intensive property is not related to the amount of substance, it is, within certain limits, independent of the scale. Imagine a point $P$ in the mixture and a spherical volume $V$ surrounding it. If the volume is large, the property considered varies throughout the volume, but as we take the average we fail to recognise them. If the volume is gradually reduced, possible variations become less and less important and the magnitude of the property will converge to a definite value, see figure 1. We define this value to be the magnitude of the quantity at $P$. But then, as we venture into the molecular domain, the amount of mixture contained in the volume becomes so small that the quantity's behaviour becomes unpredictable as a result of the disruptive effect that single atoms or molecules may have on the average value. The transition from the continuum region to the molecular region is located at volumes containing, say, $10^6$ molecules, to be on the safe side. Under standard conditions this corresponds to volumes of approximately $10^{-24}$ m$^3$ in liquids and $10^{-21}$ m$^3$ in gases.

![Figure 1](image_url)

**Figure 1.** The graph of an arbitrary extensive quantity and the corresponding volume-averaged intensive quantity as functions of the spherical volume $V$ around a point $P$. The behaviour of the extensive quantity is roughly exponential, because it is approximately inversely proportional to $V$.

Conclusion: intensive properties are suited for the continuous description of mixtures and therefore for field theories. On the other hand, extensive quantities are suited for describing bodies, which have a certain extent. So, to use extensive quantities as continuous properties, we must transform them into intensive ones by taking their average over an amount of mixture. There are several ways to do this: we can divide the properties by the mass of the sample (so-called specific quantities), by the number of moles (molar quantities) or by the volume of the sample, and we will then call them (quantity) densities.

We denote the contribution of component $i$ to the $Q$-density of a mixture by $\rho_i^Q$. The total $Q$-density of the solution is the sum of the contributions of all components: $\rho^Q = \sum_i \rho_i^Q$. Assume that, on average, each molecule of $i$ contributes an amount of $\upsilon_i^Q$ to the total $Q$-content of the...
mixture. The density \( \rho_i^Q \) then equals the molecular concentration of \( i \) times \( \psi_i^Q \): 
\[
\rho_i^Q = x_i c A \psi_i^Q = x_i \tilde{\rho}_i^Q,
\]
where \( A \) is Avogadro’s constant. The component-specific density \( \tilde{\rho}_i^Q \) is composition-dependent, and is called the partial molar \( Q \)-density of component \( i \).

\[
\rho^Q = \sum \rho_i^Q = \sum x_i^Q \rho_i^Q = x_i \tilde{\rho}_i^Q
\]

(6)

For the three most commonly used choices for \( Q \), the (partial molar) densities are listed in table 1.

**Table 1.** The (partial molar) densities for the three most common definitions of \( Q \). The symbol \( c \) stands for the total concentration, and \( \rho \) for the total (mass) density of the medium, \( M_i \) is the molar mass of component \( i \), \( c_i (= c x_i) \) its concentration, and \( \bar{v}_i \) its partial molar volume.

<table>
<thead>
<tr>
<th>( Q )</th>
<th>symbol</th>
<th>( \rho^Q )</th>
<th>( \rho_i^Q )</th>
<th>( \tilde{\rho}_i^Q )</th>
<th>( x_i^Q )</th>
<th>unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>number of moles</td>
<td>( n )</td>
<td>( c )</td>
<td>( c_i )</td>
<td>( c )</td>
<td>( x_i )</td>
<td>mol m(^{-3})</td>
</tr>
<tr>
<td>mass</td>
<td>( m )</td>
<td>( \rho )</td>
<td>( M_i c_i )</td>
<td>( M_i c )</td>
<td>( \omega_i )</td>
<td>kg m(^{-3})</td>
</tr>
<tr>
<td>volume</td>
<td>( V )</td>
<td>1</td>
<td>( \bar{v}_i )</td>
<td>( c )</td>
<td>( \varepsilon_i )</td>
<td>1</td>
</tr>
</tbody>
</table>

Now back to the main subject of this section, frames of reference. Suppose that the average velocity of the molecules of species \( i \) with respect to the arbitrary co-ordinate system is \( u_i \). This velocity will simply be called the velocity of \( i \). The displacement of \( i \) results in a flux of \( Q \), \( N_i^Q \)

\[
N_i^Q = u_i \rho_i^Q
\]

(7)

The total flux of \( Q \) is the sum of the fluxes caused by each of the species

\[
N^Q = \sum N_i^Q = p^Q \cdot u
\]

(8)

Usually, the reference frame is not arbitrary, but connected with the medium in some definite way. The most common frames are the volume-fixed, the mole-fixed, and the mass-fixed frame. In these frames no total flux exists of the quantity with respect to which they are fixed. Another frame which is sometimes encountered is the ‘solvent’-fixed frame, which is stationary relative to one of the components in the mixture. These frames are members of a whole family of possible frames. In our definition a frame is determined by a reference quantity \( q \) and a reference vector \( \mathbf{r} \). This so-called \((\mathbf{r}, q)\)-frame has a velocity \( u^{r,q} \) relative to the original, arbitrary, frame. The four common frames suggest that this reference velocity must be such that

\[
\mathbf{r} \cdot N^{r,q,q} = 0
\]

(9)

\( N^{r,q,q} \) is the \( q \)-flux vector in the new \((\mathbf{r}, q)\)-frame, which, according to equation (7), equals

\[
N_i^{r,q,q} = \tilde{\rho}_i^Q \left( u_i - u^{r,q} \right), \quad i \in \{1, \ldots, n\}
\]

(10)

The combination of the above two equations gives
For the first three frames mentioned above – volume-fixed, mole-fixed and mass-fixed –, the reference vectors are the unit vector \( \mathbf{I} \), whose elements are 1, and the reference quantities are, little surprisingly, the volume, the number of moles and the mass, respectively. For the ‘solvent’-fixed reference frame, the reference vector elements are zero except for the element that corresponds to the solvent component, which is 1. The reference quantity may be any property.

The fluxes \( \mathbf{N}^{r,q,Q} \) of a quantity \( Q \) in the \((r,q)\)-frame are, for reasons outlined in section 2.3.1, called the diffusive fluxes. These fluxes are, of course, defined by

\[
\mathbf{N}^{r,q,Q}_i = \rho^Q_i \left( \mathbf{u}_i - \mathbf{u}^{r,q}_i \right), \quad i \in \{1, \ldots, n\}
\]

It is sometimes important to know what the relation is between two diffusive fluxes of different quantities in different frames. Let us assume that for any two quantities \( Q \) and \( q \) there exists a linear transformation \( [Q<q] \) such that

\[
\rho^Q = [Q<q] \rho^q
\]

It is then easy to prove that this transformation also applies to the fluxes of \( Q \) and \( q \) relative to a common \((r,p)\)-reference frame:

\[
\mathbf{N}^{r,p,Q} = [Q<q] \mathbf{N}^{r,p,q}
\]

By extending the derivation of Toor [4], we find that two arbitrary fluxes are linked by the relation

\[
\mathbf{N}^{r,q,Q} = [Q<qP] \left( \mathbf{N}^{s,p,p} \mathbf{P}^s \mathbf{P}^r \mathbf{N}^{s,p,p} \mathbf{P}^r \mathbf{P}^s \right) \quad \text{with} \quad \mathbf{P}^r = [q<qP] \mathbf{I} \mathbf{P}^r
\]

The above definitions and expressions will be used in the next section for a concise discussion of models for mass transfer.

2.3 Two Mass Transfer Models

Mass transfer comprises two different kinds of processes: diffusive and convective transfer. Together, these two modes of transfer yield the total rate of transfer:

\[
\mathbf{N}^Q = \mathbf{N}^{r,q,Q} + \Phi^{r,q,Q}
\]

where the last term is the vector of convective fluxes. The diffusive fluxes have already been introduced in the previous section. The convective mode of transfer is relatively easily understood: it is a consequence of translation of (parts of) the medium as a whole, and the flux is simply the product of its \( Q \)-density \( \rho^Q_i \) and the reference velocity \( \mathbf{u}^{r,q}_i \) of the medium.
Although a proper description of this kind of transfer may be extremely complex and difficult for some cases such as, for instance, turbulent flow, this is never caused by a defective understanding of the theoretical basis of the transfer process. More likely is it a result of the bad times one can have in computing the flows in the medium. Diffusive mass transfer usually induces fluxes that are small in comparison with convective transfer. This often causes diffusive transfer to be rate limiting in many ‘installations’, both natural and industrial. The two modes of transfer often act simultaneously, mostly with convection as the long-distance transfer mechanism and diffusion as its close-range counterpart. This thesis is not mainly concerned with the subject of convective mass transfer, but in many cases, the two phenomena are connected.

2.3.1 Fick’s transfer model

Mass transfer in single-phase media is an important phenomenon often encountered in everyday and engineering practice, and an adequate model for its description is therefore highly desired. Ever since usable theories about diffusive mass transfer came into being, Fick’s law of diffusion has prevailed, and there seemed to be no sensible contestants. An explanation of this dominance lies in the deceptive simplicity of Fick’s law, which in its most-used form states that the diffusive mole flux of a chemical species in a homogeneous binary mixture relative to the mole-fixed reference frame is proportional to its spatial concentration gradient:

\[ N^{I,x,n} = -D \nabla \zeta, \quad n = 2 \]  

The proportionality factor \( D \) is called the diffusion coefficient or diffusivity. For several reasons the use of the concentration gradient does not seem most appropriate, and there has been some discussion over the question whether it would not be wiser to restate equation (18) as

\[ N^{I,x,n} = -c D \nabla \tilde{\chi}, \quad n = 2 \]  

which is now regarded as the better formulation. The two notations are identical only if the total concentration \( c \) does not depend on position.

Depending on the frame of reference and the flux-quantity considered, Fick’s law can take many forms. Apart from equation (19) forms in common use are (cf. Toor [4], Cussler [1]):

\[ \frac{N^{I,y,a}}{N^{I,m,m}} = -D^{I,y,a} \nabla \zeta \quad \frac{N^{I,m,m}}{N^{I,m,m}} = -\rho D^{I,m,m} \nabla \Omega, \quad n = 2 \]  

In literature diffusive mole and mass fluxes are often indicated by the symbols \( J \) and \( j \), respectively. This notation we will not adopt. It can be proven that for ideal solutions the binary diffusion coefficients in the above equations are equal. In fact, for homogeneous media there is a binary version of Fick’s equation for every combination of reference frame and flux-quantity which all have the same diffusion coefficient. If, namely, equation (19) is substituted in equation (15), we get
\[ \mathbf{N}^{r,q,Q} = -D \left( r \cdot \frac{\mathbf{p}^Q}{r} \right) \nabla \left( \frac{\rho^Q}{r} \right), \quad n = 2 \]  

where it has also been assumed that both \([Q]^{\text{ref}}\) and \([q]^{\text{ref}}\) are constant with respect to composition. As a result, \([Q]^{\text{ref}}\) is also constant (it equals \([Q]^{\text{ref}} ([q]^{\text{ref}})^{-1}\)), and so we may write

\[ \mathbf{N}^{r,q,Q} = -D \left( r \cdot \mathbf{p}^Q \right) \nabla \left( \frac{\rho^Q}{r} \right), \quad n = 2 \]  

(22)

As was mentioned above, for most frames the reference vector is the unit vector \(I\). This turns the last equation into a simple form, as could have been guessed from equation (20).

\[ \mathbf{N}^{I,q,Q} = -D \mathbf{p}^Q \nabla x^q, \quad n = 2 \]  

(23)

So, the diffusion coefficient is the same for all forms, and it satisfies the equation

\[ D \nabla x_1 = x_1 x_2 (u_2 - u_1) \]  

(24)

The fluxes appear to be functions only of the composition gradient and the velocity difference of the two components, which are both independent of the reference frame. This implies that convective transfer plays no role, which makes these fluxes diffusive by definition.

The use of Fick’s law is not limited to binary systems. There is also a multicomponent version, which reads:

\[ \mathbf{N}^{I,a,a} = -c \tilde{\mathbf{D}} \nabla x \quad \text{or} \quad \mathbf{N}^{I,a,a} = -c [\mathbf{D}] \nabla y \]  

(25)

(For the meaning of the vectors \(\nabla x\) and \(\nabla y\), see page 11.) The elements of the non-singular matrix \([\mathbf{D}]\) are related to those of \(\tilde{\mathbf{D}}\) by \(D_{ij} = \tilde{D}_{ij} - \tilde{D}_{in}\), where \(n\) is, as before, the number of the dependent component, which is usually taken to be the last component \(n\). The dimensions of the square matrices \([\mathbf{D}]\) and \(\tilde{\mathbf{D}}\) are \(n-1\) and \(n\), respectively. Analogous to the binary version, many variants of this multicomponent law exist, depending on the reference frame and the flux quantity, but here the diffusion coefficients are in general not the same. Cullinan [2] proved that these diffusivity matrices are similar – a term which is used here in the mathematical sense – and can therefore always be transformed into one another. For more information on this topic, see references [2] and [4]. As a result, all variants of Fick’s model require a minimum of \((n-1)^2\) diffusion coefficients.

Resuming, we notice that the most important features of Fick’s mass transfer model are that

• it requires \((n-1)^2\) coefficients.
• it separates diffusive and convective transfer.

Although this may not be clear at this point, there are two limitations to Fick’s law:

• it cannot incorporate other driving forces than chemical potential gradients,
• non-ideality effects of the solution are incorporated in the diffusivities.
2.3.2 The Maxwell-Stefan transfer model

An alternative to Fick’s model is presented by what is currently known as the Maxwell-Stefan or Stefan-Maxwell model, depending on which part of the world one happens to live in. It has been named after its ‘inventors’, who, in the good scientific traditions, seem to have developed it each on their own. The model takes the spatial gradient of the potential of a substance as its driving force. Further, it tells us that the opposing forces are caused by the velocity differences of the species. All this is expressed in the formula for an \( n \)-component mixture

\[
\nabla \mu_i = RT \sum_{j \neq i} x_j \frac{u_j - u_i}{D_{ij}} = \sum_{j \neq i} \zeta_{ij} x_j (u_j - u_i), \quad i \in \{1, \ldots, n\}
\]

where the \( u_i \) denote the same molar average velocities as in section 2.2.4. The coefficients \( D_{ij} \) and \( \zeta_{ij} (= RT/D_{ij}) \) are called the Maxwell-Stefan diffusion coefficients and friction coefficients, respectively. Here, the form where only the chemical potential acts as the motive force is given. More forces can be added without difficulties, and a discussion of the span of applications, as well as of the exact form the equations then take is given elsewhere [3, 5].

With subsequent use of the equations (7) and (16) these Maxwell-Stefan equations can also be written as

\[
\frac{x_i \nabla \mu_i}{RT} = \sum_{j \neq i} x_j \frac{N_j^a - x_j N_i^a}{c D_{ij}} = \sum_{j \neq i} x_j N_{ij}^{r,q,a} - x_j N_i^{r,q,a} \quad , \quad i \in \{1, \ldots, n\}
\]

This is the most common form, but other flux quantities could have been chosen just as well. Due to the Onsager equations \((D_{ij} = D_{ji})\), the sum of the right sides of all \( n \) transport equations vanishes, regardless of the fluxes \( N_{ij}^{r,q,a} \), and so we get \( \mathbf{v} \nabla \mu = 0 \), an expression which is also known as the Gibbs-Duhem equation. In other words, one of the \( n \) transport equations can be discarded without any loss of information. The remaining equations can also neatly be moulded in a matrix form:

\[
c y_i \nabla \mu_i = RT \sum_{j=1}^n \beta_{i,j} \hat{N}_{ij}^{r,q,a} , \quad i \in \{1, \ldots, n-1\}
\]

The hat indicates that we are dealing with the complete flux vector. The elements of the \((n-1 \times n)\)-matrix \([\beta]\) are given by

\[
\beta_{i,j} = \frac{x_j}{D_{ij}}, \quad \beta_{i,i} = -\sum_{j \neq i} \frac{x_j}{D_{ij}} \quad \text{if} \quad i \in \{1, \ldots, n-1\}, \quad j \neq i
\]

The chemical potentials of the various substances are \( \mu_i = \mu_{i}^{0} + RT \ln(x_i \gamma_i) \), but \( i \neq n \) and therefore \( x_i = y_i \). So, \( y_i \nabla \mu_i = RT (y_i \nabla \ln \gamma_i + \nabla y_i) \), and substitution of the expression for \( \nabla \ln \gamma_i \) yields

\[
c \Gamma \nabla y = [eta] \hat{N}^{r,q,a} \quad \text{where} \quad \Gamma_{i,j} = y_j \frac{\partial \ln \gamma_i}{\partial y_j} + \delta_{ij}
\]

Our goal is to find a solution for the \( n \) fluxes. Unfortunately, the matrix \([\beta]\), being non-square, has no inverse. Everything depends upon how the fluxes are interrelated, and thus upon the reference frame. The lack of a sufficient number of equations is made up for by an equation
relating the fluxes called the bootstrap equation. For now, let us assume that it is of the linear form†

\[ S \cdot \hat{N}^{r,q,a} = 0 \quad \Rightarrow \quad N_n^{r,q,a} = -\sum_{j \neq k} \frac{S_j}{S_n} N_j^{r,q,a} \]  
(31)

where \( \xi \) may be any scalar vector. With this expression we can rewrite eq. (30) as

\[ c[\Gamma] \nabla y = [B] \vec{N}^{r,q,a} \quad \text{with} \quad B_{i,j} = \beta_{i,j} - \frac{S_i}{S_n} \beta_{i,n} \quad \text{for} \quad i, j \in \{1, \ldots, n-1\} \]  
(32)

Note that the flux vector now has only \( n-1 \) elements: \( N_n^{r,q,a} \) has disappeared. This is the set of equations we need to solve the fluxes. It is also often written in a flux-explicit form

\[ \vec{N}^{r,q,a} = -c[G] \nabla y \quad \text{with} \quad [G] = -[B]^{-1} [\Gamma] \]  
(33)

The contents of this section can be summarised by stating that the Maxwell-Stefan model

- needs only \( n (n-1)/2 \) coefficients,
- does not necessarily separate convection and diffusion,
- can easily be extended to account for driving forces other than chemical potential gradients,
- does not lump ideality and diffusive effects.

2.3.3 Choose your weapon

So, there we are, having two models at our disposal which perform the same task: describing mass transfer. Which of the two should we choose? It is clear that if other forces than chemical potential gradients are involved, the Maxwell-Stefan model is the only possible choice. If such forces are absent the models must be essentially equivalent. Indeed, the coefficients of one model can be transmuted into those of the other model [3]. Therefore, if we choose a specific model, we do so on account of its form rather than its function. Undeniably, the Maxwell-Stefan model is the more versatile model, but in practice most diffusivity data available today are Fick diffusivities, because these can be measured directly, see also chapter 3. To obtain the Maxwell-Stefan diffusion coefficients, we must be acquainted with the thermodynamics of the system in order to compute the \( \Gamma \)-matrix. Nevertheless, this weakness of the model is also its strength: the fact that it separates the diffusivities from the thermodynamic behaviour facilitates the estimation of the coefficients at compositions at which they are not known, especially in non-ideal mixtures [3, 5]. Moreover, the Maxwell-Stefan diffusivities often appear to have a more fundamental meaning than the Fick diffusivities. Anyway, our choice for either model is determined by the type of problem and by the nature of the available data.

† Bootstrap equations by no means need to be linear: in chapter 6 we will encounter a nonlinear equation, and we will see that this gravely complicates the solution of the fluxes.
Chapter 2

NOTATION

Note: symbols for local use (which occur in a definite meaning in only one passage) are not listed here.

\[ B \] (\(n-1\times n-1\)) coefficient matrix of the Maxwell-Stefan equations (32), (s m\(^{-2}\))
\(c\) concentration, (mol m\(^{-3}\))
\(D\) Fick diffusion coefficient, (m\(^2\) s\(^{-1}\))
\(D\) Maxwell-Stefan diffusion coefficient, (m\(^2\) s\(^{-1}\))
\[ G \] Maxwell-Stefan ‘diffusivity’ matrix, equation (33), (m\(^2\) s\(^{-1}\))
\(I\) unit component vector, \(I_i = 1 \forall i \in \{1,\ldots, n\}\)
\(M\) molar mass, (kg mol\(^{-1}\))
\(n\) number of components, (–)
\(\#\) number of moles, (mol)
\(N^Q\) \(Q\)-flux, (((\(Q\)) m\(^{-2}\) s\(^{-1}\))
\(N^{r,q,Q}\) diffusive \(Q\)-flux relative to \((r, q)\)-frame, (((\(Q\)) m\(^{-2}\) s\(^{-1}\))
\(R\) gas constant, (J mol\(^{-1}\) K\(^{-1}\))
\(T\) absolute temperature, (K)
\(Q^{q}\) transformation mapping a quantity \(q\) onto a quantity \(Q\), (((\(Q\)) (\(q\))\(^{-1}\))
\(u^{r,q}\) reference velocity of \((r, q)\)-frame of reference, (m s\(^{-1}\))
\(\tilde{v}\) partial molar volume, (m\(^3\) mol\(^{-1}\))
\(x\) mole fraction, (–)
\(x^Q\) \(Q\)-fraction, (–)
\(y\) mole fraction, (–)

Greek symbols
\[ \Gamma \] non-ideality matrix, (–)
\(\delta\) Kronecker delta
\(\gamma\) activity coefficient
\(\mu\) chemical potential, (J mol\(^{-1}\))
\(\rho^Q\) \(Q\)-density, (((\(Q\)) m\(^{-3}\))
\(\tilde{\rho}^Q\) partial molar \(Q\)-density, (((\(Q\)) m\(^{-3}\))
\(\omega\) mass fraction (–)
\(\Phi^{r,q,Q}\) convective \(Q\)-flux relative to \((r, q)\)-frame, (((\(Q\)) m\(^{-2}\) s\(^{-1}\))
\(\zeta\) friction coefficient, (kg mol\(^{-1}\) s\(^{-1}\))

Subscripts
\(i\) component number \(i\)
\(j\) component number \(j\)
\(ij\) component pair \(i, j\)
\(n\) last component, number \(n\)
\(T, P\) at constant temperature and pressure
Superscripts

$m$ mass

$\#$ (number of) moles, or amount

$o$ pure component

$V$ volume

$T$ transpose of a matrix

REFERENCES

[1] Cussler E.L.
‘Diffusion. Mass transfer in fluid systems.’

‘Analysis of the flux equations of multicomponent diffusion.’

‘Multicomponent mass transfer.’

‘Reference frames in diffusion.’

‘Mass transfer.’