CHAPTER 1

MATHEMATICAL INTRODUCTION

1.1 HILBERT SPACE

1.1.1 Eigenfunction and Electronic Nonadiabatic Coupling Term

We consider a complete basis set of electronic eigenfunctions $|\zeta_k(s_e|s)|$; $k = 1, 2, \ldots, N$, which depend explicitly on the electronic coordinate $s_e$ and parametrically on the nuclear coordinate $s$. The $|\zeta_k(s_e|s)|$ functions are the eigenfunctions of the following electronic Hamiltonian $H_e(s_e|s)$

$$ (H_e(s_e|s) - u_k(s)) |\zeta_k(s_e|s)| = 0; \quad k = 1, \ldots, N \quad (1.1) $$

where $u_k(s)$ are the electronic eigenvalues [which later are recognized as the adiabatic potential energy surfaces (PESs)].

The fact that the $|\zeta_k(s_e|s)|$ functions form a complete set yields the resolution of the unity in the following way:

$$ I = \sum_{k=1}^N |\zeta_k(s_e|s)\rangle \langle \zeta_k(s_e|s)| \quad (1.2) $$

This equation (1.2) guarantees that an arbitrary function $|\xi(s_e|s)|$ can be expressed in terms of a linear combination of the $|\zeta_k(s_e|s)|$ functions with coefficients (which parametrically depend on the nuclear coordinates):

$$ |\xi(s_e|s)| = \sum_{k=1}^N |\zeta_k(s_e|s)\rangle \langle \zeta_k(s_e|s)|\xi(s_e|s)|. \quad (1.3) $$
2 MATHEMATICAL INTRODUCTION

This expansion is known as Born–Oppenheimer expansion, and the bracket notation, introduced by Dirac and used throughout this book, implies integration over the electronic coordinates.

Next we form the following magnitude

$$\tau_{jk} = \langle \zeta_j | \nabla \zeta_k \rangle; \quad k, j = \{1, 2, \ldots, N\}$$  \hspace{1cm} (1.4)

which is defined as the electronic nonadiabatic coupling term (NACT). Here the grad operator is expressed in terms of the nuclear coordinate

$$\nabla = \left\{ \frac{\partial}{\partial p}, \frac{\partial}{\partial q}, \ldots \right\}$$  \hspace{1cm} (1.5)

where $p$ and $q$ are two Cartesian (mass-scaled) coordinates. It is important to note that the elements $\tau_{jk}$ form a matrix $\tau$ with elements that are vectors.

With these definitions we form the Hilbert space defined at a given point $s$, namely, the space that contains all the functions that solve Eq. (1.1) at a given point $s$. One of the more important tasks is to connect Hilbert spaces defined at different points, for instance, at $s$ and $s + \Delta s$. It can be shown that a connection is established through the NACTs

$$|\zeta_k(s_e|s + \Delta s)\rangle = \sum_{j=1}^{N} (\delta_{kj} - \Delta s \cdot \tau_{kj})|\zeta_j(s_e|s)\rangle$$  \hspace{1cm} (1.6)

where $\delta_{jk}$ is the Kronecker delta function and the dot stands for a scalar product. Equation (1.6) is always fulfilled if the Hilbert spaces at point $s$ and at points close to it are of $N$ dimensions (viz., contain $N$ eigenfunctions). However, we show later that this relation holds, under certain conditions, also for a smaller group of states that is defined as a Hilbert subspace.

To prove Eq. (1.6) we consider the Taylor expansion for $|\zeta_k(s_e|s + \Delta s)\rangle$:

$$|\zeta_k(s_e|s + \Delta s)\rangle = |\zeta_k(s_e|s)\rangle + \Delta s \cdot |\nabla \zeta_k(s_e|s)\rangle$$  \hspace{1cm} (1.7)

Since the derivative $|\nabla \zeta_k(s_e|s)\rangle$ is also a function that belongs to the same Hilbert space, it can be presented in terms of the (electronic) basis set introduced earlier:

$$|\nabla \zeta_k(s_e|s)\rangle = \sum_{j=1}^{N} Z_{kj}(s)|\zeta_j(s_e|s)\rangle$$  \hspace{1cm} (1.8)

Multiplying both sides by $\langle \zeta_i(s_e|s) \rangle$ (which also implies integration over $s_e$) yields [see Eq. (1.4)]

$$Z_{ki} = -\tau_{ki}; \quad i, k = \{1, 2, \ldots, N\}$$  \hspace{1cm} (1.9)
1.1 HILBERT SPACE

Substituting Eqs. (1.8) and (1.9) in Eq. (1.7) yields Eq. (1.6). In what follows we consider only real eigenfunctions, for which it is easy to show that the diagonal elements of the $\tau$ matrix, namely, $\tau_{jj}$, are identically zero. As a result, Eq. (1.6) leads to an important relation known as the parallel transport law:

$$\langle \zeta_k(s_e | s) | \zeta_k(s_e | s + \Delta s) \rangle = 1 + O(\Delta s^2)$$

Equation (1.6) can also be written employing matrix notation

$$\zeta(s_e | s + \Delta s) = (I - \Delta s \cdot \tau)\zeta(s_e | s)$$

where $I$ is the unity matrix.

1.1.2 Abelian and Non-Abelian Curl Equations

In what follows we distinguish between Abelian and non-Abelian magnitudes. Abelian magnitudes are usually vectors and therefore field operations act on them in the ordinary way as, for instance, in electrodynamics. As an example, we may consider the definition of the Curl equation, for a vector $\tau$ that is to be of the form

$$H = \text{Curl} \, \tau$$

where $H$ is also a vector. The Curl operator implies that the $z$ component takes the form

$$H_z = \frac{\partial \tau_x}{\partial y} - \frac{\partial \tau_y}{\partial x}$$

The situation changes significantly in case $\tau$ (and consequently also $H$) becomes a matrix (of vectors). This situation is known as the non-Abelian case, and as will be shown next, the fact that vectors are replaced by matrices of vectors affects relations known to exist between ordinary vector functions.

As an example we consider the non-Abelian Curl equation, which is sometimes also called the extended Curl equation (to avoid confusion with the ordinary Abelian Curl equation). This equation, as mentioned earlier, is written in a matrix form and for this purpose we introduce a few definitions:

Considering two (nuclear) Cartesian coordinates $p$ and $q$, we define the following tensorial vector $F_{pq}$:

$$F_{pq} = \frac{\partial \tau_p}{\partial q} - \frac{\partial \tau_q}{\partial p} - [\tau_p, \tau_q]$$
where $\tau_\lambda; \lambda = q, p$ are the $\lambda$ components of $\tau$, defined as [see Eq. (1.4)]

$$\tau_{\lambda jk} = \left\{ \xi_j \left| \frac{\partial}{\partial \lambda} \zeta_k \right. \right\}; \{ k, j = 1, \ldots, N; \lambda = p, q \} \quad (1.14)$$

and $[\tau_p, \tau_q]$ is the commutation relation between $\tau_p$ and $\tau_q$. Equation (1.13) can be written in a more compact way

$$F = H - T \quad (1.15)$$

where

$$H_{pq} = \frac{\partial \tau_p}{\partial q} - \frac{\partial \tau_q}{\partial p} \Rightarrow H = \text{Curl } \tau \quad (1.16)$$

and

$$T_{pq} = [\tau_p, \tau_q] \Rightarrow T = [\tau \times \tau] \quad (1.17)$$

$F$ is the field tensor known also as the Yang–Mills field $^{6b, 7-10}$ Next we prove the following lemma.

**Lemma 1.1** For a Hilbert space, $F$ has to be identically zero $^{11, 12}$

$$F = 0 \quad (1.18)$$

**Proof** We consider the $p$th and $q$th components of Eq. (1.16):

$$\left( \frac{\partial \tau_p}{\partial q} \right)_{jk} = \left\{ \xi_j \left| \frac{\partial \zeta_k}{\partial q} \right. \right\} + \left\{ \xi_j \left| \frac{\partial^2 \zeta_k}{\partial q \partial p} \right. \right\} \quad (1.19a)$$

$$\left( \frac{\partial \tau_q}{\partial p} \right)_{jk} = \left\{ \xi_j \left| \frac{\partial \zeta_k}{\partial p} \right. \right\} + \left\{ \xi_j \left| \frac{\partial^2 \zeta_k}{\partial p \partial q} \right. \right\} \quad (1.19b)$$

Subtracting Eq. (1.19b) from Eq. (1.19a) and assuming that the eigenfunctions are analytic functions with respect to the nuclear coordinates yields the following result:

$$\left( H_{pq} \right)_{jk} = \left( \frac{\partial}{\partial q} \tau_p - \frac{\partial}{\partial p} \tau_q \right)_{jk} = \left\{ \frac{\partial \xi_j}{\partial q} \left| \frac{\partial \zeta_k}{\partial q} \right. \right\} + \left\{ \frac{\partial \xi_j}{\partial p} \left| \frac{\partial \zeta_k}{\partial p} \right. \right\} \quad (1.20)$$

Next we consider the $(j, k)$ element of the first term in Eq. (1.17):

$$\left( \tau_p \tau_q \right)_{jk} = \sum_{i=1}^{N} \left\{ \xi_j \left| \frac{\partial \zeta_i}{\partial p} \right. \right\} \left\{ \xi_k \left| \frac{\partial \zeta_i}{\partial q} \right. \right\} \quad (1.21)$$
and recall the following relation (that holds for real functions):

\[
\langle \xi_j | \frac{\partial \xi_i}{\partial p} \rangle = - \langle \frac{\partial \xi_j}{\partial p} | \xi_i \rangle
\]

Eq. (1.21) becomes

\[
(\tau_p \tau_q)_{jk} = - \sum_{i=1}^{N} \left( \frac{\partial \xi_j}{\partial p} | \xi_i \rangle \langle \xi_i | \frac{\partial \xi_k}{\partial q} \right) = - \left( \frac{\partial \xi_j}{\partial p} \right) \sum_{i=1}^{N} (|\xi_i \rangle \langle \xi_i | \frac{\partial \xi_k}{\partial q} \right)
\]

or, due to the resolution of the unity [see Eq. (1.2)], we finally get

\[
(\tau_p \tau_q)_{jk} = - \left( \frac{\partial \xi_j}{\partial p} | \frac{\partial \xi_k}{\partial q} \right)
\]

(1.22a)

A similar result is obtained for the second term in Eq. (1.17):

\[
(\tau_q \tau_p)_{jk} = - \left( \frac{\partial \xi_j}{\partial q} | \frac{\partial \xi_k}{\partial p} \right)
\]

(1.22b)

Subtracting Eq. (1.22b) from Eq. (1.22a) yields the following \((jk)\) element of \(T_{pq}\):

\[
(T_{pq})_{jk} = (\tau_p \tau_q)_{jk} - (\tau_q \tau_p)_{jk} = - \left( \left( \frac{\partial \xi_j}{\partial p} | \frac{\partial \xi_k}{\partial q} \right) - \left( \frac{\partial \xi_j}{\partial q} | \frac{\partial \xi_k}{\partial p} \right) \right)
\]

(1.23)

Next, comparing Eqs. (1.23) and (1.20), it is readily noted that for any arbitrary components \(p\) and \(q\) we have

\[
H_{pq} = T_{pq} \Rightarrow F_{pq} = H_{pq} - T_{pq} = 0
\]

(1.24)

which implies that Eq. (1.18) is fulfilled for a complete Hilbert space.

As a final issue in this section we consider the two-state \((N = 2)\) case. In this case the matrix \(T\) is identically zero so that Eq. (1.15) [see also Eq. (1.13)] becomes

\[
F = H = \text{Curl } \tau = 0
\]

(1.25)

which, because of the following structure of the \(2 \times 2\) \(\tau\) matrix

\[
\tau = \begin{pmatrix}
0 & \tau_{12} \\
-\tau_{12} & 0
\end{pmatrix}
\]

(1.26)

leads to the result that

\[
\text{Curl } \tau_{12} = 0
\]

(1.27)
6 MATHEMATICAL INTRODUCTION

This result implies that for any pair of coordinates \((p, q)\) we have (in case the Hilbert space is formed by two states)

\[
\frac{\partial \tau_{p12}}{\partial q} - \frac{\partial \tau_{q12}}{\partial p} = 0 \quad (1.28)
\]

which is the Abelian form of the Curl equation.

**Corollary 1.1** Equation (1.28) indicates that in case of two states \((N = 2)\) the non-Abelian case becomes Abelian.

We want to complete this section with two comments:

1. We emphasize the fact that all the derivations in this section apply as long as the electronic eigenfunctions are analytic functions at every point in the region of interest. In case certain eigenfunctions may not have first-order derivatives at some points in configuration space, for instance, at conical intersections (see Chapter 5), the respective derivative of the \(\tau\)-matrix elements are not defined at these points and therefore also the final outcome, namely, that \(F = 0\) is unlikely to hold at these points.

2. The importance of the fact that \(F = 0\) for Born–Oppenheimer systems and how it is related to other fields in physics is briefly discussed by Englman and Yahalom.9

1.1.3 Abelian and Non-Abelian Divergence Equations

To guarantee the connection between Hilbert spaces defined at relative remote points, specifically, where \(\Delta s\) is relatively large, one has to include in Eq. (1.6) terms that contain both higher powers of \(\tau_{ij}\) and higher-order derivatives of \(\tau_{ij}\). In this respect we prove next the relation between the second-order NACTs \(\tau_{jk}^{(2)}\) defined as

\[
\tau_{jk}^{(2)} = \langle \zeta_j | \nabla^2 \zeta_k \rangle; \quad k, j = \{1, 2, \ldots, N\} \quad (1.29)
\]

and the ordinary (first-order) one \(\tau_{ij}\) introduced in Eq. (1.4).

**Lemma 1.2** The second-order NACT \(\tau_{jk}^{(2)}\) can be presented in the form\(^{11–13}\)

\[
\tau_{jk}^{(2)} = \sum_{i=1}^{N} \tau_{ji} \tau_{ik} + \nabla \tau_{jk} \quad (1.30)
\]

or in matrix notation

\[
\tau^{(2)} = \tau^2 + \nabla \tau \quad (1.31)
\]
1.1 HILBERT SPACE

Proof  To derive Eq. (1.30) we consider $\nabla \tau_{kj}$. Recalling Eq. (1.4), we get

$$
\nabla \tau_{jk} = \nabla \langle \zeta_j | \nabla \zeta_k \rangle = \langle \nabla \zeta_j | \nabla \zeta_k \rangle + \langle \zeta_j | \nabla^2 \zeta_k \rangle
$$

(1.32)

To continue, we need to evaluate the first term on the right-hand side (r.h.s.) of this expression. For this purpose we employ the resolution of the unity given in Eq. (1.2)

$$
\langle \nabla \zeta_j | \nabla \zeta_k \rangle = \langle \nabla \zeta_j | \sum_{i=1}^{N} \langle \zeta_i | \zeta_i \rangle \nabla \zeta_k \rangle = \sum_{i=1}^{N} \langle \nabla \zeta_j | \zeta_i \rangle \langle \zeta_i | \nabla \zeta_k \rangle = \sum_{i=1}^{N} \tau_{ij} \cdot \tau_{ik}
$$

or, finally

$$
\langle \nabla \zeta_j | \nabla \zeta_k \rangle = -\sum_{i=1}^{N} \tau_{ji} \cdot \tau_{ik}
$$

(1.33)

Substituting Eq. (1.33) in Eq. (1.32) (and recalling the definitions given above) yields Eq. (1.30) [or Eq. (1.31)].

Corollary 1.2  Equation (1.31) was originally derived for the purpose of presenting the adiabatic nuclear Schrödinger equation in a complete and efficient form that guarantees yielding significant physical insight (see Section 2.1.1). However, it also serves another purpose and thus will be written in a slightly different form:

$$
\nabla \tau = \tau^{(2)} - \tau^2
$$

(1.34)

This equation is recognized as the extended divergence (div) equation, which together with the extended Curl equation [see Eqs. (1.15)–(1.18)] forms the Curl–Div equations that the $\tau$ matrix has to fulfill. These two equations are reminiscent of the Curl–Div equations for the vector potential in electrodynamics.\textsuperscript{14,15}

The two-state system is of special interest because in this case the non-Abelian situation becomes Abelian [i.e., the equations for the matrices become equations for the single (1,2) term]; thus, in case of the divergence equation, we have

$$
\nabla \tau_{12} = \tau_{12}^{(2)}
$$

(1.35)

(The two-state $\tau^2$ matrix produces only diagonal elements.) Equation (1.35) indicates that the divergence equation encountered in molecular physics is, in general, different from zero.

Equations (1.27) and (1.35) form the Curl–Div equations in case of a two-dimensional Hilbert space.
8 MATHEMATICAL INTRODUCTION

1.2 HILBERT SUBSPACE

Complete Hilbert spaces, in particular for realistic molecular systems, contain an infinite number of states; therefore, to be able to treat such systems numerically, we have to reveal the conditions or establish situations for which finite groups of states will behave as a quasi-complete Hilbert space. If such a group of states can be formed, we define it as a Hilbert subspace.

In the present study the breakup of the Hilbert space into Hilbert subspaces is dictated by the behavior of the previously introduced NACTs elements $\tau_{jk}$. It will be shown that the features that exist within a complete Hilbert space, for instance, the Curl–Div equations and the quantization (to be discussed in Chapters 2 and 3), are approximately valid within the respective Hilbert subspace.

In what follows we consider a group of $N$ states, and for convenience we assume them to be the $N$ lowest states. Next we assume the Hilbert space to break up into two groups: a finite group, designated as the $P$ space, which contains $N$ (lowest) states, and the complementary group, designated as the $Q$ space, which contains the rest of the Hilbert space (its dimensions can be infinite). As already mentioned, the breakup of the Hilbert space is based on the following assumption:1

$$|\tau_{jk}| \cong O(\varepsilon) \quad \text{for} \quad j \leq N; \quad k > N \quad (1.36)$$

Here $\varepsilon$ is a relatively small number. Equation (1.36) indicates that the NACTs between $P$ states and $Q$ states are all assumed to be negligibly small. This implies that the $\tau$ matrix has the following form:2

$$\tau = \begin{pmatrix} 0 & \tau_{12} & \tau_{13} & \tau_{1N} \\ -\tau_{12} & 0 & \tau_{23} & \tau_{2N} \\ -\tau_{13} & -\tau_{23} & 0 & \tau_{3N} \\ -\tau_{1N} & -\tau_{2N} & -\tau_{3N} & 0 \end{pmatrix} + O(\varepsilon)$$

$$+ \begin{pmatrix} 0 & \tau_{N+1N+2} & \tau_{N+1N+3} \\ 0 & 0 & \tau_{N+2N+3} \\ -\tau_{N+1N+2} & 0 & \tau_{N+2N+3} \\ -\tau_{N+1N+3} & -\tau_{N+2N+3} & 0 \end{pmatrix} + O(\varepsilon) \quad (1.37)$$

(Comment: In Section 6.5 this assumption is proved to exist in given regions in configuration space.)

To continue, we define the following Feshbach projection operators,3,4 namely, $P_N$, the projection operator for the $P$ space

$$P_N = \sum_{j=1}^{N} |\zeta_j\rangle\langle\zeta_j| \quad (1.38a)$$
1.2 HILBERT SUBSPACE

and \( Q_N \), the projection operator for the \( Q \) space

\[
Q_N = I - P_N
\] (1.38b)

Our aim is to use the Feshbach operators in order to show that both the extended Curl equation given in Eqs. (1.15)–(1.18) and the Div equation given in Eq. (1.34) are approximately fulfilled within this Hilbert subspace. Since the mathematical proof in both cases is similar, we present the proof just for the divergence equation. The proof for the extended Curl equation can be found in Refs. 5, 6, and 7.

Our starting point is, like before, Eq. (1.30) which, following a slight rearrangement based on Eq. (1.33), takes the form

\[
\tau_{\text{\footnotesize{(2)}}}^{jk} = -\langle \nabla \xi_j | \nabla \xi_k \rangle + \nabla \tau_{\text{\footnotesize{(1)}}}^{jk} \] (1.39)

It is important to realize that Eq. (1.39) is valid for any two states \( j \) and \( k \) and therefore also for those that fulfill \( j, k \leq N \). In what follows we consider only states belonging to the subspace. The first term on the r.h.s. is treated further as

\[
\langle \nabla \xi_j | \nabla \xi_k \rangle = \langle \nabla \xi_j | P_N + Q_N | \nabla \xi_k \rangle = \langle \nabla \xi_j | P_N | \nabla \xi_k \rangle + \langle \nabla \xi_j | Q_N | \nabla \xi_k \rangle
\] (1.40)

where the first of the two terms in Eq. (1.40) becomes

\[
\langle \nabla \xi_j | P_N | \nabla \xi_k \rangle = \sum_{i=1}^{N} \langle \nabla \xi_j | \xi_i \rangle \langle \xi_i | \nabla \xi_k \rangle = \sum_{i=1}^{N} (-\tau_{ji}) \tau_{ik}
\] (1.41a)

and the second

\[
\langle \nabla \xi_j | Q_N | \nabla \xi_k \rangle = \sum_{i>N} \langle \nabla \xi_j | \xi_i \rangle \langle \xi_i | \nabla \xi_k \rangle = \sum_{i>N} (-\tau_{ji}) \tau_{ik}
\] (1.41b)

Recalling Eq. (1.37), we note that the contribution due to Eq. (1.41b) is negligibly small (because both \( j, k \leq N \)), and therefore Eq. (1.40) becomes

\[
\langle \nabla \xi_j | \nabla \xi_k \rangle = \langle \nabla \xi_j | P_N | \nabla \xi_k \rangle = -\sum_{i=1}^{N} (\tau_{N})_{ji}(\tau_{N})_{ik}
\] (1.42)

which yields the following result for Eq. (1.39)

\[
(\tau_{N})_{\text{\footnotesize{(2)}}}^{jk} = \sum_{i=1}^{N} (\tau_{N})_{ji}(\tau_{N})_{ik} + \nabla (\tau_{N})_{jk}
\] (1.43)

Equation (1.43) is essentially identical to Eq. (1.30), which was derived for the complete Hilbert space. Equation (1.43) can now also be written as an equation between
the submatrices $\tau_N$, $\nabla \tau_N$, and $\tau^{(2)}_N$

$$\tau^{(2)}_N = \tau^2_N + \nabla \tau_N \quad (1.44)$$

which is similar to Eq. (1.31). Writing Eq. (1.44) slightly differently, as

$$\nabla \tau_N = \tau^{(2)}_N - \tau^2_N \quad (1.45)$$

yields the corresponding divergence equation.

We reiterate that one can show, employing similar considerations, that the Curl equations in Eqs. (1.15)–(1.18) can be written in the same way for a Hilbert subspace but for the corresponding $N \times N$ submatrices.

As the last subject on this issue, we would like to estimate the size of the error introduced by ignoring the contributions from the Q subspace. For this purpose we consider the Curl equation or, more specifically, the two components of the (unperturbed) vectorial $\tau$ matrix, namely, the matrices $\tau_q$ and $\tau_p$ with elements as presented in Eq. (1.14). Each of these matrices is now written [following the presentation in Eq. (1.37)] as follows:

$$\tau_x = \begin{pmatrix} \tau^{(N)}_x & \tau^{(N,L)}_x \\ \tau^{(L,N)}_x & \tau^{(L)}_x \end{pmatrix}; \quad x = q, p \quad (1.46)$$

Here $\tau^{(N)}_x$ and $\tau^{(L)}_x; x = p, q$ are the diagonal submatrices that contain the dominant NACTs, whereas $\tau^{(N,L)}_x$ [and $\tau^{(L,N)}_x]; x = p, q$ are the two off-diagonal submatrices that couple the two diagonal ones and are assumed to contain weak coupling terms, all of the order $O(\varepsilon)$. Next, following Eq. (1.13), we introduce the components of the reduced tensorial vector $\tilde{F}^{(N)}_{pq}$:

$$\tilde{F}^{(N)}_{pq} = \frac{\partial \tau^{(N)}_p}{\partial q} - \frac{\partial \tau^{(N)}_q}{\partial p} - [\tau^{(N)}_p, \tau^{(N)}_q] \quad (1.47)$$

Recalling Eqs. (1.15)–(1.17), the $N \times N$ upper diagonal block of $F$, specifically, $F^{(N)}_{pq}$, can be written in the form

$$F^{(N)}_{pq} = \tilde{F}^{(N)}_{pq} - \left\{ \tau^{(N,L)}_p \tau^{(L,N)}_q - \tau^{(N,L)}_q \tau^{(L,N)}_p \right\} \quad (1.48)$$

but because of Eq. (1.18) we have $\tilde{F}^{(N)}_{pq} = 0$ and therefore we get for $F^{(N)}_{pq}$ the following result:

$$F^{(N)}_{pq} = \left\{ \tau^{(N,L)}_p \tau^{(L,N)}_q - \tau^{(N,L)}_q \tau^{(L,N)}_p \right\} \quad (1.47')$$
1.3 VECTORIAL FIRST-ORDER DIFFERENTIAL EQUATION AND LINE INTEGRAL

Since all four matrices \( \tau_x^{(N,L)} \) and \( \tau_x^{(L,N)} \); \( x = p, q \) contain elements of the order \( O(\varepsilon) \) (or smaller), the elements of the matrix on the r.h.s. of Eq. (1.48) are of the order \( O(\varepsilon^2) \), which implies that the relevant elements of \( F_{pq}^{(N)} \) are of the following order:

\[
F_{pq}^{(N)} = O(\varepsilon^2) \tag{1.49a}
\]

In other words, the extended Curl equation within the Hilbert subspace is fulfilled up to \( O(\varepsilon^2) \). A similar equation holds for the divergence (Div) equation

\[
\nabla \tau_N - \tau_N^{(2)} + \tau_N^2 = O(\varepsilon^2) \tag{1.49b}
\]

where \( \tau_N \) is identical to the previously defined \( \tau^{(N)} \).

Because of these findings we do not distinguish any more between complete, usually infinite, Hilbert spaces and finite Hilbert subspaces. However, we do distinguish between groups of states that form a Hilbert subspace and those that do not form a Hilbert subspace.

1.3 VECTORIAL FIRST-ORDER DIFFERENTIAL EQUATION AND LINE INTEGRAL

In this section we consider various issues concerning the following first-order differential equation\(^1\)

\[
\nabla \Omega(s) + \tau(s)\Omega(s) = 0 \tag{1.50}
\]

where \( \tau \) is a vector–matrix and \( \Omega \) is a scalar–matrix and \( s \) is a point in configuration space. Equation (1.50) was probably mentioned for the first time by Hobey and McLachlan\(^3\) (although in the context of two states only). However, no attempts were made to study it. On the contrary, McLachlan, in a follow-up publication,\(^4\) termed them as “inconsistent” and concluded that they are relevant only when \( s \) stands for a (single) Cartesian coordinate.

Equation (1.50) has to be integrated from a point \( s_0 \) to a point \( s \), and this integration is usually done by assuming a contour \( \Gamma \) that combines \( s_0 \) and \( s \). However, before doing that we devote the next few sections to investigating these equations with the aim of obtaining the conditions to be satisfied by \( \tau \) in order to guarantee an analytic solution in a given region of configuration space (and if not analytic, then, at least, as close as possible to being analytic as will be discussed later). This investigation is carried out first by considering the differential equation as written above (Section 1.3.1) or its corresponding integral equation (Section 1.3.3). In order to simplify this study, we consider first the case where \( \tau \) is an ordinary vector (the Abelian case) and then extend the treatment to the case where \( \tau \) is a vector–matrix (the non-Abelian case).
12 MATHEMATICAL INTRODUCTION

1.3.1 Vectorial First-Order Differential Equation

1.3.1.1 Study of Abelian Case

In this section \( \tau \) is a vector, namely, \((\tau_x, \tau_y, \ldots)\), and consequently \( \Omega \) is a scalar function, \( f(x, y) \). The study is carried out for any two components of \( \tau \) and therefore we consider the following two equations:

\[
\begin{align*}
\frac{\partial}{\partial x} f(x, y) + \tau_x f(x, y) &= 0 \\
\frac{\partial}{\partial y} f(x, y) + \tau_y f(x, y) &= 0
\end{align*}
\]  

(1.51a, 1.51b)

Continuity and Differentiability

From the way the equations are presented, the function \( f(x, y) \) is differentiable to the first order with respect to both \( x \) and \( y \). All other features depend on the analytic characteristics of the various components of \( \tau \), in this case, the two functions \( \tau_x(x, y) \) and \( \tau_y(x, y) \). Therefore, if they are differentiable with respect to \( x \) and \( y \), then \( f(x, y) \) has second-order derivatives, namely, \( f_{xx}, f_{xy}, f_{yx}, \) and \( f_{yy} \). For instance, Eq. (1.51a) guarantees the second derivative \( f_{xx} \) in the following way

\[
f_{xx} = \frac{\partial^2 f}{\partial x^2} = - \frac{\partial \tau_x}{\partial x} f - \tau_x \frac{\partial f}{\partial x} = - \frac{\partial \tau_x}{\partial x} f + \tau_x^2 f
\]

(1.52)

and in a similar way it guarantees the existence of \( f_{xy} \). The same applies to Eq. (1.51b) and other equations.

Analyticity

It is well known that a necessary condition, for a function defined in terms of several variables, to be analytic in a given region is having derivatives (with respect to all coordinates) to all orders. However, this condition is not sufficient. The missing feature is relations between the mixed derivatives of the same order such as \( f_{xy} \) and \( f_{yx} \) and all others. Thus the analyticity is guaranteed if and only if the order of differentiation does not affect the result; for instance, we demand that \( f_{xy} = f_{yx} \), and so on. For this requirement to be fulfilled at a given point \((x, y)\), the two components of \( \tau \), namely, \( \tau_x(x, y) \) and \( \tau_y(x, y) \), cannot be arbitrary but have to be related to each other.

To find this relation, we differentiate Eq. (1.51a) with respect to \( y \)

\[
f_{xy} = \frac{\partial^2 f}{\partial y \partial x} = - \frac{\partial \tau_x}{\partial y} f - \tau_x \frac{\partial f}{\partial y} = - \frac{\partial \tau_x}{\partial y} f + \tau_x \tau_y f
\]

(1.53a)

and Eq. (1.51b) with respect to \( x \):

\[
f_{yx} = \frac{\partial^2 f}{\partial x \partial y} = - \frac{\partial \tau_y}{\partial x} f - \tau_y \frac{\partial f}{\partial x} = - \frac{\partial \tau_y}{\partial x} f + \tau_y \tau_x f
\]

(1.53b)
and subtract the first equation from the second so that we obtain

$$f_{xy} - f_{yx} = \left( \frac{\partial \tau_x}{\partial y} - \frac{\partial \tau_y}{\partial x} \right) f$$

(1.54)

Equation (1.54) implies that in order for \( f(x, y) \) to be an analytic function (viz., having \( f_{xy} = f_{yx} \)), \( \tau_x \) and \( \tau_y \) have to fulfill the following relation:

$$\frac{\partial \tau_x}{\partial y} - \frac{\partial \tau_y}{\partial x} = 0$$

(1.55)

Although Eq. (1.55) can be interpreted as the \( z \) component of a Curl equation, in fact it applies for any number of pairs of coordinates. We recall that Eq. (1.55) is identical to Eq. (1.28), which was derived earlier for two electronic eigenfunctions that form a two-state Hilbert space.

### 1.3.1.2 Study of Non-Abelian Case

Next we extend the previous treatment to the case where \( \tau \) is a matrix and concentrate again on its two components \( \tau_x \) and \( \tau_y \). As before, the elements of \( \tau_x \) and \( \tau_y \) are assumed to be analytic functions of the coordinates, and we demand that \( \Omega \) be analytic, which implies that \( \Omega_{xy} = \Omega_{yx} \). As before, we consider two components of Eq. (1.50), namely

$$\frac{\partial \Omega(x, y)}{\partial x} + \tau_x \Omega(x, y) = 0$$

(1.56a)

$$\frac{\partial \Omega(x, y)}{\partial y} + \tau_y \Omega(x, y) = 0$$

(1.56b)

and differentiate the first equation with respect to \( y \), and the second with respect to \( x \) and subtract the results:

$$\frac{\partial^2 \Omega(x, y)}{\partial y \partial x} - \frac{\partial^2 \Omega(x, y)}{\partial x \partial y} = \left[ \frac{\partial \tau_x}{\partial y} - \frac{\partial \tau_y}{\partial x} - (\tau_x \tau_y - \tau_y \tau_x) \right] \Omega(x, y)$$

Thus, in contrast to the previous case, we see that in order to guarantee the analyticity of \( \Omega \), requiring that the corresponding component of the Curl equation be zero is not enough. In fact, what is required is that the following expression become zero:

$$\frac{\partial \tau_x}{\partial y} - \frac{\partial \tau_y}{\partial x} - (\tau_x \tau_y - \tau_y \tau_x) = 0$$

(1.57)

This is the \((x, y)\) component of the tensor equation given in Eq. (1.18):}

$$\text{Curl} \tau - [\tau \times \tau] = 0$$

(1.58)
14

MATHEMATICAL INTRODUCTION

It is seen that in order for the $\tau$ matrix to yield, an analytic solution for Eq. (1.50) in a given region, it has to fulfill the extended Curl equation, a condition that only certain groups of electronic states may guarantee. In Section 1.1 we showed that these are the groups that form, in that region, a Hilbert space or, at least, a Hilbert subspace (and in this case the analyticity is fulfilled only approximately). It is important to emphasize that if Eq. (1.58) is not fulfilled, this implies not that Eq. (1.50) does not have a solution but that the solution is not analytic.

1.3.1.3 Orthogonality
This subject is related to the previous section because it is intimately associated with the first-order equation Eq. (1.50). However, because of its exceptional importance, it is treated in a separate section.

Lemma 1.3 The matrix solution $\Omega$ of Eq. (1.50) is an orthogonal matrix.

Proof To prove this lemma, we consider Eq. (1.50) and its complex conjugate

$$\nabla \Omega^\dagger(s) - \Omega^\dagger(s)\tau(s) = 0 \quad (1.50')$$

where we recall that $\tau(s)$ is an antisymmetric matrix. Next, multiplying Eq. (1.50) from the left by $\Omega^\dagger$ and Eq. (1.50') from the right by $\Omega$ and adding up the two equalities, we get

$$(\nabla \Omega^\dagger) \Omega + \Omega^\dagger \nabla \Omega = \nabla(\Omega^\dagger \Omega) = 0 \Rightarrow \Omega^\dagger \Omega = \text{const} \quad (1.59)$$

or by choosing appropriate boundary conditions, we obtain the following:

$$\Omega^\dagger \Omega = I \Rightarrow \Omega \Omega^\dagger = I \quad (1.60)$$

The orthogonality relation in Eq. (1.60) can be maintained in any region as long as the $\tau$-matrix elements are analytic functions in that region. It is important to emphasize that the orthogonality condition exists even if the Curl equation [see Eq. (1.58)] in that region is not fulfilled.

1.3.2 Integral Equation
The integral equation approach, in contrast to the differential equation approach, supplies a more general view on what to expect from the solution in a given region. The differential approach concentrates on what happens at a given point and its close neighborhood, whereas the integral approach yields information related to a given region. In the forthcoming sections we derive the relevant integral equations first for the case that $\tau(s)$ is a vector and then when it is a matrix.
1.3 VECTORIAL FIRST-ORDER DIFFERENTIAL EQUATION AND LINE INTEGRAL

\[ (x_0, y_0) \]
\[ (x, y) \]
\[ (x_0, y_0) \]
\[ (x, y) \]
\[ \Gamma' \]
\[ \Gamma'' \]

Figure 1.1 Two rectangular paths \( \Gamma' \) and \( \Gamma'' \) connecting the points \((x_0, y_0)\) and \((x, y)\) in the \((x, y)\) plane.

1.3.2.1 Integral Equation along an Open Contour

We start by considering again Eqs. (1.50), where \( \tau(s) \) is an ordinary vector and \( \Omega(s) \) is a scalar function \( f(s) \). In order to convert Eq. (1.50) into an integral equation that connects an initial point \( s_0 \) with a final point \( s \), we have also to assume a contour \( \Gamma \) that contains these two points and along which the integral equation has to be solved. Thus symbolically the relevant integral equation to be considered is

\[ f(s|\Gamma) = f(s_0) - \int_{s_0}^{s} ds \cdot \tau(s|\Gamma) f(s|\Gamma) \quad (1.61) \]

In order to simplify our discussion, we assume a specific contour, \( \Gamma' \) (see Fig. 1.1(a)) made up of two straight lines as

\[ \Gamma' \equiv \{(x_0, y_0) \rightarrow (x, y_0) \rightarrow (x, y)\} \quad (1.62a) \]

so that the relevant integral equation (1.61) becomes

\[ f(x, y) = f(x_0, y_0) - \int_{x_0}^{x} dx' \tau_x(x', y_0) f(x', y_0) - \int_{y_0}^{y} dy' \tau_y(x, y') f(x, y') \quad (1.63a) \]

In order to verify that this expression is a solution, it has to be substituted in Eqs. (1.51) and examined accordingly. As an example, we consider Eq. (1.51a) and start by evaluating \( (\partial/\partial x) f(x, y) \):

\[ \frac{\partial f(x, y)}{\partial x} = -\tau_x(x, y_0) f(x, y_0) - \int_{y_0}^{y} dy' \left( \frac{\partial \tau_x(x, y')}{\partial x} f(x, y') + \tau_y(x, y') \frac{\partial f(x, y')}{\partial x} \right) \quad (1.64) \]
MATHEMATICAL INTRODUCTION

Next we assume that the Curl equation is fulfilled (at least) along $\Gamma_1$ so that we can replace $(\partial/\partial x)\tau_z(x, y')$ in the first term under the integral sign by $(\partial/\partial y')\tau_x(x, y')$, and then replace the resulting expression $[\tau_z(x, y')f(x, y')]\partial\tau_x(x, y')/\partial y'$ by

$$\frac{\partial\tau_z(x, y')}{\partial y'}f(x, y') = \frac{\partial[\tau_z(x, y')f(x, y')]}{\partial y'} - \tau_x(x, y')\frac{\partial f(x, y')}{\partial y'}$$

Carrying out both steps yields, for Eq. (1.64), the following result:

$$\frac{\partial f(x, y)}{\partial x} = -\tau_z(x, y)f(x, y) - \int_{y_0}^{y} dy' \left( -\tau_z(x, y')\frac{\partial f(x, y')}{\partial y'} + \tau_x(x, y')\frac{\partial f(x, y')}{\partial x} \right)$$

(1.64')

However, it can be seen that, due to Eqs. (1.51), the term under the integral sign is identically zero so that, in fact, Eq. (1.64') becomes identical to (1.51a). Thus the expression in (1.63a), indeed, satisfies Eq. (1.51).

In a similar way we can show that Eq. (1.63a) is also a solution of Eq. (1.51b). In fact, the derivation is straightforward and does not require even fulfillment of the Curl condition.

Short Summary We showed that Eq. (1.63a) satisfies Eqs. (1.51) if the Curl condition is satisfied for a given $x$ value and for any $y'$ value defined in the interval. However, if at a given point, say, $P(x_z, y_z)$, the Curl condition is not fulfilled, then an integral equation for a contour that contains $P(x_z, y_z)$ cannot be formed. Still we are allowed to employ any other contour as long as it surrounds $P(x_z, y_z)$.

Having derived Eq. (1.61) for a function $f(x, y)$, where $\tau$ is a vector, we are in a position to extend it for the case where $\tau$ is a matrix (which causes $\Omega$ to be a matrix as well):

$$\Omega(s|\Gamma) = \Omega(s_0) - \int_{s_0}^{s} ds \cdot \tau(s|\Gamma)\Omega(s|\Gamma)$$

(1.65)

We do not derive Eq. (1.65) as it is essentially similar to the derivation of Eq. (1.61). The only difference is that in this case we have to consider, in the appropriate instance, the extended Curl equation [see Eq. (1.58)] instead of Eq. (1.55).

1.3.2.2 Integral Equation along a Closed Contour

The equation to be treated in this case is of the form

$$f(s_0|\Gamma) = f(s_0) - \int_{\Gamma} ds \cdot \tau(s|\Gamma)f(s|\Gamma)$$

(1.66)
1.3 VECTORIAL FIRST-ORDER DIFFERENTIAL EQUATION AND LINE INTEGRAL

Equation (1.66) does not look like an ordinary integral equation; rather, it appears to be more a result of an integration along a closed contour. In the present section we study the condition(s) for which the final value $f(s_0|\Gamma)$ differs from the initial value $f(s_0)$. If the two are equal, this implies that the value of the integral along the closed contour is identically zero. In principle, the two values are expected to be the same, but as is shown next, this is not always be the case. In what follows we continue to call Eq. (1.66) an integral equation because, after all, it is the equation used to calculate the value of $f(s)$ at the endpoint: $s = s_0$.

The integral equation along a closed contour is best studied by considering two open contours with identical initial and final points\(^2\,^5\) (see Fig. 1.1). For this purpose we consider the integral equation along a second contour (see Fig. (1.1(b)) \(\Gamma'' \equiv \{(x_0, y_0) \to (x_0, y) \to (x, y)\}\) (1.62b)

so that the relevant integral equation is

$$f(x, y) = f(x_0, y_0) - \int_{y_0}^{y} dy \, \tau_y(x_0, y) f(x_0, y) - \int_{x_0}^{x} dx \, \tau_x(x, y) f(x, y) \quad (1.63b)$$

It is important to emphasize that Eqs. (1.63) are both legitimate solutions of Eq. (1.50) or, more specifically, of Eq. (1.51), at the point \((x, y)\), although the results as obtained by the two independent calculations may not be the same.

Having the two equations, Eqs. (1.63a) and (1.63b), we examine what happens at the point of contact \((x, y) = (x_f, y_f)\) — the point where the two contours $\Gamma'$ and $\Gamma''$ cross and form the closed contour $\Gamma$, which can be symbolically written as

$$\Gamma = \Gamma' - \Gamma'' \quad (1.67)$$

At that point Eq. (1.63a) yields the value $f(x_f, y_f|\Gamma')$ and Eq. (1.63b), the value $f(x_f, y_f|\Gamma'')$ — the two values are not necessarily the same. Thus, if $\Delta f(x_f, y_f|\Gamma)$ is defined as the difference

$$\Delta f(x_f, y_f|\Gamma') = f(x_f, y_f|\Gamma') - f(x_f, y_f|\Gamma'') \quad (1.67')$$

then, due to Eq. (1.67), $\Delta f(x_f, y_f|\Gamma)$ is the value of the integral along the closed contour $\Gamma$. Therefore, if the two integral equations yield the same result at \((x_f, y_f)\), this implies that the value of the integral equation along the closed contour is zero. In what follows we designate the value of $\Delta f(x_f, y_f|\Gamma)$ by $\Delta f(\Gamma)$.

In order to continue, we consider a situation where the initial point \((x_0, y_0)\) and the final point \((x_f, y_f)\) are close to each other, so that if $x_f = x_0 + \Delta x$ and $y_f = y_0 + \Delta y$, then both $\Delta x$ and $\Delta y$, are small enough to justify the approximations to be employed.
MATHEMATICAL INTRODUCTION

Thus \( \Gamma' \) is assumed to be the following open contour:

\[
\Gamma' \equiv \{(x_0, y_0) \rightarrow (x_0 + \Delta x, y_0) \} \rightarrow (x_0 + \Delta x, y_0 + \Delta y) \}
\] (1.68a)

and, in the same way, \( \Gamma'' \) is assumed to be

\[
\Gamma'' \equiv \{(x_0, y_0) \rightarrow (x_0, y_0 + \Delta y) \} \rightarrow (x_0 + \Delta x, y_0 + \Delta y) \}
\] (1.68b)

Substitute Eqs. (1.63a) and (1.63b) in Eq. (1.67'), replace \((x_f, y_f) \) by \((x_0 + \Delta x, y_0 + \Delta y) \), and perform the required algebraic changes leading to the result

\[
\Delta f(\Gamma) = \int_{x_0}^{x_0+\Delta x} dx \{ \tau_x(x, y_0) f(x, y_0) - \tau_x(x, y_0 + \Delta y) f(x, y + \Delta y) \}
\]

\[
+ \int_{y_0}^{y_0+\Delta y} dy \{ \tau_y(x_0, y) f(x_0, y) - \tau_y(x_0 + \Delta x, y) f(x_0 + \Delta x, y) \}
\]

which can be written (because \( \Delta x \) and \( \Delta y \) are small enough) as

\[
\Delta f(\Gamma') = \Delta y \int_{x_0}^{x_0+\Delta x} dx \left[ \frac{\partial}{\partial y} (\tau_x(x, \tilde{y}) f(x, \tilde{y})) \right] + \Delta x \int_{y_0}^{y_0+\Delta y} dy \left[ \frac{\partial}{\partial x} (\tau_y(\tilde{x}, y) f(\tilde{x}, y)) \right]
\] (1.69)

Next we replace each integral by a product between the relevant integrand (at some intermediate point) and the respective interval length so that Eq. (1.69) becomes

\[
\Delta f(\Gamma) = \Delta y \Delta x \left[ \frac{\partial}{\partial y} (\tau_x(\tilde{x}, \tilde{y}) f(\tilde{x}, \tilde{y})) - \frac{\partial}{\partial x} (\tau_y(\tilde{x}, \tilde{y}) f(\tilde{x}, \tilde{y})) \right]
\] (1.70)

Recalling again Eqs. (1.51) and assuming that \( f(x, y) \) is a continuous function, Eq. (1.70) can be further simplified so that we finally have

\[
\Delta f(\Gamma) = \Delta y \Delta x \left[ \frac{\partial}{\partial y} (\tau_x(x, y)) - \frac{\partial}{\partial x} (\tau_y(x, y)) \right] f(x, y)
\] (1.71)

or

\[
\Delta f(\Gamma) = (\text{Curl} \, \tau)_{xy} f(x, y) \Delta x \Delta y
\] (1.72)
1.3 VECTORIAL FIRST-ORDER DIFFERENTIAL EQUATION AND LINE INTEGRAL

where \((x, y)\) is some point in the interval \(x_0 \leq x \leq x_0 + \Delta x\) and \(y_0 \leq y \leq y_0 + \Delta y\) and \(\Delta x\) and \(\Delta y\) both \(\to 0\).

It is readily noted that the value of \(\Delta f(x_0, y_0)\) depends solely on \(\text{Curl} \ \tau\). So, if \(\tau\) fulfills Eq. (1.55), then the results as calculated along the two different \textit{infinitesimal} contours are identical but differ if \(\text{Curl} \ \tau \neq 0\).

Although \(\Gamma'\) and \(\Gamma''\) are infinitesimal contours, this does not retract from the generality of the derivation. In case we are interested in a regular-size region that contains one such point [which causes \(\Delta f(x_f, y_f) \neq 0\)], it has been shown (see Appendix C in Ref. 2) that any closed contour that surrounds this point can be presented as a sum of all the infinitesimal closed contours in the region surrounded by this regular-size contour, and therefore the conclusions regarding the value of the \textit{closed-contour-integration} depends solely on what happens at this one point inside that region immaterial of the size of that region.

**Short Summary**  
If in a given region \(\text{Curl} \ \tau = 0\) at \textit{every} point in that region, the \textit{integral} in Eq. (1.66), along any closed contour in that region, is equal to zero. If, in a given region, \(\text{Curl} \ \tau\) is not defined at some points in the region (these points can be considered, at this stage, as \textit{pathological} points), then the \textit{integral} in Eq. (1.66) is zero as long as the closed contour does not surround any of these points. If, however, it surrounds one (or more) points of this kind, it becomes mathematically undefined (this problem and how it is related to Extended Stokes theorem is discussed further in Section 6.2.3).

Having studied Eq. (1.66), for a function \(f(x, y, \ldots)\) where \(\tau\) is a vector, we can now extend it for the case that both \(\tau\) and \(\Omega\) are matrices:

\[
\Omega(s_0|\Gamma) = \Omega(s_0) - \oint_{\Gamma} ds \cdot \tau(s|\Gamma)\Omega(s|\Gamma) \quad (1.73)
\]

The procedure that led to the analysis and the conclusions that ended with Eq. (1.72) applies also for the matrix equation in Eq. (1.73) with only one modification. The condition to be fulfilled in this case is now not by the \(\text{Curl} \ \tau\) but by its extended non-Abelian version, namely, tensorial field \(F\). Thus, Eq. (1.72) has to be replaced in this case by \(^5\)

\[
\Delta \Omega(\Gamma) = \Delta y \ \Delta x F_{xy} \Omega(x, y) \quad (1.74)
\]

where both \(\Delta x\) and \(\Delta y\) \(\to 0\).

We do not repeat the relevant derivation (it can be found in Ref. 5) but just present the following summary. If in a given region \(F = 0\) at \textit{every} point in that region the \textit{integral} in Eq. (1.73), along any closed contour in that region, is equal to zero. If, in a given region, \(F\) differs from zero at some points, then the closed line \textit{integral}, in Eq. (1.73), is zero as long as the closed contour does not surround any of these points, but differs from zero when the contour surrounds one of these points. If, in a given region, \(F\) is not well defined at some points, then any closed line \textit{integral} that surrounds one of these points is mathematically undefined.
1.3.3 Solution of Differential Vector Equation

In this section we derive the solution of the differential equation for the matrix $\Omega$ as given Eq. (1.50). Since we are considering matrices and not functions, the derivation is not straightforward. For instance, an expansion in terms of some power series is not likely to succeed because the product of two different matrices is not commutative (unless both are diagonal). This is the reason why the derivation has to be done in a different way, known as propagation.\(^6\)-\(^8\) Propagation is characterized by the fact that the value of the unknown function at one point is calculated by its value, at a nearby point employing an approximation based on finite differences. As mentioned earlier, in order to solve Eq. (1.50), we have to assign a contour $\Gamma_1$ and solve the equations along this contour. Therefore the propagation has to be done along that contour.

As an example, we consider a planar case characterized by the coordinates $(x, y)$ and show how to solve Eqs. (1.56). Defining a grid of $L$ points along $\Gamma_1$, namely, $\{ P_0, P_1, \ldots, P_n, \ldots, P_L \}$, where each point is defined as $P_n = (x_n, y_n)$ and $P_0$ is the initial point for which is assigned an initial value in terms of the matrix $\Omega(x_0, y_0)$.

Assuming that we reached, by propagation, the point $P_n$, we show how to continue to the point $P_{n+1}$. This we do by carrying out two consecutive steps $(x_n, y_n) \rightarrow (x_n + \Delta x, y_n + \Delta y)$ where $\Delta x$ and $\Delta y$ are both small enough. Next we consider a point $(\tilde{x}_n, \tilde{y}_n)$ in the planar interval:

$$x_n \leq \tilde{x}_n \leq x_n + \Delta x \quad \text{and} \quad y_n \leq \tilde{y}_n \leq y_n + \Delta y \quad (1.75)$$

To carry out the first step, we employ the approximate equation [see Eq. (1.56a)]

$$\frac{\partial}{\partial x} \Omega(x, y) + \tau_x(\tilde{x}_n, \tilde{y}_n) \Omega(x, y) = 0 \quad (1.76a)$$

where $\tau_x(x, y)$ is replaced by $\tau_x$ calculated at some intermediate point $(\tilde{x}_n, \tilde{y}_n)$, namely, $\tau_x(x, y) \sim \tau_x(\tilde{x}_n, \tilde{y}_n)$.

Next, we introduce the orthogonal matrix $G(x, y)$ and the corresponding diagonal matrix $t_x(x, y)$, both defined through the relation

$$t_x(\tilde{x}_n, \tilde{y}_n) = G(\tilde{x}_n, \tilde{y}_n) \tau_x(\tilde{x}_n, \tilde{y}_n) G(\tilde{x}_n, \tilde{y}_n)^T \quad (1.77)$$

Here $G$ and $t_x$ contain, respectively, the eigenvectors and the eigenvalues of the $\tau_x$ matrix. Multiplying Eq. (1.76a) from the left by $G(\tilde{x}_n, \tilde{y}_n)$ and from the right by $G(\tilde{x}_n, \tilde{y}_n)^T$ and defining

$$\tilde{\Omega}(x, y) = G(\tilde{x}_n, \tilde{y}_n) \Omega(x, y) G(\tilde{x}_n, \tilde{y}_n)^T \quad (1.78)$$

we get

$$\frac{\partial}{\partial x} \tilde{\Omega}(x, y) + t_x(\tilde{x}_n, \tilde{y}_n) \tilde{\Omega}(x, y) = 0 \quad (1.79)$$
1.3 VECTORIAL FIRST-ORDER DIFFERENTIAL EQUATION AND LINE INTEGRAL

Since \( t_x \) is a diagonal (constant) matrix, the solution of Eq. (1.79) at \( x = x_n+1 \) is straightforward:

\[
\tilde{\Omega}(x_n+1, \tilde{y}_n) = \exp\{-\Delta x \, t_x(\tilde{x}_n, \tilde{y}_n)\} \Omega(x_n, y_n) \tag{1.80}
\]

To return to the original matrices \( \tau_x \) and \( \Omega \), Eq. (1.80) is multiplied from the left and right by \( G^\dagger(\tilde{x}_n, \tilde{y}_n) \) and \( G(\tilde{x}_n, \tilde{y}_n) \), respectively, and following other minor modifications, we get

\[
\Omega(x_n+1, y_n) = \exp \left\{ - \int_{x_n}^{x_{n+1}} \tau_x(x, \tilde{y}_n) \, dx \right\} \Omega(x_n, y_n) \tag{1.81}
\]

To perform the second step, in the \( y \) direction, we employ Eq. (1.56b) and consider, as before, the approximate equation

\[
\frac{\partial}{\partial y} \Omega(x, y) + \tau_y(\tilde{x}_n, \tilde{y}_n) \Omega(x, y) = 0 \tag{1.76b}
\]

where \( \tau_y(x, y) \) is replaced by \( \tau_y \) calculated at some (other) intermediate point \((\tilde{x}_n, \tilde{y}_n)\), namely, \( \tau_y(x, y) = \tau_y(\tilde{x}_n, \tilde{y}_n) \).

Continuing in the same manner as before [see Eq. (1.81)], we get

\[
\Omega(x_n+1, y_{n+1}) = \exp \left\{ - \int_{y_n}^{y_{n+1}} \tau_y(\tilde{x}_n, y) \, dy \right\} \Omega(x_{n+1}, \tilde{y}_n) \tag{1.82}
\]

Substituting Eq. (1.81) in Eq. (1.82) and recalling the definition of the scalar product

\[
\tau(s) \cdot ds \sim \tau_x(x, \tilde{y}_n) dx + \tau_y(\tilde{x}_n, y) dy \tag{1.83}
\]

where \( s \equiv (x, y) \), we finally obtain

\[
\Omega(s_{n+1}) = \exp \left\{ - \int_{s_n}^{s_{n+1}} \tau(s) \cdot ds \right\} \Omega(s_n) \tag{1.84}
\]

Equation (1.84) can be now be extended to any number of steps so that the final result is

\[
\Omega(s) = \varphi \exp \left\{ - \int_{s_0}^{s} \tau(s') \cdot ds' \right\} \Omega(s_0) \tag{1.85}
\]
MATHEMATICAL INTRODUCTION

where \( \wp \) is the ordering operator that tells us to carry out the integration, in Eq. (1.85), in the following order \( \{P_0 \rightarrow P_1 \cdots \rightarrow P_n \rightarrow P_{n+1} \cdots \rightarrow P_L\} \) all this along the assigned contour \( \Gamma \).

As a final subject in this section we derive the numerical representation of the solution in Eq. (1.85) for a given contour \( \Gamma \). This is done according to the following recipe:

1. Divide \( \Gamma \) into \( L \) sections where the \( n \)th section is defined in terms of its two endpoints \( [s_{n-1}, s_n] \).

2. Write the exponentiated integral as an ordered product of \( L \) exponentiated integrals, each related to one section:

\[
\Omega(s_L) = \left[ \prod_{n=1}^{L} \exp \left\{ -\int_{s_{n-1}}^{s_n} \tau(s') \cdot ds' \right\} \right] \Omega(s_0) \tag{1.86}
\]

3. Approximate the integrals in each exponential by \( \tau(s_n) \Delta_n \), where \( s_n \) is some intermediate point and \( \Delta_n \) is defined as vectorial length \( \Delta_n = s_n - s_{n-1} \), so that Eq. (1.86) becomes

\[
\Omega(s_L) = \left[ \prod_{n=1}^{L} \exp \left\{ -\tau(s_n) \cdot \Delta_n \right\} \right] \Omega(s_0) \tag{1.87}
\]

4. Replace each exponential by

\[
\exp \left\{ -\tau(s_n) \cdot \Delta_n \right\} = G(s_n)E(s_n)G^\dagger(s_n) \tag{1.88}
\]

where \( G(s_n) \) is a matrix that diagonalizes \( \tau(s_n) \), namely

\[
\tau(s_n) = G(s_n)t(s_n)G^\dagger(s_n) \tag{1.89}
\]

[here \( t(s_n) \) is a diagonal matrix that contains the eigenvalues of \( \tau(s_n) \) and \( E(s_n) \) is given in the form]

\[
E(s_n) = \exp \left( -t(s_n) \cdot \Delta_n \right) \tag{1.90}
\]

5. Substitute Eq. (1.88) in Eq. (1.87) so that

\[
\Omega(s_L) = \left\{ \prod_{n=1}^{L} [G(s_n)E(s_n)G^\dagger(s_n)] \right\} \Omega(s_0) \tag{1.91}
\]

This completes the numerical presentation of the solution to Eq. (1.50).
1.4 SUMMARY AND CONCLUSIONS

In Section 1.3 two approaches are presented to treat first-order differential equations as presented in Eq. (1.50): (1) by converting it into an integral equation as presented in Eq. (1.65) and (2) by employing the propagation technique and deriving an analytic expression as presented in Eq. (1.85). This expression is then “translated” into a numerical language as given in Eq. (1.91). The two approaches are expected to yield the same results. In this section we discuss to some extent results, for closed contours.

Equation (1.73) [see also (1.66)] is the integral equation for a closed contour, and we have shown that the value of the integral is zero when the contour \( \Gamma \) does not surround any of the points for which the tensorial field \( \mathbf{F} \) is not defined. In case it surrounds one or more points, the integral becomes undefined. The same is expected for the explicit closed contour that follows from Eq. (1.85):

\[
\Omega(s_0|\Gamma) = \varphi \exp \left\{ - \oint_{\Gamma} \tau(s'|\Gamma) \cdot ds' \right\} \Omega(s_0)
\]  

(1.92)

Defining the abovementioned exponentiated closed-contour integral as \( \mathbf{D}(\Gamma) \)

\[
\mathbf{D}(\Gamma) = \varphi \exp \left\{ - \oint_{\Gamma} \tau(s'|\Gamma) \cdot ds' \right\}
\]  

(1.93)

we can state that the \( \mathbf{D} \) matrix is the unit matrix \( \mathbf{I} \) if the contour does not surround any of the pathological points.

The \( \mathbf{D} \) matrix can also be defined in terms of the line integral [see Eq.1.73)] as follows

\[
\mathbf{D}(\Gamma) = \mathbf{I} - \oint_{\Gamma} ds \cdot \tau(s|\Gamma)\Omega(s|\Gamma)
\]  

(1.94)

where \( \mathbf{I} \) is the unit matrix.

PROBLEM

1.1 Consider a planar system described in terms of two polar coordinates \( s \equiv (\varphi, q) \) and the corresponding components of \( \tau \): \( \tau = (\tau_\varphi/q, \tau_q) \). Assume \( \tau_q \) to be identically zero and \( \tau_\varphi \) to be a 2 \( \times \) 2 matrix of the type

\[
\tau_\varphi = \begin{pmatrix}
0 & \frac{1}{2} \\
-\frac{1}{2} & 0
\end{pmatrix}
\]  

(1.95)
Employing Eq. (1.94) for a circular contour, prove that the corresponding two-state $D$ matrix takes the following form:

$$D = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$$ (1.96)

**Solution** Our first step is to derive $\Omega(\phi, q)$, and this is done using Eqs. (1.56) after transforming them to polar coordinates:

$$\frac{\partial}{\partial q} \Omega(\phi, q) + \tau_q \Omega(\phi, q) = 0$$ (1.97a)

$$\frac{1}{q} \left( \frac{\partial}{\partial \phi} \Omega(\phi, q) + \tau_\phi \Omega(\phi, q) \right) = 0$$ (1.97b)

Recalling that $\tau_q$ is identically zero (and consequently $q = q_0$) and $\tau_\phi$ is given in Eq. (1.95), we find that $\Omega(\phi, q = q_0)$ has to fulfill the following (matrix) first-order differential equation:

$$\frac{\partial}{\partial \phi} \Omega(\phi, q_0) + \begin{pmatrix} 0 & 1/2 \\ -1/2 & 0 \end{pmatrix} \Omega(\phi, q_0) = 0$$ (1.98)

To derive the solution of Eq. (1.98), we assume that

$$\Omega(\phi, q_0) = \begin{pmatrix} \cos \gamma(\phi) & \sin \gamma(\phi) \\ -\sin \gamma(\phi) & \cos \gamma(\phi) \end{pmatrix}$$ (1.99)

Substituting Eq. (1.99) in Eq. (1.98) yields for $\gamma(\phi)$ the result $\gamma(\phi) = \phi/2$. Next, substituting Eq. (1.99) in Eq. (1.98), we get for $D(\Gamma) = D(q_0)$

$$D(q_0) = I - \int_0^{2\pi} d\phi \begin{pmatrix} 0 & 1/2 \\ -1/2 & 0 \end{pmatrix} \begin{pmatrix} \cos(\phi/2) & \sin(\phi/2) \\ -\sin(\phi/2) & \cos(\phi/2) \end{pmatrix}$$

or following the integration

$$D(q_0) = I - 2I = -I$$ (1.100)

which is identical to Eq. (1.96).
REFERENCES

Section 1.1

Section 1.2

Section 1.3