# 1
# Examples for Chapter 1, Linear Algebra

## 1.1 Numbers

If $z = 1 + i$, then

$$\frac{1}{z} = \frac{1}{1+i} = \frac{1-i}{(1+i)(1-i)} = \frac{1-i}{2} = \frac{1}{2} - \frac{i}{2}. \tag{1.1}$$

In other symbols,

$$\frac{1}{z} = \frac{z^*}{z\, z^*} = \frac{z^*}{|z|^2}. \tag{1.2}$$

Don't worry about Grassmann numbers. Suppose $a, b$ are complex numbers and $\theta$ is a Grassmann number or equivalently a Grassmann variable. Then because $\theta^2 = 0$, the inverse of $\alpha = a + b\theta$ is

$$\frac{1}{\alpha} = \frac{1}{a+b\theta} = \frac{a-b\theta}{(a+b\theta)(a-b\theta)} = \frac{a-b\theta}{a^2} = \frac{1}{a} - \frac{b\theta}{a^2} \tag{1.3}$$

and $e^{a\theta}$ is

$$e^{a\theta} = 1 + a\theta. \tag{1.4}$$

## 1.2 Arrays

You know about inner products like

$$\boldsymbol{k} \cdot \boldsymbol{x} \equiv \vec{k} \cdot \vec{x} = k_1 x_1 + k_2 x_2 + k_3 x_3. \tag{1.5}$$

Here's a Lorentz-invariant inner product of two 4-vectors

$$p \cdot x = \vec{p} \cdot \vec{x} - p^0 x^0. \tag{1.6}$$

Particle physicists often add a minus sign and write

$$p \cdot x = p^0 x^0 - \vec{p} \cdot \vec{x}. \tag{1.7}$$

They do that to have

$$p^2 = p \cdot p = p^0 p^0 - \vec{p} \cdot \vec{p} = \left(\frac{E}{c}\right)^2 - (\vec{p})^2 = m^2 c^2. \tag{1.8}$$

## 1.3 Matrices

Vectors must have the same number of components to have an inner product, but any two vectors can have two kinds of outer product. For instance,

$$\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \begin{pmatrix} 4 & 5 & 6 & 7 \end{pmatrix} = \begin{pmatrix} 4 & 5 & 6 & 7 \\ 8 & 10 & 12 & 14 \\ 12 & 15 & 18 & 21 \end{pmatrix} = \begin{pmatrix} 4 & 5 & 6 & 7 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \tag{1.9}$$

and

$$\begin{pmatrix} 4 \\ 5 \\ 6 \\ 7 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \end{pmatrix} = \begin{pmatrix} 4 & 8 & 12 \\ 5 & 10 & 15 \\ 6 & 12 & 18 \\ 7 & 14 & 21 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \end{pmatrix} \begin{pmatrix} 4 \\ 5 \\ 6 \\ 7 \end{pmatrix}. \tag{1.10}$$

The second of these two equations is the transpose (denoted by a $\mathsf{T}$) of the first:

$$\begin{aligned} \begin{pmatrix} 4 \\ 5 \\ 6 \\ 7 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \end{pmatrix} &= \begin{pmatrix} 4 & 8 & 12 \\ 5 & 10 & 15 \\ 6 & 12 & 18 \\ 7 & 14 & 21 \end{pmatrix} = \begin{pmatrix} 4 & 5 & 6 & 7 \\ 8 & 10 & 12 & 14 \\ 12 & 15 & 18 & 21 \end{pmatrix}^{\mathsf{T}} \\ &= \left( \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \begin{pmatrix} 4 & 5 & 6 & 7 \end{pmatrix} \right)^{\mathsf{T}} = \begin{pmatrix} 4 & 5 & 6 & 7 \end{pmatrix}^{\mathsf{T}} \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}^{\mathsf{T}} \\ &= \begin{pmatrix} 4 \\ 5 \\ 6 \\ 7 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \end{pmatrix}. \end{aligned} \tag{1.11}$$

**Example 1.1** (Cross-products) Three-dimensional space is special in that it has a cross-product. The cross-product $\boldsymbol{A} \times \boldsymbol{B}$ of two 3-vectors $\boldsymbol{A}$ and $\boldsymbol{B}$ is the 3-vector whose $i$th component is the sum

$$\left(\boldsymbol{A} \times \boldsymbol{B}\right)_i = \sum_{j,k=1}^{3} \epsilon_{ijk} A_j \, B_k \tag{1.12}$$

in which $\epsilon_{ijk}$ is totally antisymmetric with $\epsilon_{123} = 1$. So $\epsilon_{213} = -1$, $\epsilon_{113} = 0$, $\epsilon_{231} = 1$, etc. □

The trace of a matrix is the sum of its diagonal elements. If

$$A = \begin{pmatrix} 4 & 8 & 12 \\ 5 & 10 & 15 \\ 6 & 12 & 18 \end{pmatrix}, \tag{1.13}$$

then its trace is

$$\mathrm{Tr} A = 32 = \mathrm{Tr}\big(A^{\mathsf{T}}\big). \tag{1.14}$$

If $V$ is a vector with complex components

$$V = \begin{pmatrix} 1+i \\ -i \\ 3+i \end{pmatrix}, \tag{1.15}$$

then its complex conjugate and adjoint (or hermitian adjoint) are

$$V^* = \begin{pmatrix} 1-i \\ i \\ 3-i \end{pmatrix} \quad \text{and} \quad V^\dagger = \begin{pmatrix} 1-i & i & 3-i \end{pmatrix}. \tag{1.16}$$

The complex conjugate of the matrix

$$B = \begin{pmatrix} i & 2 & i \\ 3 & i & 4 \\ 1 & -i & 6 \end{pmatrix}, \tag{1.17}$$

is

$$B^* = \begin{pmatrix} -i & 2 & -i \\ 3 & -i & 4 \\ 1 & i & 6 \end{pmatrix}. \tag{1.18}$$

The adjoint or hermitian adjoint of the matrix

$$B = \begin{pmatrix} i & 2 & i \\ 3 & i & 4 \\ 1 & -i & 6 \end{pmatrix} \tag{1.19}$$

is the complex conjugate of its transpose

$$B^\dagger = \begin{pmatrix} -i & 3 & 1 \\ 2 & -i & i \\ -i & 4 & 6 \end{pmatrix}. \tag{1.20}$$

The matrix

$$C = \begin{pmatrix} i & 2 & i \\ 2 & i & 4 \\ i & 4 & 6 \end{pmatrix} \tag{1.21}$$

is symmetric.

A matrix is hermitian if it is equal to its adjoint. The matrix

$$D = \begin{pmatrix} 1 & i & -i \\ -i & 2 & 2i \\ i & -2i & -7 \end{pmatrix} \tag{1.22}$$

is hemitian. Its diagonal elements are 1, 2, and -7, and they are real. All the diagonal elements of every hermitian matrix are real.

The inverse $A^{-1}$ of a matrix $A$ is a matrix that does this

$$A\,A^{-1} = A^{-1}\,A = I \tag{1.23}$$

in which the identity matrix $I$ is a diagonal matrix with 1 on its main diagonal and 0's elsewhere

$$I = 1 \quad \text{or} \quad I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{or} \quad I = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \text{etc.} \tag{1.24}$$

The Matlab command

```
A =[  1 2 3 ;   4 5 6 ;  -1 2 -3 ]
```

generates the matrix

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ -1 & 2 & -3 \end{pmatrix}, \tag{1.25}$$

and the Matlab command

```
inv(A)
```

generates its inverse

$$A^{-1} = \begin{pmatrix} -1.1250 & 0.5000 & -0.1250 \\ 0.2500 & 0 & 0.2500 \\ 0.5417 & -0.1667 & -0.1250 \end{pmatrix}. \tag{1.26}$$

A matrix is unitary if its adjoint is its inverse

$$U\,U^\dagger = U^\dagger\,U = I. \tag{1.27}$$

A real unitary matrix is orthogonal, and

$$O\,O^\dagger = O\,O^\mathsf{T} = O^\dagger\,O = O^\mathsf{T}\,O = I. \tag{1.28}$$

## 1.4 Matrix Multiplication and Commutation Relations

```
>> X = [ 0 -1 0 0 ; -1 0 0 1; 0 0 0 0; 0 -1 0 0 ]

X =

     0    -1     0     0
    -1     0     0     1
     0     0     0     0
     0    -1     0     0

>> Y = [0 0 -1 0; 0 0 0 0; -1 0 0 1; 0 0 -1 0]

Y =

     0     0    -1     0
     0     0     0     0
    -1     0     0     1
     0     0    -1     0

>> J = [0 0 0 0; 0 0 -1 0; 0 1 0 0; 0 0 0 0]

J =

     0     0     0     0
     0     0    -1     0
     0     1     0     0
     0     0     0     0

>> J*X -X*J - Y

ans =

     0     0     0     0
     0     0     0     0
     0     0     0     0
```

```
     0     0     0     0
```

So $[J, X] = Y$.

```
>> J*Y - Y*J + X

ans =

     0     0     0     0
     0     0     0     0
     0     0     0     0
     0     0     0     0
```

So $[J, Y] = -X$. Thanks to Yu Chia Lin for pointing out a typo here.

The polarization vectors for a massless particle moving in the 3-direction are $\epsilon_\pm$

```
>> epp =  [ 0; 1; -1i; 0]

epp =

   0.0000 + 0.0000i
   1.0000 + 0.0000i
   0.0000 - 1.0000i
   0.0000 + 0.0000i

>> epm = [ 0; 1; 1i; 0]

epm =

   0.0000 + 0.0000i
   1.0000 + 0.0000i
   0.0000 + 1.0000i
   0.0000 + 0.0000i

>> syms a b

>> (a*X + b*Y)*epp

ans =
```

```
 - a + b*1i
          0
          0
 - a + b*1i

>> (a*X + b*Y)*epm

ans =

 - a - b*1i
          0
          0
 - a - b*1i
```

So

$$(aX + bY)^i{}_k \epsilon^k_\pm = (-a \pm ib)\, p^i/p^0. \tag{1.29}$$

## 1.5 Vector Spaces

You can multiply vectors by real and by complex numbers

$$2\begin{pmatrix} 3 \\ i \\ -4 \end{pmatrix} = \begin{pmatrix} 6 \\ 2i \\ -8 \end{pmatrix} \quad \text{and} \quad (1+i)\begin{pmatrix} 3 \\ i \\ -4 \end{pmatrix} = \begin{pmatrix} 3+3i \\ -1+i \\ -4-4i \end{pmatrix}. \tag{1.30}$$

And you can add vectors that are multiplied by (real or) complex numbers $z$ and $w$

$$z\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} + w\begin{pmatrix} 4i \\ 5 \\ -6i \end{pmatrix} = \begin{pmatrix} z+4iw \\ 2z+5w \\ 3z-6wi \end{pmatrix}. \tag{1.31}$$

## 1.6 Vector Spaces and Dimension

In quantum mechanics, we often represent systems by states, e.g., $|\ell, m\rangle$ represents a state that has angular momentum $\ell\hbar$ and angular momentum $m\hbar$ in the z direction. Here $m$ can be $-\ell$, $-\ell + 1$, $\ldots \ell - 1$, $\ell$. So there are $2\ell + 1$ states $|\ell, m\rangle$ for a given $\ell$. We can add states in quantum mechanics, so the possible states of angular momentum $\ell$ are a sum from $m = -\ell$ to

$m = \ell$ of the states $|\ell, m\rangle$, each multiplied by an arbitrary complex number $z_m$

$$|\ell, z\rangle = \sum_{m=-\ell}^{\ell} z_m |\ell, m\rangle. \tag{1.32}$$

The states $|\ell, m\rangle$ span a vector space of dimension $2\ell+1$. They form a basis for this space because every state in it can be written as a linear combination of the $|\ell, m\rangle$ as in equation (1.32). They are the orthonormal

$$\langle \ell, m | \ell', m' \rangle = \delta_{mm'} \tag{1.33}$$

eigenstates of the z component of the angular momentum $\boldsymbol{L}$

$$L_z |\ell, m\rangle = m\hbar \, |\ell, m\rangle. \tag{1.34}$$

If $\ell = 1$, then in a basis in which $L_z$ is diagonal

$$L_z = \hbar \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \tag{1.35}$$

its eigenstates are

$$|1, 1\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad |1, 0\rangle = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad |1, -1\rangle = \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix}. \tag{1.36}$$

These states span a space of dimension 3. The state $|1, m\rangle$ is an eigenstate of $L_z$ with eigenvalue $m\hbar$

$$L_z |1, m\rangle = m\hbar |1, m\rangle \tag{1.37}$$

for $m = -1, 0$, and 1.

If $\ell = 2$, the diagonal form of $L_z$ is

$$L_z = \hbar \begin{pmatrix} 2 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & -2 \end{pmatrix}. \tag{1.38}$$

In this basis, the states $|2,m\rangle$ are

$$|2,2\rangle = \begin{pmatrix}1\\0\\0\\0\\0\end{pmatrix}, \quad |2,1\rangle = \begin{pmatrix}0\\1\\0\\0\\0\end{pmatrix}, \quad |2,0\rangle = \begin{pmatrix}0\\0\\1\\0\\0\end{pmatrix},$$
$$|2,-1\rangle = \begin{pmatrix}0\\0\\0\\1\\0\end{pmatrix}, \quad \text{and} \quad |2,-2\rangle = \begin{pmatrix}0\\0\\0\\0\\1\end{pmatrix}. \tag{1.39}$$

The state $|2,m\rangle$ is an eigenstate of $L_z$ with eigenvalue $m\hbar$

$$L_z|2,m\rangle = m\hbar|2,m\rangle \tag{1.40}$$

for $m = -2, -1, 0, 1$, and 2. The states (1.39) are orthonormal and complete, i.e., they span the space of their dimension, 5.

If $\ell = 1/2$, the angular-momentum matrices $\boldsymbol{S}$ are the Pauli matrices

$$\sigma_1 = \begin{pmatrix}0 & 1\\1 & 0\end{pmatrix}, \quad \sigma_2 = \begin{pmatrix}0 & -i\\i & 0\end{pmatrix}, \quad \text{and} \quad \sigma_3 = \begin{pmatrix}1 & 0\\0 & -1\end{pmatrix} \tag{1.41}$$

multiplied by $\hbar/2$

$$\boldsymbol{S} = \frac{\hbar}{2}\boldsymbol{\sigma}. \tag{1.42}$$

So the states $|\frac{1}{2}, m\rangle$

$$|\tfrac{1}{2}, \tfrac{1}{2}\rangle = \begin{pmatrix}1\\0\end{pmatrix} \quad \text{and} \quad |\tfrac{1}{2}, -\tfrac{1}{2}\rangle = \begin{pmatrix}0\\1\end{pmatrix} \tag{1.43}$$

are eigenstates of $S_z$ with eigenvalues $\pm\hbar/2$

$$S_z|\tfrac{1}{2}, m\rangle = m\hbar|\tfrac{1}{2}, m\rangle. \tag{1.44}$$

## 1.7 Eigenstates and Eigenvalues

What about the eigenstates and eigenvalues of $S_x = (\hbar/2)\,\sigma_1$? We call these states $|\frac{1}{2}, m, x\rangle$, set

$$|\tfrac{1}{2}, m, x\rangle = \begin{pmatrix}a\\b\end{pmatrix}, \tag{1.45}$$

and require that

$$\sigma_1|\tfrac{1}{2}, m, x\rangle = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = 2m \begin{pmatrix} a \\ b \end{pmatrix} = \pm \begin{pmatrix} a \\ b \end{pmatrix}. \tag{1.46}$$

So if $m = 1/2$, then $a = b$, while if $m = -1/2$, then $a = -b$. Normalizing these states and setting their arbitrary phases equal to unity, we find that

$$|\tfrac{1}{2}, \tfrac{1}{2}, x\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad \text{and} \quad |\tfrac{1}{2}, -\tfrac{1}{2}, x\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}. \tag{1.47}$$

Similarly, we call $|\tfrac{1}{2}, m, y\rangle$ the eigenstates and eigenvalues of $S_y = (\hbar/2)\,\sigma_2$, set

$$|\tfrac{1}{2}, m, y\rangle = \begin{pmatrix} c \\ d \end{pmatrix}, \tag{1.48}$$

and require that

$$\sigma_2|\tfrac{1}{2}, m, y\rangle = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \begin{pmatrix} c \\ d \end{pmatrix} = 2m \begin{pmatrix} c \\ d \end{pmatrix} = \pm \begin{pmatrix} c \\ d \end{pmatrix}. \tag{1.49}$$

So if $m = 1/2$, then $c = -id$, while if $m = -1/2$, then $c = id$. Normalizing these states and setting their arbitrary phases equal to unity, we find that

$$|\tfrac{1}{2}, \tfrac{1}{2}, y\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ i \end{pmatrix} \quad \text{and} \quad |\tfrac{1}{2}, -\tfrac{1}{2}, y\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -i \end{pmatrix}. \tag{1.50}$$

Matlab knows how to do this faster than we can do it with our fingers. The Matlab command

```
s1 = [ 0 1; 1 0]
```

makes $\sigma_1$, and the command

```
[V,D] = eig(s1)
```

returns the eigenvectors of $\sigma_1$ as

```
V =
   -0.7071    0.7071
    0.7071    0.7071
```

and its eigenvalues $2m$ as

```
D =
    -1      0
     0      1 .
```

These answers are equivalent to those (1.47) we just computed because Matlab reports eigenvalues in increasing order and because eigenvectors are defined only up to an overall factor or up to an overall phase if they are normalized.

The Matlab command

```
s2 = [ 0 -i; i 0]
```

makes $\sigma_2$, and the command

```
[V,D] = eig(s2)
```

returns the eigenvectors of $\sigma_2$ as

```
V =
   0.0000 - 0.7071i   0.0000 - 0.7071i
  -0.7071 + 0.0000i   0.7071 + 0.0000i
```

and its eigenvalues $2m$ as

```
D =
    -1      0
     0      1 .
```

As the size of the matrix increases, so does the utility of Matlab. The Matlab command

```
A =[  1 2 3 ;   4 5 6 ;  -1 2 -3 ]
```

generates the matrix

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ -1 & 2 & -3 \end{pmatrix}, \tag{1.51}$$

and the Matlab command

```
[V,D] = eig(A)
```

returns the eigenvectors as

```
V =
   -0.3530   -0.8654   -0.3682
   -0.9246    0.0842   -0.4126
   -0.1431    0.4940    0.8332
```

and their eigenvalues as

```
D =
    7.4556         0         0
         0   -0.9072         0
         0         0   -3.5484 .
```

To have Matlab check this we enter

```
A*V
V*D
```

and get

```
ans =
   -2.6315    0.7851    1.3064
   -6.8936   -0.0764    1.4639
   -1.0670   -0.4481   -2.9566

ans =
   -2.6315    0.7851    1.3064
   -6.8936   -0.0764    1.4639
   -1.0670   -0.4481   -2.9566
```

which verifies Matlab's computation.

## 1.8 Linear Independence

The usual unit vectors $\hat{\boldsymbol{x}}$, $\hat{\boldsymbol{y}}$, and $\hat{\boldsymbol{z}}$ of 3-space are linearly independent. We can write every point in 3-space uniquely as

$$\boldsymbol{r} = r_x \hat{\boldsymbol{x}} + r_y \hat{\boldsymbol{y}} + r_z \hat{\boldsymbol{z}}. \tag{1.52}$$

In terms of $\hat{\boldsymbol{x}}$, $\hat{\boldsymbol{y}}$, and $\hat{\boldsymbol{z}}$, the only way to write the origin is

$$\boldsymbol{0} = 0\,\hat{\boldsymbol{x}} + 0\,\hat{\boldsymbol{y}} + 0\,\hat{\boldsymbol{z}}. \tag{1.53}$$

This silly equation is nearly the definition of linear independence: The $n$ vectors $V_1, \ldots, V_n$ (all with the same number $m$ of components) are linearly independent if (and only if) the only way to write the zero $m$-vector is

$$0 = 0\,V_1 + 0\,V_2 + \cdots + 0\,V_{n-1} + 0\,V_n. \tag{1.54}$$

The vectors $V_i$ might be linearly independent if

$$m \ge n \tag{1.55}$$

but can't be linearly independent if

$$m < n. \tag{1.56}$$

The $n$ real (complex) vectors $V_1, \ldots, V_n$ (all with the same number $m$ of components) are linearly dependent if (and only if) one can find $n$ real (complex) numbers $r_i$ ($z_i$), not all of which are zero, with which to write the zero $m$-vector as

$$\begin{aligned} 0 &= r_1 V_1 + r_2 V_2 + \cdots + r_{n-1} V_{n-1} + r_n V_n \\ &\text{or as} \\ 0 &= z_1 V_1 + z_2 V_2 + \cdots + z_{n-1} V_{n-1} + z_n V_n \end{aligned} \tag{1.57}$$

So any three 2-vectors are linearly dependent.

## 1.9 Determinants

Levi-Civita invented various totally antisymmetric symbols such as the 2-index symbol

$$1 = \epsilon_{12} = -\epsilon_{21} \quad 0 = \epsilon_{11} = \epsilon_{22} \tag{1.58}$$

and the 3-index symbol

$$\begin{aligned} 1 &= \epsilon_{123} = \epsilon_{231} = \epsilon_{312} \\ -1 &= \epsilon_{213} = \epsilon_{132} = \epsilon_{321}. \end{aligned} \tag{1.59}$$

The Levi-Civita symbols are zero whenever any index occurs more than once

$$\begin{aligned} 0 &= \epsilon_{111} = \epsilon_{112} = \epsilon_{113} \\ 0 &= \epsilon_{221} = \epsilon_{222} = \epsilon_{223} \\ &\vdots \end{aligned} \tag{1.60}$$

In terms of the L-C symbols, the **determinant** of a $2 \times 2$ matrix $A$ is

$$\det A = |A| = A_{11}A_{22} - A_{21}A_{12} = \sum_{i,j=1}^{2} \epsilon_{ij} A_{i1} A_{j2}, \tag{1.61}$$

that of a $3 \times 3$ matrix is

$$\det A = \sum_{i,j,k=1}^{3} \epsilon_{ijk}\, A_{i1} A_{j2} A_{k3} = \sum_{i,j,k=1}^{3} \epsilon_{ijk}\, A_{1i} A_{2j} A_{3k}, \tag{1.62}$$

that of a $4 \times 4$ matrix is

$$\det A = \sum_{i,j,k,\ell=1}^{4} \epsilon_{ijk\ell}\, A_{i1} A_{j2} A_{k3} A_{\ell 4} = \sum_{i,j,k=1}^{3} \epsilon_{ijk\ell}\, A_{1i} A_{2j} A_{3k} A_{4\ell} \tag{1.63}$$

and so forth.

If the columns of a $2 \times 2$ matrix $A$ are two 2-vectors $V^1$ and $V^2$, then its determinant is

$$\det A = \sum_{i,j=1}^{2} \epsilon_{ij}\, V_i^1 V_j^2 = V_1^1 V_2^2 - V_2^1 V_1^2. \tag{1.64}$$

Note that because the L-C symbol is antisymmetric

$$\sum_{i,j=1}^{2} \epsilon_{ij}\, V_i^2 V_j^2 = V_1^2 V_2^2 - V_2^2 V_1^2 = 0. \tag{1.65}$$

So the antisymmetry of the L-C symbol means that $\det A$ does not change if we add to $V^1$ a multiple of $V^2$:

$$\sum_{i,j=1}^{2} \epsilon_{ij}\, (V_i^1 + w V_i^2) V_j^2 = \sum_{i,j=1}^{2} \epsilon_{ij}\, V_i^1 V_j^2 = \det A. \tag{1.66}$$

If the 2-vectors $V^1$ and $V^2$ are linearly dependent, then we can find numbers $z_1$ and $z_2$ such that

$$0 = z_1 V^1 + z_2 V^2 \quad \text{and so too} \quad 0 = V^1 + (z_2/z_1) V^2. \tag{1.67}$$

So setting $V^1 = -z_1/z_2 V^2 \equiv -w V^2$, we see that the determinant of a matrix made of two linearly dependent 2-vectors vanishes

$$\det A = \sum_{i,j=1}^{2} \epsilon_{ij}\, V_i^1 V_j^2 = \sum_{i,j=1}^{2} \epsilon_{ij}\, (-w V_i^2) V_j^2 = 0. \tag{1.68}$$

This result generalizes to every number of dimensions: The determinant of an $n \times n$ matrix whose columns are linearly dependent vanishes.

## 1.10 Eigenstates and Eigenvalues

Now let's consider eigenvector problem

$$AZ = \lambda Z \tag{1.69}$$

in which $A$ is a $2 \times 2$ matrix, $Z$ is a 2-vector, and $\lambda$ is an eigenvalue. Subtracting $\lambda$ times the $2 \times 2$ identity matrix from both sides of this equation, we have

$$(A - \lambda I)Z = 0. \tag{1.70}$$

Now defining $V^1$ and $V^2$ to be the two columns of the $2 \times 2$ matrix $A - \lambda I$

$$V^1 = \begin{pmatrix} A_{11} - \lambda \\ A_{21} \end{pmatrix} \quad \text{and} \quad V^2 = \begin{pmatrix} A_{12} \\ A_{22} - \lambda \end{pmatrix}, \tag{1.71}$$

we see that the eigenvector equation

$$0 = (A - \lambda I)Z = z_1 V^1 + z_2 V^2 \tag{1.72}$$

is the statement that the two vectors $V^1$ and $V^2$ are linearly dependent. But if $V^1$ and $V^2$ are linearly dependent, then the determinant of the matrix $A - \lambda I$ must vanish

$$0 = \det(A - \lambda I). \tag{1.73}$$

This is a quadratic equation for $\lambda$. It has two solutions, $\lambda_1$ and $\lambda_2$. Sometimes the solutions are the same, and $\lambda_1 = \lambda_2$.

In terms of the eigenvalues $\lambda_1$ and $\lambda_2$, one has two equations for the components $z_1$ and $z_2$ of the eigenvectors:

$$\begin{aligned} 0 &= z_1(A_{11} - \lambda_i) + z_2 A_{12} \\ 0 &= z_1 A_{21} + z_2(A_{22} - \lambda_i). \end{aligned} \tag{1.74}$$

These two equations say that

$$z_2 = -\frac{A_{11} - \lambda_i}{A_{12}} z_1 \quad \text{and} \quad z_2 = -\frac{A_{21}}{A_{22} - \lambda_i} z_1, \tag{1.75}$$

and so are consistent only if

$$(A_{11} - \lambda_i)(A_{22} - \lambda_i) = A_{12} A_{21} \tag{1.76}$$

which is to say, only if $\det(A - \lambda I) = 0$. You might think that we have 6 unknowns here — 2 $\lambda$'s, 2 $z_1$'s, and 2 $z_2$'s. But eigenvectors are defined only up to a complex factor. So the two equations (1.75) don't determine $z_1$ or $z_2$ but only their ratio $w = z_1/z_2$. So we have 4 unknowns: $w_1$, $w_2$, $\lambda_1$, and $\lambda_2$ and three equations, the two of (1.75) and (1.76).

We can illustrate this by dividing the two equations (1.74) by $z_2$ so as to get 2 equations for 2 unknowns $w = z_1/z_2$ and $\lambda$

$$\begin{aligned} 0 &= w(A_{11} - \lambda) + A_{12} \\ 0 &= w A_{21} + A_{22} - \lambda. \end{aligned} \tag{1.77}$$

The second equation gives $\lambda = wA_{21}+A_{22}$. When we substitute this formula for $w$ into the first equation, we get the quadratic equation

$$A_{21}w^2 + (A_{22} - A_{11})w - A_{12} = 0 \tag{1.78}$$

which we can solve for its two roots $w_i = (z_1/z_2)_i$. The eigenvalues then are $\lambda_i = w_i A_{21} + A_{22}$.

Such results hold for $n \times n$ eigenvector problems: The $n \times n$ determinant $\det(A - \lambda I)$ vanishes if and only if $\lambda$ is an eigenvalue of the $n \times n$ matrix $A$

$$AZ = \lambda Z \iff \det(A - \lambda I) = 0. \tag{1.79}$$

This is an $n$th-order polynomial equation

$$0 = z_n \lambda^n + z_{n-1}\lambda^{n-1} + \cdots + z_1\lambda + z_0 \tag{1.80}$$

also called an $n$th-degree polynomial equation. We can easily solve quadratic equations. Cubic ones are more complicated; quartic ones are even more complicated; and equations of higher order in general have only numerical solutions.

So the computer programs that find the eigenvalues and eigenvectors of $n \times n$ matrices don't compute the determinant and set it equal to zero. Indeed, the determinant of an $n \times n$ matrix has $n!$ terms, which becomes unmanageable for $n \gtrsim 10$ and hopeless for $n \gtrsim 20$ since

```
factorial(10) = 3628800 and factorial(20) = 2.4329e+18.
```

So the computer programs that solve for eigenvalues and eigenvectors use efficient algorithms such as the $LU$ decomposition, $A = LU$ in which the matrix $L$ is lower triangular with 1 on its main diagonal, and $U$ is upper triangular (Tadeusz Banachiewicz 1882–1954). For example, the $LU$ decomposition of the matrix $A$ is

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ z & 1 \end{pmatrix} \begin{pmatrix} \alpha & \beta \\ 0 & \gamma \end{pmatrix} = \begin{pmatrix} \alpha & \beta \\ z\alpha & z\beta + \gamma \end{pmatrix} \tag{1.81}$$

in which $\alpha = a$, $\beta = b$, $z = c/\alpha = c/a$, and $\gamma = d - z\beta = d - c\beta/a = d - bc/a$. So the $LU$ decomposition of $A$ is

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ c/a & 1 \end{pmatrix} \begin{pmatrix} a & b \\ 0 & d - bc/a \end{pmatrix}. \tag{1.82}$$

Matlab can do this for you:

```
>> syms a b c d
>> A = [ a b ; c d]
```

```
A =

[ a, b]
[ c, d]

>> [L,U] = lu(A)

L =

[   1, 0]
[ c/a, 1]


U =

[ a,           b]
[ 0, d - (b*c)/a].
>> L*U

ans =

[ a, b]
[ c, d] .
```

Matlab's webpages describe this and other ways to decompose a matrix and to find its eigenvectors and eigenvalues `https://www.mathworks.com/help/matlab/linear-algebra.html`.

The $LU$ decomposition lets one compute determinants much more easily than with Levi-Civita's symbols. The reason is that

$$\det A = \det(LU) = \det L \det U, \tag{1.83}$$

and the determinant of a triangular matrix is just the product of its diagonal elements. So

$$\det L = 1, \tag{1.84}$$

and

$$\det U = U_{11} U_{22} \cdots U_{nn}. \tag{1.85}$$

So in the $2 \times 2$ example (1.82)

$$\det A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \det a(d - bc/a) = ad - bc. \tag{1.86}$$

## 1.11 Eigenvectors and Eigenvalues from Matlab

First, we make a real $3 \times 3$ random matrix

```
>> M = rand(3)

M =

    0.3171    0.4387    0.7952
    0.9502    0.3816    0.1869
    0.0344    0.7655    0.4898. .
```

Matlab's eig gives its eigenvalues

```
>> e=eig(M)

e =

  -0.1327 + 0.4941i
  -0.1327 - 0.4941i
   1.4539 + 0.0000i .
```

They are complex even though $M$ is real because $M$ is not symmetric, $M \neq M^{\mathsf{T}}$. Matlab gives $M$'s eigenvectors as

```
>> [V, D] = eig(M)

V =

  -0.2541 + 0.4076i  -0.2541 - 0.4076i   0.5967 + 0.0000i
   0.6370 + 0.0000i   0.6370 + 0.0000i   0.6180 + 0.0000i
  -0.4610 - 0.3885i  -0.4610 + 0.3885i   0.5120 + 0.0000i


D =

  -0.1327 + 0.4941i   0.0000 + 0.0000i   0.0000 + 0.0000i
   0.0000 + 0.0000i  -0.1327 - 0.4941i   0.0000 + 0.0000i
```

```
   0.0000 + 0.0000i   0.0000 + 0.0000i   1.4539 + 0.0000i .
```

Now we generate two $4 \times 4$ random matrices, multiply one by $i$, and add them to get a $4 \times 4$ complex matrix $C$. Then we get its eigenvalues as eig(C)

```
>> R = rand(4)

R =

    0.0975    0.9649    0.4854    0.9157
    0.2785    0.1576    0.8003    0.7922
    0.5469    0.9706    0.1419    0.9595
    0.9575    0.9572    0.4218    0.6557

>> C = R + i*rand(4)

C =

   0.0975 + 0.0357i   0.9649 + 0.7577i   0.4854 + 0.1712i   0.9157 + 0.0462i
   0.2785 + 0.8491i   0.1576 + 0.7431i   0.8003 + 0.7060i   0.7922 + 0.0971i
   0.5469 + 0.9340i   0.9706 + 0.3922i   0.1419 + 0.0318i   0.9595 + 0.8235i
   0.9575 + 0.6787i   0.9572 + 0.6555i   0.4218 + 0.2769i   0.6557 + 0.6948i

>> e = eig(C)

e =

   2.5716 + 1.9648i
  -0.0536 - 0.4794i
  -0.7540 + 0.3692i
  -0.7112 - 0.3491i  .
```

To also get the eigenvectors, we use [V,D] = eig(A):

```
>> [V,D] = eig(C)

V =

   0.4030 - 0.0924i  -0.4470 - 0.0128i  -0.2896 + 0.2366i   0.6414 + 0.0000i
   0.4741 + 0.0716i   0.0277 + 0.3895i  -0.5716 + 0.0382i  -0.2405 + 0.2058i
   0.5317 + 0.0069i   0.7387 + 0.0000i   0.4358 - 0.0069i   0.0443 - 0.6354i
   0.5625 + 0.0000i  -0.0478 - 0.3155i   0.5847 + 0.0000i  -0.2823 + 0.0554i
```

```
D =

   2.5716 + 1.9648i   0.0000 + 0.0000i   0.0000 + 0.0000i   0.0000 + 0.0000i
   0.0000 + 0.0000i  -0.0536 - 0.4794i   0.0000 + 0.0000i   0.0000 + 0.0000i
   0.0000 + 0.0000i   0.0000 + 0.0000i  -0.7540 + 0.3692i   0.0000 + 0.0000i
   0.0000 + 0.0000i   0.0000 + 0.0000i   0.0000 + 0.0000i  -0.7112 - 0.3491i

>> C*V - V*D

ans =

   1.0e-14 *

   0.2442 + 0.1443i  -0.0507 - 0.0111i  -0.0389 + 0.0056i   0.0444 + 0.0222i
  -0.0888 + 0.0000i   0.0167 - 0.0860i   0.0000 + 0.0194i   0.0111 - 0.0125i
   0.1110 + 0.1332i   0.0139 - 0.0167i   0.0000 - 0.0194i   0.0000 - 0.0222i
   0.1998 + 0.1332i   0.0083 + 0.0014i  -0.0500 + 0.0250i   0.0611 - 0.0160i .
```

This last result says that apart from roundoff errors of order 1e-14

```
C V = V D
```

which says that the columns of V are the eigenvectors of C with eigenvalues that are the nonzero elements of the diagonal matrix D. In detail, this is

$$\sum_{k=1}^{4} C_{ik} V_{kj} = D_{jj} V_{ij}. \tag{1.87}$$

## 1.12 Linear Least Squares

Matlab says:

## 1.13 States and Density Operators

In quantum mechanics, we often represent systems by states, e.g., $|1,0,0\rangle$ for the ground state of hydrogen with energy $E_1$ and angular momentum $\ell = 0$ (here spin is neglected) and $|2,1,1\rangle$ for the $n = 2$ state with energy $E_2$ and angular momentum $\hbar$. We can add states

$$|\psi\rangle = z|1,0,0\rangle + w|2,1,1\rangle. \tag{1.88}$$

This state $|\psi\rangle$ is entangled because if we measure its energy to be $E_1$ or $E_2$ then we know that its angular momentum is $\ell = 0$ or $\ell = \hbar$.

We use density operators to describe systems about which we know less. For instance, the density operator

$$\rho = \tfrac{1}{2}|1,0,0\rangle\langle 1,0,0| + \tfrac{1}{2}|2,1,1\rangle\langle 2,1,1| \tag{1.89}$$

represents a system whose energy is equally likely to be $E_1$ or $E_2$.

The operators we use most often in quantum mechanics are linear operators. For example, the hamiltonian $H$ for a hydrogen atom maps the state $|\psi\rangle$ into

$$H|\psi\rangle = zH|1,0,0\rangle + wH|2,1,1\rangle = zE_1|1,0,0\rangle + wE_2|2,1,1\rangle. \tag{1.90}$$

Here we used the fact that the state $|1,0,0\rangle$ is an eigenstate of $H$ with eigenvalue $E_1$ and that $|2,0,0\rangle$ is an eigenstate of $H$ with eigenvalue $E_2$

$$H|1,0,0\rangle = E_1|1,0,0\rangle \quad \text{and} \quad H|2,1,1\rangle = E_2|2,1,1\rangle. \tag{1.91}$$

These eigenstate equations are extensions to vector spaces of infinite dimension of the concepts we learned in Sections (1.6–1.10).

## 1.14 Notes on Some Problems

Problem 1.40 corrected: The coherent state $|\{\alpha(\boldsymbol{k},\ell)\}\rangle$ is an eigenstate of the annihilation operator $a(\boldsymbol{k},\ell)$ with eigenvalue $\alpha(\boldsymbol{k},\ell)$ for each mode of the electromagnetic field with wavenumber $\boldsymbol{k}$ and polarization $\ell$

$$a(\boldsymbol{k},\ell)|\{\alpha(\boldsymbol{k},\ell)\}\rangle = \alpha(\boldsymbol{k},\ell)|\{\alpha(\boldsymbol{k},\ell)\}\rangle. \tag{1.92}$$

The positive-frequency part $E_i^{(+)}(t,\boldsymbol{x})$ of the electric field is a linear combination of the annihilation operators

$$E_i^{(+)}(t,\boldsymbol{x}) = \sum_{\boldsymbol{k}} \sum_{\ell=1}^{2} a(\boldsymbol{k},\ell)\, e_i(\boldsymbol{k},\ell)\, e^{i(\boldsymbol{k}\cdot\boldsymbol{x}-\omega t)} \tag{1.93}$$

in which the wavenumber $\boldsymbol{k}$ is summed over $\boldsymbol{k} = 2\pi\boldsymbol{n}/L$ where $L$ is an appropriate length such as the length of a laser cavity or the width of the universe, and $\boldsymbol{n}$ is a vector of integers. Show that $|\{\alpha_k\}\rangle$ is an eigenstate of $E_i^{(+)}(t,\boldsymbol{x})$

$$E_i^{(+)}(t,\boldsymbol{x})|\{\alpha(\boldsymbol{k},\ell)\}\rangle = \mathcal{E}_i^{(+)}(t,\boldsymbol{x})|\{\alpha(\boldsymbol{k},\ell)\}\rangle \tag{1.94}$$

and find its eigenvalue $\mathcal{E}_i^{(+)}(t,\boldsymbol{x})$.

# 2
# Examples for Chapter 2, Vector Calculus

## 2.1 Helmholtz Decomposition

We can use the delta-function formula (2.34) to write any suitably smooth 3-dimensional vector field $\boldsymbol{V}(\boldsymbol{x})$ as

$$\boldsymbol{V}(\boldsymbol{x}) = -\int \boldsymbol{V}(\boldsymbol{r}) \,\triangle \left( \frac{1}{|\boldsymbol{r}-\boldsymbol{x}|} \right) d^3r \tag{2.1}$$

in which the derivatives $\boldsymbol{\nabla}\cdot\boldsymbol{\nabla} = \boldsymbol{\nabla}^2 = \triangle$ can be both with respect to $\boldsymbol{x}$ or both with respect to $\boldsymbol{r}$. Taking them to be with respect to $\boldsymbol{x}$, we have

$$\boldsymbol{V}(\boldsymbol{x}) = -\boldsymbol{\nabla}^2 \int \frac{\boldsymbol{V}(\boldsymbol{r})}{|\boldsymbol{r}-\boldsymbol{x}|} \, d^3r. \tag{2.2}$$

We now use our formula (2.49) for the curl of a curl

$$\boldsymbol{\nabla}^2 \boldsymbol{V} = \boldsymbol{\nabla}(\boldsymbol{\nabla}\cdot\boldsymbol{V}) - \boldsymbol{\nabla}\times(\boldsymbol{\nabla}\times\boldsymbol{V}) \tag{2.3}$$

to write $\boldsymbol{V}(\boldsymbol{x})$ as

$$\boldsymbol{V}(\boldsymbol{x}) = -\boldsymbol{\nabla}\left( \boldsymbol{\nabla}\cdot \int \frac{\boldsymbol{V}(\boldsymbol{r})}{|\boldsymbol{r}-\boldsymbol{x}|} \, d^3r \right) + \boldsymbol{\nabla}\times\left( \boldsymbol{\nabla}\times \int \frac{\boldsymbol{V}(\boldsymbol{r})}{|\boldsymbol{r}-\boldsymbol{x}|} \, d^3r \right). \tag{2.4}$$

Thus any suitably smooth 3-dimensional vector field $\boldsymbol{V}(\boldsymbol{x})$ can be written as the sum

$$\boldsymbol{V}(\boldsymbol{x}) = \boldsymbol{\nabla}\phi(\boldsymbol{x}) + \boldsymbol{\nabla}\times\boldsymbol{A}(\boldsymbol{x}) \tag{2.5}$$

of the gradient of a scalar field

$$\phi(\boldsymbol{x}) = -\boldsymbol{\nabla}\cdot \int \frac{\boldsymbol{V}(\boldsymbol{r})}{|\boldsymbol{r}-\boldsymbol{x}|} \, d^3r \tag{2.6}$$

and the curl of a vector field

$$\boldsymbol{A}(\boldsymbol{x}) = \boldsymbol{\nabla}\times \int \frac{\boldsymbol{V}(\boldsymbol{r})}{|\boldsymbol{r}-\boldsymbol{x}|} \, d^3r \tag{2.7}$$

(Hermann von Helmholtz, 1821–1894).

# 3
# Examples for Chapter 3, Fourier Series

## 3.1 Fourier and Dirac

If we combine the Fourier series (3.2)

$$f(x) = \sum_{n=-\infty}^{\infty} f_n \, \frac{e^{inx}}{\sqrt{2\pi}} \tag{3.1}$$

with the formula (3.3) for the Fourier coefficients

$$f_n = \int_0^{2\pi} \frac{e^{-inx}}{\sqrt{2\pi}} \, f(x) \, dx, \tag{3.2}$$

then we get equation (3.120):

$$f(x) = \sum_{n=-\infty}^{\infty} f_n \, \frac{e^{inx}}{\sqrt{2\pi}} = \sum_{n=-\infty}^{\infty} \int_0^{2\pi} \frac{e^{-iny}}{\sqrt{2\pi}} \, f(y) \frac{e^{inx}}{\sqrt{2\pi}} \, dy. \tag{3.3}$$

If the function $f(x)$ is suitably smooth, then we may change the order of summation and the integration and get for $0 \le x \le 2\pi$

$$f(x) = \int_0^{2\pi} \left( \sum_{n=-\infty}^{\infty} \frac{e^{in(x-y)}}{2\pi} \right) f(y) \, dy \tag{3.4}$$

which is (3.121). This equation says that for $0 \le x, y \le 2\pi$

$$\sum_{n=-\infty}^{\infty} \frac{e^{in(x-y)}}{2\pi} = \delta(x - y). \tag{3.5}$$

But the right-hand side of (2.4) is periodic in $x$ with period $2\pi$, and it defines the periodic extension $f_p(x)$ of the function $f(x)$ from the interval $[0, 2\pi]$ to the whole real line

$$f_p(x) = f_p(x + 2\pi m) = \int_0^{2\pi} \left( \sum_{n=-\infty}^{\infty} \frac{e^{in(x-y)}}{2\pi} \right) f(y)\, dy \tag{3.6}$$

in which $m$ is an integer. So the sum of phases is a sum of delta functions

$$\sum_{n=-\infty}^{\infty} \frac{e^{in(x-y)}}{2\pi} = \sum_{m=-\infty}^{\infty} \delta(x - y - 2\pi m) \tag{3.7}$$

which is (3.123). This is the Dirac comb. It is illustrated in Fig. 3.11.

## 3.2 Hilbert and Dirac

The Fourier series is an example of a much more general class of series. Suppose $H_n(x), n = 0, 1, 2, \ldots, \infty$ is a set of orthonormal functions

$$\int_a^b H_n^*(x)\, H_m(x)\, dx = \delta_{nm}. \tag{3.8}$$

These functions span a vector space $S$ of functions

$$f(x) = \sum_{n=0}^{\infty} f_n\, H_n(x). \tag{3.9}$$

The orthonormality (3.8) of these functions implies that the coefficients $f_n$ of the expansion (3.9) are

$$\int_a^b f(x)\, H_n^*(x)\, dx = \int_a^b \left( \sum_{m=0}^{\infty} f_m\, H_m(x) \right) H_n^*(x)\, dx = \sum_{m=0}^{\infty} f_m\, \delta_{nm} = f_n. \tag{3.10}$$

These three equations (3.8–3.10) are analogous to the three basic equations (3.1–3.3) of Fourier series.

By combining the expansion (3.9) of $f(x)$ with the formula (3.10) for its coefficients $f_n$, we have

$$\begin{aligned} f(x) &= \sum_{n=0}^{\infty} f_n\, H_n(x) = \sum_{n=0}^{\infty} \int_a^b f(y)\, H_n^*(y)\, dy\, H_n(x) \\ &= \int_a^b \left( \sum_{n=0}^{\infty} H_n^*(y)\, H_n(x) \right) f(y)\, dy \end{aligned} \tag{3.11}$$

for all points $a \le x \le b$. Thus the orthonormal functions $H_n(x)$ provide a representation for the delta function suitable for functions in the space $S$ for $a \le x, y \le b$

$$\delta(x-y) = \sum_{n=0}^{\infty} H_n^*(y)\, H_n(x) \tag{3.12}$$

which is analogous to the expansion (3.5) of the delta function. It differs from the Dirac comb (3.7) because the orthonormal functions $H_n(x)$ may not be periodic.

## 3.3 Example 3.8 again

Example 3.8 derived the Fourier series for the function that is $1 + \cos 2x$ for $|x| \le \pi/2$ and zero otherwise. The series for the function that is $1 + \cos 2x$ for all $x$ is much simpler:

$$1 + \cos 2x = 1 + \frac{1}{2}\left(e^{2ix} + e^{-2ix}\right). \tag{3.13}$$

Its coefficients $f_n$ are nonzero only for $n = -2, 0, 2$.

# 4
# Examples for Chapter 4, Fourier Transforms

## 4.1 Fourier derivation of Helmholtz decomposition

The Levi-Civita identity (2.56)

$$\sum_{i=1}^{3} \epsilon_{ijk}\epsilon_{imn} = \delta_{jm}\delta_{kn} - \delta_{jn}\delta_{km} \tag{4.1}$$

implies that

$$\begin{aligned} \boldsymbol{k} \times (\boldsymbol{k} \times \boldsymbol{V}) &= \sum_{j,k,\ell=1}^{3} \epsilon_{ijk}\epsilon_{k\ell m}\, k_j\, k_\ell\, V_m \\ &= \sum_{j,\ell=1}^{3} (\delta_{i\ell}\delta_{jm} - \delta_{im}\delta_{j\ell})\, k_j\, k_\ell\, V_m \\ &= (\boldsymbol{k}\cdot\boldsymbol{V})\, k_i - (\boldsymbol{k}\cdot\boldsymbol{k})\, V_i. \end{aligned} \tag{4.2}$$

Thus we can use any nonzero 3-vector $\boldsymbol{k}$ to write every 3-vector $\boldsymbol{V}$ as

$$\boldsymbol{V} = \frac{1}{\boldsymbol{k}\cdot\boldsymbol{k}}\Big( (\boldsymbol{k}\cdot\boldsymbol{V})\,\boldsymbol{k} - \boldsymbol{k}\times(\boldsymbol{k}\times\boldsymbol{V})\Big). \tag{4.3}$$

This expansion lets us write the Fourier transform $\boldsymbol{V}(\boldsymbol{k})$ of any square-integrable 3-vector field $\boldsymbol{V}(\boldsymbol{x})$ as

$$\begin{aligned} \boldsymbol{V}(\boldsymbol{x}) &= \int \boldsymbol{V}(\boldsymbol{k})\, e^{i\boldsymbol{k}\cdot\boldsymbol{x}}\, d^3x = \int \Big( \frac{\boldsymbol{k}(\boldsymbol{k}\cdot\boldsymbol{V}(\boldsymbol{k}))}{\boldsymbol{k}^2} - \frac{\boldsymbol{k}\times(\boldsymbol{k}\times\boldsymbol{V}(\boldsymbol{k}))}{\boldsymbol{k}^2}\Big)\, e^{i\boldsymbol{k}\cdot\boldsymbol{x}}\, d^3x \\ &= \int \frac{\boldsymbol{k}(\boldsymbol{k}\cdot\boldsymbol{V}(\boldsymbol{k}))}{\boldsymbol{k}^2}\, e^{i\boldsymbol{k}\cdot\boldsymbol{x}} - \frac{\boldsymbol{k}\times(\boldsymbol{k}\times\boldsymbol{V}(\boldsymbol{k}))}{\boldsymbol{k}^2}\, e^{i\boldsymbol{k}\cdot\boldsymbol{x}}\, d^3x \\ &= \nabla \int \frac{-i\,\boldsymbol{k}\cdot\boldsymbol{V}(\boldsymbol{k})}{\boldsymbol{k}^2}\, e^{i\boldsymbol{k}\cdot\boldsymbol{x}}\, d^3x + \nabla\times \int \frac{i\,\boldsymbol{k}\times\boldsymbol{V}(\boldsymbol{k})}{\boldsymbol{k}^2}\, e^{i\boldsymbol{k}\cdot\boldsymbol{x}}\, d^3x \end{aligned} \tag{4.4}$$

which is the sum of the gradient of a scalar field plus the curl of a vector field.

## 4.2 3D Delta Function

The 3-dimensional delta function is

$$\delta(\boldsymbol{x}-\boldsymbol{y}) = \int e^{\pm i(\boldsymbol{x}-\boldsymbol{y})\cdot\boldsymbol{k}} \, \frac{d^3k}{(2\pi)^3} \tag{4.5}$$

in which you can use either + or −. The $n$-dimensional delta function is

$$\delta(\boldsymbol{x}-\boldsymbol{y}) = \int e^{\pm i(\boldsymbol{x}-\boldsymbol{y})\cdot\boldsymbol{k}} \, \frac{d^nk}{(2\pi)^n}. \tag{4.6}$$

The Laplace transform of $t^{s-1}$ is related to the gamma function (5.58)

$$s^{-z}\,\Gamma(s) = \int_0^\infty dt\, e^{-st}\, t^{s-1}. \tag{4.7}$$

# 5

# Examples for Chapter 5, Series

## 5.1 Convergence

The trace is cyclic so we expect that

$$\mathrm{Tr}(q\,p) = \mathrm{Tr}(p\,q). \tag{5.1}$$

But then we'd have

$$\mathrm{Tr}\big([q,p]\big) = 0. \tag{5.2}$$

But we know that

$$[q,p] = i\hbar, \tag{5.3}$$

so now we have

$$0 = \mathrm{Tr}\big([q,p]\big) = \mathrm{Tr}(i\hbar) = i\hbar\mathrm{Tr}(I) = i\hbar\,\infty. \tag{5.4}$$

The raising and lowering operators explained in Section 3.12 offer an equivalent paradox

$$0 = \mathrm{Tr}\big([a,a^\dagger]\big) = \mathrm{Tr}(1) = \mathrm{Tr}(I) = \infty. \tag{5.5}$$

These equations don't represent a breakdown of quantum mechanics. They present us with divergent series

$$\begin{aligned} \mathrm{Tr}\big([a,a^\dagger]\big) &= \mathrm{Tr}\big(aa^\dagger\big) - \mathrm{Tr}\big(a^\dagger a\big) \\ &= \sum_{n=0}^{\infty}\langle n|aa^\dagger|n\rangle - \sum_{n=0}^{\infty}\langle n|a^\dagger a|n\rangle \\ &= \sum_{n=0}^{\infty} n+1 - \sum_{n=0}^{\infty} n = \sum_{n=0}^{\infty} 1 = \infty. \end{aligned} \tag{5.6}$$

They tell us to keep in mind that not all series converge appropriately in every equation.

## 5.2 Geometric Series

**Example 5.1** (Geometric series) The sum of the first $n$ terms of a geometric series (5.10) is

$$S_n(z) = \sum_{k=0}^{n} z^k = \frac{1 - z^{k+1}}{1 - z}. \tag{5.7}$$

This Matlab script shows how this series goes for $z = 1/2$ in blue and $z = 2$ in red:

```
n=0:1:200;
z = 2;
s = (1-z.^(n+1))./(1-z);
plot(n,s,'-r','LineWidth',2)
hold on
z = 1.1;
s = (1-z.^(n+1))./(1-z);
plot(n,s,'-m','LineWidth',2)
z = 1.01;
s = (1-z.^(n+1))./(1-z);
plot(n,s,'-','LineWidth',2,'Color',[.8 0 .5])
z = 1.001;
s = (1-z.^(n+1))./(1-z);
plot(n,s,'-','LineWidth',2,'Color',[.5 0 .5])
z = 0.5;
s = (1-z.^(n+1))./(1-z);
plot(n,s,'-b','LineWidth',2)
z = 0.9;
s = (1-z.^(n+1))./(1-z);
plot(n,s,'-g','LineWidth',2)
z = 0.99;
s = (1-z.^(n+1))./(1-z);
plot(n,s,'-','LineWidth',2,'Color',[0 .5 .5])
z = 0.999;
s = (1-z.^(n+1))./(1-z);
plot(n,s,'-','LineWidth',2,'Color',[.05,0,.55])
axis([0 100 0 100])
textx='$n$';
xlabel(textx,'Interpreter','latex')
texty='Sum';
ylabel(texty)
```

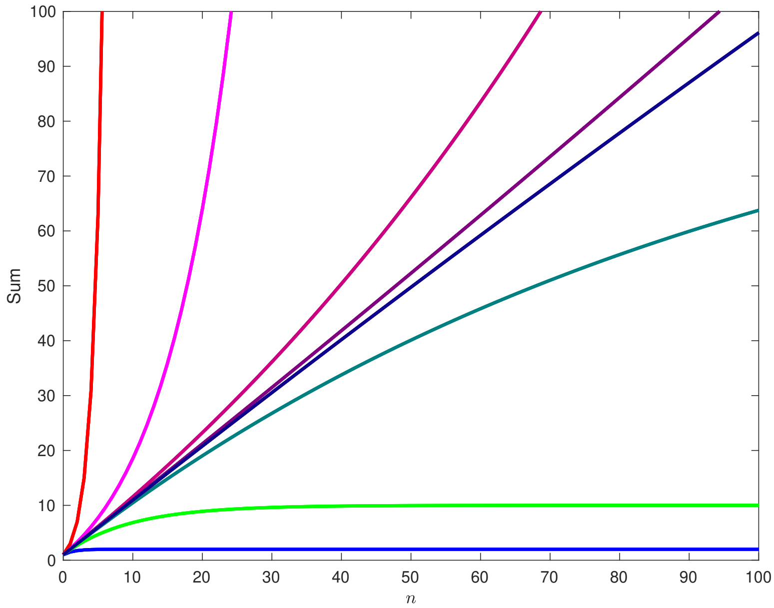


Figure 5.1 The geometric series (5.7) for $z = 2$ red, 1.1 magenta, 1.01 dark red, 1.001 0.9 purple, 0.5 blue, 0.9 green, 0.99 blue green, and 0.999 navy blue.

```
print -dpdf ~/papers/math/PowerSum
print -depsc ~/papers/math/PowerSum
```

The resulting plot is Fig. 5.1.

## 5.3 Leibniz's Rule

In the notation

$$f^{(n)}(x) \equiv \frac{d^n}{dx^n} f(x) \tag{5.8}$$

for derivatives Leibniz's rule for differentiating the product of two functions is

$$\frac{d^n}{dx^n} [f(x)\, g(x)] = \sum_{k=0}^{n} \binom{n}{k} f^{(k)}(x)\, g^{(n-k)}(x) \tag{5.9}$$

(Gottfried Leibniz, 1646–1716).

The rule is obviously true for $n = 0$ and $n = 1$, and one may use mathematical induction to prove it. Keeping in mind that $(-1)! = \infty$, we find for

$0 \le k \le n$

$$\begin{aligned}\binom{n-1}{k} + \binom{n-1}{k-1} &= \frac{(n-1)!}{k!(n-k-1)!} + \frac{(n-1)!}{(k-1)!(n-k)!} \\ &= \frac{(n-1)!}{(k-1)!(n-k-1)!}\left[\frac{1}{k} + \frac{1}{n-k}\right] \\ &= \frac{(n-1)!}{(k-1)!(n-k-1)!}\left[\frac{n}{k(n-k)}\right] \\ &= \frac{n!}{k!(n-k)!} = \binom{n}{k}.\end{aligned} \tag{5.10}$$

Let's assume that for $\ell = 0 \ldots n-1$ that

$$\frac{d^\ell}{dx^\ell}\left[f(x)\, g(x)\right] = \sum_{k=0}^{\ell} \binom{\ell}{k} f^{(k)}(x)\, g^{(\ell-k)}(x).$$

So for $\ell = n-1$, we have

$$\frac{d^{n-1}}{dx^{n-1}}\left[f(x)\, g(x)\right] = \sum_{k=0}^{n-1} \binom{n-1}{k} f^{(k)}(x)\, g^{(n-1-k)}(x).$$

So using the result (5.10) of the first part of this exercise, we get

$$\begin{aligned}\frac{d^n}{dx^n}(f\, g) &= \frac{d}{dx}\sum_{k=0}^{n-1} \binom{n-1}{k} f^{(k)}\, g^{(n-1-k)} \\ &= \sum_{k=0}^{n-1} \binom{n-1}{k}\left[f^{(k+1)}\, g^{(n-1-k)} + f^{(k)}\, g^{(n-k)}\right] \\ &= \sum_{k=0}^{n-1} \binom{n-1}{k} f^{(k+1)}\, g^{(n-1-k)} + \sum_{k=0}^{n-1} \binom{n-1}{k} f^{(k)}\, g^{(n-k)}.\end{aligned}$$

Since

$$\binom{n-1}{n} = 0, \tag{5.11}$$

we can replace $n-1$ with $n$ in the second sum

$$\frac{d^n}{dx^n}(f\, g) = \sum_{k=0}^{n-1} \binom{n-1}{k} f^{(k+1)}\, g^{(n-1-k)} + \sum_{k=0}^{n} \binom{n-1}{k} f^{(k)}\, g^{(n-k)}.$$

In the first sum, we set $j = k+1$ and so $k = j-1$, and in the second sum, we set $j = k$

$$\frac{d^n}{dx^n}(f\,g) = \sum_{j=1}^{n} \binom{n-1}{j-1} f^{(j)}\,g^{(n-j)} + \sum_{j=0}^{n} \binom{n-1}{j} f^{(j)}\,g^{(n-j)}.$$

Since $(-1)! = \infty$, the binomial coefficient

$$\binom{n-1}{-1} = 0,$$

and so we can start the first sum at $j = 0$ and use the identity (5.10) to write

$$\begin{aligned}\frac{d^n}{dx^n}(f\,g) &= \sum_{j=0}^{n} \binom{n-1}{j-1} f^{(j)}\,g^{(n-j)} + \sum_{j=0}^{n} \binom{n-1}{j} f^{(j)}\,g^{(n-j)} \\ &= \sum_{j=0}^{n} \left[\binom{n-1}{j-1} + \binom{n-1}{j}\right] f^{(j)}\,g^{(n-j)} = \sum_{j=0}^{n} \binom{n}{j} f^{(j)}\,g^{(n-j)}\end{aligned}$$

which is Leibniz's rule. □

# 6

# Examples for Chapter 6, Complex-Variable Theory

## 6.1 Analyticity

**Example 6.1** ($z = x + iy$) Is the function $f(x, y) = x + iy = z$ analytic? If we compute its derivative at $(x, y) = (0, 0)$ by setting $x = \epsilon$ and $y = 0$, then the limit is

$$\lim_{\epsilon \to 0} \frac{f(\epsilon, 0) - f(0, 0)}{\epsilon} = \lim_{\epsilon \to 0} \frac{\epsilon}{\epsilon} = 1, \tag{6.1}$$

while if we instead set $x = 0$ and $y = \epsilon$, then the limit is

$$\lim_{\epsilon \to 0} \frac{f(0, \epsilon) - f(0, 0)}{i\epsilon} = \lim_{\epsilon \to 0} \frac{i\epsilon}{i\epsilon} = 1. \tag{6.2}$$

So $f(x, y) = x - iy$ may be differentiable at $z = 0$. □

**Example 6.2** ($z^*$) Is the function $f(x, y) = x - iy = z^*$ analytic? If we compute its derivative at $(x, y) = (0, 0)$ by setting $x = \epsilon$ and $y = 0$, then the limit is

$$\lim_{\epsilon \to 0} \frac{f(\epsilon, 0) - f(0, 0)}{\epsilon} = \lim_{\epsilon \to 0} \frac{\epsilon}{\epsilon} = 1, \tag{6.3}$$

while if we instead set $x = 0$ and $y = \epsilon$, then the limit is

$$\lim_{\epsilon \to 0} \frac{f(0, \epsilon) - f(0, 0)}{i\epsilon} = \lim_{\epsilon \to 0} \frac{-i\epsilon}{i\epsilon} = -1. \tag{6.4}$$

So $f(x, y) = x - iy = z^*$ is not an analytic function of $z$. □

We can't apply the definition (6.1) of differentiability to every point $z$ for every function $f(z)$. We need a better test of analyticity.

## 6.2 Cauchy-Riemann Conditions

If $f(x, y) = u(x, y) + iv(x, y)$ with $u$ and $v$ real, is analytic then $df = f'(z)dz$ with $dz = dx + idy$ and

$$df = \left(\frac{\partial u}{\partial x} + i\frac{\partial v}{\partial x}\right) dx = f'(z)\, dx \;\text{ and }\; df = \left(\frac{\partial u}{\partial y} + i\frac{\partial v}{\partial y}\right) dy = f'(z)\, idy, \tag{6.5}$$

or

$$f'(z) = \left(\frac{\partial u}{\partial x} + i\frac{\partial v}{\partial x}\right) = \frac{1}{i}\left(\frac{\partial u}{\partial y} + i\frac{\partial v}{\partial y}\right). \tag{6.6}$$

These complex equations imply the two real equations

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y} \quad \text{and} \quad \frac{\partial v}{\partial x} = -\frac{\partial u}{\partial y} \tag{6.7}$$

or more succinctly

$$u_x = v_y \quad \text{and} \quad v_x = -u_y \tag{6.8}$$

which are the Cauchy-Riemann conditions.

**Example 6.3** ($f(x, y) = x^2 - y^2$) For the function $f(x, y) = u(x, y) + iv(x, y)$, the real and imaginary parts are $u = x^2 - y^2$ and $v = 0$, and so the Cauchy-Riemann conditions require that

$$2x = 0 \quad \text{and} \quad 0 = 2y. \tag{6.9}$$

So $f(x, y) = x^2 - y^2$ is not analytic. □

## 6.3 Calculus of residues

**Example 6.4** (Isolated pole) Let's consider the integral

$$I = \oint_{\mathcal{C}} a_n(w)(z - w)^n \, dz \tag{6.10}$$

along a closed, counterclockwise contour $\mathcal{C}$ around the point $w$. Setting $z = w + \epsilon e^{i\theta}$, we find

$$\begin{aligned} I &= a_n(w) \int_0^{2\pi} \left(\epsilon e^{i\theta}\right)^n i\epsilon\, e^{i\theta}\, d\theta = i a_n(w) \epsilon^{n+1} \int_0^{2\pi} e^{i(n+1)\theta}\, d\theta \\ &= 2\pi i\, a_{-1}(w)\, \delta_{n,-1}. \end{aligned} \tag{6.11}$$

This is why if $f(z)$ has a Laurent series (6.89-6.90), only the $n = -1$ term $a_{-1}(z_0)$ contributes

$$\oint_{\mathcal{C}} f(z)\,dz = \oint_{\mathcal{C}} \sum_{n=-\infty}^{\infty} a_n(z_0)\,(z-z_0)^n\,dz = 2\pi i\,a_{-1}(z_0). \tag{6.12}$$

□

## 6.4 Ghost Contours

**Example 6.5** (Yukawa potential) An intermediate formula in the derivation of the Yukawa potential (6.154) was

$$\begin{aligned} G_Y(\boldsymbol{x}) &= \int \frac{d^3k}{(2\pi)^3}\,\frac{e^{i\boldsymbol{k}\cdot\boldsymbol{x}}}{\boldsymbol{k}^2+m^2} \\ &= \frac{1}{ir}\int_{-\infty}^{\infty}\frac{dk}{(2\pi)^2}\,\frac{k}{(k-im)(k+im)}\,e^{ikr} \end{aligned} \tag{6.13}$$

in which $k = |\boldsymbol{k}|$ and $r = |\boldsymbol{x}|$. Since $r > 0$, we add a ghost contour that goes over the north pole of the upper half plane wherein $ikr = i(k_r + ik_i)r$ has a negative real part

$$G_Y(\boldsymbol{x}) = \frac{1}{ir}\oint_{\mathcal{C}}\frac{dk}{(2\pi)^2}\,\frac{k}{(k-im)(k+im)}\,e^{ikr}. \tag{6.14}$$

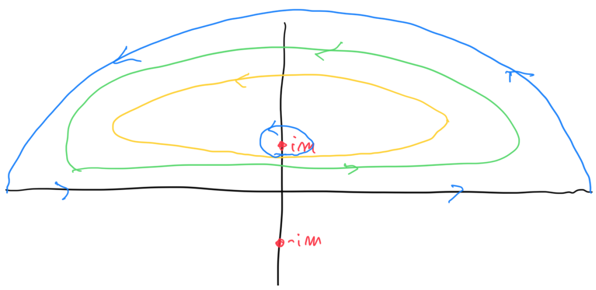


Figure 6.1 We shrink the ghost contour to a tiny loop about $k = im$.

The contour $\mathcal{C}$ encircles the point $k = im$ in a counter-clockwise sense, and so Cauchy's integral formula (6.40) gives

$$G_Y(\boldsymbol{x}) = \frac{1}{ir}\,2\pi i\left[\frac{1}{(2\pi)^2}\,\frac{k}{k+im}\,e^{ikr}\right]\Bigg|_{k=im} = \frac{1}{4\pi r}\,e^{-mr} \tag{6.15}$$

in units with $\hbar = c = 1$. Since $\hbar = c = 1$ in this formula (6.16), we can season it with factors of $\hbar$ and $c$ until the argument of the exponential becomes dimensionless standard units. We then get

$$G_Y(\boldsymbol{x}) = \frac{1}{4\pi r}\, e^{-cmr/\hbar}. \tag{6.16}$$

This potential is the Green's function for the differential operator $-\triangle + m^2$ because

$$\begin{aligned}(-\triangle + m^2)\, G_Y(\boldsymbol{x}) &= (-\triangle + m^2) \int \frac{d^3k}{(2\pi)^3} \frac{e^{i\boldsymbol{k}\cdot\boldsymbol{x}}}{\boldsymbol{k}^2 + m^2} \\ &= \int \frac{d^3k}{(2\pi)^3} \frac{(\boldsymbol{k}^2 + m^2) e^{i\boldsymbol{k}\cdot\boldsymbol{x}}}{\boldsymbol{k}^2 + m^2} = \int \frac{d^3k}{(2\pi)^3}\, e^{i\boldsymbol{k}\cdot\boldsymbol{x}} = \delta(\boldsymbol{x}).\end{aligned} \tag{6.17}$$

If we turn this last equation into a convolution

$$(-\triangle + m^2) \int d^3y\, G_Y(\boldsymbol{x} - \boldsymbol{y})\, j(\boldsymbol{y}) = \int d^3y\, \delta(\boldsymbol{x} - \boldsymbol{y}) j(\boldsymbol{y}) = j(\boldsymbol{x}), \tag{6.18}$$

we find that the solution to the equation

$$(-\triangle + m^2)\, f(\boldsymbol{x}) = j(\boldsymbol{x}) \tag{6.19}$$

is

$$f(\boldsymbol{x}) = \int d^3y\, G_Y(\boldsymbol{x} - \boldsymbol{y})\, j(\boldsymbol{y}). \tag{6.20}$$

□

**Example 6.6** (Ghost countour)

$$I = \int_{-\infty}^{\infty} \frac{dx}{1 + x^2} = \int_{-\infty}^{\infty} \frac{dx}{(x+i)(x-i)} \tag{6.21}$$

Add a ghost contour in UHP

$$I = \oint \frac{dx}{(x+i)(x-i)} = \frac{1}{2\pi i} \frac{1}{2i} \tag{6.22}$$

□

# 7
# Examples for Chapter 7, Differential Equations

A differential is exact if it is the change $dG$ in some function

$$dG = G_x dx + G_y dy. \tag{7.1}$$

If $F(x, y) = U(x, y) + iV(x, y)$, then

$$\begin{aligned} dF &= dU + idV \\ &= U_x dx + U_y dy + i(V_x dx + V_y dy). \end{aligned} \tag{7.2}$$

If, further, $dF$ is proportional to $dz = dx + idy$, then

$$\begin{aligned} dF &= (u + iv)(dx + idy) \\ &= (udx - vdy) + i(vdx + udy). \end{aligned} \tag{7.3}$$

So now we have

$$\begin{aligned} u = U_x \quad &\text{and} \quad v = -U_y \\ v = V_x \quad &\text{and} \quad u = V_y \end{aligned} \tag{7.4}$$

which imply that

$$\begin{gathered} U_x = V_y \quad \text{and} \quad V_x = -U_y \\ \text{as well as} \\ u_y = -\, v_x \quad \text{and} \quad v_y = u_x \end{gathered} \tag{7.5}$$

which are the Cauchy-Riemann conditions for both $F$ and its derivative

$$F = U + iV \quad \text{and} \quad F' = \frac{dF}{dz} = u + iv. \tag{7.6}$$

All integrals such as

$$'F = \int_{z_0}^{z} dz' F(z') \tag{7.7}$$

and derivatives of a function that is analytic in a simply connected region are also analytic there.

# 8

# Examples for Chapter 11, Group Theory

## 8.1 Little Group

A state of a particle of momentum $p$ can be defined in terms of a standard Lorentz transformation of a state of a standard fiducial momentum $k$. For a particle of mass $m > 0$, the standard fiducial momentum is

$$k = \begin{pmatrix} m \\ 0 \\ 0 \\ 0 \end{pmatrix} \tag{8.1}$$

in units with $\hbar = c = 1$. For a massless particle, the standard fiducial momentum is

$$k = \begin{pmatrix} k \\ 0 \\ 0 \\ k \end{pmatrix} \tag{8.2}$$

in which $k$ is an arbitrary momentum.

A state of a particle of fiducial momentum $k$ may carry another label $s$ related to the intrinsic spin of the particle

$$|k, s\rangle. \tag{8.3}$$

A state of momentum $p$ is then defined in terms of a standard Lorentz transformation $L(p)$ that takes $k$ to $p$

$$p = L(p)\, k \tag{8.4}$$

as

$$|p, s\rangle = n(p)\, U(L(p))\, |k, s\rangle \tag{8.5}$$

in which $n(p)$ is a factor of normalization, and $U(L(p))$ is a unitary operator that implements standard Lorentz transformation $L(p)$.

An arbitrary Lorentz transformation $\Lambda$ takes $p$ to $\Lambda\, p$ and $U(\Lambda)$ takes the state $|p, s\rangle$ to

$$U(\Lambda)\, |p, s\rangle = n(p)\, U(\Lambda)\, U(L(p))\, |k, s\rangle \tag{8.6}$$

The Wigner rotation $W(\lambda, p)$ is defined by the equation

$$\Lambda\, L(p) = L(\Lambda p) W(\lambda, p) \tag{8.7}$$

as

$$W(\Lambda, p) = \big(L(\Lambda p)\big)^{-1}\, \Lambda\, L(p). \tag{8.8}$$

We see that $W(\Lambda, p)$ takes a fiducial momentum $k$ to

$$W(\Lambda, p)\, k = \big(L(\Lambda p)\big)^{-1}\, \Lambda\, L(p)\, k = k. \tag{8.9}$$

The **little group** of the fiducial momentum $k$ is the group of Lorentz transformations that leave $k$ invariant. For particles of mass $m > 0$, the fiducial momentum $k$ (8.1 is unchanged by the group of rotations. So for massive particles, the little group is the group of rotations. For massless particles, the little group is the group of Lorentz transformations that leave the fiducial momentum $k$ (8.2 unchanged. Rotations about the $z$ axis leave the fiducial momentum $k$ (8.2 unchanged. The $4 \times 4$ matrices $J, X, Y$ of Section 1.4 leave $k$ invariant.