that is nonnegative when the matrices are the same
\[(A, A) = \text{Tr} A^\dagger A = \sum_{i=1}^N \sum_{j=1}^L A_{ij}^* A_{ij} = \sum_{i=1}^N \sum_{j=1}^L |A_{ij}|^2 \geq 0 \quad (1.87)\]
which is zero only when \( A = 0 \). So this inner product is positive definite.

A vector space with a positive-definite inner product (1.73–1.76) is called an inner-product space, a metric space, or a pre-Hilbert space.

A sequence of vectors \( f_n \) is a Cauchy sequence if for every \( \epsilon > 0 \) there is an integer \( N(\epsilon) \) such that \( \|f_n - f_m\| < \epsilon \) whenever both \( n \) and \( m \) exceed \( N(\epsilon) \). A sequence of vectors \( f_n \) converges to a vector \( f \) if for every \( \epsilon > 0 \) there is an integer \( N(\epsilon) \) such that \( \|f - f_n\| < \epsilon \) whenever \( n \) exceeds \( N(\epsilon) \). An inner-product space with a norm defined as in (1.80) is complete if each of its Cauchy sequences converges to a vector in that space. A Hilbert space is a complete inner-product space. Every finite-dimensional inner-product space is complete and so is a Hilbert space. But the term Hilbert space more often is used to describe infinite-dimensional complete inner-product spaces, such as the space of all square-integrable functions (David Hilbert, 1862–1943).

**Example 1.17** (The Hilbert Space of Square-Integrable Functions) For the vector space of functions (1.55), a natural inner product is
\[(f, g) = \int_a^b dx f^*(x)g(x). \quad (1.88)\]
The squared norm \( \| f \| \) of a function \( f(x) \) is
\[\| f \|^2 = \int_a^b dx |f(x)|^2. \quad (1.89)\]
A function is square integrable if its norm is finite. The space of all square-integrable functions is an inner-product space; it also is complete and so is a Hilbert space.

**Example 1.18** (Minkowski Inner Product) The Minkowski or Lorentz inner product \((p, x)\) of two 4-vectors \( p = (E/c, p_1, p_2, p_3) \) and \( x = (ct, x_1, x_2, x_3) \) is \( p \cdot x = Et \). It is indefinite, nondegenerate (1.79), and invariant under Lorentz transformations, and often is written as \( p \cdot x \) or as \( px \). If \( p \) is the 4-momentum of a freely moving physical particle of mass \( m \), then
\[p \cdot p = p \cdot p - E^2/c^2 = -c^2m^2 \leq 0. \quad (1.90)\]
The Minkowski inner product satisfies the rules (1.73, 1.74, and 1.79), but
and the formula (1.118) for the \( k \)th orthonormal linear combination of the vectors \( |V_k⟩ \) is

\[
|U_k⟩ = \frac{|u_k⟩}{\sqrt{⟨u_k|u_k⟩}}.
\]

The vectors \( |U_k⟩ \) are not unique; they vary with the order of the \( |V_k⟩ \).

Vectors and linear operators are abstract. The numbers we compute with are inner products like \( ⟨g|f⟩ \) and \( ⟨g|A|f⟩ \). In terms of \( N \) orthonormal basis vectors \( |n⟩ \) with \( f_n = ⟨n|f⟩ \) and \( g_n^* = ⟨g|n⟩ \), we can use the expansion (1.131) to write these inner products as

\[
⟨g|f⟩ = ⟨g|I|f⟩ = \sum_{n=1}^{N} ⟨g|n⟩⟨n|f⟩ = \sum_{n=1}^{N} g_n^* f_n
\]

\[
⟨g|A|f⟩ = ⟨g|IAI|f⟩ = \sum_{n,\ell=1}^{N} ⟨g|n⟩⟨n|A|\ell⟩⟨\ell|f⟩ = \sum_{n,\ell=1}^{N} g_n^* A_{n\ell} f_\ell
\]

in which \( A_{n\ell} = ⟨n|A|\ell⟩ \). We often gather the inner products \( f_\ell = ⟨\ell|f⟩ \) into a column vector \( f \) with components \( f_\ell = ⟨\ell|f⟩ \)

\[
f = \begin{pmatrix}
⟨1|f⟩ \\
⟨2|f⟩ \\
... \\
⟨N|f⟩
\end{pmatrix} = \begin{pmatrix}
f_1 \\
f_2 \\
... \\
f_N
\end{pmatrix}
\]

and the \( ⟨n|A|\ell⟩ \) into a matrix \( A \) with matrix elements \( A_{n\ell} = ⟨n|A|\ell⟩ \). If we also line up the inner products \( g_n^* = ⟨n|g⟩^* \) in a row vector that is the transpose of the complex conjugate of the column vector \( g \)

\[
g^\dagger = (⟨1|g⟩^*, ⟨2|g⟩^*, ..., ⟨N|g⟩^*) = (g_1^*, g_2^*, ..., g_N^*)
\]

then we can write inner products in matrix notation as \( ⟨g|f⟩ = g^\dagger f \) and as \( ⟨g|A|f⟩ = g^\dagger A f \).

If we switch to a different basis, say from \( |n⟩ \)’s to \( |α_n⟩ \)’s, then the components of the column vectors change from \( f_n = ⟨n|f⟩ \) to \( f_n' = ⟨α_n|f⟩ \), and similarly those of the row vectors \( g^\dagger \) and of the matrix \( A \) change, but the bras, the kets, the linear operators, and the inner products \( ⟨g|f⟩ \) and \( ⟨g|A|f⟩ \)
Remarkably, this translation operator is an exponential of the momentum operator $U(a) = \exp(-i p \cdot a / \hbar)$ in which $\hbar = h / 2\pi = 1.054 \times 10^{-34}$ Js is Planck’s constant divided by $2\pi$.

In two-dimensions, with basis states $|x, y\rangle$ that are orthonormal in Dirac’s sense, $\langle x, y| x', y' \rangle = \delta(x - x')\delta(y - y')$, the unitary operator

$$U(\theta) = \int |x \cos \theta - y \sin \theta, x \sin \theta + y \cos \theta\rangle \langle x, y| \, dx \, dy$$

(1.171)

rotates a system in space by the angle $\theta$. This rotation operator is the exponential $U(\theta) = \exp(-i \theta L_z / \hbar)$ in which the $z$ component of the angular momentum is $L_z = x p_y - y p_x$.

We may carry most of our intuition about matrices over to these unitary transformations that change from one infinite basis to another. But we must use common sense and keep in mind that infinite sums and integrals do not always converge.

### 1.18 Antunitary, Antilinear Operators

Certain maps on states $|\psi\rangle \rightarrow |\psi'\rangle$, such as those involving time reversal, are implemented by operators $K$ that are **antilinear**

$$K(z \psi + w \phi) = K(z |\psi\rangle + w |\phi\rangle) = z^* K |\psi\rangle + w^* K |\phi\rangle = z^* K \psi + w^* K \phi$$

(1.172)

and **antiunitary**

$$(K \phi, K \psi) = \langle K \phi | K \psi \rangle = (\phi, \psi)^* = (\langle \phi | \psi \rangle)^* = \langle \psi | \phi \rangle = (\psi, \phi).$$

(1.173)

In Dirac notation, these rules are $K(z |\psi\rangle) = z^* |\psi\rangle$ and $K(w |\phi\rangle) = w^* |\phi\rangle$.

### 1.19 Symmetry in Quantum Mechanics

In quantum mechanics, a symmetry is a one-to-one map of states $|\psi\rangle \leftrightarrow |\psi'\rangle$ and $|\phi\rangle \leftrightarrow |\phi'\rangle$ that preserves probabilities

$$|\langle \phi' | \psi' \rangle|^2 = |\langle \phi | \psi \rangle|^2.$$  

(1.174)

Eugene Wigner (1902–1995) showed that every symmetry in quantum mechanics can be represented either by an operator $U$ that is linear and unitary or by an operator $K$ that is antilinear and antiunitary. The antilinear, antiunitary case seems to occur only when the symmetry involves time reversal. Most symmetries are represented by operators that are linear and unitary. Unitary operators are of great importance in quantum mechanics.
shows that the operations (1.189) on columns that don’t change the value of the determinant can be written as matrix multiplication from the right by a matrix that has unity on its main diagonal and zeros below. Now consider the matrix product

$$\begin{pmatrix} A & 0 \\ -I & B \end{pmatrix} \begin{pmatrix} I & B \\ 0 & I \end{pmatrix} = \begin{pmatrix} A & AB \\ -I & 0 \end{pmatrix}$$

(1.202)

in which $A$ and $B$ are $N \times N$ matrices, $I$ is the $N \times N$ identity matrix, and $0$ is the $N \times N$ matrix of all zeros. The second matrix on the left-hand side has unity on its main diagonal and zeros below, and so it does not change the value of the determinant of the matrix to its left, which then must equal that of the matrix on the right-hand side:

$$\det \begin{pmatrix} A & 0 \\ -I & B \end{pmatrix} = \det \begin{pmatrix} A & AB \\ -I & 0 \end{pmatrix}.$$ 

(1.203)

By using Laplace’s expansion (1.183) along the first column to evaluate the determinant on the left-hand side and his expansion along the last row to compute the determinant on the right-hand side, one finds that the determinant of the product of two matrices is the product of the determinants

$$\det A \det B = \det AB.$$ 

(1.204)

**Example 1.27 (Two $2 \times 2$ Matrices)** When the matrices $A$ and $B$ are both $2 \times 2$, the two sides of (1.203) are

$$\det \begin{pmatrix} A & 0 \\ -I & B \end{pmatrix} = \det \begin{pmatrix} a_{11} & a_{12} & 0 & 0 \\ a_{21} & a_{22} & 0 & 0 \\ -1 & 0 & b_{11} & b_{12} \\ 0 & -1 & b_{21} & b_{22} \end{pmatrix}$$

(1.205)

$$= a_{11}a_{22}\det B - a_{21}a_{12}\det B = \det A \det B$$

and

$$\det \begin{pmatrix} A & AB \\ -I & 0 \end{pmatrix} = \det \begin{pmatrix} a_{11} & a_{12} & (ab)_{11} & (ab)_{12} \\ a_{21} & a_{22} & (ab)_{21} & (ab)_{22} \\ -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}$$

(1.206)

$$= (-1)^4 C_{42} = (-1)(-1) \det AB = \det AB$$

and so they give the product rule $\det A \det B = \det AB$. $\square$
remains true when the matrix $A$ replaces the variable $\lambda$

$$P(A, A) = \sum_{k=0}^{N} p_k A^k = 0. \quad (1.258)$$

To see why, we use the formula (1.196)

$$\delta_{k\ell} \det A = \sum_{i=1}^{N} A_{ik} C_{i\ell} \quad (1.259)$$

to write the determinant $|A - \lambda I| = P(\lambda, A)$ as the product of the matrix $A - \lambda I$ and the transpose of its matrix of cofactors

$$(A - \lambda I) C(\lambda, A)^T = |A - \lambda I| I = P(\lambda, A) I. \quad (1.260)$$

The transpose of the matrix of cofactors of the matrix $A - \lambda I$ is a polynomial in $\lambda$ with matrix coefficients

$$C(\lambda, A)^T = C_0 + C_1 \lambda + \cdots + C_{N-1} \lambda^{N-1}. \quad (1.261)$$

The left-hand side of equation (1.260) is then

$$(A - \lambda I)C(\lambda, A)^T = AC_0 + (AC_1 - C_0) \lambda + (AC_2 - C_1) \lambda^2 + \ldots + (AC_{N-1} - C_{N-2}) \lambda^{N-1} - C_{N-1} \lambda^{N}. \quad (1.262)$$

Equating equal powers of $\lambda$ on both sides of (1.260), we have using (1.257) and (1.262)

$$AC_0 = p_0 I$$
$$AC_1 - C_0 = p_1 I$$
$$AC_2 - C_1 = p_2 I$$
$$\ldots = \ldots$$
$$AC_{N-1} - C_{N-2} = p_{N-1} I$$
$$-C_{N-1} = p_N I. \quad (1.263)$$

We now multiply from the left the first of these equations by $I$, the second by $A$, the third by $A^2$, \ldots, and the last by $A^N$ and then add the resulting equations. All the terms on the left-hand sides cancel, while the sum of those on the right gives $P(A, A)$. Thus a square matrix $A$ obeys its characteristic equation $0 = P(A, A)$ or

$$0 = \sum_{k=0}^{N} p_k A^k = |A| I + p_1 A + \cdots + (-1)^{N-1} \langle \text{Tr} A \rangle A^{N-1} + (-1)^N A^N \quad (1.264)$$
1.27 Functions of Matrices


Because every \( N \times N \) matrix \( A \) obeys its characteristic equation, its \( N \)th power \( A^N \) can be expressed as a linear combination of its lesser powers

\[
A^N = (-1)^{N-1}(|A|I + p_1A + p_2A^2 + \cdots + (-1)^{N-1}(\text{Tr}A)A^{N-1}). \tag{1.265}
\]

For instance, the square \( A^2 \) of every \( 2 \times 2 \) matrix is given by

\[
A^2 = -|A|I + (\text{Tr}A)A. \tag{1.266}
\]

**Example 1.35** (Spin-one-half rotation matrix) If \( \theta \) is a real 3-vector and \( \sigma \) is the 3-vector of Pauli matrices (1.32), then the square of the traceless \( 2 \times 2 \) matrix \( A = \theta \cdot \sigma \) is

\[
(\theta \cdot \sigma)^2 = -|\theta \cdot \sigma|I - \begin{vmatrix}
\theta_3 & \theta_1 - i\theta_2 \\
\theta_1 + i\theta_2 & -\theta_3
\end{vmatrix}I = \theta^2 I \tag{1.267}
\]

in which \( \theta^2 = \theta \cdot \theta \). One may use this identity to show (exercise (1.28)) that

\[
\exp(-i\theta \cdot \sigma/2) = \cos(\theta/2)I - i\hat{\theta} \cdot \sigma \sin(\theta/2) \tag{1.268}
\]

in which \( \hat{\theta} \) is a unit 3-vector. For a spin-one-half object, this matrix represents a right-handed rotation of \( \theta \) radians about the axis \( \hat{\theta} \). \( \square \)

**1.27 Functions of Matrices**

What sense can we make of a function \( f \) of an \( N \times N \) matrix \( A \)? and how would we compute it? One way is to use the characteristic equation (1.265) to express every power of \( A \) in terms of \( I, A, \ldots, A^{N-1} \) and the coefficients \( p_0 = |A|, p_1, p_2, \ldots, p_{N-2} \), and \( p_{N-1} = (-1)^{N-1}\text{Tr}A \). Then if \( f(x) \) is a polynomial or a function with a convergent power series

\[
f(x) = \sum_{k=0}^{\infty} c_k x^k \tag{1.269}
\]

in principle we may express \( f(A) \) in terms of \( N \) functions \( f_k(p) \) of the coefficients \( p \equiv (p_0, \ldots, p_{N-1}) \) as

\[
f(A) = \sum_{k=0}^{N-1} f_k(p) A^k. \tag{1.270}
\]

The identity (1.268) for \( \exp(-i\theta \cdot \sigma/2) \) is an \( N = 2 \) example of this technique which can become challenging when \( N > 3 \).
Example 1.46  Suppose $A$ is the $3 \times 2$ matrix
\[
A = \begin{pmatrix} r_1 & p_1 \\ r_2 & p_2 \\ r_3 & p_3 \end{pmatrix}
\] (1.385)

and the vector $|y\rangle$ is the cross-product $|y\rangle = L = r \times p$. Then no solution $|x\rangle$ exists to the equation $A|x\rangle = |y\rangle$ (unless $r$ and $p$ are parallel) because $A|x\rangle$ is a linear combination of the vectors $r$ and $p$ while $|y\rangle = L$ is perpendicular to both $r$ and $p$.

Even when the matrix $A$ is square, the equation (1.381) sometimes has no solutions. For instance, if $A$ is a square matrix that vanishes, $A = 0$, then (1.381) has no solutions whenever $|y\rangle \neq 0$. And when $N > M$, as in for instance
\[
\begin{pmatrix} a & b & c \\ d & e & f \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}
\] (1.386)

the solution (1.384) is never unique, for we may add to it any linear combination of the vectors $|n\rangle$ that $A$ annihilates for $M < n \leq N$
\[
|x\rangle = \sum_{n=1}^{\min(M,N)} \frac{\langle m_n | y \rangle}{S_n} |n\rangle + \sum_{n=M+1}^{N} x_n |n\rangle.
\] (1.387)

These are the vectors $|n\rangle$ for $M < n \leq N$ that $A$ maps to zero since they do not occur in the sum (1.362) which stops at $n = \min(M,N) < N$.

Example 1.47 (The CKM Matrix)  In the standard model, the mass matrices of the $u, c, t$ and $d, s, b$ quarks are $3 \times 3$ complex matrices $M_u$ and $M_d$ with singular-value decompositions $M_u = U_u \Sigma_u V_u$ and $M_d = U_d \Sigma_d V_d$ whose singular-values are the quark masses. The unitary CKM matrix $U_u^T U_d$ (Cabibbo, Kobayashi, Maskawa) describes transitions among the quarks mediated by the $W^\pm$ gauge bosons. By redefining the quark fields, one may make the CKM matrix real, apart from a phase that violates charge-conjugation-parity ($CP$) symmetry.

The adjoint of a complex symmetric matrix $M$ is its complex conjugate, $M^\dagger = M^*$. So by (1.351), its right singular vectors $|n\rangle$ are the eigenstates of $M^* M$
\[
M^* M |n\rangle = S_n^2 |n\rangle
\] (1.388)
and by (1.366) its left singular vectors $|m_n\rangle$ are the eigenstates of $MM^*$

$$MM^*|m_n\rangle = (M^*M)^*|m_n\rangle = S^2_n|m_n\rangle. \tag{1.389}$$

Thus its left singular vectors are the complex conjugates of its right singular vectors, $|m_n\rangle = |n\rangle^*$. So the unitary matrix $V$ is the complex conjugate of the unitary matrix $U$, and the SVD of $M$ is (Autonne, 1915)

$$M = U\Sigma U^T. \tag{1.390}$$

### 1.32 The Moore-Penrose Pseudoinverse

Although a matrix $A$ has an inverse $A^{-1}$ if and only if it is square and has a nonzero determinant, one may use the singular-value decomposition to make a pseudoinverse $A^+$ for an arbitrary $M \times N$ matrix $A$. If the singular-value decomposition of the matrix $A$ is

$$A = U\Sigma V^\dagger \tag{1.391}$$

then the Moore-Penrose pseudoinverse (Eliakim H. Moore 1862–1932, Roger Penrose 1931–) is

$$A^+ = V\Sigma^+ U^\dagger \tag{1.392}$$

in which $\Sigma^+$ is the transpose of the matrix $\Sigma$ with every nonzero entry replaced by its inverse (and the zeros left as they are). One may show that the pseudoinverse $A^+$ satisfies the four relations

$$AA^+ A = A \quad \text{and} \quad A^+ A A^+ = A^+$$

$$(AA^+)^\dagger = AA^+ \quad \text{and} \quad (A^+ A)^\dagger = A^+ A. \tag{1.393}$$

and that it is the only matrix that does so.

Suppose that all the singular values of the $M \times N$ matrix $A$ are positive. In this case, if $A$ has more rows than columns, so that $M > N$, then the product $AA^+$ is the $N \times N$ identity matrix $I_N$

$$A^+ A = V^\dagger \Sigma^+ \Sigma V = V^\dagger I_N V = I_N \tag{1.394}$$

and $AA^+$ is an $M \times M$ matrix that is not the identity matrix $I_M$. If instead $A$ has more columns than rows, so that $N > M$, then $AA^+$ is the $M \times M$ identity matrix $I_M$

$$AA^+ = U\Sigma\Sigma^+ U^\dagger = U I_M U^\dagger = I_M \tag{1.395}$$

but $A^+ A$ is an $N \times N$ matrix that is not the identity matrix $I_N$. If the
by setting its derivatives with respect to \( \rho_n, \lambda_1, \) and \( \lambda_2 \) equal to zero
\[
\frac{\partial L}{\partial \rho_n} = -k (\ln \rho_n + 1) - \lambda_1 - \lambda_2 E_n = 0 \tag{1.339}
\]
\[
\frac{\partial L}{\partial \lambda_1} = \sum_n \rho_n - 1 = 0 \tag{1.340}
\]
\[
\frac{\partial L}{\partial \lambda_2} = \sum_n \rho_n E_n - E = 0. \tag{1.341}
\]
The first (1.339) of these conditions implies that
\[
\rho_n = \exp\left[-(\lambda_1 + \lambda_2 E_n + k)/k\right] \tag{1.342}
\]
We satisfy the second condition (1.340) by choosing \( \lambda_1 \) so that
\[
\rho_n = \frac{\exp(-\lambda_2 E_n/k)}{\sum_n \exp(-\lambda_2 E_n/k)}. \tag{1.343}
\]
Setting \( \lambda_2 = 1/T \), we define the temperature \( T \) so that \( \rho \) satisfies the third condition (1.341). Its eigenvalue \( \rho_n \) then is
\[
\rho_n = \frac{\exp(-E_n/kT)}{\sum_n \exp(-E_n/kT)}. \tag{1.344}
\]
In terms of the inverse temperature \( \beta \equiv 1/(kT) \), the density operator is
\[
\rho = \frac{e^{-\beta H}}{\text{Tr}(e^{-\beta H})} \tag{1.345}
\]
which is the Boltzmann distribution.

1.31 The Singular-Value Decomposition

Every complex \( M \times N \) rectangular matrix \( A \) is the product of an \( M \times M \) unitary matrix \( U \), an \( M \times N \) rectangular matrix \( \Sigma \) that is zero except on its main diagonal which consists of \( A \)'s nonnegative singular values \( S_k \), and an \( N \times N \) unitary matrix \( V^\dagger \)
\[
A = U \Sigma V^\dagger. \tag{1.346}
\]
This singular-value decomposition is a key theorem of matrix algebra.

Suppose \( A \) is a linear operator that maps vectors in an \( N \)-dimensional vector space \( V_N \) into vectors in an \( M \)-dimensional vector space \( V_M \). The spaces \( V_N \) and \( V_M \) will have infinitely many orthonormal bases \( \{ |n,a\rangle \in V_N \} \) and \( \{ |m,b\rangle \in V_M \} \) labeled by continuous parameters \( a \) and \( b \). Each pair of
The singular-value decomposition of $A$ then is

$$A = I_M A I_N = \sum_{n'=1}^M |m_{n'}\rangle \langle m_{n'}| \sum_{n=1}^N |n\rangle \langle n|. \quad (1.359)$$

There are $\min(M,N)$ singular values $S_n$ all nonnegative. For the positive singular values, equations (1.352 & 1.354) show that the matrix element $\langle m_{n'}|A|n\rangle$ vanishes unless $n' = n$

$$\langle m_{n'}|A|n\rangle = \frac{1}{S_{n'}} \langle An'|An\rangle = S_n \delta_{n'n}. \quad (1.360)$$

For the $Z$ vanishing singular values, equation (1.353) shows that $A|n\rangle = 0$ and so

$$\langle m_{n'}|A|n\rangle = 0. \quad (1.361)$$

Thus only the $\min(M,N) - Z$ singular values that are positive contribute to the singular-value decomposition (1.359). If $N > M$, then there can be at most $M$ nonzero eigenvalues $e_n$. If $N \leq M$, there can be at most $N$ nonzero $e_n$'s. The final form of the singular-value decomposition then is a sum of dyadics weighted by the positive singular values

$$A = \sum_{n=1}^{\min(M,N)} |m_n\rangle S_n \langle n| = \sum_{n=1}^{\min(M,N) - Z} |m_n\rangle S_n \langle n|. \quad (1.362)$$

The vectors $|m_n\rangle$ and $|n\rangle$ respectively are the left and right singular vectors. The nonnegative numbers $S_n$ are the singular values.

The linear operator $A$ maps the $\min(M,N)$ right singular vectors $|n\rangle$ into the $\min(M,N)$ left singular vectors $S_n|m_n\rangle$ scaled by their singular values

$$A|n\rangle = S_n|m_n\rangle \quad (1.363)$$

and its adjoint $A^\dagger$ maps the $\min(M,N)$ left singular vectors $|m_n\rangle$ into the $\min(M,N)$ right singular vectors $|n\rangle$ scaled by their singular values

$$A^\dagger|m_n\rangle = S_n|n\rangle. \quad (1.364)$$

The $N$-dimensional vector space $V_N$ is the **domain** of the linear operator $A$. If $N > M$, then $A$ annihilates $N - M + Z$ of the basis vectors $|n\rangle$. The **null space** or **kernel** of $A$ is the space spanned by the basis vectors $|n\rangle$ that
Linear Algebra

A annihilates. The vector space spanned by the left singular vectors $|m_n\rangle$ with nonzero singular values $S_n > 0$ is the range or image of $A$. It follows from the singular value decomposition (1.362) that the dimension $N$ of the domain is equal to the dimension of the kernel $N - M + Z$ plus that of the range $M - Z$, a result called the rank-nullity theorem.

Incidentally, the vectors $|m_n\rangle$ are the eigenstates of the hermitian matrix $AA^\dagger$ as one may see from the explicit product of the expansion (1.362) with its adjoint

$$AA^\dagger = \sum_{n=1}^{\min(M,N)} |m_n\rangle S_n \langle n| \sum_{n' = 1}^{\min(M,N)} |n'\rangle S_{n'} \langle m_{n'}|$$

$$= \sum_{n=1}^{\min(M,N)} \sum_{n' = 1}^{\min(M,N)} |m_n\rangle S_n \delta_{n n'} S_{n'} \langle m_{n'}|$$

$$= \sum_{n=1}^{\min(M,N)} |m_n\rangle S_n^2 \langle m_n|$$  \hspace{1cm} (1.365)

which shows that $|m_n\rangle$ is an eigenvector of $AA^\dagger$ with eigenvalue $e_n = S_n^2$

$$AA^\dagger|m_n\rangle = S_n^2|m_n\rangle.$$  \hspace{1cm} (1.366)

The SVD expansion (1.362) usually is written as a product of three explicit matrices, $A = U \Sigma V^\dagger$. The middle matrix $\Sigma$ is an $M \times N$ matrix with the $\min(M, N)$ singular values $S_n = \sqrt{e_n}$ on its main diagonal and zeros elsewhere. By convention, one writes the $S_n$ in decreasing order with the biggest $S_n$ as entry $\Sigma_{11}$. The first matrix $U$ and the third matrix $V^\dagger$ depend upon the bases one uses to represent the linear operator $A$. If these basis vectors are $|\alpha_k\rangle \& |\beta_\ell\rangle$, then

$$A_{k\ell} = \langle \alpha_k | A | \beta_\ell \rangle = \sum_{n=1}^{\min(M,N)} \langle \alpha_k | m_n \rangle S_n \langle n | \beta_\ell \rangle$$  \hspace{1cm} (1.367)

so that the $k, n$th entry in the matrix $U$ is $U_{kn} = \langle \alpha_k | m_n \rangle$. The columns of the matrix $U$ are the left singular vectors of the matrix $A$:

$$\begin{pmatrix}
U_{1n} \\
U_{2n} \\
\vdots \\
U_{Mn}
\end{pmatrix} = \begin{pmatrix}
\langle \alpha_1 | m_n \rangle \\
\langle \alpha_2 | m_n \rangle \\
\vdots \\
\langle \alpha_M | m_n \rangle
\end{pmatrix}.$$  \hspace{1cm} (1.368)

Similarly, the $n, \ell$th entry of the matrix $V^\dagger$ is $(V^\dagger)_{n,\ell} = \langle n | \beta_\ell \rangle$. Thus $V_{\ell,n} = $..
Exercises

The eight states of the system $|t,u,v\rangle \equiv (a_1^\dagger)^t(a_2^\dagger)^u(a_3^\dagger)^v|0,0,0\rangle$. We can represent them by eight 8-vectors each of which has seven 0’s with a 1 in position $4t + 2u + v + 1$. How big should the matrices that represent the creation and annihilation operators be? Write down the three matrices that represent the three creation operators.

1.38 Show that the Schwarz inner product (1.430) is degenerate because it can violate (1.79) for certain density operators and certain pairs of states.

1.39 Show that the Schwarz inner product (1.431) is degenerate because it can violate (1.79) for certain density operators and certain pairs of operators.

1.40 The coherent state $|\{\alpha_k\}\rangle$ is an eigenstate of the annihilation operator $a_k$ with eigenvalue $\alpha_k$ for each mode $k$ of the electromagnetic field, $a_k|\{\alpha_k\}\rangle = \alpha_k|\{\alpha_k\}\rangle$. The positive-frequency part $E^{(+)}_i(x)$ of the electric field is a linear combination of the annihilation operators

$$E^{(+)}_i(x) = \sum_k a_k E^{(+)}_i(k) e^{i(kx-\omega t)}. \quad (1.453)$$

Show that $|\{\alpha_k\}\rangle$ is an eigenstate of $E^{(+)}_i(x)$ as in (1.442) and find its eigenvalue $E_i(x)$. 
Moreover if $f^{(k+1)}$ is piecewise continuous, then

$$f_n = \int_{-\pi}^{\pi} \left\{ \frac{d}{dx} \left[ f^{(k)}(x) \frac{e^{-in x}}{(in)^{k+1}} \right] - f^{(k+1)}(x) \frac{e^{-in x}}{(in)^{k+1}} \right\} dx$$

$$= \int_{-\pi}^{\pi} f^{(k+1)}(x) \frac{e^{-in x}}{(in)^{k+1}} dx.$$  (2.55)

Since $f^{(k+1)}(x)$ is piecewise continuous on the closed interval $[-\pi, \pi]$, it is bounded there in absolute value by, let us say, $M$. So the Fourier coefficients of a $C^k$ periodic function with $f^{(k+1)}$ piecewise continuous are bounded by

$$|f_n| \leq \frac{1}{n^{k+1}} \int_{-\pi}^{\pi} |f^{(k+1)}(x)| \, dx \leq \frac{2\pi M}{n^{k+1}}.$$  (2.56)

We often can carry this derivation one step further. In most simple examples, the piecewise continuous periodic function $f^{(k+1)}(x)$ actually is piecewise continuously differentiable between its successive jumps at $x_j$. In this case, the derivative $f^{(k+2)}(x)$ is a piecewise continuous function plus a sum of a finite number of delta functions with finite coefficients. Thus we can integrate once more by parts. If for instance the function $f^{(k+1)}(x)$ jumps $J$ times between $-\pi$ and $\pi$ by $\Delta f^{(k+1)}_j$, then its Fourier coefficients are

$$f_n = \int_{-\pi}^{\pi} f^{(k+2)}(x) \frac{e^{-in x}}{(in)^{k+2}} dx$$

$$= \sum_{j=1}^{J} \int_{x_j}^{x_{j+1}} f^{(k+2)}(x) \frac{e^{-in x}}{(in)^{k+2}} dx + \sum_{j=1}^{J} \Delta f^{(k+1)}_j \frac{e^{-in x_j}}{(in)^{k+2}}$$  (2.57)

in which the subscript $s$ means that we’ve separated out the delta functions. The Fourier coefficients then are bounded by

$$|f_n| \leq \frac{2\pi M}{n^{k+2}}$$  (2.58)

in which $M$ is related to the maximum absolute values of $f^{(k+2)}_s(x)$ and of the $\Delta f^{(k+1)}_j$. The Fourier series of periodic $C^k$ functions converge very rapidly if $k$ is big.

Example 2.6 (Fourier Series of a $C^0$ Function) The function defined by

$$f(x) = \begin{cases} 
0 & -\pi \leq x < 0 \\
x & 0 \leq x < \pi/2 \\
\pi - x & \pi/2 \leq x \leq \pi 
\end{cases}$$  (2.59)

is continuous on the interval $[-\pi, \pi]$ and its first derivative is piecewise
with coefficients
\[ f_n = \frac{2}{L} \int_0^L \sin \frac{n\pi x}{L} f(x) \, dx \quad (2.145) \]
and the representation
\[ \sum_{m=-\infty}^{\infty} \delta(x - z - 2mL) = \frac{2}{L} \sum_{n=1}^{\infty} \sin \frac{n\pi x}{L} \sin \frac{n\pi z}{L} \quad (2.146) \]
for the Dirac comb on \( S_L \).

### 2.13 Periodic Boundary Conditions

Periodic boundary conditions are often convenient. For instance, rather than study an infinitely long one-dimensional system, we might study the same system, but of length \( L \). The ends cause effects not present in the infinite system. To avoid them, we imagine that the system forms a circle and impose the periodic boundary condition
\[ \psi(x \pm L, t) = \psi(x, t). \quad (2.147) \]

In three dimensions, the analogous conditions are
\[ \psi(x \pm L, y, z, t) = \psi(x, y, z, t) \]
\[ \psi(x, y \pm L, z, t) = \psi(x, y, z, t) \]
\[ \psi(x, y, z \pm L, t) = \psi(x, y, z, t). \quad (2.148) \]

The eigenstates \( |p\rangle \) of the free hamiltonian \( H = p^2/2m \) have wave functions
\[ \psi_p(x) = \langle x|p\rangle = e^{ixp/\hbar}/(2\pi\hbar)^{3/2}. \quad (2.149) \]
The periodic boundary conditions (2.148) require that each component \( p_i \) of momentum satisfy \( LP_i/\hbar = 2\pi n_i \) or
\[ p = \frac{2\pi \hbar n}{L} = \frac{\hbar n}{L} \quad (2.150) \]
where \( n \) is a vector of integers, which may be positive or negative or zero.

Periodic boundary conditions arise naturally in the study of solids. The atoms of a perfect crystal are at the vertices of a Bravais lattice
\[ x_i = x_0 + \sum_{i=1}^{3} n_i a_i \quad (2.151) \]
in which the three vectors \( a_i \) are the primitive vectors of the lattice and
the \( n_i \) are three integers. The Hamiltonian of such an infinite crystal is invariant under translations in space by

\[
\sum_{i=1}^{3} n_i a_i. \tag{2.152}
\]

To keep the notation simple, let’s restrict ourselves to a cubic lattice with lattice spacing \( a \). Then since the momentum operator \( p \) generates translations in space, the invariance of \( H \) under translations by \( an \)

\[
\exp(ian \cdot p) H \exp(-ian \cdot p) = H \tag{2.153}
\]

implies that \( \exp(ian \cdot p) \) and \( H \) are compatible observables \( [\exp(ian \cdot p), H] = 0 \). As explained in section 1.30, it follows that we may choose the eigenstates of \( H \) also to be eigenstates of \( p \)

\[
e^{iap \cdot n/\hbar} |\psi\rangle = e^{iak \cdot n} |\psi\rangle \tag{2.154}
\]

which implies that

\[
\psi(x + an, t) = e^{iak \cdot n} \psi(x, t). \tag{2.155}
\]

Setting

\[
\psi(x) = e^{ik \cdot x} u(x) \tag{2.156}
\]

we see that condition (2.155) implies that \( u(x) \) is periodic

\[
u(x + an) = u(x). \tag{2.157}
\]

For a general Bravais lattice, this \textbf{Born–von Karman} periodic boundary condition is

\[
u \left(x + \sum_{i=1}^{3} n_i a_i, t \right) = u(x, t). \tag{2.158}
\]

Equations (2.155) and (2.157) are known as \textbf{Bloch’s theorem}.

\textbf{Exercises}

2.1 Show that \( \sin \omega_1 x + \sin \omega_2 x \) is the same as (2.9).
2.2 Find the Fourier series for the function \( \exp(ax) \) on the interval \( -\pi < x \leq \pi \).
2.3 Find the Fourier series for the function \( (x^2 - \pi^2)^2 \) on the same interval \( (-\pi, \pi) \).
2.4 Find the Fourier series for the function \( (1 + \cos x) \sin ax \) on the interval \( (-\pi, \pi) \).
One often needs to relate a function’s Fourier series to its Fourier transform. So let’s compare the Fourier series (3.1) for the function \( f(x) \) on the interval \([-L/2, L/2]\) with its Fourier transform (3.9) in the limit of large \( L \)

\[
f(x) = \sum_{n=-\infty}^{\infty} f_n \frac{e^{i2\pi nx/L}}{\sqrt{L}} = \sum_{n=-\infty}^{\infty} f_n \frac{e^{i k_n x}}{\sqrt{L}} = \int_{-\infty}^{\infty} \tilde{f}(k) e^{ikx} \frac{dk}{\sqrt{2\pi}} \quad (3.12)
\]

in which \( k_n = 2\pi n/L = 2\pi y/L \). Now \( f_n = \hat{f}(y) \), and so by the definition (3.6) of \( \tilde{f}(k) \), we have \( f_n = \hat{f}(Lk/2\pi) = \sqrt{2\pi/L} \tilde{f}(k) \). Thus, to get the Fourier series from the Fourier transform, we multiply the series by \( 2\pi/L \) and use the Fourier transform at \( k_n \) divided by \( \sqrt{2\pi} \)

\[
f(x) = \frac{1}{\sqrt{L}} \sum_{n=-\infty}^{\infty} f_n e^{ik_n x} = \frac{2\pi}{L} \sum_{n=-\infty}^{\infty} \frac{\tilde{f}(k_n)}{\sqrt{2\pi}} e^{ik_n x}. \quad (3.13)
\]

Going the other way, we have

\[
f(x) = \int_{-\infty}^{\infty} \tilde{f}(k) e^{ikx} \frac{dk}{\sqrt{2\pi}} = \frac{L}{2\pi} \int_{-\infty}^{\infty} \tilde{f}(Lk/2\pi) e^{ikx} dk. \quad (3.14)
\]

**Example 3.1 (The Fourier Transform of a Gaussian Is a Gaussian)** The Fourier transform of the gaussian \( f(x) = \exp(-m^2 x^2) \) is

\[
\tilde{f}(k) = \int_{-\infty}^{\infty} dx \sqrt{\frac{2\pi}{m}} e^{-ikx} e^{-m^2 x^2}.
\]

We complete the square in the exponent:

\[
\tilde{f}(k) = e^{-k^2/4m^2} \int_{-\infty}^{\infty} dx \sqrt{\frac{2\pi}{m}} e^{-m^2(x+ik/2m)^2}.
\]

As we shall see in Sec. 5.14 when we study analytic functions, we may shift \( x \) to \( x - ik/2m^2 \), so the term \( ik/2m^2 \) in the exponential has no effect on the value of the \( x \)-integral.

\[
\tilde{f}(k) = e^{-k^2/4m^2} \int_{-\infty}^{\infty} dx \sqrt{\frac{2\pi}{m}} e^{-m^2 x^2} = \frac{1}{\sqrt{2m}} e^{-k^2/4m^2}.
\]

Thus, the Fourier transform of a gaussian is another gaussian

\[
\tilde{f}(k) = \int_{-\infty}^{\infty} dx \sqrt{\frac{2\pi}{m}} e^{-ikx} e^{-m^2 x^2} = \frac{1}{\sqrt{2m}} e^{-k^2/4m}.
\]

But the two gaussians are very different: if the gaussian \( f(x) = \exp(-m^2 x^2) \) decreases slowly as \( x \to \infty \) because \( m \) is small (or quickly because \( m \) is big), then its gaussian Fourier transform \( \tilde{f}(k) = \exp(-k^2/4m^2)/m\sqrt{2} \) decreases quickly as \( k \to \infty \) because \( m \) is small (or slowly because \( m \) is big).
3.6 Convolutions

If we generalize the relations (3.12–3.14) between Fourier series and transforms from one to \(n\) dimensions, then we find that the Fourier series corresponding to the Fourier transform (3.94) is

\[
f(x_1, \ldots, x_n) = \left(\frac{2\pi}{L}\right)^n \sum_{j_1=-\infty}^{\infty} \cdots \sum_{j_n=-\infty}^{\infty} e^{i(k_{j_1}x_1 + \cdots + k_{j_n}x_n)} \frac{\hat{f}(k_{j_1}, \ldots, k_{j_n})}{(2\pi)^{n/2}}
\]

(3.95)

in which \(k_{j_\ell} = \frac{2\pi j_{\ell}}{L}\). Thus, for \(n = 3\) we have

\[
f(x) = \frac{(2\pi)^3}{V} \sum_{j_1=-\infty}^{\infty} \sum_{j_2=-\infty}^{\infty} \sum_{j_3=-\infty}^{\infty} e^{i k_{j_1} \cdot x} \frac{\hat{f}(k_j)}{(2\pi)^{3/2}}
\]

(3.96)

in which \(k_j = (k_{j_1}, k_{j_2}, k_{j_3})\) and \(V = L^3\) is the volume of the box.

**Example 3.8** (The Feynman Propagator) For a spinless quantum field of mass \(m\), Feynman’s propagator is the four-dimensional Fourier transform

\[
\Delta_F(x) = \int \frac{\exp(ik \cdot x)}{k^2 + m^2 - i\epsilon} \frac{d^4k}{(2\pi)^4}
\]

(3.97)

where \(k \cdot x = k \cdot x - k^0 x^0\), all physical quantities are in natural units \((c = h = 1)\), and \(x^0 = ct = t\). The tiny imaginary term \(-i\epsilon\) makes \(\Delta_F(x - y)\) proportional to the mean value in the vacuum state \(|0\rangle\) of the time-ordered product of the fields \(\phi(x)\) and \(\phi(y)\) (section 5.34)

\[
-i \Delta_F (x - y) = \langle 0| T[\phi(x)\phi(y)] |0\rangle \equiv \theta(x^0 - y^0) \langle 0|\phi(x)\phi(y)|0\rangle + \theta(y^0 - x^0) \langle 0|\phi(y)\phi(x)|0\rangle
\]

(3.98)

in which \(\theta(a) = (a + |a|)/2|a|\) is the Heaviside function (2.166).

3.6 Convolutions

The convolution of \(f(x)\) with \(g(x)\) is the integral

\[
f \ast g(x) = \int_{-\infty}^{\infty} \frac{dy}{\sqrt{2\pi}} f(x - y) g(y).
\]

(3.99)

The convolution product is symmetric

\[
f \ast g(x) = g \ast f(x)
\]

(3.100)
because setting \( z = x - y \), we have

\[
 f * g(x) = \int_{-\infty}^{\infty} \frac{dy}{\sqrt{2\pi}} f(x - y) g(y) = - \int_{-\infty}^{\infty} \frac{dz}{\sqrt{2\pi}} f(z) g(x - z)
 = \int_{-\infty}^{\infty} \frac{dz}{\sqrt{2\pi}} g(x - z) f(z) = g * f(x).
\]

(3.101)

Convolutions may look strange at first, but they often occur in physics in the three-dimensional form

\[
 F(x) = \int G(x - x') S(x') \, d^3x
\]

(3.102)
in which \( G \) is a Green’s function and \( S \) is a source.

**Example 3.9** (Gauss’s Law) The divergence of the electric field \( E \) is the microscopic charge density \( \rho \) divided by the electric permittivity of the vacuum \( \epsilon_0 = 8.854 \times 10^{-12} \) F/m, that is, \( \nabla \cdot E = \rho/\epsilon_0 \). This constraint is known as Gauss’s law. If the charges and fields are independent of time, then the electric field \( E \) is the gradient of a scalar potential \( E = -\nabla \phi \).

These last two equations imply that \( \phi \) obeys Poisson’s equation

\[
 -\nabla^2 \phi = \frac{\rho}{\epsilon_0}.
\]

(3.103)

We may solve this equation by using Fourier transforms as described in Sec. 3.13. If \( \tilde{\phi}(k) \) and \( \tilde{\rho}(k) \) respectively are the Fourier transforms of \( \phi(x) \) and \( \rho(x) \), then Poisson’s differential equation (3.103) gives

\[
 -\nabla^2 \phi(x) = -\nabla^2 \int e^{ik \cdot x} \tilde{\phi}(k) \, d^3k = \int k^2 e^{ik \cdot x} \tilde{\phi}(k) \, d^3k = \frac{\rho(x)}{\epsilon_0} = \int e^{ik \cdot x} \frac{\tilde{\rho}(k)}{\epsilon_0} \, d^3k
\]

(3.104)

which implies the algebraic equation \( \tilde{\phi}(k) = \frac{\tilde{\rho}(k)}{\epsilon_0 k^2} \) which is an instance of (3.163). Performing the inverse Fourier transformation, we find for the scalar potential

\[
 \phi(x) = \int e^{ik \cdot x} \tilde{\phi}(k) \, d^3k = \int e^{ik \cdot x} \frac{\tilde{\rho}(k)}{\epsilon_0 k^2} \, d^3k
 = \int e^{ik \cdot x} \frac{1}{k^2} \int e^{-ik \cdot x'} \frac{\rho(x')}{\epsilon_0} \, d^3x' \, d^3k = \int G(x - x') \frac{\rho(x')}{\epsilon_0} \, d^3x'
\]

in which

\[
 G(x - x') = \int \frac{d^3k}{(2\pi)^3} \frac{1}{k^2} e^{ik \cdot (x - x')}. \quad (3.106)
\]
3.6 Convolutions

This function $G(x - x')$ is the Green’s function for the differential operator $-\nabla^2$ in the sense that

$$-\nabla^2 G(x - x') = \int \frac{d^3 k}{(2\pi)^3} e^{ik \cdot (x - x')} = \delta^{(3)}(x - x'). \quad (3.107)$$

This Green’s function ensures that expression (3.105) for $\phi(x)$ satisfies Poisson’s equation (3.103). To integrate (3.106) and compute $G(x - x')$, we use spherical coordinates with the $z$-axis parallel to the vector $x - x'$.

$$G(x - x') = \int \frac{dk}{(2\pi)^2} \frac{1}{k^2} e^{ik |x - x'|} = \int_0^\infty \frac{dk}{(2\pi)^2} \int_{-1}^1 d\cos \theta \, e^{ik |x - x'| \cos \theta} = \int_0^\infty \frac{dk}{2\pi^2 |x - x'|} \int_0^{\infty} \frac{\sin k |x - x'|}{k} \cos \theta \, dk.$$ \quad (3.108)

In example 5.35 of section 5.34 on Cauchy’s principal value, we’ll show that

$$\int_0^\infty \frac{\sin k}{k} \, dk = \frac{\pi}{2}. \quad (3.109)$$

Using this result, we have

$$\int \frac{d^3 k}{(2\pi)^3} \frac{1}{k^2} e^{ik \cdot (x - x')} = G(x - x') = \frac{1}{4\pi |x - x'|}. \quad (3.110)$$

Finally by substituting this formula for $G(x - x')$ into Eq. (3.105), we find that the Fourier transform $\phi(x)$ of the product $\tilde{\rho}(k)/k^2$ of the functions $\tilde{\rho}(k)$ and $1/k^2$ is the convolution

$$\phi(x) = \frac{1}{4\pi \epsilon_0} \int \frac{\rho(x')}{|x - x'|} \, d^3 x' \quad (3.111)$$

of their Fourier transforms $1/|x - x'|$ and $\rho(x')$. The Fourier transform of the product of any two functions is the convolution of their Fourier transforms, as we’ll see in the next section. (George Green 1793–1841)

Example 3.10 (The Magnetic Vector Potential) The magnetic induction $B$ has zero divergence (as long as there are no magnetic monopoles) and so may be written as the curl $B = \nabla \times A$ of a vector potential $A$. For time-independent currents, Ampère’s law is $\nabla \times B = \mu_0 J$ in which $\mu_0 = 1/(\epsilon_0 c^2) = 4\pi \times 10^{-7}$ N A$^{-2}$ is the permeability of the vacuum. It follows
that in the Coulomb gauge $\nabla \cdot A = 0$, the magnetostatic vector potential $A$ satisfies the equation
\[ \nabla \times B = \nabla \times (\nabla \times A) = \nabla (\nabla \cdot A) - \nabla^2 A = -\nabla^2 A = \mu_0 J. \] (3.112)

Applying the Fourier-transform technique (3.103–3.111), we find that the Fourier transforms of $A$ and $J$ satisfy the algebraic equation
\[ \tilde{A}(k) = \mu_0 \frac{\tilde{J}(k)}{k^2} \] (3.113)
which is an instance of (3.163). Performing the inverse Fourier transform, we see that $A$ is the convolution
\[ A(x) = \frac{\mu_0}{4\pi} \int d^3x' \frac{J(x')}{|x-x'|}. \] (3.114)

If in the solution (3.111) of Poisson’s equation, $\rho(x)$ is translated by $a$, then so is $\phi(x)$. That is, if $\rho'(x) = \rho(x+a)$ then $\phi'(x) = \phi(x+a)$. Similarly, if the current $J(x)$ in (3.114) is translated by $a$, then so is the potential $A(x)$. **Convolutions respect translational invariance.** That’s one reason why they occur so often in the formulas of physics.

### 3.7 The Fourier Transform of a Convolution

The Fourier transform of the convolution $f \ast g$ is the product of the Fourier transforms $\tilde{f}$ and $\tilde{g}$:
\[ \tilde{f \ast g}(k) = \tilde{f}(k) \tilde{g}(k). \] (3.115)

To see why, we form the Fourier transform $\tilde{f \ast g}(k)$ of the convolution $f \ast g(x)$
\[ \tilde{f \ast g}(k) = \int_{-\infty}^{\infty} \frac{dx}{\sqrt{2\pi}} e^{-ikx} f \ast g(x) \]
\[ = \int_{-\infty}^{\infty} \frac{dx}{\sqrt{2\pi}} e^{-ikx} \int_{-\infty}^{\infty} \frac{dy}{\sqrt{2\pi}} f(x-y) g(y). \] (3.116)

Now we write $f(x-y)$ and $g(y)$ in terms of their Fourier transforms $\tilde{f}(p)$ and $\tilde{g}(q)$
\[ \tilde{f \ast g}(k) = \int_{-\infty}^{\infty} \frac{dx}{\sqrt{2\pi}} e^{-ikx} \int_{-\infty}^{\infty} \frac{dy}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \frac{dp}{\sqrt{2\pi}} \tilde{f}(p) e^{ip(x-y)} \int_{-\infty}^{\infty} \frac{dq}{\sqrt{2\pi}} \tilde{g}(q) e^{iqy} \] (3.117)
and use the representation (3.36) of Dirac’s delta function twice to get
\begin{align*}
\tilde{f} \ast g(k) &= \int_{-\infty}^{\infty} \frac{dy}{2\pi} \int_{-\infty}^{\infty} dp \int_{-\infty}^{\infty} dq \delta(p - k) \tilde{f}(p) \tilde{g}(q) e^{i(q-p)y} \\
&= \int_{-\infty}^{\infty} dp \int_{-\infty}^{\infty} dq \delta(p - k) \delta(q - p) \tilde{f}(p) \tilde{g}(q) \\
&= \int_{-\infty}^{\infty} dp \delta(p - k) \tilde{f}(p) \tilde{g}(p) = \tilde{f}(k) \tilde{g}(k) \quad (3.118)
\end{align*}
which is (3.115). Examples 3.9 and 3.10 were illustrations of this result.

3.8 Fourier Transforms and Green’s Functions

A Green’s function \( G(x) \) for a differential operator \( P \) turns into a delta function when acted upon by \( P \), that is, \( PG(x) = \delta(x) \). If the differential operator is a polynomial \( P(\partial) \equiv P(\partial_1, \ldots, \partial_n) \) in the derivatives \( \partial_1, \ldots, \partial_n \) with constant coefficients, then a suitable Green’s function \( G(x) \equiv G(x_1, \ldots, x_n) \) will satisfy
\begin{equation}
P(\partial)G(x) = \delta^{(n)}(x). \quad (3.119)
\end{equation}
Expressing both \( G(x) \) and \( \delta^{(n)}(x) \) as Fourier transforms, we get
\begin{equation}
P(\partial)G(x) = \int d^n k P(ik) e^{ik \cdot x} \tilde{G}(k) = \delta^{(n)}(x) = \int \frac{d^n k}{(2\pi)^n} e^{ik \cdot x} \quad (3.120)
\end{equation}
which gives us the algebraic equation
\begin{equation}
\tilde{G}(k) = \frac{1}{(2\pi)^n P(ik)}. \quad (3.121)
\end{equation}
Thus the Green’s function \( G_P \) for the differential operator \( P(\partial) \) is
\begin{equation}
G_P(x) = \int \frac{d^n k}{(2\pi)^n} \frac{e^{ik \cdot x}}{P(ik)}. \quad (3.122)
\end{equation}

**Example 3.11 (Green and Yukawa)** In 1935, Hideki Yukawa (1907–1981) proposed the partial differential equation
\begin{equation}
P_Y(\partial)G_Y(x) \equiv (-\Delta + m^2)G_Y(x) = (-\nabla^2 + m^2)G_Y(x) = \delta(x). \quad (3.123)
\end{equation}
Our (3.122) gives as the Green’s function for \( P_Y(\partial) \) the Yukawa potential
\begin{equation}
G_Y(x) = \int \frac{d^3 k}{(2\pi)^3} \frac{e^{ik \cdot x}}{P_Y(ik)} = \int \frac{d^3 k}{(2\pi)^3} \frac{e^{ik \cdot x}}{k^2 + m^2} = \frac{e^{-mr}}{4\pi r} \quad (3.124)
\end{equation}
an integration done in example 5.21.
One may extend the definition (4.36) of \( n \)-factorial from positive integers to complex numbers by means of the integral formula
\[
\Gamma(z) = \int_0^\infty e^{-t} t^z \, dt
\]
for \( \Re z > -1 \). In particular
\[
0! = \int_0^\infty e^{-t} \, dt = 1
\]
which explains the definition (4.37). The factorial function \((z - 1)!\) in turn defines the \textit{gamma function} for \( \Re z > 0 \) as
\[
\Gamma(z) = \int_0^\infty e^{-t} t^{z-1} \, dt = (z - 1)!
\]
as may be seen from (4.53). By differentiating this formula and integrating it by parts, we see that the gamma function satisfies the key identity
\[
\Gamma(z + 1) = \int_0^\infty \left( -\frac{d}{dt} e^{-t} \right) t^z \, dt = \int_0^\infty e^{-t} \left( \frac{d}{dt} t^z \right) \, dt = \int_0^\infty e^{-t} z t^{z-1} \, dt = z \Gamma(z).
\]
Since \( \Gamma(1) = 0! = 1 \), we may use this identity (4.56) to extend the definition (5.102) of the gamma function in unit steps into the left half-plane
\[
\Gamma(z) = \frac{1}{z} \Gamma(z + 1) = \frac{1}{z} \frac{1}{z+1} \Gamma(z+2) = \frac{1}{z} \frac{1}{z+1} \frac{1}{z+2} \Gamma(z+3) = \ldots
\]
as long as we avoid the negative integers and zero. This extension leads to Euler’s definition
\[
\Gamma(z) = \lim_{n \to \infty} \frac{1 \cdot 2 \cdot 3 \cdots n}{z(z+1)(z+2)\cdots(z+n)} n^z
\]
and to Weierstrass’s (exercise 4.6)
\[
\Gamma(z) = \frac{1}{z} e^{-\gamma z} \left[ \prod_{n=1}^\infty \left( 1 + \frac{z}{n} \right) e^{-z/n} \right]^{-1}
\]
(Karl Theodor Wilhelm Weierstrass, 1815–1897), and is an example of analytic continuation (section 5.12).

One may show (exercise 4.8) that another formula for \( \Gamma(z) \) is
\[
\Gamma(z) = 2 \int_0^\infty e^{-t^2} t^{2z-1} \, dt
\]
for $\text{Re } z > 0$ and that
\[
\Gamma(n + \frac{1}{2}) = \frac{(2n)!}{n!2^{2n}} \sqrt{\pi} \quad (4.61)
\]
which implies (exercise 4.11) that
\[
\Gamma(n + \frac{1}{2}) = \frac{(2n - 1)!!}{2^n} \sqrt{\pi}. \quad (4.62)
\]

**Example 4.7** (Bessel Function of nonintegral index) We can use the gamma-function formula (4.55) for $n!$ to extend the definition (4.49) of the Bessel function of the first kind $J_n(\rho)$ to nonintegral values $\nu$ of the index $n$. Replacing $n$ by $\nu$ and $(m + n)!$ by $\Gamma(m + \nu + 1)$, we get
\[
J_{\nu}(\rho) = \left(\frac{\rho}{2}\right)^{\nu} \sum_{m=0}^{\infty} \frac{(-1)^m}{m!\Gamma(m + \nu + 1)} \left(\frac{\rho}{2}\right)^{2m} \quad (4.63)
\]
which makes sense even for complex values of $\nu$.

**Example 4.8** (Spherical Bessel Function) The spherical Bessel function is defined as
\[
j_\ell(\rho) \equiv \sqrt{\frac{\pi}{2\rho}} J_{\ell+1/2}(\rho). \quad (4.64)
\]
For small values of its argument $|\rho| \ll 1$, the first term in the series (4.63) dominates and so (exercise 4.7)
\[
j_\ell(\rho) \approx \frac{\sqrt{\pi}}{2} \left(\frac{\rho}{2}\right)^{\ell} \frac{1}{\Gamma(\ell + 3/2)} = \frac{\ell! (2\rho)^{\ell}}{(2\ell + 1)!} = \frac{\rho^{\ell}}{(2\ell + 1)!!} \quad (4.65)
\]
as one may show by repeatedly using the key identity $\Gamma(z + 1) = z \Gamma(z)$.

### 4.6 Taylor Series

If the function $f(x)$ is a real-valued function of a real variable $x$ with a continuous $N$th derivative, then Taylor’s expansion for it is
\[
f(x + a) = f(x) + af'(x) + \frac{a^2}{2}f''(x) + \cdots + \frac{a^{N-1}}{(N-1)!}f^{(N-1)}(x) + E_N
\]
\[
= \sum_{n=0}^{N-1} \frac{a^n}{n!} f^{(n)}(x) + E_N \quad (4.66)
\]
Example 4.12 (Planck’s Distribution) Max Planck (1858–1947) showed that the electromagnetic energy in a closed cavity of volume $V$ at a temperature $T$ in the frequency interval $d\nu$ about $\nu$ is

$$dU(\beta, \nu, V) = \frac{8\pi hV}{c^3} \frac{\nu^3}{e^{\beta \hbar \nu} - 1} \, d\nu$$

(4.94)

in which $\beta = 1/(kT)$, $k = 1.3806503 \times 10^{-23} \text{ J/K}$ is Boltzmann’s constant, and $h = 6.626068 \times 10^{-34} \text{ Js}$ is Planck’s constant. The total energy then is the integral

$$U(\beta, V) = \frac{8\pi hV}{c^3} \int_0^\infty \frac{\nu^3}{e^{\beta \hbar \nu} - 1} \, d\nu$$

(4.95)

which we may do by letting $x = \beta \hbar \nu$ and using the geometric series (4.31)

$$U(\beta, V) = \frac{8\pi (kT)^4 V}{(hc)^3} \int_0^\infty \frac{x^3}{e^x - 1} \, dx$$

$$= \frac{8\pi (kT)^4 V}{(hc)^3} \int_0^\infty \frac{x^3 e^{-x}}{1 - e^{-x}} \, dx$$

$$= \frac{8\pi (kT)^4 V}{(hc)^3} \int_0^\infty x^3 e^{-x} \sum_{n=0}^\infty e^{-nx} \, dx.$$  

(4.96)

The geometric series is absolutely and uniformly convergent for $x > 0$, and we may interchange the limits of summation and integration. After another change of variables, the Gamma-function formula (5.102) gives

$$U(\beta, V) = \frac{8\pi (kT)^4 V}{(hc)^3} \sum_{n=0}^\infty \int_0^\infty x^3 e^{-(n+1)x} \, dx$$

$$= \frac{8\pi (kT)^4 V}{(hc)^3} \sum_{n=0}^\infty \frac{1}{(n+1)^4} \int_0^\infty y^3 e^{-y} \, dy$$

$$= \frac{8\pi (kT)^4 V}{(hc)^3} 3! \zeta(4) = \frac{8\pi^5 (kT)^4 V}{15(hc)^3}.$$  

(4.97)

It follows that the power radiated by a “black body” is proportional to the fourth power of its temperature and to its area $A$

$$P = \sigma A T^4$$

(4.98)

in which

$$\sigma = \frac{2\pi^5 k^4}{15 h^3 c^2} = 5.670400(40) \times 10^{-8} \text{ W m}^{-2} \text{ K}^{-4}$$

(4.99)

is Stefan’s constant.
The number of photons in the black-body distribution (4.94) at inverse temperature \( \beta \) in the volume \( V \) is

\[
N(\beta, V) = \frac{8\pi V}{c^3} \int_0^\infty \frac{\nu^2}{e^{\beta \hbar \nu} - 1} d\nu = \frac{8\pi V}{(c\beta h)^3} \int_0^\infty \frac{x^2}{e^x - 1} dx
\]

\[
= \frac{8\pi V}{(c\beta h)^3} \int_0^\infty \frac{x^2 e^{-x}}{1 - e^{-x}} dx = \frac{8\pi V}{(c\beta h)^3} \int_0^\infty x^2 e^{-x} \sum_{n=0}^{\infty} e^{-nx} dx
\]

\[
= \frac{8\pi V}{(c\beta h)^3} \zeta(3) 2! = \frac{8\pi (kT)^3 V}{(ch)^3} \zeta(3) 2!.
\]

(4.100)

The mean energy \( \langle E \rangle \) of a photon in the black-body distribution (4.94) is the energy \( U(\beta, V) \) divided by the number of photons \( N(\beta, V) \)

\[
\langle E \rangle = \langle h\nu \rangle = \frac{3! \zeta(4)}{2! \zeta(3)} kT = \frac{\pi^4}{30 \zeta(3)} kT
\]

(4.101)

or \( \langle E \rangle \approx 2.70118 kT \) since Apéry’s constant \( \zeta(3) \) is 1.2020569032 … (Roger Apéry, 1916–1994).

\[\square\]

**Example 4.13 (The Lerch Transcendent)**  The **Lerch transcendent** is the series

\[
\Phi(z, s, \alpha) = \sum_{n=0}^{\infty} \frac{z^n}{(n + \alpha)^s}.
\]

(4.102)

It converges when \( |z| < 1 \) and \( \text{Re } s > 0 \) and \( \text{Re } \alpha > 0 \).

\[\square\]

### 4.11 Bernoulli Numbers and Polynomials

The **Bernoulli numbers** \( B_n \) are defined by the infinite series

\[
\frac{x}{e^x - 1} = \sum_{n=0}^{\infty} \frac{x^n}{n!} = \sum_{n=0}^{\infty} B_n \frac{x^n}{n!}
\]

(4.103)

for the **generating function** \( x/(e^x - 1) \). They are the successive derivatives

\[
B_n = \frac{d^n}{dx^n} \frac{x}{e^x - 1} \bigg|_{x=0}.
\]

(4.104)
So $B_0 = 1$ and $B_1 = -1/2$. The remaining odd Bernoulli numbers vanish

$$B_{2n+1} = 0 \quad \text{for } n > 0 \quad (4.105)$$

and the remaining even ones are given by Euler’s zeta function (4.92) formula

$$B_{2n} = \frac{(-1)^{n-1}2(2n)!}{(2\pi)^{2n}} \zeta(2n) \quad \text{for } n > 0. \quad (4.106)$$

The Bernoulli numbers occur in the power series for many transcendental functions, for instance

$$\coth x = \frac{1}{x} + \sum_{k=1}^{\infty} \frac{2^{2k}B_{2k}}{(2k)!} x^{2k-1} \quad \text{for } x^2 < \pi^2. \quad (4.107)$$

**Bernoulli’s polynomials** $B_n(y)$ are defined by the series

$$\frac{xe^{xy}}{e^x - 1} = \sum_{n=0}^{\infty} B_n(y) \frac{x^n}{n!} \quad (4.108)$$

for the generating function $xe^{xy}/(e^x - 1)$.

Some authors (Whittaker and Watson, 1927, p. 125–127) define Bernoulli’s numbers instead by

$$B_n = \frac{2(2n)!}{(2\pi)^{2n}} \zeta(2n) = 4n \int_0^{\infty} \frac{t^{2n-1} dt}{e^{2\pi t} - 1} \quad (4.109)$$

a result due to Carda.

### 4.12 Asymptotic Series

A series

$$s_n(x) = \sum_{k=0}^{n} \frac{a_k}{x^k} \quad (4.110)$$

is an **asymptotic** expansion for a real function $f(x)$ if the remainder $R_n$

$$R_n(x) = f(x) - s_n(x) \quad (4.111)$$

satisfies the condition

$$\lim_{x \to \infty} x^n R_n(x) = 0 \quad (4.112)$$

for fixed $n$. In this case, one writes

$$f(x) \approx \sum_{k=0}^{\infty} \frac{a_k}{x^k} \quad (4.113)$$
where the wavy equal sign indicates equality in the sense of (4.112). Some authors add the condition:

$$\lim_{n \to \infty} x^n R_n(x) = \infty$$

(4.114)

for fixed $x$.

**Example 4.14** (The Asymptotic Series for $E_1$) Let’s develop an asymptotic expansion for the function

$$E_1(x) = \int_x^\infty e^{-y} \frac{dy}{y}$$

(4.115)

which is related to the exponential-integral function

$$Ei(x) = \int_{-\infty}^x e^y \frac{dy}{y}$$

(4.116)

by the tricky formula $E_1(x) = -Ei(-x)$. Since

$$\frac{e^{-y}}{y} = -\frac{d}{dy} \left( \frac{e^{-y}}{y} \right) - \frac{e^{-y}}{y^2}$$

(4.117)

we may integrate by parts, getting

$$E_1(x) = \frac{e^{-x}}{x} - \int_x^\infty e^{-y} \frac{dy}{y^2}.$$  

(4.118)

Integrating by parts again, we find

$$E_1(x) = \frac{e^{-x}}{x} - \frac{e^{-x}}{x^2} + 2 \int_x^\infty e^{-y} \frac{dy}{y^3}.$$  

(4.119)

Eventually, we develop the series

$$E_1(x) = e^{-x} \left( \frac{0!}{x} - \frac{1!}{x^2} + \frac{2!}{x^3} - \frac{3!}{x^4} + \frac{4!}{x^5} - \ldots \right)$$

(4.120)

with remainder

$$R_n(x) = (-1)^n n! \int_x^\infty e^{-y} \frac{dy}{y^{n+1}}.$$  

(4.121)

Setting $y = u + x$, we have

$$R_n(x) = (-1)^n \frac{n!}{x^{n+1}} \int_0^\infty e^{-u} \frac{du}{(1 + \frac{u}{x})^{n+1}}$$

(4.122)
which satisfies the condition (4.112) that defines an asymptotic series
\[
\lim_{x \to \infty} x^n R_n(x) = \lim_{x \to \infty} (-1)^n \frac{n! e^{-x}}{x} \int_0^\infty e^{-u} \frac{du}{(1 + \frac{u}{x})^{n+1}}
\]
\[
= \lim_{x \to \infty} (-1)^n \frac{n! e^{-x}}{x} \int_0^\infty e^{-u} \, du
\]
\[
= \lim_{x \to \infty} (-1)^n \frac{n! e^{-x}}{x} = 0
\]
(4.123)
for fixed \(n\).

Asymptotic series often occur in physics. In such physical problems, a small parameter \(\lambda\) usually plays the role of \(1/x\). A perturbative series
\[
S_n(\lambda) = \sum_{k=0}^n a_k \lambda^k
\]
(4.124)
is an asymptotic expansion of the physical quantity \(S(\lambda)\) if the remainder
\[
R_n(\lambda) = S(\lambda) - S_n(\lambda)
\]
(4.125)
satisfies for fixed \(n\)
\[
\lim_{\lambda \to 0} \lambda^{-n} R_n(\lambda) = 0.
\]
(4.126)
The WKB approximation and the Dyson series for quantum electrodynamics are asymptotic expansions in this sense.

4.13 Some Electrostatic Problems

Gauss’s law \(\nabla \cdot D = \rho\) equates the divergence of the electric displacement \(D\) to the density \(\rho\) of free charges (charges that are free to move in or out of the dielectric medium—as opposed to those that are part of the medium and bound to it by molecular forces). In electrostatic problems, Maxwell’s equations reduce to Gauss’s law and the static form \(\nabla \times E = 0\) of Faraday’s law which implies that the electric field \(E\) is the gradient of an electrostatic potential \(E = -\nabla V\). (James Maxwell 1831–1879, Michael Faraday 1791–1867)

Across an interface with normal vector \(\hat{n}\) between two dielectrics, the tangential electric field is continuous while the normal electric displacement jumps by the surface density of free charge \(\sigma\)
\[
\hat{n} \times (E_2 - E_1) = 0 \quad \text{and} \quad \sigma = \hat{n} \cdot (D_2 - D_1).
\]
(4.127)

In a linear dielectric, the electric displacement \(D\) is proportional to the
of permittivity $\varepsilon_2$, and the lower region $z < -t/2$ is a uniform linear dielectric of permittivity $\varepsilon_3$. Suppose the lower infinite $x$-$y$-plane $z = -t/2$ has a uniform surface charge density $-\sigma$, while the upper plane $z = t/2$ has a uniform surface charge density $\sigma$. What is the energy per unit area of this system? What is the pressure on the second dielectric? What is the capacitance per unit area of the stack?
of analyticity. Since $I_M = -I$, the integral of $f(z)$ along this closed contour vanishes:

$$\oint f(z) \, dz = I + I_M = I - I = 0 \quad (5.25)$$

and we have again derived Cauchy’s integral theorem.

Since every polynomial $P(z) = c_0 + c_1 z + \cdots + c_n z^n$ is entire (everywhere analytic), it follows that its integral along any closed contour must vanish

$$\oint P(z) \, dz = 0. \quad (5.26)$$

**Example 5.3 (A pole)** The derivative of the function $f(z) = 1/(z - z_0)$

$$f'(z) = \lim_{dz \to 0} \left( \frac{1}{z + dz - z_0} - \frac{1}{z - z_0} \right) \frac{1}{dz} = -\frac{1}{(z - z_0)^2} \quad (5.27)$$

exists everywhere except at $z = z_0$, a region that is not simply connected.

### 5.3 Cauchy’s Integral Formula

Let $f(z)$ be analytic in a simply connected region $\mathcal{R}$ and $z_0$ a point inside this region. We first will integrate the function $f(z)/(z - z_0)$ along a tiny closed counterclockwise contour around the point $z_0$. The contour is a circle of radius $\epsilon$ with center at $z_0$ with points $z = z_0 + \epsilon e^{i\theta}$ for $0 \leq \theta \leq 2\pi$, and $dz = i\epsilon e^{i\theta} d\theta$. Since $z - z_0 = \epsilon e^{i\theta}$, the contour integral in the limit $\epsilon \to 0$ is

$$\oint_{\epsilon} \frac{f(z)}{z - z_0} \, dz = \int_0^{2\pi} \left[ \frac{f(z_0) + f'(z_0) (z - z_0)}{z - z_0} \right] i\epsilon e^{i\theta} d\theta$$

$$= \int_0^{2\pi} \left[ \frac{f(z_0) + f'(z_0) \epsilon e^{i\theta}}{\epsilon e^{i\theta}} \right] i\epsilon e^{i\theta} d\theta$$

$$= \int_0^{2\pi} \left[ f(z_0) + f'(z_0) \epsilon e^{i\theta} \right] i d\theta. \quad (5.28)$$

The $\theta$-integral involving $f'(z_0)$ vanishes, and so we have

$$f(z_0) = \frac{1}{2\pi i} \oint_{\epsilon} \frac{f(z)}{z - z_0} \, dz \quad (5.29)$$

which is a miniature version of Cauchy’s integral formula.

Now consider the counterclockwise contour $\mathcal{C}'$ in Fig. 5.3 which is a big counterclockwise circle, a small clockwise circle, and two parallel straight lines, all within a simply connected region $\mathcal{R}$ in which $f(z)$ is analytic. The
Third-Harmonic Microscopy

Figure 5.6 In the limit in which the distance $L$ is much larger than the wavelength $\lambda$, the integral (5.128) is non-zero when an edge (solid line) lies where the beam is focused but not when a feature (dots) lies where the beam is not focused. Only features within the focused region are visible.

which is in the UHP since the length $b > 0$, but no singularity in the LHP $y < 0$. So the integral of $f(z)$ along the closed contour from $z = -R$ to $z = R$ and then along the ghost contour vanishes. But since the integral along the ghost contour vanishes, so does the integral from $-R$ to $R$. Thus when the dispersion is normal, the third-harmonic signal vanishes, $E_3 = 0$, as long as the medium with constant $\chi^{(3)}(z)$ effectively extends from $-\infty$ to $\infty$ so that its edges are in the unfocused region like the dotted lines of Fig. 5.6. But an edge with varying $\chi^{(3)}(z)$ in the focused region like the solid line of the figure does make a third-harmonic signal $E_3$. Third-harmonic microscopy lets us see features instead of background.
We let $\epsilon \to 0$ and find
\[ I = -2i \int_{-\infty}^{\infty} dx \int_{0}^{\infty} dy \frac{1}{(1 + x^2 + y^2)^2}. \] (5.181)
Changing variables to $\rho^2 = x^2 + y^2$, we have
\[ I = -4\pi i \int_{0}^{\infty} d\rho \frac{\rho}{(1 + \rho^2)^2} = 2\pi i \int_{0}^{\infty} d\rho \frac{d}{d\rho} \frac{1}{1 + \rho^2} = -2\pi i. \] (5.182)
which is simpler than evaluating the integral (5.178) directly.

5.15 Logarithms and Cuts

By definition, a function $f$ is single valued; it maps every number $z$ in its domain into a unique image $f(z)$. A function that maps only one number $z$ in its domain into each $f(z)$ in its range is said to be one to one. A one-to-one function $f(z)$ has a well-defined inverse function $f^{-1}(z)$.

The exponential function is one to one when restricted to the real numbers. It maps every real number $x$ into a positive number $\exp(x)$. It has an inverse function $\ln(x)$ that maps every positive number $\exp(x)$ back into $x$. But the exponential function is not one to one on the complex numbers because $\exp(z + 2\pi ni) = \exp(z)$ for every integer $n$. The exponential function is many to one. Thus on the complex numbers, the exponential function has no inverse function. Its would-be inverse function $\ln$ maps it to $\ln(\exp(z))$ or $z + 2\pi ni$ which is not unique. It has in it an arbitrary integer $n$.

In other words, when exponentiated, the logarithm of a complex number $z$ returns $\exp(\ln z) = z$. So if $z = r \exp(i\theta)$, then a suitable logarithm is $\ln z = \ln r + i\theta$. But what is $\theta$? In the polar representation of $z$, the argument $\theta$ can just as well be $\theta + 2\pi n$ because both give $z = r \exp(i\theta) = r \exp(i\theta + i2\pi n)$. So $\ln r + i\theta + i2\pi n$ is a correct value for $\ln[r \exp(i\theta)]$ for every integer $n$.

People usually want one of the correct values of a logarithm, rather than all of them. Two conventions are common. In the first convention, the angle $\theta$ is zero along the positive real axis and increases continuously as the point $z$ moves counterclockwise around the origin, until at points just below the positive real axis, $\theta = 2\pi - \epsilon$ is slightly less than $2\pi$. In this convention, the value of $\theta$ drops by $2\pi$ as one crosses the positive real axis moving counterclockwise. This discontinuity on the positive real axis is called a cut.

The second common convention puts the cut on the negative real axis. Here the value of $\theta$ is the same as in the first convention when the point $z$ is in the upper half-plane. But in the lower half-plane, $\theta$ decreases from
The second common convention puts the cut on the negative real axis. Here the value of \( \theta \) is the same as in the first convention when the point \( z \) is in the upper half-plane. But in the lower half-plane, \( \theta \) decreases from 0 to \(-\pi\) as the point \( z \) moves clockwise from the positive real axis to just below the negative real axis, where \( \theta = -\pi + \epsilon \). As one crosses the negative real axis moving clockwise or up, \( \theta \) jumps by \( 2\pi \) while crossing the cut. The two conventions agree in the upper half-plane but differ by \( 2\pi \) in the lower half-plane.

Sometimes it is convenient to place the cut on the positive or negative imaginary axis — or along a line that makes an arbitrary angle with the real axis. In any particular calculation, we are at liberty to define the polar angle \( \theta \) by placing the cut anywhere we like, but we must not change from one convention to another in the same computation.

## 5.16 Powers and Roots

The logarithm is the key to many other functions to which it passes its arbitrariness. For instance, any power \( a \) of \( z = r \exp(i\theta) \) is defined as

\[
z^a = \exp(a \ln z) = \exp [a (\ln r + i\theta + i2\pi n)] = r^a e^{ia\theta} e^{i2\pi na}.
\] (5.183)

So \( z^a \) is not unique unless \( a \) is an integer. The square-root, for example,

\[
\sqrt{z} = \exp \left[ \frac{1}{2} (\ln r + i\theta + i2\pi n) \right] = \sqrt{r} e^{i\theta/2} e^{in\pi} = (-1)^n \sqrt{r} e^{i\theta/2}
\] (5.184)
changes sign when we change \( \theta \) by \( 2\pi \) as we cross a cut. The \( m \)-th root

\[
\sqrt[m]{z} = z^{1/m} = \exp \left( \frac{\ln z}{m} \right)
\] (5.185)
changes by \( \exp(\pm 2\pi i/m) \) when we cross a cut and change \( \theta \) by \( 2\pi \). And when \( a = u + iv \) is a complex number, \( z^a \) is

\[
z^a = e^{a \ln z} = e^{(u+iv)(\ln r + i\theta + i2\pi n)} = r^{u+iv} e^{(-v+iu)(\theta+2\pi n)}
\] (5.186)
which changes by \( \exp[2\pi(-v + iu)] \) as we cross a cut.

**Example 5.27** \((i^i)\) The number \( i = \exp(i\pi/2 + i2\pi n) \) for any integer \( n \). So the general value of \( i^i \) is \( i^i = \exp[i(i\pi/2 + i2\pi n)] = \exp(-\pi/2 - 2\pi n) \).

One can define a sequence of \( m \)-th-root functions

\[
\left( z^{1/m} \right)_n = \exp \left( \frac{\ln r + i(\theta + 2\pi n)}{m} \right)
\] (5.187)
one for each integer \( n \). These functions are the **branches** of the \( m \)-th-root.
interval \([-1, 1]\). Let’s promote \(x\) to a complex variable \(z\) and write the square root as \(\sqrt{1-x^2} = -i\sqrt{x^2-1} = -i\sqrt{(z-1)(z+1)}\). As in the last example (5.29), if in both of the square roots we put the cut on the negative (or the positive) real axis, then the function \(f(z) = 1/((z-k)(-i)\sqrt{(z-1)(z+1)})\) will be analytic everywhere except along a cut on the interval \([-1, 1]\) and at \(z = k\). The circle \(z = Re^{i\theta}\) for \(0 \leq \theta \leq 2\pi\) is a ghost contour as \(R \to \infty\). If we shrink-wrap this ccw contour around the pole at \(z = k\) and the interval \([-1, 1]\), then we get \(0 = -2I + 2\pi i/((-i)\sqrt{k-1}\sqrt{k+1})\) or

\[
I = -\frac{\pi}{\sqrt{k-1}\sqrt{k+1}}. \tag{5.193}
\]

So if \(k = -2\), then \(I = \pi/\sqrt{3}\), while if \(k = 2\), then \(I = -\pi/\sqrt{3}\). \(\square\)

**Example 5.31 (Contour Integral with a Cut)** Let’s compute the integral

\[
I = \int_{0}^{\infty} \frac{x^a}{(x+1)^2} \, dx \tag{5.194}
\]

for \(-1 < a < 1\). We promote \(x\) to a complex variable \(z\) and put the cut on the positive real axis. Since

\[
\lim_{|z| \to \infty} \frac{|z|^{a+1}}{|z+1|^2} = 0, \tag{5.195}
\]

the integrand vanishes faster than \(1/|z|\), and we may add two ghost contours, \(G_+\) counter-clockwise around the upper half-plane and \(G_-\) counter-clockwise around the lower half-plane, as shown in Fig. 5.8.

We add a contour \(C_-\) that runs from \(-\infty\) to the double pole at \(z = -1\), loops around that pole, and then runs back to \(-\infty\); the two long contours along the negative real axis cancel because the cut in \(\theta\) lies on the positive real axis. So the contour integral along \(C_-\) is just the clockwise integral around the double pole which by Cauchy’s integral formula (5.34) is

\[
\oint_{C_-} \frac{z^a}{(z-(-1))^2} \, dz = -2\pi i \frac{dz^a}{dz} \bigg|_{z=-1} = 2\pi i a e^{\pi ai}. \tag{5.196}
\]

We also add the integral \(I_-\) from \(\infty\) to 0 just below the real axis

\[
I_- = \int_{\infty}^{0} \frac{(x-ix)^a}{(x-ix+1)^2} \, dx = \int_{\infty}^{0} \frac{\exp(a(\ln(x) + 2\pi i))}{(x+1)^2} \, dx \tag{5.197}
\]

which is

\[
I_- = -e^{2\pi ai} \int_{0}^{\infty} \frac{x^a}{(x+1)^2} \, dx = -e^{2\pi ai} I. \tag{5.198}
\]
in the lower half plane. The delta function in the second integral then gives \( \pi/2 \), so that

\[
I = \oint dk \frac{e^{ik}}{2i(k + i\epsilon)} + \frac{\pi}{2} = \frac{\pi}{2}
\]  

as stated in (3.109).

---

**Example 5.36 (The Feynman Propagator)** Adding \( \pm i\epsilon \) to the denominator of a pole term of an integral formula for a function \( f(x) \) can slightly shift the pole into the upper or lower half plane, causing the pole to contribute if a ghost contour goes around the upper half-plane or the lower half-plane. Such an \( i\epsilon \) can impose a boundary condition on a Green’s function.

The Feynman propagator \( \Delta_F(x) \) is a Green’s function for the Klein-Gordon differential operator (Weinberg, 1995, pp. 274–280)

\[
(m^2 - \Box)\Delta_F(x) = \delta^4(x)
\]  

in which \( x = (x^0, \mathbf{x}) \) and

\[
\Box = \nabla^2 = \nabla \cdot \nabla
\]  

is the four-dimensional version of the Laplacian \( \Delta \equiv \nabla \cdot \nabla \). Here \( \delta^4(x) \) is the four-dimensional Dirac delta function (3.36)

\[
\delta^4(x) = \int \frac{d^4q}{(2\pi)^4} \exp[i(q \cdot x - q^0 x^0)] = \int \frac{d^4q}{(2\pi)^3} e^{iqx}
\]  

in which \( qx = q \cdot x - q^0 x^0 \) is the Lorentz-invariant inner product of the 4-vectors \( q \) and \( x \). There are many Green’s functions that satisfy Eq. (5.230). Feynman’s propagator \( \Delta_F(x) \) is the one that satisfies boundary conditions that will become evident when we analyze the effect of its \( i\epsilon \)

\[
\Delta_F(x) = \int \frac{d^4q}{(2\pi)^4} \frac{\exp(iqx)}{q^2 + m^2 - i\epsilon} = \int \frac{d^3q}{(2\pi)^3} \int_{-\infty}^{\infty} \frac{dq^0}{2\pi} \frac{e^{iqx - iq^0x^0}}{q^2 + m^2 - i\epsilon}.
\]  

The quantity \( E_q = \sqrt{q^2 + m^2} \) is the energy of a particle of mass \( m \) and momentum \( q \) in natural units with the speed of light \( c = 1 \). Using this abbreviation and setting \( \epsilon' = \epsilon/2E_q \), we may write the denominator as

\[
q^2 + m^2 - i\epsilon = q \cdot q - (q^0)^2 + m^2 - i\epsilon = (E_q - i\epsilon' - q^0) (E_q - i\epsilon' + q^0) + \epsilon'^2
\]  

(5.234)
5.20 Kramers-Kronig Relations

If we use $\sigma E$ for the current density $J$ and $E(t) = e^{-i\omega t} E$ for the electric field, then Maxwell’s equation $\nabla \times B = \mu J + \epsilon \mu \dot{E}$ becomes

$$\nabla \times B = -i\omega \epsilon \mu \left( 1 + i \frac{\sigma}{\epsilon \omega} \right) E \equiv -i\omega n^2 \epsilon_0 \mu_0 E \quad (5.268)$$

and reveals the squared index of refraction as

$$n^2(\omega) = \frac{\epsilon \mu}{\epsilon_0 \mu_0} \left( 1 + i \frac{\sigma}{\epsilon \omega} \right). \quad (5.269)$$

The imaginary part of $n^2$ represents the scattering of light mainly by electrons. At high frequencies in nonmagnetic materials $n^2(\omega) \rightarrow 1$, and so Kramers and Kronig applied the Hilbert-transform relations (5.267) to the function $n^2(\omega) - 1$ in order to satisfy condition (5.255). Their relations are

$$\text{Re}(n^2(\omega_0)) = 1 + \frac{2}{\pi} P \int_0^\infty \frac{\omega \text{Im}(n^2(\omega))}{\omega^2 - \omega_0^2} d\omega \quad (5.270)$$

and

$$\text{Im}(n^2(\omega_0)) = -\frac{2\omega_0}{\pi} P \int_0^\infty \frac{\text{Re}(n^2(\omega)) - 1}{\omega^2 - \omega_0^2} d\omega. \quad (5.271)$$

What Kramers and Kronig actually wrote was slightly different from these dispersion relations (5.270 & 5.271). H. A. Lorentz had shown that the index of refraction $n(\omega)$ is related to the forward scattering amplitude $f(\omega)$ for the scattering of light by a density $N$ of scatterers (Sakurai, 1982)

$$n(\omega) = 1 + \frac{2\pi e^2}{\omega_0^2} N f(\omega). \quad (5.272)$$

They used this formula to infer that the real part of the index of refraction approached unity in the limit of infinite frequency and applied the Hilbert transform (5.267)

$$\text{Re}[n(\omega)] = 1 + \frac{2}{\pi} P \int_0^\infty \frac{\omega \text{Im}[n(\omega)']}{\omega'^2 - \omega^2} d\omega'. \quad (5.273)$$

The Lorentz relation (5.272) expresses the imaginary part $\text{Im}[n(\omega)]$ of the index of refraction in terms of the imaginary part of the forward scattering amplitude $f(\omega)$

$$\text{Im}[n(\omega)] = 2\pi (c/\omega)^2 N \text{Im}[f(\omega)]. \quad (5.274)$$

And the optical theorem relates $\text{Im}[f(\omega)]$ to the total cross-section

$$\sigma_{\text{tot}} = \frac{4\pi}{|k|} \text{Im}[f(\omega)] = \frac{4\pi c}{\omega} \text{Im}[f(\omega)]. \quad (5.275)$$
and that the real part of the forward scattering amplitude is given by the Kramers-Kronig integral (5.276) of the total cross-section

$$\text{Re}(f(\omega)) = \frac{\omega^2}{2\pi^2 c} P \int_0^\infty \frac{\sigma_{\text{tot}}(\omega') d\omega'}{\omega'^2 - \omega^2}. \quad (5.286)$$

So the real part of the index of refraction is

$$n_r(\omega) = 1 + \frac{cN}{\pi} P \int_0^\infty \frac{\sigma_{\text{tot}}(\omega') d\omega'}{\omega'^2 - \omega^2}. \quad (5.287)$$

If the amplitude for forward scattering is of the Breit-Wigner form

$$f(\omega) = f_0 \frac{\Gamma/2}{\omega_0 - \omega - i\Gamma/2} \quad (5.288)$$

then by (5.285) the real part of the index of refraction is

$$n_r(\omega) = 1 + \frac{\pi c^2 N f_0 \Gamma (\omega_0 - \omega)}{\omega^2 [(\omega - \omega_0)^2 + \Gamma^2/4]} \quad (5.289)$$

and by (5.283) the group velocity is

$$v_g = c \left[ 1 + \frac{\pi c^2 N f_0 \Gamma \omega_0}{\omega^2} \left( \frac{(\omega - \omega_0)^2 - \Gamma^2/4}{[(\omega - \omega_0)^2 + \Gamma^2/4]^2} \right) \right]^{-1}. \quad (5.290)$$

This group velocity $v_g$ is less than $c$ whenever $(\omega - \omega_0)^2 > \Gamma^2/4$. But we get fast light $v_g > c$, if $(\omega - \omega_0)^2 < \Gamma^2/4$, and even backwards light, $v_g < 0$, if $\omega \approx \omega_0$ with $4\pi c^2 N f_0 / \Gamma \omega_0 \gg 1$. Robert W. Boyd’s papers explain how to make slow and fast light (Bigelow et al., 2003) and backwards light (Gehring et al., 2006).

We can use the principal-part identity (5.224) to subtract

$$0 = \frac{cN}{\pi} P \int_0^\infty \frac{\sigma_{\text{tot}}(\omega') d\omega'}{\omega'^2 - \omega^2} \quad (5.291)$$

from the Kramers-Kronig integral (5.287) so as to write the index of refraction in the regularized form

$$n_r(\omega) = 1 + \frac{cN}{\pi} P \int_0^\infty \frac{\sigma_{\text{tot}}(\omega') - \sigma_{\text{tot}}(\omega)}{\omega'^2 - \omega^2} d\omega' \quad (5.292)$$

which we can differentiate and use in the group-velocity formula (5.283)

$$v_g(\omega) = c \left[ 1 + \frac{cN}{\pi} P \int_0^\infty \frac{\sigma_{\text{tot}}(\omega') - \sigma_{\text{tot}}(\omega)}{(\omega'^2 - \omega^2)^2} d\omega' \right]^{-1}. \quad (5.293)$$
\(T(z)\) are defined by its Laurent series
\[
T(z) = \sum_{n=-\infty}^{\infty} \frac{L_n}{z^{n+2}}
\]
and the inverse relation
\[
L_n = \frac{1}{2\pi i} \oint z^{n+1} T(z) \, dz.
\]
Thus the commutator of two modes involves two loop integrals
\[
[L_m, L_n] = \left[ \frac{1}{2\pi i} \oint z^{m+1} T(z) \, dz, \frac{1}{2\pi i} \oint w^{n+1} T(w) \, dw \right]
\]
which we may deform as long as we cross no poles. Let’s hold \(w\) fixed and deform the \(z\) loop so as to keep the \(T\)’s radially ordered when \(z\) is near \(w\) as in Fig. 5.10. The operator-product expansion of the radially ordered product \(\mathcal{R}\{T(z)T(w)\}\) is
\[
\mathcal{R}\{T(z)T(w)\} = \frac{c/2}{(z-w)^4} + \frac{2}{(z-w)^2} T(w) + \frac{1}{z-w} T'(w) + \ldots
\]
in which the prime means derivative, \(c\) is a constant, and the dots denote terms that are analytic in \(z\) and \(w\). The commutator introduces a minus sign that cancels most of the two contour integrals and converts what remains into an integral along a tiny circle \(C_w\) about the point \(w\) as in Fig. 5.10
\[
[L_m, L_n] = \oint \frac{dw}{2\pi i} w^{n+1} \oint \frac{dz}{2\pi i} z^{m+1} \left[ \frac{c/2}{(z-w)^4} + \frac{2T(w)}{(z-w)^2} + \frac{T'(w)}{z-w} \right]
\]
After doing the \(z\)-integral, which is left as a homework exercise (5.43), one may use the Laurent series (5.336) for \(T(w)\) to do the \(w\)-integral, which one may choose to be along a tiny circle about \(w = 0\), and so find the commutator
\[
[L_m, L_n] = (m-n) L_{m+n} + \frac{c}{12} m(m^2 - 1) \delta_{m+n,0}
\]
of the Virasoro algebra.

**Exercises**

5.1 Compute the two limits (5.6) and (5.7) of example 5.2 but for the function \(f(x, y) = x^2 - y^2 + 2ixy\). Do the limits now agree? Explain.

5.2 Show that if \(f(z)\) is analytic in a disk, then the integral of \(f(z)\) around a tiny (isosceles) triangle of side \(\epsilon \ll 1\) inside the disk is zero to order \(\epsilon^2\).
5.12 Evaluate the contour integral of the function $f(z) = \sin wz/(z - 5)^3$ along the curve $z = 6 + 4(\cos t + i \sin t)$ for $0 \leq t \leq 2\pi$.

5.13 Evaluate the contour integral of the function $f(z) = \sin wz/(z - 5)^3$ along the curve $z = -6 + 4(\cos t + i \sin t)$ for $0 \leq t \leq 2\pi$.

5.14 Is the function $f(x, y) = x^2 + iy^2$ analytic?

5.15 Is the function $f(x, y) = x^3 - 3xy^2 + 3i x^2 y - iy^3$ analytic? Is the function $x^3 - 3xy^2$ harmonic? Does it have a minimum or a maximum? If so, what are they?

5.16 Is the function $f(x, y) = x^2 + y^2 + i(x^2 + y^2)$ analytic? Is $x^2 + y^2$ a harmonic function? What is its minimum, if it has one?

5.17 Derive the first three nonzero terms of the Laurent series for $f(z) = 1/(e^z - 1)$ about $z = 0$.

5.18 Assume that a function $g(z)$ is meromorphic in $\mathbb{R}$ and has a Laurent series (5.97) about a point $w \in \mathbb{R}$. Show that as $z \to w$, the ratio $g'(z)/g(z)$ becomes (5.95).

5.19 Find the poles and residues of the functions $1/\sin z$ and $1/\cos z$.

5.20 Derive the integral formula (5.122) from (5.121).

5.21 Show that if $\text{Re} w < 0$, then for arbitrary complex $z$

$$
\int_{-\infty}^{\infty} e^{w(x+z)^2} \, dx = \sqrt{\frac{\pi}{-w}}.
$$

(5.347)

5.22 Use a ghost contour to evaluate the integral

$$
\int_{-\infty}^{\infty} \frac{x \sin x}{x^2 + a^2} \, dx.
$$

Show your work; do not just quote the result of a commercial math program.

5.23 For $a > 0$ and $b^2 - 4ac < 0$, use a ghost contour to do the integral

$$
\int_{-\infty}^{\infty} \frac{dx}{ax^2 + bx + c}.
$$

(5.348)

5.24 Show that

$$
\int_{0}^{\infty} \cos ax e^{-x^2} \, dx = \frac{1}{2} \sqrt{\pi} e^{-a^2/4}.
$$

(5.349)

5.25 Show that

$$
\int_{-\infty}^{\infty} \frac{dx}{1 + x^4} = \frac{\pi}{\sqrt{2}}.
$$

(5.350)

5.26 Evaluate the integral

$$
\int_{0}^{\infty} \frac{\cos x}{1 + x^4} \, dx.
$$

(5.351)
5.27 Show that the Yukawa Green’s function (5.151) reproduces the Yukawa potential (5.141) when \( n = 3 \). Use \( K_{1/2}(x) = \sqrt{\frac{\pi}{2x}} e^{-x} \) (9.99).

5.28 Derive the two explicit formulas (5.188) and (5.189) for the square-root of a complex number.

5.29 What is \((-i)^i\)? What is the most general value of this expression?

5.30 Use the indefinite integral (5.223) to derive the principal-part formula (5.224).

5.31 The Bessel function \( J_n(x) \) is given by the integral

\[
J_n(x) = \frac{1}{2\pi i} \oint_C e^{(x/2)(z^{-1}/z)} \frac{dz}{z^{n+1}}
\]

along a counter-clockwise contour about the origin. Find the generating function for these Bessel functions, that is, the function \( G(x, z) \) whose Laurent series has the \( J_n(x) \)'s as coefficients

\[
G(x, z) = \sum_{n=-\infty}^{\infty} J_n(x) z^n.
\]

5.32 Show that the Heaviside function \( \theta(y) = (y + |y|)/(2|y|) \) is given by the integral

\[
\theta(y) = \frac{1}{2\pi i} \int_{-\infty}^{\infty} e^{iyx} \frac{dx}{x - i\epsilon}
\]

in which \( \epsilon \) is an infinitesimal positive number.

5.33 Show that the integral of \( \exp(ik)/k \) along the contour from \( k = L \) to \( k = L + iH \) and then to \( k = -L + iH \) and then down to \( k = -L \) vanishes in the double limit \( L \to \infty \) and \( H \to \infty \).

5.34 Use a ghost contour and a cut to evaluate the integral

\[
I = \int_{-1}^{1} \frac{dx}{(x^2 + 1)\sqrt{1 - x^2}}
\]

by imitating example 5.30. Be careful when picking up the poles at \( z = \pm i \). If necessary, use the explicit square-root formulas (5.188) and (5.189).

5.35 Redo the previous exercise (5.34) by defining the square roots so that the cuts run from \(-\infty\) to \(-1\) and from \(1\) to \(\infty\). Take advantage of the evenness of the integrand and integrate on a contour that is slightly above the whole real axis. Then add a ghost contour around the upper half plane.

5.36 Show that if \( u \) is even and \( v \) is odd, then the Hilbert transforms (5.265) imply (5.267).
Example 6.6 (The Helmholtz Equation in Three Dimensions) In three
dimensions and in rectangular coordinates \( r = (x, y, z) \), the function
\( f(x, y, z) = X(x)Y(y)Z(z) \) is a solution of the ODE \(-\Delta f = k^2 f\) as long
as \( X, Y, \) and \( Z \) satisfy \(-X'' = a^2 X, -Y'' = b^2 Y, \) and \(-Z'' = c^2 Z\)
with \( a^2 + b^2 + c^2 = k^2 \). We set \( X_n(x) = \alpha \sin \alpha x + \beta \cos \alpha x \) and so forth.
Arbitrary linear combinations of the products \( X_n Y_b Z_c \) also are solutions of
Helmholtz’s equation \(-\Delta f = k^2 f\) as long as \( a^2 + b^2 + c^2 = k^2 \).

In cylindrical coordinates \((\rho, \phi, z)\), the laplacian (6.34) is

\[
\nabla \cdot \nabla f = \Delta f = \frac{1}{\rho} \left[ (\rho f_{\rho})_{\rho} + \frac{1}{\rho} f_{\phi\phi} + \rho f_{zz} \right]
\]

and so if we substitute \( f(\rho, \phi, z) = P(\rho) \Phi(\phi) Z(z) \) into Helmholtz’s equation
\(-\Delta f = k^2 f\) and multiply both sides by \(-\rho^2/P \Phi Z\), then we get

\[
\frac{\rho^2}{f} \Delta f = \frac{\rho^2 P'' + \rho P'}{P} + \frac{\Phi''}{\Phi} + \frac{\rho^2 Z''}{Z} = -\alpha^2 \rho^2.
\]

If we set \( Z_k(z) = e^{kz} \), then this equation becomes (6.46) with \( k^2 \) replaced
by \( \alpha^2 + k^2 \). Its solution then is

\[
f(\rho, \phi, z) = J_n(\sqrt{\alpha^2 + k^2} \rho) e^{i\phi} e^{kz}
\]

in which \( n \) must be an integer if the solution is to apply to the full range of \( \phi \)
from 0 to 2\( \pi \). The case in which \( \alpha = 0 \) corresponds to Laplace’s equation
with solution \( f(\rho, \phi, z) = J_n(k\rho) e^{i\phi} e^{kz} \). We could have required \( Z \) to satisfy
\( Z'' = -k^2 Z \). The solution (6.51) then would be

\[
f(\rho, \phi, z) = J_n(\sqrt{\alpha^2 - k^2} \rho) e^{i\phi} e^{ikz}.
\]

But if \( \alpha^2 - k^2 < 0 \), we write this solution in terms of the modified Bessel
function \( I_n(x) = i^{-n} J_n(ix) \) (section 9.3) as

\[
f(\rho, \phi, z) = I_n(\sqrt{k^2 - \alpha^2} \rho) e^{i\phi} e^{ikz}.
\]

In spherical coordinates, the laplacian (6.35) is

\[
\Delta f = \frac{1}{r^2} \frac{\partial}{\partial r} \left( r^2 \frac{\partial f}{\partial r} \right) + \frac{1}{r^2 \sin \theta} \frac{\partial}{\partial \theta} \left( \sin \theta \frac{\partial f}{\partial \theta} \right) + \frac{1}{r^2 \sin^2 \theta} \frac{\partial^2 f}{\partial \phi^2}.
\]

If we set \( f(r, \theta, \phi) = R(r) \Theta(\theta) \Phi_m(\phi) \) where \( \Phi_m = e^{im\phi} \) and multiply both
sides of the Helmholtz equation \(-\Delta f = k^2 f\) by \(-r^2/R \Theta \Phi\), then we get

\[
\frac{(r^2 R')'}{R} + \frac{(\sin \theta \Theta')'}{\sin \theta \Theta} - \frac{m^2}{\sin^2 \theta} = -k^2 r^2.
\]
The first term is a function of $r$, the next two terms are functions of $\theta$, and the last term is a constant. So we set the $r$-dependent terms equal to a constant $\ell(\ell + 1) - k^2$ and the $\theta$-dependent terms equal to $-\ell(\ell + 1)$, and we require the associated Legendre function $\Theta_{\ell,m}(\theta)$ to satisfy (8.91)

\[
(\sin \theta \Theta_{\ell,m})' / \sin \theta + [\ell(\ell + 1) - m^2 / \sin^2 \theta] \Theta_{\ell,m} = 0. \tag{6.56}
\]

If $\Phi(\phi) = e^{im\phi}$ is to be single valued for $0 \leq \phi \leq 2\pi$, then the parameter $m$ must be an integer. The constant $\ell$ also must be an integer with $-\ell \leq m \leq \ell$ (example 6.29, section 8.12) if $\Theta_{\ell,m}(\theta)$ is to be single valued and finite for $0 \leq \theta \leq \pi$. The product $f = R \Theta \Phi$ then will obey Helmholtz's equation 

\[
-\triangle f = k^2 f \quad \text{if the radial function } \quad R_k,\ell(r) = j_{\ell}(kr) \quad \text{satisfies}
\]

\[
(r^2 R_k,\ell)' + [k^2 r^2 - \ell(\ell + 1)] R_k,\ell = 0 \tag{6.57}
\]

which it does because the spherical Bessel function $j_{\ell}(x)$ obeys Bessel's equation (9.63)

\[
(x^2 j_{\ell}')' + [x^2 - \ell(\ell + 1)] j_{\ell} = 0. \tag{6.58}
\]

In three dimensions, Helmholtz’s equation separates in 11 standard coordinate systems (Morse and Feshbach, 1953, pp. 655–664).

### 6.6 Wave Equations

You can easily solve some of the linear homogeneous partial differential equations of electrodynamics (exercise 6.6) and quantum field theory.

**Example 6.7** (The Klein-Gordon Equation) In Minkowski space, the analog of the laplacian in natural units ($h = c = 1$) is (summing over $a$ from 0 to 3)

\[
\Box = \partial_a \partial^a = \Delta - \frac{\partial^2}{\partial x_0^2} = \Delta - \frac{\partial^2}{\partial t^2} \tag{6.59}
\]

and the Klein-Gordon wave equation is

\[
(\Box - m^2) A(x) = \left(\Delta - \frac{\partial^2}{\partial t^2} - m^2\right) A(x) = 0. \tag{6.60}
\]

If we set $A(x) = B(px)$ where $px = p_a x^a = p \cdot x - p^0 x^0$, then the $k$th partial derivative of $A$ is $p_k$ times the first derivative of $B$

\[
\frac{\partial}{\partial x_k} A(x) = \frac{\partial}{\partial x_k} B(px) = p_k B'(px) \tag{6.61}
\]
and so the Klein-Gordon equation (6.60) becomes
\[(\Box - m^2) A = \left(\mathbf{p}^2 - (p^0)^2\right) B'' - m^2 B = \mathbf{p}^2 B'' - m^2 B = 0 \tag{6.62}\]
in which \(p^2 = \mathbf{p}^2 - (p^0)^2\). Thus if \(B(p \cdot x) = \exp(ip \cdot x)\) so that \(B'' = -B\), and if the energy-momentum 4-vector \((p^0, \mathbf{p})\) satisfies \(p^2 + m^2 = 0\), then \(A(x)\) will satisfy the Klein-Gordon equation. The condition \(p^2 + m^2 = 0\) relates the energy \(p^0 = \sqrt{\mathbf{p}^2 + m^2}\) to the momentum \(\mathbf{p}\) for a particle of mass \(m\). □

**Example 6.8 (Field of a Spinless Boson)** The quantum field
\[
\phi(x) = \int \frac{d^3 p}{\sqrt{2p^0(2\pi)^3}} \left[a(p)e^{ipx} + a^\dagger(p)e^{-ipx}\right] \tag{6.63}
\]
describes spinless bosons of mass \(m\). It satisfies the Klein-Gordon equation \((\Box - m^2) \phi(x) = 0\) because \(p^0 = \sqrt{\mathbf{p}^2 + m^2}\). The operators \(a(p)\) and \(a^\dagger(p)\) respectively represent the annihilation and creation of the bosons and obey the commutation relations
\[
[a(p), a^\dagger(p')] = \delta^3(p - p') \quad \text{and} \quad [a(p), a(p')] = [a^\dagger(p), a^\dagger(p')] = 0 \tag{6.64}
\]
in units with \(\hbar = c = 1\). These relations make the field \(\phi(x)\) and its time derivative \(\dot{\phi}(y)\) satisfy the **canonical equal-time commutation relations**
\[
[\phi(x, t), \dot{\phi}(y, t)] = i \delta^3(x - y) \quad \text{and} \quad [\phi(x, t), \phi(y, t)] = [\dot{\phi}(x, t), \dot{\phi}(y, t)] = 0 \tag{6.65}
\]
in which the dot means time derivative. □

**Example 6.9 (Field of the Photon)** The electromagnetic field has four components, but in the Coulomb or radiation gauge \(\nabla \cdot A(x) = 0\), the component \(A_0\) is a function of the charge density, and the vector potential \(A\) in the absence of charges and currents satisfies the wave equation \(\Box A(x) = 0\) for a spin-one massless particle. We write it as
\[
A(x) = \sum_{s=1}^{2} \int \frac{d^3 p}{\sqrt{2p^0(2\pi)^3}} \left[\mathbf{e}(p, s) a(p, s) e^{ipx} + e^*(p, s) a^\dagger(p, s) e^{-ipx}\right] \tag{6.66}
\]
in which the sum is over the two possible polarizations \(s\). The energy \(p^0\) is equal to the modulus \(|\mathbf{p}|\) of the momentum because the photon is massless, \(p^2 = 0\). The dot-product of the polarization vectors \(\mathbf{e}(p, s)\) with the momentum vanishes \(\mathbf{p} \cdot \mathbf{e}(p, s) = 0\) so as to respect the gauge condition.
(6.150) was linear in $y$. So we could set $P = \alpha (ry - s)$ and $Q = \alpha$. When $P$ and $Q$ are more complicated, integrating factors are harder to find or nonexistent.

**Example 6.23** (Bodies Falling in Air) The downward speed $v$ of a mass $m$ in a gravitational field of constant acceleration $g$ is described by the inhomogeneous first-order ODE $mv_t = mg - bv$ in which $b$ represents air resistance. This equation is like (6.150) but with $t$ instead of $x$ as the independent variable, $r = b/m$, and $s = g$. Thus by (6.157), its solution is

$$v(t) = \frac{mg}{b} + \left( v(0) - \frac{mg}{b} \right) e^{-bt/m}. \quad (6.158)$$

The terminal speed $mg/b$ is nearly 200 km/h for a falling man. A diving Peregrine falcon can exceed 320 km/h; so can a falling bullet. But mice can fall down mine shafts and run off unhurt, and insects and birds can fly.

If the falling bodies are microscopic, a statistical model is appropriate. The potential energy of a mass $m$ at height $h$ is $V = mgh$. The heights of particles at temperature $T$ K follow Boltzmann’s distribution (1.345)

$$P(h) = P(0) e^{-mgh/kT} \quad (6.159)$$

in which $k = 1.3806504 \times 10^{-23}$ J/K = 8.617343 \times 10^{-5} $eV/K$ is his constant. The probability depends exponentially upon the mass $m$ and drops by a factor of $e$ with the scale height $S = kT/mg$, which can be a few kilometers for a small molecule.

**Example 6.24** (R-C Circuit) The capacitance $C$ of a capacitor is the charge $Q$ it holds (on each plate) divided by the applied voltage $V$, that is, $C = Q/V$. The current $I$ through the capacitor is the time derivative of the charge $I = \dot{Q} = CV$. The voltage across a resistor of $R$ $\Omega$ (Ohms) through which a current $I$ flows is $V = IR$ by Ohm’s law. So if a time-dependent voltage $V(t)$ is applied to a capacitor in series with a resistor, then $V(t) = Q/C + IR$. The current $I$ therefore obeys the first-order differential equation

$$\dot{I} + I/RC = \dot{V}/R \quad (6.160)$$

or (6.150) with $x \to t$, $y \to I$, $r \to 1/RC$, and $s \to \dot{V}/R$. Since $r$ is a constant, the integrating factor $\alpha(x) \to \alpha(t)$ is

$$\alpha(t) = \alpha(t_0) e^{(t-t_0)/RC}. \quad (6.161)$$

Our general solution (6.157) of linear first-order ODEs gives us the expres-
in which \( a_0 \neq 0 \) is the coefficient of the lowest power of \( x \) in \( y(x) \). Differentiating, we have

\[
y'(x) = \sum_{n=0}^{\infty} (r + n) a_n x^{r+n-1}
\]

(6.183)

and

\[
y''(x) = \sum_{n=0}^{\infty} (r + n)(r + n - 1) a_n x^{r+n-2}.
\]

(6.184)

When we substitute the three series (6.182–6.184) into our differential equation \( x^2 y'' + xp(x)y' + q(x)y = 0 \), we find

\[
\sum_{n=0}^{\infty} [(n+r)(n+r-1) + (n+r)p(x) + q(x)] a_n x^{n+r} = 0.
\]

(6.185)

If this equation is to be satisfied for all \( x \), then the coefficient of every power of \( x \) must vanish. The lowest power of \( x \) is \( x^r \), and it occurs when \( n = 0 \) with coefficient \( [r(r-1 + p(0)) + q(0)] a_0 \). Thus since \( a_0 \neq 0 \), we have

\[
r(r-1 + p(0)) + q(0) = 0.
\]

(6.186)

This quadratic \textbf{indicial equation} has two roots \( r_1 \) and \( r_2 \).

To analyze higher powers of \( x \), we introduce the notation

\[
p(x) = \sum_{j=0}^{\infty} p_j x^j \quad \text{and} \quad q(x) = \sum_{j=0}^{\infty} q_j x^j
\]

(6.187)

in which \( p_0 = p(0) \) and \( q_0 = q(0) \). The requirement (exercise 6.16) that the coefficient of \( x^{r+k} \) vanish gives us a \textbf{recurrence relation}

\[
a_k = - \left[ \frac{1}{(r+k)(r+k-1+p_0) + q_0} \right] \sum_{j=0}^{k-1} [(j+r)p_{k-j} + q_{k-j}] a_j
\]

(6.188)

that expresses \( a_k \) in terms of \( a_0, a_1, \ldots a_{k-1} \). When \( p(x) \) and \( q(x) \) are polynomials of low degree, these equations become much simpler.

**Example 6.28 (Sines and Cosines)** To apply Frobenius’s method the ODE \( y'' + \omega^2 y = 0 \), we first write it in the form \( x^2 y'' + xp(x)y' + q(x)y = 0 \) in which \( p(x) = 0 \) and \( q(x) = \omega^2 x^2 \). So both \( p(0) = p_0 = 0 \) and \( q(0) = q_0 = 0 \), and the indicial equation (6.186) is \( r(r-1) = 0 \) with roots \( r_1 = 0 \) and \( r_2 = 1 \).

We first set \( r = r_1 = 0 \). Since the \( p \)’s and \( q \)’s vanish except for \( q_0 = \omega^2 \), the recurrence relation (6.188) is \( a_k = -q_0 a_{k-2}/k(k-1) = -\omega^2 a_{k-2}/k(k-1) \). Thus \( a_2 = -\omega^2 a_0/2 \), and \( a_{2n} = (-1)^n \omega^{2n} a_0/(2n)! \). The recurrence relation (6.188) gives no information about \( a_1 \), so to find the simplest solution, we
set $a_1 = 0$. The recurrence relation $a_k = -\omega^2 a_{k-2}/k(k-1)$ then makes all the terms $a_{2n+1}$ of odd index vanish. Our solution for the first root $r_1 = 0$ then is

$$y(x) = \sum_{n=0}^{\infty} a_n x^n = a_0 \sum_{n=0}^{\infty} (-1)^n \frac{(\omega x)^{2n}}{(2n)!} = a_0 \cos \omega x.$$ \hspace{1cm} (6.189)

Similarly, the recurrence relation (6.188) for the second root $r_2 = 1$ is $a_k = -\omega^2 a_{k-2}/k(k+1)$, so that $a_{2n} = (-1)^n \omega^{2n} a_0/(2n+1)!$, and we again set all the terms of odd index equal to zero. Thus we have

$$y(x) = x \sum_{n=0}^{\infty} a_n x^n = \frac{a_0}{\omega} \sum_{n=0}^{\infty} (-1)^n \frac{(\omega x)^{2n+1}}{(2n+1)!} = \frac{a_0}{\omega} \sin \omega x$$ \hspace{1cm} (6.190)

as our solution for the second root $r_2 = 1$.

Frobenius’s method sometimes shows that solutions exist only when a parameter in the ODE assumes a special value called an eigenvalue.

**Example 6.29 (Legendre’s Equation)** If one rewrites Legendre’s equation $(1 - x^2)y'' - 2xy' + \lambda y = 0$ as $x^2 y'' + xpy' + qy = 0$, then one finds $p(x) = -2x^2/(1 - x^2)$ and $q(x) = x^2 \lambda/(1 - x^2)$, which are analytic but not polynomials. In this case, it is simpler to substitute the expansions (6.182–6.184) directly into Legendre’s equation $(1 - x^2)y'' - 2xy' + \lambda y = 0$. We then find

$$\sum_{n=0}^{\infty} [(n + r)(n + r - 1)(1 - x^2)x^{n+r-2} - 2(n + r)x^{n+r} + \lambda x^{n+r}] a_n = 0.$$ 

The coefficient of the lowest power of $x$ is $r(r-1)a_0$, and so the indicial equation is $r(r-1) = 0$. For $r = 0$, we shift the index $n$ on the term $n(n-1)x^{n-2}a_n$ to $n = j + 2$ and replace $n$ by $j$ in the other terms:

$$\sum_{j=0}^{\infty} \{ (j + 2)(j + 1) a_{j+2} - [j(j - 1) + 2j - \lambda] a_j \} x^j = 0.$$ \hspace{1cm} (6.191)

Since the coefficient of $x^j$ must vanish, we get the recursion relation

$$a_{j+2} = \frac{j(j + 1) - \lambda}{(j + 2)(j + 1)} a_j$$ \hspace{1cm} (6.192)

which for big $j$ says that $a_{j+2} \approx a_j$. Thus the series (6.182) does not converge for $|x| \geq 1$ unless $\lambda = j(j + 1)$ for some integer $j$ in which case the series (6.182) is a Legendre polynomial (chapter 8).\hspace{1cm} $\Box$
6.27 Self-Adjoint Differential Operators

If \( p(x) \) and \( q(x) \) are real, then the differential operator

\[
L = -\frac{d}{dx} \left( p(x) \frac{d}{dx} \right) + q(x) \tag{6.233}
\]

is formally self adjoint. Such operators are interesting because if we take any two functions \( u \) and \( v \) that are twice differentiable on an interval \([a, b]\) and integrate \( vL u \) twice by parts over the interval, we get

\[
(v, L u) = \int_a^b v L u \, dx = \int_a^b v \left[ -(pu')' + qu \right] \, dx
= \int_a^b [pu'v' + uv] \, dx - [vpu']_a^b
= \int_a^b [-pu'v' + quv] \, dx + [pW(u, v)]_a^b
= \int_a^b (L v) u \, dx + [p(uv' - vv')]'_a^b \tag{6.234}
\]

which is Green’s formula

\[
\int_a^b (vL u - u L v) \, dx = [pW(u, v)]'_a^b = [pW(u, v)]_a^b \tag{6.235}
\]

(George Green, 1793–1841). Its differential form is Lagrange’s identity

\[
vL u - u L v = [pW(u, v)]' \tag{6.236}
\]

(Joseph-Louis Lagrange, 1736–1813). Thus if the twice-differentiable functions \( u \) and \( v \) satisfy boundary conditions at \( x = a \) and \( x = b \) that make the boundary term (6.235) vanish

\[
[p(uv' - vv')]_a^b = [pW(u, v)]_a^b = 0 \tag{6.237}
\]

then the real differential operator \( L \) is symmetric

\[
(v, L u) = \int_a^b v L u \, dx = \int_a^b u L v \, dx = (u, L v). \tag{6.238}
\]

A real linear operator \( A \) that acts in a real vector space and satisfies the analogous relation (1.161)

\[
(g, A f) = (f, A g) \tag{6.239}
\]
Differential Equations

We see that the real part of $\psi_i$ satisfies them, and by subtracting them, we see that the imaginary part of $\psi_i$ also satisfies them. So it might seem that $\psi_i = u_i + iv_i$ is made of two real eigenfunctions with the same eigenvalue.

But each eigenfunction $u_i$ in the domain $D$ satisfies two homogeneous boundary conditions as well as its second-order differential equation

$$-(p u_i')' + q u_i = \lambda_i \rho u_i$$

and so $u_i$ is the unique solution in $D$ to this equation. There can be no other eigenfunction in $D$ with the same eigenvalue. In a regular Sturm-Liouville system, there is no degeneracy. All the eigenfunctions $u_i$ are orthogonal and can be normalized on the interval $[a, b]$ with weight function $\rho(x)$

$$\int_a^b u_i^* \rho u_i \, dx = \delta_{ij}.$$  

They may be taken to be real.

It is true that the eigenfunctions of a second-order differential equation come in pairs because one can use Wronski’s formula (6.268)

$$y_2(x) = y_1(x) \int^x \frac{dx'}{p(x') y_1^2(x')}$$

(6.327)

to find a linearly independent second solution with the same eigenvalue. But the second solutions don’t obey the boundary conditions of the domain. Bessel functions of the second kind, for example, are infinite at the origin.

A set of eigenfunctions $u_i$ is complete in the mean in a space $S$ of functions if every function $f \in S$ can be represented as a series

$$f(x) = \sum_{i=1}^{\infty} a_i u_i(x)$$

(6.328)

called a Fourier series) that converges in the mean, that is

$$\lim_{N \to \infty} \int_a^b \left| f(x) - \sum_{i=1}^{N} a_i u_i(x) \right|^2 \rho(x) \, dx = 0.$$  

(6.329)

The natural space $S$ is the space $L_2(a, b)$ of all functions $f$ that are square-integrable on the interval $[a, b]$

$$\int_a^b |f(x)|^2 \rho(x) \, dx < \infty.$$  

(6.330)

The orthonormal eigenfunctions of every regular Sturm-Liouville system on an interval $[a, b]$ are complete in the mean in $L_2(a, b)$. The completeness of
we write \( G(x - y) \) in terms of the complete set of eigenfunctions \( u_k \) as

\[
G(x - y) = \sum_{k=1}^{\infty} \frac{u_k(x)u_k(y)}{\lambda_k}
\]  

(6.410)

so that the action \( Lu_k = \lambda_k \rho u_k \) turns \( G \) into

\[
L G(x - y) = \sum_{k=1}^{\infty} \frac{Lu_k(x)u_k(y)}{\lambda_k} = \sum_{k=1}^{\infty} \rho(x) u_k(x) u_k(y) = \delta(x - y)
\]

(6.411)

our \( \alpha = 1 \) series expansion (6.374) of the delta function.

### 6.39 Green’s Functions in One Dimension

In one dimension, we can explicitly solve the inhomogeneous ordinary differential equation \( L f(x) = g(x) \) in which

\[
L = -\frac{d}{dx} \left( p(x) \frac{d}{dx} \right) + q(x)
\]

(6.412)

is formally self adjoint. We’ll build a Green’s function from two solutions \( u \) and \( v \) of the homogeneous equation \( L u(x) = L v(x) = 0 \) as

\[
G(x, y) = \frac{1}{A} [\theta(x - y)u(y)v(x) + \theta(y - x)u(x)v(y)]
\]

(6.413)

in which \( \theta(x) = (x + |x|)/(2|x|) \) is the Heaviside step function (Oliver Heaviside 1850–1925), and \( A \) is a constant which we’ll presently identify. We’ll show that the expression

\[
f(x) = \int_a^b G(x, y) g(y) \, dy = \frac{v(x)}{A} \int_a^x u(y) g(y) \, dy + \frac{u(x)}{A} \int_x^b v(y) g(y) \, dy
\]

solves our inhomogeneous equation. Differentiating, we find after a cancellation

\[
f'(x) = \frac{v'(x)}{A} \int_a^x u(y) g(y) \, dy + \frac{u'(x)}{A} \int_x^b v(y) g(y) \, dy.
\]

(6.414)
Differentiating again, we have

\[
f''(x) = \frac{v''(x)}{A} \int_a^x u(y) \, g(y) \, dy + \frac{u''(x)}{A} \int_x^b v(y) \, g(y) \, dy
\]
\[
+ \frac{v'(x)u(x)g(x)}{A} - \frac{u'(x)v(x)g(x)}{A}
\]
\[
= \frac{v''(x)}{A} \int_a^x u(y) \, g(y) \, dy + \frac{u''(x)}{A} \int_x^b v(y) \, g(y) \, dy
\]
\[
+ \frac{W(x)}{A} g(x) \quad (6.415)
\]

in which \( W(x) \) is the wronskian \( W(x) = u(x)v'(x) - u'(x)v(x) \). The result (6.266) for the wronskian of two linearly independent solutions of a self-adjoint homogeneous ODE gives us \( W(x) = W(x_0) p(x_0) / p(x) \). We set the constant \( A = -W(x_0) p(x_0) \) so that the last term in (6.415) is \( -g(x)/p(x) \).

It follows that

\[
Lf(x) = \left[ \frac{Lv(x)}{A} \right] \int_a^x u(y) \, g(y) \, dy + \left[ \frac{Lu(x)}{A} \right] \int_x^b v(y) \, g(y) \, dy + g(x) = g(x).
\]

But \( Lu(x) = Lv(x) = 0 \), so we see that \( f \) satisfies our inhomogeneous equation \( Lf(x) = g(x) \).

**6.40 Nonlinear Differential Equations**

The field of nonlinear differential equations is too vast to cover here, but we may hint at some of its features by considering some examples from cosmology and particle physics.

The Friedmann equations of general relativity (11.410 & 11.412) for the scale factor \( a(t) \) of a homogeneous, isotropic universe are

\[
\frac{\ddot{a}}{a} = -\frac{4\pi G}{3} (\rho + 3p) \quad \text{and} \quad \left( \frac{\dot{a}}{a} \right)^2 = \frac{8\pi G}{3} \rho - \frac{k}{a^2} \quad (6.417)
\]

in which \( k \) respectively is 1, 0, and \(-1\) for closed, flat, and open geometries. (The scale factor \( a(t) \) tells how much space has expanded or contracted by the time \( t \).) These equations become more tractable when the energy density \( \rho \) is due to a single constituent whose pressure \( p \) is related to it by an equation of state \( p = \omega \rho \). Conservation of energy \( \dot{\rho} = -3(\rho + p)/a \) (11.426–11.431) then ensures (exercise 6.30) that the product \( \rho a^{3(1+\omega)} \) is...
Figure 6.4 The field $\phi(x)$ of the soliton (6.435) at rest ($v = 0$) at position $x_0 = 0$ for $\lambda = 1 = \phi_0$. The energy density of the field vanishes when $\phi = \pm \phi_0 = \pm 1$. The energy of this soliton is concentrated at $x = 0$.

in which $C$ is a constant of integration.

The equations of particle physics are nonlinear. Physicists usually use perturbation theory to cope with the nonlinearities. But occasionally they focus on the nonlinearities and treat the fields classically or semi-classically. To keep things relatively simple, we’ll work in a space-time of only two dimensions and consider a model field theory described by the action density

$$
\mathcal{L} = \frac{1}{2} \left( \dot{\phi}^2 - \phi'^2 \right) - V(\phi) \tag{6.427}
$$

in which $V$ is a simple function of the field $\phi$. Lagrange’s equation for this theory is

$$
\ddot{\phi} - \phi'' = -\frac{dV}{d\phi}. \tag{6.428}
$$
We can convert this partial differential equation to an ordinary one by making the field \( \phi \) depend only upon the combination \( u = x - vt \) rather than upon both \( x \) and \( t \). We then have \( \dot{\phi} = -v \phi_u \). With this restriction to traveling-wave solutions, Lagrange’s equation reduces to

\[
(1 - v^2) \phi_{uu} = \frac{dV}{d\phi}. \tag{6.429}
\]

We multiply both sides of this equation by \( \phi_u \)

\[
(1 - v^2) \phi_u \phi_{uu} = \frac{dV}{d\phi} \phi_u \tag{6.430}
\]

and integrate both sides to get \( (1 - v^2) \frac{1}{2} \phi_u^2 = V + E \) in which \( E \) is a constant of integration

\[
E = \frac{1}{2} (1 - v^2) \phi_u^2 - V(\phi). \tag{6.431}
\]

We can convert (exercise 6.37) this equation into a problem of integration

\[
u - u_0 = \pm \int \frac{\sqrt{1 - v^2} \phi}{\sqrt{2(E + V(\phi))}} d\phi. \tag{6.432}
\]

By inverting the resulting equation relating \( u \) to \( \phi \), we may find the soliton solution \( \phi(u - u_0) \), which is a lump of energy traveling with speed \( v \).

**Example 6.48 (Soliton of the \( \phi^4 \) Theory)** To simplify the integration (6.432), we take as the action density

\[
\mathcal{L} = \frac{1}{2} \left( \dot{\phi}^2 - \phi^2 \right) - \left[ \frac{\lambda^2}{2} (\phi^2 - \phi_0^2)^2 - E \right]. \tag{6.433}
\]

Our formal solution (6.432) gives

\[
u - u_0 = \pm \int \frac{\sqrt{1 - v^2}}{\lambda (\phi^2 - \phi_0^2)} d\phi = \pm \frac{\sqrt{1 - v^2}}{\lambda \phi_0} \tanh^{-1}(\phi/\phi_0) \tag{6.434}
\]

or

\[
\phi(x - vt) = \mp \phi_0 \tanh \left[ \lambda \phi_0 \frac{x - x_0 - v(t - t_0)}{\sqrt{1 - v^2}} \right] \tag{6.435}
\]

which is a soliton (or an antisoliton) at \( x_0 + v(t - t_0) \). A unit soliton at rest is plotted in Fig. 6.4. Its energy is concentrated at \( x = 0 \) where \( |\phi^2 - \phi_0^2| \) is maximal.
In rectangular coordinates, the curl of a curl is by definition (6.40)

\[(\nabla \times (\nabla \times \mathbf{E}))_i = \sum_{j,k=1}^{3} \epsilon_{ijk} \partial_j (\nabla \times \mathbf{E})_k = \sum_{j,k,l,m=1}^{3} \epsilon_{ijk} \epsilon_{klm} \partial_l E_m.\]

Use Levi-Civita’s identity (1.449) to show that

\[\nabla \times (\nabla \times \mathbf{E}) = \nabla(\nabla \cdot \mathbf{E}) - \Delta \mathbf{E}.\]

This formula defines \(\Delta \mathbf{E}\) in any system of orthogonal coordinates.

Show that since the Bessel function \(J_n(x)\) satisfies Bessel’s equation (6.48), the function \(P_n(\rho) = J_n(k\rho)\) satisfies (6.47).

Show that (6.58) implies that \(R_{k,\ell}(r) = j_{\ell}(k r)\) satisfies (6.57).

Use (6.56, 6.57), and \(\Phi''_m = -m^2 \Phi_m\) to show in detail that the product \(f(r, \theta, \phi) = R_{k,\ell}(r) \Theta_{\ell,m}(\theta) \Phi_m(\phi)\) satisfies \(-\Delta f = k^2 f\).

Replacing Helmholtz’s \(k^2\) by \(2m(E - V(r))/\hbar^2\), we get Schrödinger’s equation

\[-(h^2/2m)\Delta \psi(r, \theta, \phi) + V(r) \psi(r, \theta, \phi) = E \psi(r, \theta, \phi).\]

Let \(\psi(r, \theta, \phi) = R_{n,\ell}(r) \Theta_{\ell,m}(\theta)e^{im\phi}\) in which \(\Theta_{\ell,m}\) satisfies (6.56) and show that the radial function \(R_{n,\ell}\) must obey

\[-(r^2 R''_{n,\ell})/r^2 + [(\ell(\ell + 1))/r^2 + 2mV/r^2] R_{n,\ell} = 2mE_{n,\ell} R_{n,\ell}/\hbar^2.\]

Use the empty-space Maxwell’s equations \(\nabla \cdot \mathbf{B} = 0, \nabla \times \mathbf{E} + \mathbf{B} = 0,\)

\(\nabla \cdot \mathbf{E} = 0,\) and \(\nabla \times \mathbf{B} = \mathbf{E}/c^2 = 0\) and the formula (6.437) to show that in vacuum \(\Delta \mathbf{E} = \mathbf{E}/c^2\) and \(\Delta \mathbf{B} = \mathbf{B}/c^2\).

Argue from symmetry and anti-symmetry that \([\gamma^a, \gamma^b] \partial_a \partial_b = 0\) in which the sums over \(a\) and \(b\) run from 0 to 3.

Suppose a voltage \(V(t) = V \sin(\omega t)\) is applied to a resistor of \(R\) (\(\Omega\)) in series with a capacitor of capacitance \(C\) (\(F\)). If the current through the circuit at time \(t = 0\) is zero, what is the current at time \(t\)?

(a) Is \((1 + x^2 + y^2)^{-3/2} \left[(1 + y^2)y\, dx + (1 + x^2)x\, dy\right] = 0\) exact? (b) Find its general integral and solution \(y(x)\). Use section 6.11.

(a) Separate the variables of the ODE \((1 + y^2)y\, dx + (1 + x^2)x\, dy = 0.\)

(b) Find its general integral and solution \(y(x)\).

Find the general solution to the differential equation \(y' + y/x = c/x\).

Find the general solution to the differential equation \(y' + xy = ce^{-x^2/2}\).
James Bernoulli studied ODEs of the form $y' + p y = q y^n$ in which $p$ and $q$ are functions of $x$. Division by $y^n$ and the substitution $v = y^{1-n}$ gives us the equation $v' + (1-n)p v = (1-n) q$ which is soluble as shown in section (6.16). Use this method to solve the ODE $y' - y/2x = 5x^2 y^5$.

Integrate the ODE $(xy + 1) \, dx + 2x^2 (2xy - 1) \, dy = 0$. Hint: Use the variable $v(x) = xy(x)$ instead of $y(x)$.

Show that the points $x = \pm 1$ and $\infty$ are regular singular points of Legendre’s equation (6.181).

Use the vanishing of the coefficient of every power of $x$ in (6.185) and the notation (6.187) to derive the recurrence relation (6.188).

In example 6.29, derive the recursion relation for $r = 1$ and discuss the resulting eigenvalue equation.

In example 6.29, show that the solutions associated with the roots $r = 0$ and $r = 1$ are the same.

For a hydrogen atom, we set $V(r) = -e^2/4\pi\epsilon_0 r \equiv -q^2/r$ in (6.439) and get $(r^2 R_{n,\ell}')' = - \left[ (2m/\hbar^2) \left( E_{n,\ell} + Zq^2/r \right) \right] r^2 - \ell(\ell+1) R_{n,\ell} = 0$.

At tiny $r$, $(r^2 R_{n,\ell}')' \equiv \ell(\ell+1)R_{n,\ell}$ and $R_{n,\ell}(r) \sim r^{\ell}$.

Set $R_{n,\ell}(r) = r^\ell \exp(-\sqrt{-2mE_{n,\ell}\hbar}/\hbar) P_{n,\ell}(r)$ and apply the method of Frobenius to find the values of $E_{n,\ell}$ for which $R_{n,\ell}$ is suitably normalizable.

Show that as long as the matrix $\mathcal{Y}_{kj} = y^{(\ell_j)}(x_j)$ is nonsingular, the $n$ boundary conditions

$$b_j = y^{(\ell_j)}(x_j) = \sum_{k=1}^{n} c_k \, y^{(\ell_j)}(x_j) \quad \text{(6.440)}$$

determine the $n$ coefficients $c_k$ of the expansion (6.222) to be

$$C^T = B^T \mathcal{Y}^{-1} \quad \text{or} \quad C_k = \sum_{j=1}^{n} b_j \mathcal{Y}_{jk}^{-1}. \quad \text{(6.441)}$$

Show that if the real and imaginary parts $u_1$, $u_2$, $v_1$, and $v_2$ of $\psi$ and $\chi$ satisfy boundary conditions at $x = a$ and $x = b$ that make the boundary term (6.235) vanish, then its complex analog (6.242) also vanishes.

Show that if the real and imaginary parts $u_1$, $u_2$, $v_1$, and $v_2$ of $\psi$ and $\chi$ satisfy boundary conditions at $x = a$ and $x = b$ that make the boundary term (6.235) vanish, and if the differential operator $L$ is real and self adjoint, then (6.238) implies (6.243).

Show that if $D$ is the set of all twice-differentiable functions $u(x)$ on $[a, b]$ that satisfy Dirichlet’s boundary conditions (6.245) and if the function $p(x)$ is continuous and positive on $[a, b]$, then the adjoint set
The generating function \( g(t, x) \) is even under the reflection of both independent variables, so

\[
g(t, x) = \sum_{n=0}^{\infty} t^n P_n(x) = \sum_{n=0}^{\infty} (-t)^n P_n(-x) = g(-t, -x) \quad (8.51)
\]

which implies that

\[
P_n(-x) = (-1)^n P_n(x) \quad \text{whence} \quad P_{2n+1}(0) = 0. \quad (8.52)
\]

With more effort, one can show that

\[
P_{2n}(0) = (-1)^n \frac{(2n - 1)!!}{(2n)!!} \quad \text{and that} \quad |P_n(x)| \leq 1. \quad (8.53)
\]

### 8.7 Schlaefli’s Integral

Schlaefli used Cauchy’s integral formula (5.36) and Rodrigues’s formula

\[
P_n(x) = \frac{1}{2^n n!} \left( \frac{d}{dx} \right)^n (x^2 - 1)^n \quad (8.54)
\]

to express \( P_n(z) \) as a counterclockwise contour integral around the point \( z \)

\[
P_n(z) = \frac{1}{2^n 2\pi i} \oint \frac{(z^2 - 1)^n}{(z' - z)^{n+1}} dz'. \quad (8.55)
\]

### 8.8 Orthogonal Polynomials

Rodrigues’s formula (8.8) generates other families of orthogonal polynomials. The \( n \)-th order polynomials \( R_n \) in which the \( e_n \) are constants

\[
R_n(x) = \frac{1}{e_n w(x)} \frac{d^n}{dx^n} [w(x) Q^n(x)] \quad (8.56)
\]

are orthogonal on the interval from \( a \) to \( b \) with weight function \( w(x) \)

\[
\int_a^b R_n(x) R_k(x) w(x) dx = N_n \delta_{nk} \quad (8.57)
\]

as long as \( Q(x) \) vanishes at \( a \) and \( b \) (exercise 8.8)

\[
Q(a) = Q(b) = 0. \quad (8.58)
\]
Instead, we define \( P_{\ell,m} \) in terms of the \( m \)th derivative \( P^{(m)}_{\ell}(x) \) as

\[
P_{\ell,m}(x) \equiv (1 - x^2)^{m/2} P^{(m)}_{\ell}(x)
\]

and compute the derivatives

\[
P^{(m+1)}_{\ell}(x) = \left( P'_{\ell,m} + \frac{m x P_{\ell,m}}{1 - x^2} \right) (1 - x^2)^{-m/2}
\]

\[
P^{(m+2)}_{\ell}(x) = \left[ P''_{\ell,m} + \frac{2m x P'_{\ell,m}}{1 - x^2} + \frac{m P_{\ell,m}}{1 - x^2} + \frac{m(m+2)x^2 P_{\ell,m}}{(1 - x^2)^2} \right] (1 - x^2)^{-m/2}.
\]

When we put these three expressions in equation (8.94), we get the desired ODE (8.92).

Thus the associated Legendre functions are

\[
P_{\ell,m}(x) = (1 - x^2)^{m/2} P^{(m)}_{\ell}(x) = (1 - x^2)^{m/2} \frac{d^m}{dx^m} P_{\ell}(x)
\]

They are simple polynomials in \( x = \cos \theta \) and \( \sqrt{1 - x^2} = \sin \theta \)

\[
P_{\ell,m}(\cos \theta) = \sin^m \theta \frac{d^m}{d \cos^m \theta} P_{\ell}(\cos \theta).
\]

It follows from Rodrigues’s formula (8.8) for the Legendre polynomial \( P_{\ell}(x) \) that \( P_{\ell,m}(x) \) is given by the similar formula

\[
P_{\ell,m}(x) = \frac{(1 - x^2)^{m/2}}{2^\ell \ell!} \frac{d^{\ell+m}}{dx^{\ell+m}} (x^2 - 1)^\ell
\]

which tells us that under parity \( P^m_{\ell}(x) \) changes by \((-1)^{\ell+m}\)

\[
P_{\ell,m}(-x) = (-1)^{\ell+m} P_{\ell,m}(x).
\]

Rodrigues’s formula (8.101) for the associated Legendre function makes sense as long as \( \ell + m \geq 0 \). This last condition is the requirement in quantum mechanics that \( m \) not be less than \(-\ell\). And if \( m \) exceeds \( \ell \), then \( P_{\ell,m}(x) \) is given by more than \( 2\ell \) derivatives of a polynomial of degree \( 2\ell \); so \( P_{\ell,m}(x) = 0 \) if \( m > \ell \). This last condition is the requirement in quantum mechanics that \( m \) not be greater than \( \ell \). So we have

\[
-\ell \leq m \leq \ell.
\]

One may show that

\[
P_{\ell,-m}(x) = (-1)^m \frac{(\ell - m)!}{(\ell + m)!} P_{\ell,m}(x).
\]
Figure 8.2 CMB temperature fluctuations over the celestial sphere as measured by the Planck satellite. The average temperature is 2.7255 K. White regions are warmer and black ones colder by about 0.0005 degrees. © ESA and the Planck Collaboration.

\[ P_\ell(\mathbf{n} \cdot \mathbf{n}') = \frac{4\pi}{2\ell + 1} \sum_{m=-\ell}^{\ell} Y_{\ell,m}(\theta, \phi) Y_{\ell,m}^*(\theta', \phi') \]

\[ \text{(8.121)} \]

**Example 8.8** (CMB Radiation) Instruments on the Wilkinson Microwave Anisotropy Probe (WMAP) and Planck satellites in orbit at the Lagrange point L2 (in the Earth’s shadow, 1.5×10^6 km farther from the Sun) have measured the temperature \( T(\theta, \phi) \) of the cosmic microwave background (CMB) radiation as a function of the polar angles \( \theta \) and \( \phi \) in the sky as shown in Fig. 8.2. This radiation is photons last scattered when the visible universe became transparent at an age of 380,000 years and a temperature (3,000 K) cool enough for hydrogen atoms to be stable. This initial transparency is usually (and inexplicably) called recombination.

Since the spherical harmonics \( Y_{\ell,m}(\theta, \phi) \) are complete on the sphere, we
Figure 8.3 The power spectrum $D_\ell = \ell (\ell + 1) C_\ell / 2\pi$ of the CMB temperature fluctuations in $\mu$K$^2$ as measured by the Planck Collaboration (arXiv:1303.5062) is plotted against the angular size and the multipole moment $\ell$. The solid curve is the $\Lambda$CDM prediction.

can expand the temperature as

$$T(\theta, \phi) = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} a_{\ell,m} Y_{\ell,m}(\theta, \phi)$$  \hspace{1cm} (8.122)

in which the coefficients are by (8.117)

$$a_{\ell,m} = \int_0^{2\pi} d\phi \int_0^\pi \sin \theta \, d\theta \, Y_{\ell,m}^*(\theta, \phi) T(\theta, \phi).$$  \hspace{1cm} (8.123)

The average temperature $\overline{T}$ contributes only to $a_{0,0} = \overline{T} = 2.7255$ K. The other coefficients describe the difference $\Delta T(\theta, \phi) = T(\theta, \phi) - \overline{T}$. The angular power spectrum is

$$C_\ell = \frac{1}{2\ell + 1} \sum_{m=-\ell}^{\ell} |a_{\ell,m}|^2.$$  \hspace{1cm} (8.124)

If we let the unit vector $\hat{n}$ point in the direction $\theta, \phi$ and use the addition theorem (8.121), then we can write the angular power spectrum as

$$C_\ell = \frac{1}{4\pi} \int d^2\hat{n} \int d^2\hat{n}' P_\ell(\hat{n} \cdot \hat{n}') T(\hat{n}) T(\hat{n}').$$  \hspace{1cm} (8.125)
In Fig. 8.3, the measured values (arXiv:1303.5062) of the power spectrum \( D_\ell = \ell(\ell + 1) C_\ell / 2\pi \) are plotted against \( \ell \) for \( 1 < \ell < 1300 \) with the angles and distances decreasing with \( \ell \). The power spectrum is a snapshot at the moment of initial transparency of the temperature distribution of the plasma of photons, electrons, and nuclei undergoing acoustic oscillations. In these oscillations, gravity opposes radiation pressure, and \( |\Delta T(\theta, \phi)| \) is maximal both when the oscillations are most compressed and when they are most rarefied. Regions that gravity has squeezed to maximum compression at transparency form the first and highest peak. Regions that have bounced off their first maximal compression and that radiation pressure has expanded to minimum density at transparency form the second peak. Those at their second maximum compression at transparency form the third peak, and so forth.

The solid curve is the prediction of a model with inflation, cold dark matter, and a cosmological constant \( \Lambda \). In this model, the age of the visible universe is 13.817 Gyr; the Hubble constant is \( H_0 = 67.3 \text{ km/sMpc} \); the energy density of the universe is enough to make the universe flat; and the fractions of the energy density due to baryons, dark matter, and dark energy are 4.9\%, 26.6\%, and 68.5\% (Edwin Hubble 1889–1953).

Much is known about Legendre functions. The books *A Course of Modern Analysis* (Whittaker and Watson, 1927, chap. XV) and *Methods of Mathematical Physics* (Courant and Hilbert, 1955) are outstanding.

**Exercises**

8.1 Use conditions (8.6) and (8.7) to find \( P_0(x) \) and \( P_1(x) \).

8.2 Using the Gram-Schmidt method (section 1.10) to turn the functions \( x^n \) into a set of functions \( L_n(x) \) that are orthonormal on the interval \([-1, 1]\) with inner product (8.2), find \( L_n(x) \) for \( n = 0, 1, 2, \) and 3. Isn’t Rodrigues’s formula (8.8) easier to use?

8.3 Derive the conditions (8.6–8.7) on the coefficients \( a_k \) of the Legendre polynomial \( P_n(x) = a_0 + a_1 x + \cdots + a_n x^n \).

8.4 Use equations (8.6–8.7) to find \( P_3(x) \) and \( P_4(x) \).

8.5 In superscript notation (6.19), Leibniz’s rule (4.46) for derivatives of products \( uv \) of functions is

\[
(uv)^{(n)} = \sum_{k=0}^{n} \binom{n}{k} u^{(n-k)} v^{(k)}. \tag{8.126}
\]
These integrals (exercise 9.8) give $J_n(0) = 0$ for $n \neq 0$, and $J_0(0) = 1$.

By differentiating the generating function (9.5) with respect to $u$ and identifying the coefficients of powers of $u$, one finds the recursion relation

$$J_{n-1}(z) + J_{n+1}(z) = \frac{2n}{z} J_n(z).$$

(9.8)

Similar reasoning after taking the $z$ derivative gives (exercise 9.10)

$$J_{n-1}(z) - J_{n+1}(z) = 2 J'_n(z).$$

(9.9)

By using the gamma function (section 5.12), one may extend Bessel’s equation (9.4) and its solutions $J_n(z)$ to non-integral values of $n$

$$J_\nu(z) = \left(\frac{z}{2}\right)^\nu \sum_{m=0}^{\infty} \frac{(-1)^m}{m! \Gamma(m + \nu + 1)} \left(\frac{z}{2}\right)^{2m}.$$

(9.10)

Letting $z = ax$ in (9.4), we arrive (exercise 9.11) at the self-adjoint form (6.307) of Bessel’s equation

$$-\frac{d}{dx} \left( x \frac{d}{dx} J_n(ax) \right) + \frac{n^2}{x} J_n(ax) = a^2 x J_n(ax).$$

(9.11)

In the notation of equation (6.287), $p(x) = x$, $a^2$ is an eigenvalue, and $\rho(x) = x$ is a weight function. To have a self-adjoint system (section 6.28) on an interval $[0, b]$, we need the boundary condition (6.247)

$$0 = \left[ p(J_n v' - J'_n v) \right]_0^b = \left[ x(J_n v' - J'_n v) \right]_0^b$$

(9.12)

for all functions $v(x)$ in the domain $D$ of the system. Since $p(x) = x$, $J_0(0) = 1$, and $J_n(0) = 0$ for integers $n > 0$, the terms in this boundary condition vanish at $x = 0$ as long as the domain consists of functions $v(x)$ that are twice differentiable on the interval $[0, b]$. To make these terms vanish at $x = b$, we require that $J_n(ab) = 0$ and that $v(b) = 0$. So $ab$ must be a zero $z_{n,m}$ of $J_n(z)$, that is $J_n(ab) = J_n(z_{n,m}) = 0$. With $a = z_{n,m}/b$, Bessel’s equation (9.11) is

$$-\frac{d}{dx} \left( x \frac{d}{dx} J_n(z_{n,m} x/b) \right) + \frac{n^2}{x} J_n(z_{n,m} x/b) = \frac{z_{n,m}^2}{b^2} x J_n(z_{n,m} x/b).$$

(9.13)

For fixed $n$, the eigenvalue $a^2 = z_{n,m}^2/b^2$ is different for each positive integer $m$. Moreover as $m \to \infty$, the zeros $z_{n,m}$ of $J_n(x)$ rise as $m\pi$ as one might expect since the leading term of the asymptotic form (9.3) of $J_n(x)$ is proportional to $\cos(x - n\pi/2 - \pi/4)$ which has zeros at $m\pi + (n+1)\pi/2 + \pi/4$. It follows that the eigenvalues $a^2 \approx (m\pi)^2/b^2$ increase without limit as $m \to \infty$ in accordance with the general result of section 6.34. It follows then from
When $\alpha = 0$, the Helmholtz equation reduces to the Laplace equation $\triangle V = 0$ of electrostatics which the simpler functions

$$V_{k,n}(\rho, \phi, z) = J_n(k\rho)e^{\pm in\phi}e^{\pm kz} \quad \text{and} \quad V_{k,n}(\rho, \phi, z) = J_n(ik\rho)e^{\pm in\phi}e^{\pm ikz}$$

satisfy.

The product $i^{-\nu} J_\nu(ik\rho)$ is real and is known as the modified Bessel function

$$I_\nu(k\rho) \equiv i^{-\nu} J_\nu(ik\rho).$$

It occurs in various solutions of the diffusion equation $D\triangle \phi = \dot{\phi}$. The function $V(\rho, \phi, z) = B(\rho)\Phi(\phi)Z(z)$ satisfies

$$\triangle V = \frac{1}{\rho} \left[ (\rho V_\rho)_\rho + \frac{1}{\rho} V_{\rho\phi} + \rho V_{zz} \right] = \alpha^2 V$$

if $B(\rho)$ obeys Bessel’s equation

$$\rho\frac{d}{d\rho} \left( \rho \frac{dB}{d\rho} \right) - ((\alpha^2 - k^2)\rho^2 + n^2) B = 0$$

and $\Phi$ and $Z$ respectively satisfy

$$-\frac{d^2\Phi}{d\phi^2} = n^2\Phi(\phi) \quad \text{and} \quad \frac{d^2Z}{dz^2} = k^2Z(z)$$

or if $B(\rho)$ obeys the Bessel equation

$$\rho\frac{d}{d\rho} \left( \rho \frac{dB}{d\rho} \right) - ((\alpha^2 + k^2)\rho^2 + n^2) B = 0$$

and $\Phi$ and $Z$ satisfy

$$-\frac{d^2\Phi}{d\phi^2} = n^2\Phi(\phi) \quad \text{and} \quad \frac{d^2Z}{dz^2} = -k^2Z(z).$$

In the first case (9.37 & 9.38), the solution $V$ is

$$V_{k,n}(\rho, \phi, z) = I_n(\sqrt{\alpha^2 - k^2} \rho)e^{\pm in\phi}e^{\pm kz}$$

while in the second case (9.39 & 9.40), it is

$$V_{k,n}(\rho, \phi, z) = I_n(\sqrt{\alpha^2 + k^2} \rho)e^{\pm in\phi}e^{\pm ikz}.$$

In both cases, $n$ must be an integer if the solution is to be single valued on the full range of $\phi$ from 0 to $2\pi$. 
Since \( p = (\epsilon_w - \epsilon_t)/(\epsilon_w + \epsilon_t) > 0 \), the principal image charge \( pq \) at \((0, 0, -h)\) has the same sign as the charge \( q \) and so contributes a positive term proportional to \( pq^2 \) to the energy. So a lipid slab repels a nearby charge in water no matter what the sign of the charge.

A cell membrane is a phospholipid bilayer. The lipids avoid water and form a 4-nm-thick layer that lies between two 0.5-nm layers of phosphate groups which are electric dipoles. These electric dipoles cause the cell membrane to weakly attract ions that are within 0.5 nm of the membrane.

**Example 9.3** (Cylindrical Wave Guides) An electromagnetic wave traveling in the \( z \)-direction down a cylindrical wave guide looks like

\[
E e^{i\phi} e^{i(kz-\omega t)} \quad \text{and} \quad B e^{i\phi} e^{i(kz-\omega t)}
\]

in which \( E \) and \( B \) depend upon \( \rho \)

\[
E = E_\rho \hat{\rho} + E_\phi \hat{\phi} + E_z \hat{z} \quad \text{and} \quad B = B_\rho \hat{\rho} + B_\phi \hat{\phi} + B_z \hat{z}
\]

in cylindrical coordinates (section 6.4). If the wave guide is an evacuated, perfectly conducting cylinder of radius \( r \), then on the surface of the wave guide the parallel components of \( E \) and the normal component of \( B \) must vanish which leads to the boundary conditions

\[
E_z(r) = 0, \quad E_\phi(r) = 0, \quad \text{and} \quad B_\rho(r) = 0.
\]

Since the \( E \) and \( B \) fields have subscripts, we will use commas to denote derivatives as in \( \partial(\rho E_\phi)/\partial \rho \equiv (\rho E_\phi)_\rho \) and so forth. In this notation, the vacuum forms \( \nabla \times E = -B \) and \( \nabla \times B = \hat{E}/c^2 \) of the Faraday and Maxwell-Ampère laws give us (exercise 9.14) the field equations

\[
E_{z,\phi}/\rho - ikE_\phi = i\omega B_\rho \quad \text{\( ikE_\rho - E_{z,\rho} = i\omega B_\phi \)} \quad \text{\( ikB_\rho - B_{z,\rho} = -i\omega E_\phi/c^2 \)}
\]

\[
[(\rho E_\phi)_\rho - inE_\rho]/\rho = i\omega B_z \quad [(\rho B_\phi)_\rho - i\omega B_\rho]/\rho = -i\omega E_z/c^2.
\]

Solving them for the \( \rho \) and \( \phi \) components of \( E \) and \( B \) in terms of their \( z \) components (exercise 9.15), we find

\[
E_\rho = \frac{-ikE_{z,\rho} + n\omega B_z/\rho}{k^2 - \omega^2/c^2} \quad E_\phi = \frac{nkE_z/\rho + i\omega B_{z,\rho}}{k^2 - \omega^2/c^2}
\]

\[
B_\rho = \frac{-ikB_{z,\rho} - n\omega E_z/c^2\rho}{k^2 - \omega^2/c^2} \quad B_\phi = \frac{nkB_z/\rho - i\omega E_{z,\rho}/c^2}{k^2 - \omega^2/c^2}.
\]

The fields \( E_z \) and \( B_z \) obey the wave equations (11.91, exercise 6.6)

\[
-\nabla E_z = -\nabla \phi/c^2 = \omega^2 E_z/c^2 \quad \text{and} \quad -\nabla B_z = -\nabla B_\phi/c^2 = \omega^2 B_z/c^2.
\]
9.2 Spherical Bessel functions of the first kind

at \( \rho = r \) as well as the separable wave equations (9.57). The frequencies of the resonant TE modes then are \( \omega_{n,m,\ell} = c \sqrt{z_{n,m}^2/r^2 + \pi^2 \ell^2/h^2} \).

The TM modes are \( B_z = 0 \) and

\[
E_z = J_n(z_{n,m}, \rho/r) e^{im\phi} \sin(\pi \ell z/h) e^{-i\omega t}
\]

with resonant frequencies \( \omega_{n,m,\ell} = c \sqrt{z_{n,m}^2/r^2 + \pi^2 \ell^2/h^2} \).

\[\square\]

9.2 Spherical Bessel functions of the first kind

If in Bessel’s equation (9.4), one sets \( n = \ell + 1/2 \) and \( j_\ell = \sqrt{\pi/2x} J_{\ell+1/2} \), then one may show (exercise 9.21) that

\[
x^2 j_\ell''(x) + 2x j_\ell'(x) + [x^2 - \ell(\ell + 1)] j_\ell(x) = 0 \tag{9.63}
\]

which is the equation for the spherical Bessel function \( j_\ell \).

We saw in example 6.6 that by setting \( V(r, \theta, \phi) = R_k,\ell(r) \Theta_\ell,m(\theta) \Phi_m(\phi) \) we could separate the variables of Helmholtz’s equation \(-\Delta V = k^2 V\) in spherical coordinates

\[
\frac{r^2 \Delta V}{V} = \frac{(r^2 R_k,\ell)'}{R_k,\ell} + \frac{(\sin \theta \Theta_\ell,m)'}{\sin \theta \Theta_\ell,m} + \frac{\Phi''}{\sin^2 \theta \Phi} = -k^2 r^2. \tag{9.64}
\]

Thus if \( \Phi_m(\phi) = e^{im\phi} \) so that \( \Phi_m'' = -m^2 \Phi_m \), and if \( \Theta_\ell,m \) satisfies the associated Legendre equation (8.91)

\[
\sin \theta (\sin \theta \Theta_\ell,m)' + [\ell(\ell + 1) \sin^2 \theta - m^2] \Theta_\ell,m = 0 \tag{9.65}
\]

then the product \( V(r, \theta, \phi) = R_k,\ell(r) \Theta_\ell,m(\theta) \Phi_m(\phi) \) will obey (9.64) because in view of (9.63) the radial function \( R_k,\ell(r) = j_\ell(kr) \) satisfies

\[
(r^2 R_k',\ell)' + [k^2 r^2 - \ell(\ell + 1)] R_k,\ell = 0. \tag{9.66}
\]

In terms of the spherical harmonic \( Y_{\ell,m}(\theta, \phi) = \Theta_\ell,m(\theta) \Phi_m(\phi) \), the solution is \( V(r, \theta, \phi) = j_\ell(kr) Y_{\ell,m}(\theta, \phi) \).

Rayleigh’s formula gives the spherical Bessel function

\[
j_\ell(x) \equiv \sqrt{\frac{\pi}{2x}} J_{\ell+1/2}(x) \tag{9.67}
\]

as the \( \ell \)th derivative of \( \sin x/x \)

\[
j_\ell(x) = (-1)^\ell x^\ell \left( \frac{1}{x} \frac{d}{dx} \right)^\ell \left( \frac{\sin x}{x} \right) \tag{9.68}
\]
9.2 Spherical Bessel functions of the first kind

\[ j_1(x) = \sin \frac{x}{x^2} - \cos \frac{x}{x}. \]

Rayleigh’s formula leads to the recursion relation (exercise 9.22)

\[ j_{\ell+1}(x) = \frac{\ell}{x} j_\ell(x) - j'_\ell(x) \quad (9.69) \]

with which one can show (exercise 9.23) that the spherical Bessel functions as defined by Rayleigh’s formula do satisfy their differential equation (9.66) with \( x = kr \).

The spherical Bessel functions \( j_\ell(kr) \) satisfy the self-adjoint Sturm-Liouville (6.333) equation (9.66)

\[ -r^2 j''_\ell - 2r j'_\ell + \ell(\ell + 1) j_\ell = k^2 r^2 j_\ell \quad (9.70) \]

with eigenvalue \( k^2 \) and weight function \( \rho = r^2 \). If \( j_\ell(z_{\ell,n}) = 0 \), then the functions \( j_\ell(kr) = j_\ell(z_{\ell,n}r/a) \) vanish at \( r = a \) and form an orthogonal basis

\[ \int_0^a j_\ell(z_{\ell,n}r/a) j_\ell(z_{\ell,m}r/a) r^2 \, dr = \frac{a^3}{2} j_{\ell+1}(z_{\ell,n}) \delta_{n,m} \quad (9.71) \]

for a self-adjoint system on the interval \([0,a]\). Moreover, since as \( n \to \infty \) the eigenvalues \( k_{\ell,n}^2 = z_{\ell,n}^2/a^2 \approx [(n + \ell/2)\pi]^2/a^2 \to \infty \), the eigenfunctions \( j_\ell(z_{\ell,n}r/a) \) also are complete in the mean (section 6.35).

On an infinite interval, the analogous relation is

\[ \int_0^\infty j_\ell(kr) j_\ell(k'r) r^2 \, dr = \frac{\pi}{2k^2} \delta(k-k'). \quad (9.72) \]

If we write the spherical Bessel function \( j_0(x) \) as the integral

\[ j_0(z) = \frac{\sin z}{z} = \frac{1}{2} \int_{-1}^1 e^{izx} \, dx \quad (9.73) \]

and use Rayleigh’s formula (9.68), we may find an integral for \( j_\ell(z) \)

\[ j_\ell(z) = (-1)^\ell z^\ell \left( \frac{1}{z} \frac{d}{dz} \right)^\ell \left( \frac{\sin z}{z} \right) = (-1)^\ell z^\ell \left( \frac{1}{z} \frac{d}{dz} \right)^\ell \frac{1}{2} \int_{-1}^1 e^{izx} \, dx \]

\[ = \frac{x^\ell}{2} \int_{-1}^1 (1 - x^2)^\ell \frac{d}{dx} e^{izx} \, dx = \frac{(-i)^\ell}{2} \int_{-1}^1 (1 - x^2)^\ell \frac{d^\ell}{dx^\ell} e^{izx} \, dx \]

\[ = \frac{(-i)^\ell}{2} \int_{-1}^1 e^{izx} \frac{d^\ell}{dx^\ell} (x^2 - 1)^\ell \frac{d^\ell}{dx^\ell} e^{izx} \, dx = \frac{(-i)^\ell}{2} \int_{-1}^1 P_\ell(x) e^{izx} \, dx \quad (9.74) \]

(exercise 9.24) that contains Rodrigues’s formula (8.8) for the Legendre polynomial \( P_\ell(x) \). With \( z = kr \) and \( x = \cos \theta \), this formula

\[ i^\ell j_\ell(kr) = \frac{1}{2} \int_{-1}^1 P_\ell(\cos \theta) e^{ikr \cos \theta} \, d\cos \theta \quad (9.75) \]
The boundary condition \( u_\ell(kr_0) = 0 \) fixes the ratio \( v_\ell = d_\ell/c_\ell \) of the constants \( c_\ell \) and \( d_\ell \). Thus for \( \ell = 0 \), Rayleigh’s formulas (9.68 & 9.104) and the boundary condition say that \( kr_0 u_0(r_0) = c_0 \sin(kr_0) - d_0 \cos(kr_0) = 0 \) or \( d_0/c_0 = \tan kr_0 \). The \( s \)-wave then is \( u_0(kr) = c_0 \sin(kr - kr_0)/(kr \cos kr_0) \), which tells us that the tangent of the phase shift is \( \tan \delta_0(k) = -kr_0 \). By (9.90), the cross-section at low energy is \( \sigma \approx 4\pi r_0^2 \) or four times the classical value.

Similarly, one finds (exercise 9.28) that the tangent of the \( p \)-wave phase shift is
\[
\tan \delta_1(k) = \frac{kr_0 \cos kr_0 - \sin kr_0}{\cos kr_0 - kr_0 \sin kr_0}.
\] (9.108)

For \( kr_0 \ll 1 \), we have \( \delta_1(k) \approx -(kr_0)^3/3 \); more generally, the \( \ell \)th phase shift is \( \delta_\ell(k) \approx -(kr_0)^{2\ell+1}/\{(2\ell+1)[(2\ell-1)!!]^2\} \) for a potential of range \( r_0 \) at low energy \( k \ll 1/r_0 \).

Further Reading

A great deal is known about Bessel functions. Students may find Mathematical Methods for Physics and Engineering (Riley et al., 2006) as well as the classics A Treatise on the Theory of Bessel Functions (Watson, 1995), A Course of Modern Analysis (Whittaker and Watson, 1927, chap. XVII), and Methods of Mathematical Physics (Courant and Hilbert, 1955) of special interest.

Exercises

9.1 Show that the series (9.1) for \( J_n(\rho) \) satisfies Bessel’s equation (9.4).

9.2 Show that the generating function \( \exp(z(u - 1/u)/2) \) for Bessel functions is invariant under the substitution \( u \rightarrow -1/u \).

9.3 Use the invariance of \( \exp(z(u - 1/u)/2) \) under \( u \rightarrow -1/u \) to show that \( J_{-n}(z) = (-1)^n J_n(z) \).

9.4 By writing the generating function (9.5) as the product of the exponentials \( \exp(zu/2) \) and \( \exp(-z/2u) \), derive the expansion
\[
\exp \left[ \frac{z}{2} (u - u^{-1}) \right] = \sum_{n=0}^{\infty} \sum_{m=-n}^{\infty} \left( \frac{z}{2} \right)^{m+n} \frac{u^{m+n}}{(m+n)!} \left( -\frac{z}{2} \right)^n \frac{u^{-n}}{n!}.
\] (9.109)

9.5 From this expansion (9.109) of the generating function (9.5), derive the power-series expansion (9.1) for \( J_n(z) \).
In the formula (9.5) for the generating function \( \exp(z(u - 1/u)/2) \), replace \( u \) by \( \exp i \theta \) and then derive the integral representation (9.6) for \( J_n(z) \). Start with the interval \([-\pi, \pi]\).

From the general integral representation (9.6) for \( J_n(z) \), derive the two integral formulas (9.7) for \( J_0(z) \).

Show that the integral representations (9.6 & 9.7) imply that for any integer \( n \neq 0 \), \( J_n(0) = 0 \), while \( J_0(0) = 1 \).

By differentiating the generating function (9.5) with respect to \( u \) and identifying the coefficients of powers of \( u \), derive the recursion relation

\[
J_{n-1}(z) + J_{n+1}(z) = \frac{2n}{z} J_n(z). \tag{9.110}
\]

By differentiating the generating function (9.5) with respect to \( z \) and identifying the coefficients of powers of \( u \), derive the recursion relation

\[
J_{n-1}(z) - J_{n+1}(z) = 2 J'_n(z). \tag{9.111}
\]

Change variables to \( z = ax \) and turn Bessel’s equation (9.4) into the self-adjoint form (9.11).

If \( y = J_n(ax) \), then equation (9.11) is \( (xy')' + (xa^2 - n^2/x)y = 0 \). Multiply this equation by \( xy' \), integrate from 0 to \( b \), and so show that if \( ab = z_{n,m} \) and \( J_n(z_{n,m}) = 0 \), then

\[
2 \int_0^b x J_n^2(ax) \, dx = b^2 J'_n^2(z_{n,m}) \tag{9.112}
\]

which is the normalization condition (9.14).

Show that with \( \lambda \equiv z^2/r_d^2 \), the change of variables \( \rho = zr/r_d \) and \( u(r) = J_n(\rho) \) turns \(- (r u')' + n^2 u/r = \lambda r^2 u \) into (9.25).

Use the formula (6.42) for the curl in cylindrical coordinates and the vacuum forms \( \nabla \times \mathbf{E} = - \mathbf{B} \) and \( \nabla \times \mathbf{B} = \mathbf{E}/c^2 \) of the laws of Faraday and Maxwell-Ampère to derive the field equations (9.55).

Derive equations (9.56) from (9.55).

Show that \( J_n(\sqrt{\omega^2/c^2 - k^2} \rho)e^{in\phi}e^{i(kz-\omega t)} \) is a traveling-wave solution (9.52) of the wave equations (9.57).

Find expressions for the nonzero TM fields in terms of the formula (9.59) for \( E_z \).

Show that the TE field \( E_z = 0 \) and \( B_z = J_n(\sqrt{\omega^2/c^2 - k^2} \rho)e^{in\phi}e^{i(kz-\omega t)} \) will satisfy the boundary conditions (9.54) if \( \sqrt{\omega^2/c^2 - k^2 r} \) is a zero \( z'_{n,m} \) of \( J'_n \).
9.19 Show that if \( \ell \) is an integer and if \( \omega^2/c^2 - \pi^2\ell^2/h^2 \) is a zero \( z_{n,m}' \) of \( J_n' \), then the fields \( E_z = 0 \) and \( B_z = J_n(z_{n,m}'\rho/r) e^{i\omega \phi} \sin(\ell \pi z/h) e^{-i\omega t} \) satisfy both the boundary conditions (9.54) at \( \rho = r \) and those (9.60) at \( z = 0 \) and \( h \) as well as the wave equations (9.57). Hint: Use Maxwell’s equations \( \nabla \times \mathbf{E} = -\frac{\delta \mathbf{B}}{\delta t} \) and \( \nabla \times \mathbf{B} = \frac{\delta \mathbf{E}}{\delta t}/c^2 \) as in (9.55).

9.20 Show that the resonant frequencies of the TM modes of the cavity of example 9.4 are \( \omega_{n,m,\ell} = c\sqrt{z_{n,m}^2/r^2 + \pi^2\ell^2/h^2} \).

9.21 By setting \( n = \ell + 1/2 \) and \( j_\ell = \sqrt{\pi/2} J_{\ell+1/2} \), show that Bessel’s equation (9.4) implies that the spherical Bessel function \( j_\ell \) satisfies (9.63).

9.22 Show that Rayleigh’s formula (9.68) implies the recursion relation (9.69).

9.23 Use the recursion relation (9.69) to show by induction that the spherical Bessel functions \( j_\ell(x) \) as given by Rayleigh’s formula (9.68) satisfy their differential equation (9.66) which with \( x = kr \) is

\[
-x^2 j''_\ell - 2x j'_\ell + \ell(\ell + 1) j_\ell = x^2 j_\ell.
\]  

(9.113)

Hint: start by showing that \( j_0(x) = \sin(x)/x \) satisfies this equation. This problem involves some tedium.

9.24 Iterate the trick

\[
\frac{d}{dz} \int_{-1}^{1} e^{ixz} dx = \frac{i}{z} \int_{-1}^{1} xe^{ixz} dx = \frac{i}{2z} \int_{-1}^{1} e^{ixz} d(x^2 - 1)
\]

\[
= -\frac{i}{2z} \int_{-1}^{1} (x^2 - 1)e^{ixz} = \frac{1}{2} \int_{-1}^{1} (x^2 - 1)e^{ixz} dx
\]

(9.114)

to show that (Schwinger et al., 1998, p. 227)

\[
\left( \frac{d}{dz} \right)^\ell \int_{-1}^{1} e^{ixz} dx = \int_{-1}^{1} \frac{(x^2 - 1)^\ell}{2^\ell \ell!} e^{ixz} dx.
\]  

(9.115)

9.25 Use the expansions (9.76 & 9.77) to show that the inner product of the ket \( |r\rangle \) that represents a particle at \( r \) with polar angles \( \theta \) and \( \phi \) and the one \( |k\rangle \) that represents a particle with momentum \( p = \hbar k \) with polar angles \( \theta' \) and \( \phi' \) is with \( k \cdot r = kr \cos \theta \)

\[
\langle r | k \rangle = \frac{1}{(2\pi)^{3/2}} e^{i\mathbf{k} \cdot \mathbf{r}} = \frac{1}{(2\pi)^{3/2}} \sum_{\ell=0}^{\infty} \frac{(2\ell + 1) P_\ell(\cos \theta)}{\ell!} i^\ell j_\ell(kr)
\]

\[
= \frac{1}{(2\pi)^{3/2}} e^{i\mathbf{k} \cdot \mathbf{r}} = \sqrt{\frac{2}{\pi}} \sum_{\ell=0}^{\infty} i^\ell j_\ell(kr) Y_{\ell,m}(\theta, \phi) Y_{\ell,m}^*(\theta', \phi').
\]  

(9.116)
Thus $U(g)$ cannot change $E$ or $j$, and so
\[
\langle E', j', m' | U(g) | E, j, m \rangle = \delta_{E'E} \delta_{j'j} \delta_{m'm} D^{(j)}_{m'm}(g).
\]

The matrix element (10.11) is a single sum over $E$ and $j$ in which the irreducible representations $D^{(j)}_{m'm}(g)$ of the rotation group $SU(2)$ appear
\[
\langle \phi | U(g) | \psi \rangle = \sum_{E,j,m} \langle \phi | E, j, m | \rangle D^{(j)}_{m'm}(g) \langle E, j, m | \psi \rangle.
\]

This is how the block-diagonal form (10.7) usually appears in calculations. The matrices $D^{(j)}_{m'm}(g)$ inherit the unitarity of the operator $U(g)$.

\[ \square \]

10.4 Subgroups

If all the elements of a group $S$ also are elements of a group $G$, then $S$ is a subgroup of $G$. Every group $G$ has two trivial subgroups—the identity element $e$ and the whole group $G$ itself. Many groups have more interesting subgroups. For example, the rotations about a fixed axis is an abelian subgroup of the group of all rotations in 3-dimensional space.

A subgroup $S \subset G$ is an invariant subgroup if every element $s$ of the subgroup $S$ is left inside the subgroup under the action of every element $g$ of the whole group $G$, that is, if
\[
g^{-1}s g = s' \in S \quad \text{for all} \quad g \in G.
\]

This condition often is written as $g^{-1} S g = S$ for all $g \in G$ or as
\[
S g = g S \quad \text{for all} \quad g \in G.
\]

Invariant subgroups also are called normal subgroups.

A set $C \subset G$ is called a conjugacy class if it’s invariant under the action of the whole group $G$, that is, if $C g = g C$ or
\[
g^{-1}C g = C \quad \text{for all} \quad g \in G.
\]

A subgroup that is the union of a set of conjugacy classes is invariant.

The center $C$ of a group $G$ is the set of all elements $c \in G$ that commute with every element $g$ of the group, that is, their commutators
\[
[c, g] \equiv cg - gc = 0
\]
vanish for all $g \in G$. 

Example 10.8 (Centers Are Abelian Subgroups) Does the center $C$ always form an abelian subgroup of its group $G$? The product $c_1c_2$ of any two elements $c_1$ and $c_2$ of the center commutes with every element $g$ of $G$ since $c_1c_2g = c_1gc_2 = gc_1c_2$. So the center is closed under multiplication. The identity element $e$ commutes with every $g \in G$, so $e \in C$. If $c' \in C$, then $c'g = gc'$ for all $g \in G$, and so multiplication of this equation from the left and the right by $c'^{-1}$ gives $gc'^{-1} = c'^{-1}g$, which shows that $c'^{-1} \in C$. The subgroup $C$ is abelian because each of its elements commutes with all the elements of $G$ including those of $C$ itself.

So the center of any group always is one of its abelian invariant subgroups. The center may be trivial, however, consisting either of the identity or of the whole group. But a group with a nontrivial center can not be simple or semisimple (section 10.23).

10.5 Cosets

If $H$ is a subgroup of a group $G$, then for every element $g \in G$ the set of elements $Hg \equiv \{hg | h \in H, g \in G\}$ is a right coset of the subgroup $H \subset G$. (Here $\subset$ means is a subset of or equivalently is contained in.)

If $H$ is a subgroup of a group $G$, then for every element $g \in G$ the set of elements $gH$ is a left coset of the subgroup $H \subset G$.

The number of elements in a coset is the same as the number of elements of $H$, which is the order of $H$.

An element $g$ of a group $G$ is in one and only one right coset (and in one and only one left coset) of the subgroup $H \subset G$. For suppose instead that $g$ were in two right cosets $g \in Hg_1$ and $g \in Hg_2$, so that $g = h_1g_1 = h_2g_2$ for suitable $h_1, h_2 \in H$ and $g_1, g_2 \in G$. Then since $H$ is a (sub)group, we have $g_2 = h_2^{-1}h_1g_1 = h_3g_1$, which says that $g_2 \in Hg_1$. But this means that every element $h_2g_2 \in Hg_2$ is of the form $hg_2 = hh_3g_1 = h_4g_1 \in Hg_1$. So every element $hg_2 \in Hg_2$ is in $Hg_1$: the two right cosets are identical, $Hg_1 = Hg_2$.

The right (or left) cosets are the points of the quotient coset space $G/H$.

If $H$ is an invariant subgroup of $G$, then by definition (10.16) $Hg = gH$ for all $g \in G$, and so the left cosets are the same sets as the right cosets. In this case, the coset space $G/H$ is itself a group with multiplication defined
and generate the elements of the group $SU(2)$

$$\exp \left( i \theta \cdot \vec{\sigma} \right) = I \cos \frac{\theta}{2} + i \hat{\theta} \cdot \vec{\sigma} \sin \frac{\theta}{2}$$  \hspace{1cm} (10.118)

in which $I$ is the $2 \times 2$ identity matrix, $\theta = \sqrt{\theta^2}$ and $\hat{\theta} = \theta / \theta$.

It follows from (10.117) that the spin operators satisfy

$$[S_a, S_b] = i \hbar \epsilon_{abc} S_c.$$  \hspace{1cm} (10.119)

The raising and lowering operators

$$S_\pm = S_1 \pm i S_2$$  \hspace{1cm} (10.120)

have simple commutators with $S_3$

$$[S_3, S_\pm] = \pm \hbar S_\pm.$$  \hspace{1cm} (10.121)

This relation implies that if the state $|j, m\rangle$ is an eigenstate of $S_3$ with eigenvalue $\hbar m$, then the states $S_\pm |j, m\rangle$ either vanish or are eigenstates of $S_3$ with eigenvalues $\hbar(m \pm 1)$

$$S_3 S_\pm |j, m\rangle = S_\pm S_3 |j, m\rangle \pm \hbar S_\pm |j, m\rangle = \hbar(m \pm 1) S_\pm |j, m\rangle.$$  \hspace{1cm} (10.122)

Thus the raising and lowering operators raise and lower the eigenvalues of $S_3$. When $j = 1/2$, the possible values of $m$ are $m = \pm 1/2$, and so with the usual sign and normalization conventions

$$S_+ |-\rangle = \hbar |+\rangle \text{ and } S_- |+\rangle = \hbar |-\rangle$$  \hspace{1cm} (10.123)

while

$$S_+ |+\rangle = 0 \text{ and } S_- |-\rangle = 0.$$  \hspace{1cm} (10.124)

The square of the total spin operator is simply related to the raising and lowering operators and to $S_3$

$$S^2 = S_1^2 + S_2^2 + S_3^2 = \frac{1}{2} S_+ S_- + \frac{1}{2} S_- S_+ + S_3^2.$$  \hspace{1cm} (10.125)

But the squares of the Pauli matrices are unity, and so $S_a^2 = (\hbar/2)^2$ for all three values of $a$. Thus

$$S^2 = \frac{3}{4} \hbar^2$$  \hspace{1cm} (10.126)

is a Casimir operator (10.105) for a spin one-half system.

**Example 10.20** (Two Spin 1/2’s) Consider two spin operators $\mathbf{S}^{(1)}$ and
Example 10.24 (Lack of Analyticity) One may define a function \( f(q) \) of a quaternionic variable and then ask what functions are analytic in the sense that the (one-sided) derivative

\[
f'(q) = \lim_{q' \to 0} \left[ f(q + q') - f(q) \right] q'^{-1}
\]

exists and is independent of the direction through which \( q' \to 0 \). This space of functions is extremely limited and does not even include the function \( f(q) = q^2 \) (exercise 10.22). \( \square \)

10.28 The Symplectic Group \( Sp(2n) \)

The symplectic group \( Sp(2n) \) consists of \( 2n \times 2n \) matrices \( W \) that map \( n \)-tuples \( q \) of quaternions into \( n \)-tuples \( q' = Wq \) of quaternions with the same value of the quadratic quaternionic form

\[
||q'||^2 = ||q'_1||^2 + ||q'_2||^2 + \cdots + ||q'_n||^2 = ||q_1||^2 + ||q_2||^2 + \cdots + ||q_n||^2 = ||q||^2.
\]

(10.185)

By (10.177), the quadratic form \( ||q'||^2 \) times the \( 2 \times 2 \) identity matrix \( I \) is equal to the hermitian form \( q'^\dagger q' \)

\[
||q'||^2 I = q'^\dagger q' = q'_1^\dagger q'_1 + \cdots + q'_n^\dagger q'_n = q^\dagger W^\dagger Wq
\]

(10.186)

and so any matrix \( W \) that is both a \( 2n \times 2n \) unitary matrix and an \( n \times n \) matrix of quaternions keeps \( ||q'||^2 = ||q||^2 \)

\[
||q'||^2 I = q^\dagger W^\dagger Wq = q^\dagger q = ||q||^2 I.
\]

(10.187)

The group \( Sp(2n) \) thus consists of all \( 2n \times 2n \) unitary matrices that also are \( n \times n \) matrices of quaternions. (This last requirement is needed so that \( q' = Wq \) is an \( n \)-tuple of quaternions.)

The generators \( t_a \) of the symplectic group \( Sp(2n) \) are \( 2n \times 2n \) direct-product matrices of the form

\[
I \otimes A, \quad \sigma_1 \otimes S_1, \quad \sigma_2 \otimes S_2, \quad \text{and} \quad \sigma_3 \otimes S_3
\]

(10.188)

in which \( I \) is the \( 2 \times 2 \) identity matrix, the three \( \sigma_i \)'s are the Pauli matrices, \( A \) is an imaginary \( n \times n \) anti-symmetric matrix, and the \( S_i \) are \( n \times n \) real symmetric matrices. These generators \( t_a \) close under commutation

\[
[t_a, t_b] = i\alpha_{abc}t_c.
\]

(10.189)

Any imaginary linear combination \( i\alpha_a t_a \) of these generators is not only a
is called a **right-handed Weyl spinor**. One may show (exercise 10.38) that the action density

\[ L_r(x) = i \zeta^\dagger(x) (\partial_0 I + \nabla \cdot \sigma) \zeta(x) \]  

is Lorentz covariant

\[ U(L) L_r(x) U^{-1}(L) = L_r(Lx). \]  

**Example 10.33 (Why \( \zeta \) Is Right Handed)**  
An argument like that of example (10.31) shows that the field \( \zeta(x) \) satisfies the wave equation

\[ (\partial_0 I + \nabla \cdot \sigma) \zeta(x) = 0 \]  

or in momentum space

\[ (E - p \cdot \sigma) \zeta(p) = 0. \]  

Thus, \( E = |p| \), and \( \zeta(p) \) is an eigenvector of \( \hat{p} \cdot \hat{J} \)

\[ \hat{p} \cdot \hat{J} \zeta(p) = \frac{1}{2} \zeta(p) \]  

with eigenvalue 1/2. A particle whose spin is parallel to its momentum is said to have **positive helicity** or to be **right handed**. Nearly massless antineutrinos are nearly right handed.

The Majorana mass term

\[ L_M(x) = \frac{1}{2} \left[ i m \zeta^\dagger(x) \sigma_2 \zeta(x) + (i m \zeta^\dagger(x) \sigma_2 \zeta(x))^\dagger \right] \]  

like (10.266) is Lorentz covariant.

### 10.33 The Dirac Representation of the Lorentz Group

Dirac’s representation of \( SO(3,1) \) is the direct sum \( D^{(1/2,0)} \oplus D^{(0,1/2)} \) of \( D^{(1/2,0)} \) and \( D^{(0,1/2)} \). Its generators are the 4 \( \times \) 4 matrices

\[ J = \frac{1}{2} \begin{pmatrix} \sigma & 0 \\ 0 & -\sigma \end{pmatrix} \quad \text{and} \quad K = \frac{i}{2} \begin{pmatrix} -\sigma & 0 \\ 0 & \sigma \end{pmatrix}. \]  

Dirac’s representation uses the **Clifford algebra** of the gamma matrices \( \gamma^a \) which satisfy the anticommutation relation

\[ \{ \gamma^a, \gamma^b \} \equiv \gamma^a \gamma^b + \gamma^b \gamma^a = 2 \eta^{ab} \]  

in which \( \eta \) is the 4 \( \times \) 4 diagonal matrix (10.222) with \( \eta^{00} = -1 \) and \( \eta^{jj} = 1 \) for \( j = 1, 2, \) and 3.
Suppose \( T(y) \) is a translation that takes a 4-vector \( x \) to \( x + y \) and \( T(z) \) is a translation that takes a 4-vector \( x \) to \( x + z \). Then \( T(z)T(y) \) and \( T(y)T(z) \) both take \( x \) to \( x + y + z \). So if a translation \( T(y) = T(t, y) \) is represented by a unitary operator \( U(t, y) = \exp(iHt - iP \cdot y) \), then the Hamiltonian \( H \) and the momentum operator \( P \) commute with each other

\[
[H, P^j] = 0 \quad \text{and} \quad [P^i, P^j] = 0. \quad (10.299)
\]

We can figure out the commutation relations of \( H \) and \( P \) with the angular-momentum \( J \) and boost \( K \) operators by realizing that \( P^a \) is a 4-vector. Let

\[
U(\theta, \lambda) = e^{-i\theta \cdot J - i\lambda \cdot K} \quad (10.300)
\]

be the (infinite-dimensional) unitary operator that represents (in Hilbert space) the infinitesimal Lorentz transformation

\[
L = I + \theta \cdot R + \lambda \cdot B \quad (10.301)
\]

where \( R \) and \( B \) are the six \( 4 \times 4 \) matrices (10.231 & 10.232). Then because \( P \) is a 4-vector under Lorentz transformations, we have

\[
U^{-1}(\theta, \lambda)PU(\theta, \lambda) = e^{+i\theta \cdot J + i\lambda \cdot K} P e^{-i\theta \cdot J - i\lambda \cdot K} = (I + \theta \cdot R + \lambda \cdot B) P \quad (10.302)
\]

or using (10.272)

\[
(I + i\theta \cdot J + i\lambda \cdot K) H (I - i\theta \cdot J - i\lambda \cdot K) = H + \lambda \cdot P \quad (10.303)
\]

\[
(I + i\theta \cdot J + i\lambda \cdot K) P (I - i\theta \cdot J - i\lambda \cdot K) = P + H \lambda + \theta \wedge P.
\]

Thus, one finds (exercise 10.42) that \( H \) is invariant under rotations, while \( P \) transforms as a 3-vector

\[
[J_i, H] = 0 \quad \text{and} \quad [J_i, P_j] = i\epsilon_{ijk} P_k \quad (10.304)
\]

and that

\[
[K_i, H] = -iP_i \quad \text{and} \quad [K_i, P_j] = -i\delta_{ij} H. \quad (10.305)
\]

By combining these equations with (10.285), one may write (exercise 10.44) the Lie algebra of the Poincaré group as

\[
i[J^{ab}, J^{cd}] = \eta^{bc} J^{ad} - \eta^{ac} J^{bd} - \eta^{da} J^{cb} + \eta^{db} J^{ca}
\]

\[
i[P^a, J^{bc}] = \eta^{ab} P^c - \eta^{ac} P^b
\]

\[
[P^a, P^b] = 0. \quad (10.306)
\]
Exercises

10.1 Show that all $n \times n$ (real) orthogonal matrices $O$ leave invariant the quadratic form $x_1^2 + x_2^2 + \cdots + x_n^2$, that is, that if $x' = O x$, then $x'^2 = x^2$.

10.2 Show that the set of all $n \times n$ orthogonal matrices forms a group.

10.3 Show that all $n \times n$ unitary matrices $U$ leave invariant the quadratic form $|x_1|^2 + |x_2|^2 + \cdots + |x_n|^2$, that is, that if $x' = U x$, then $|x'|^2 = |x|^2$.

10.4 Show that the set of all $n \times n$ unitary matrices forms a group.

10.5 Show that the set of all $n \times n$ unitary matrices with unit determinant forms a group.

10.6 Show that the matrix $D^{(j)}_{m'm}(g) = \langle j, m'|U(g)|j, m \rangle$ is unitary because the rotation operator $U(g)$ is unitary $\langle j, m'|U^\dagger(g)U(g)|j, m \rangle = \delta_{m'm}$.

10.7 Invent a group of order 3 and compute its multiplication table. For extra credit, prove that the group is unique.

10.8 Show that the relation (10.20) between two equivalent representations is an isomorphism.

10.9 Suppose that $D_1$ and $D_2$ are equivalent, finite-dimensional, irreducible representations of a group $G$ so that $D_2(g) = S D_1(g) S^{-1}$ for all $g \in G$. What can you say about a matrix $A$ that satisfies $D_2(g) A = A D_1(g)$ for all $g \in G$?

10.10 Find all components of the matrix $\exp(i\alpha A)$ in which

$$A = \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix}. \quad (10.307)$$

10.11 If $[A, B] = B$, find $e^{i\alpha A} Be^{-i\alpha A}$. Hint: what are the $\alpha$-derivatives of this expression?

10.12 Show that the tensor-product matrix (10.31) of two representations $D_1$ and $D_2$ is a representation.

10.13 Find a $4 \times 4$ matrix $S$ that relates the tensor-product representation $D_{1 \otimes 2}$ to the direct sum $D_1 \oplus D_2$.

10.14 Find the generators in the adjoint representation of the group with structure constants $f_{abc} = \epsilon_{abc}$ where $a, b, c$ run from 1 to 3. Hint: The answer is three $3 \times 3$ matrices $t_a$, often written as $L_a$.

10.15 Show that the generators (10.90) satisfy the commutation relations (10.93).

The classic *Lie Algebras in Particle Physics* (Georgi, 1999), which inspired much of this chapter, is outstanding.
10.16 Show that the demonstrated equation (10.98) implies the commutation relation (10.99).

10.17 Use the Cayley-Hamilton theorem (1.264) to show that the $3 \times 3$ matrix (10.96) that represents a right-handed rotation of $\theta$ radians about the axis $\theta$ is given by (10.97).

10.18 Verify the mixed Jacobi identity (10.142).

10.19 For the group $SU(3)$, find the structure constants $f_{123}$ and $f_{231}$.

10.20 Show that every $2 \times 2$ unitary matrix of unit determinant is a quaternion of unit norm.

10.21 Show that the quaternions as defined by (10.175) are closed under addition and multiplication and that the product $xq$ is a quaternion if $x$ is real and $q$ is a quaternion.

10.22 Show that the one-sided derivative $f'(q)$ (10.184) of the quaternionic function $f(q) = q^2$ depends upon the direction along which $q' \to 0$.

10.23 Show that the generators (10.188) of $Sp(2n)$ obey commutation relations of the form (10.189) for some real structure constants $f_{abc}$ and a suitably extended set of matrices $A, A', \ldots$ and $S_k, S'_k, \ldots$.

10.24 Show that for $0 < \epsilon \ll 1$, the real $2n \times 2n$ matrix $T = \exp(\epsilon JS)$ in which $S$ is symmetric satisfies $T^TJT = J$ (at least up to terms of order $\epsilon^2$) and so is in $Sp(2n, R)$.

10.25 Show that the matrix $T$ of (10.197) is in $Sp(2, R)$.

10.26 Use the parametrization (10.217) of the group $SU(2)$, show that the parameters $a(c, b)$ that describe the product $g(a(c, b)) = g(c)g(b)$ are those of (10.219).

10.27 Use formulas (10.219) and (10.212) to show that the left-invariant measure for $SU(2)$ is given by (10.220).

10.28 In tensor notation, which is explained in chapter 11, the condition (10.229) that $I + \omega$ be an infinitesimal Lorentz transformation reads $(\omega^T)_b^a = \omega^a_b = -\eta_{bc} \omega^c_d \eta^{da}$ in which sums over $c$ and $d$ from 0 to 3 are understood. In this notation, the matrix $\eta_{cf}$ lowers indices and $\eta^{ph}$ raises them, so that $\omega^a_b = -\omega_{bd} \eta^{da}$. (Both $\eta_{cf}$ and $\eta^{ph}$ are numerically equal to the matrix $\eta$ displayed in equation (10.222).) Multiply both sides of the condition (10.229) by $\eta_{ae} = \eta_{ea}$ and use the relation $\eta^{da} \eta_{ae} = \eta^d_e = \delta^d_e$ to show that the matrix $\omega_{ab}$ with both indices lowered (or raised) is antisymmetric, that is,

$$\omega_{ba} = -\omega_{ab} \quad \text{and} \quad \omega^{ba} = -\omega^{ab}. \quad (10.308)$$

10.29 Show that the six matrices (10.231) and (10.232) satisfy the $SO(3, 1)$ condition (10.229).
10.30 Show that the six generators $J$ and $K$ obey the commutations relations (10.234–10.236).

10.31 Show that if $J$ and $K$ satisfy the commutation relations (10.234–10.236) of the Lie algebra of the Lorentz group, then so do $J$ and $-K$.

10.32 Show that if the six generators $J$ and $K$ obey the commutation relations (10.234–10.236), then the six generators $J^+$ and $J^-$ obey the commutation relations (10.243).

10.33 Relate the parameter $\alpha$ in the definition (10.253) of the standard boost $B(p)$ to the 4-vector $p$ and the mass $m$.

10.34 Derive the formulas for $D^{(1/2,0)}(0, \alpha \hat{p})$ given in equation (10.254).

10.35 Derive the formulas for $D^{(0,1/2)}(0, \alpha \hat{p})$ given in equation (10.271).

10.36 For infinitesimal complex $z$, derive the 4-vector properties (10.255 & 10.272) of $(-I, \sigma)$ under $D^{(1/2,0)}$ and of $(I, \sigma)$ under $D^{(0,1/2)}$.

10.37 Show that under the unitary Lorentz transformation (10.257), the action density (10.258) is Lorentz covariant (10.259).

10.38 Show that under the unitary Lorentz transformation (10.273), the action density (10.274) is Lorentz covariant (10.275).

10.39 Show that under the unitary Lorentz transformations (10.257 & 10.273), the Majorana mass terms (10.266 & 10.279) are Lorentz covariant.

10.40 Show that the definitions of the gamma matrices (10.281) and of the generators (10.283) imply that the gamma matrices transform as a 4-vector under Lorentz transformations (10.284).

10.41 Show that (10.283) and (10.284) imply that the generators $J^{ab}$ satisfy the commutation relations (10.285) of the Lorentz group.

10.42 Show that the spinor $\zeta = \sigma_2 \xi^*$ defined by (10.295) is right handed (10.273) if $\xi$ is left handed (10.257).

10.43 Use (10.303) to get (10.304 & 10.305).

The coefficients $e'_i \cdot e_j$ form an orthogonal matrix, and the linear operator
\begin{equation}
\sum_{i=1}^{n} e_i e'_i = \sum_{i=1}^{n} |e_i \rangle \langle e'_i | \tag{11.21}
\end{equation}
is an orthogonal (real, unitary) transformation. The change $x \to x'$ is a rotation plus a possible reflection (exercise 11.2).

**Example 11.2 (A Euclidean Space of Two Dimensions)** In two-dimensional euclidean space, one can describe the same point by euclidean $(x, y)$ and polar $(r, \theta)$ coordinates. The derivatives
\begin{equation}
\frac{\partial r}{\partial x} = \frac{x}{r} = \frac{\partial x}{\partial r} \quad \text{and} \quad \frac{\partial r}{\partial y} = \frac{y}{r} = \frac{\partial y}{\partial r} \tag{11.22}
\end{equation}
respect the symmetry (11.18), but (exercise 11.1) these derivatives
\begin{equation}
\frac{\partial \theta}{\partial x} = -\frac{y}{r^2} \neq \frac{\partial x}{\partial \theta} = -y \quad \text{and} \quad \frac{\partial \theta}{\partial y} = \frac{x}{r^2} \neq \frac{\partial y}{\partial \theta} = x \tag{11.23}
\end{equation}
do not.

### 11.6 Summation Conventions

When a given index is repeated in a product, that index usually is being summed over. So to avoid distracting summation symbols, one writes
\begin{equation}
A_i B_i \equiv \sum_{i=1}^{n} A_i B_i. \tag{11.24}
\end{equation}
The sum is understood to be over the relevant range of indices, usually from 0 or 1 to 3 or $n$. Where the distinction between covariant and contravariant indices matters, an index that appears twice in the same monomial, once as a subscript and once as a superscript, is a dummy index that is summed over as in
\begin{equation}
A_i B^i \equiv \sum_{i=1}^{n} A_i B^i. \tag{11.25}
\end{equation}
These summation conventions make tensor notation almost as compact as matrix notation. They make equations easier to read and write.
How weak are the static gravitational fields we know about? The dimensionless ratio $\phi/c^2$ is $10^{-39}$ on the surface of a proton, $10^{-9}$ on the Earth, $10^{-6}$ on the surface of the sun, and $10^{-4}$ on the surface of a white dwarf.

### 11.41 Gravitational Time Dilation

Suppose we have a system of coordinates $x^i$ with a metric $g_{ik}$ and a clock at rest in this system. Then the proper time $d\tau$ between ticks of the clock is

$$d\tau = (1/c)\sqrt{-g_{ij}dx^i dx^j} = \sqrt{-g_{00}} dt$$  \hspace{1cm} (11.340)

where $dt$ is the time between ticks in the $x^i$ coordinates, which is the laboratory frame in the gravitational field $g_{00}$. By the principle of equivalence (section 11.39), the proper time $d\tau$ between ticks is the same as the time between ticks when the same clock is at rest deep in empty space.

If the clock is in a weak static gravitational field due to a mass $M$ at a distance $r$, then

$$-g_{00} = 1 + 2\phi/c^2 = 1 - 2GM/c^2 r$$  \hspace{1cm} (11.341)

is a little less than unity, and the interval of proper time between ticks

$$d\tau = \sqrt{-g_{00}} dt = \sqrt{1 - 2GM/c^2 r} dt$$  \hspace{1cm} (11.342)

is slightly less than the interval $dt$ between ticks in the coordinate system of an observer at $x$ in the rest frame of the clock and the mass, and in its gravitational field. Since $dt > d\tau$, the laboratory time $dt$ between ticks is greater than the proper or intrinsic time $d\tau$ between ticks of the clock unaffected by any gravitational field. Clocks near big masses run slow.

Now suppose we have two identical clocks in at different heights above sea level. The time $T_l$ for the lower clock to make $N$ ticks will be longer than the time $T_u$ for the upper clock to make $N$ ticks. The ratio of the clock times will be

$$\frac{T_l}{T_u} = \frac{\sqrt{1 - 2GM/c^2(r + h)}}{\sqrt{1 - 2GM/c^2 r}} \approx 1 + \frac{gh}{c^2}. \hspace{1cm} (11.343)$$

Now imagine that a photon going down passes the upper clock which measures its frequency as $\nu_u$ and then passes the lower clock which measures its frequency as $\nu_l$. The slower clock will measure a higher frequency. The ratio of the two frequencies will be the same as the ratio of the clock times

$$\frac{\nu_l}{\nu_u} = 1 + \frac{gh}{c^2}. \hspace{1cm} (11.344)$$

As measured by the lower, slower clock, the photon is blue shifted.
Black holes are not really black. Stephen Hawking (1942–) has shown that the intense gravitational field of a black hole of mass $M$ radiates at temperature

$$T = \frac{\hbar c^3}{8\pi k G M}$$

(11.389)

in which $k = 8.617343 \times 10^{-5}$ eV K$^{-1}$ is Boltzmann’s constant, and $\hbar$ is Planck’s constant $\hbar = 6.6260693 \times 10^{-34}$ Js divided by $2\pi$, $\hbar = \hbar / (2\pi)$.

The black hole is entirely converted into radiation after a time

$$t = \frac{5120 \pi G^2}{\hbar c^4} M^3$$

(11.390)

proportional to the cube of its mass.

### 11.48 Cosmology

Astrophysical observations tell us that on the largest observable scales, space is flat or very nearly flat; that the visible universe contains at least $10^{90}$ particles; and that the cosmic microwave background radiation is isotropic to one part in $10^5$ apart from a Doppler shift due the motion of the Earth. These and other observations suggest that potential energy expanded our universe by $\exp(60) = 10^{26}$ during an era of inflation that could have been as brief as $10^{-35}$ s. The potential energy that powered inflation became the radiation of the Big Bang.

During the first three minutes, some of that radiation became hydrogen, helium, neutrinos, and dark matter. But for 65,000 years after the Big Bang, most of the energy of the visible universe was radiation. Because the momentum of a particle but not its mass falls with the expansion of the universe, this era of radiation gradually gave way to an era of matter. This transition happened when the temperature $kT$ of the universe fell to 1.28 eV.

The era of matter lasted for 8.8 billion years. After 380,000 years, the universe had cooled to $kT = 0.26$ eV, and less than 1% of the atoms were ionized. Photons no longer scattered off a plasma of electrons and ions. The universe became transparent. The photons that last scattered just before this initial transparency became the cosmic microwave background radiation or CMBR that now surrounds us, red-shifted to 2.7255 $\pm 0.0006$ K.

The era of matter has been followed by the current era of dark energy during which the energy of the visible universe is mostly a potential energy
called **dark energy** (something like a **cosmological constant**). Dark energy has been accelerating the expansion of the universe for the past 5 billion years and may continue to do so forever.

It is now $13.817 \pm 0.048$ billion years after the Big Bang, and the dark-energy density is $\rho_{de} = 5.827 \times 10^{-30} \text{ g cm}^{-3}$ or 68.5 percent ($\pm 1.8\%$) of the **critical energy density** $\rho_c = \frac{3H_0^2}{8\pi G} = 1.87837 \times 10^{-29} \text{ g cm}^{-3}$ needed to make the universe flat. Here $H_0 = 100 \text{ km s}^{-1} \text{ Mpc}^{-1}$ is the **Hubble constant**, one parsec is 3.262 light-years, the Hubble time is $1/H_0 = 9.778 h^{-1} \times 10^9$ years, and $h = 0.673 \pm 0.012$ is not to be confused with Planck’s constant.

Matter makes up $31.5 \pm 1.8\%$ of the critical density, and baryons only $4.9 \pm 0.06\%$ of it. Baryons are 15% of the total matter in the visible universe. The other 85% does not interact with light and is called **dark matter**.

Einstein’s equations (11.373) are second-order, non-linear partial differential equations for 10 unknown functions $g_{ij}(x)$ in terms of the energy-momentum tensor $T_{ij}(x)$ throughout the universe, which of course we don’t know. The problem is not quite hopeless, however. The ability to choose arbitrary coordinates, the appeal to symmetry, and the choice of a reasonable form for $T_{ij}$ all help.

Hubble showed us that the universe is expanding. The cosmic microwave background radiation looks the same in all spatial directions (apart from a Doppler shift due to the motion of the Earth relative to the local supercluster of galaxies). Observations of clusters of galaxies reveal a universe that is homogeneous on suitably large scales of distance. So it is plausible that the universe is **homogeneous** and **isotropic** in space, but not in time. One may show (Carroll, 2003) that for a universe of such symmetry, the line element in **comoving coordinates** is

$$ds^2 = -dt^2 + a^2 \left[ \frac{dr^2}{1 - kr^2} + r^2 \left( d\theta^2 + \sin^2 \theta \, d\phi^2 \right) \right]. \quad (11.391)$$

Whitney’s embedding theorem tells us that any smooth four-dimensional manifold can be embedded in a flat space of eight dimensions with a suitable **signature**. We need only four or five dimensions to embed the space-time described by the line element (11.391). If the universe is closed, then the signature is $(-1,1,1,1,1)$, and our three-dimensional space is the **3-sphere** which is the surface of a four-dimensional sphere in four space dimensions. The points of the universe then are

$$p = (t, a \sin \chi \sin \theta \cos \phi, a \sin \chi \sin \theta \sin \phi, a \sin \chi \cos \theta, a \cos \chi) \quad (11.392)$$

The points of the universe then are
The other nonzero $\Gamma$’s are
\[
\Gamma^1_{22} = -r (1 - kr^2) \quad \Gamma^1_{33} = -r (1 - kr^2) \sin^2 \theta \quad (11.399)
\]
\[
\Gamma^2_{12} = \Gamma^3_{13} = \frac{1}{r} = \Gamma^2_{21} = \Gamma^3_{31} \quad (11.400)
\]
\[
\Gamma^2_{33} = -\sin \theta \cos \theta \quad \Gamma^3_{23} = \cot \theta = \Gamma^3_{32}. \quad (11.401)
\]

Our formulas (11.350 & 11.348) for the Ricci and curvature tensors give
\[
R_{00} = R^{n}_{0n0} = [\partial_0 + \Gamma_0, \partial_n + \Gamma_n]_0^n. \quad (11.402)
\]
Clearly the commutator of $\Gamma_0$ with itself vanishes, and one may use the formulas (11.397–11.401) for the other connections to check that
\[
[\Gamma_0, \Gamma_n]_0^n = \Gamma^n_{0k} \Gamma^k_{n0} - \Gamma^n_{nk} \Gamma^k_{00} = 3 \left( \frac{\dot{a}}{a} \right)^2 \quad (11.403)
\]
and that
\[
\partial_0 \Gamma^n_{n0} = 3 \partial_0 \left( \frac{\dot{a}}{a} \right) = 3 \frac{\ddot{a}}{a} - 3 \left( \frac{\dot{a}}{a} \right)^2 \quad (11.404)
\]
while $\partial_n \Gamma^n_{00} = 0$. So the 00-component of the Ricci tensor is
\[
R_{00} = 3 \frac{\ddot{a}}{a}. \quad (11.405)
\]

Similarly, one may show that the other non-zero components of Ricci’s tensor are
\[
R_{11} = -\frac{A}{1 - kr^2} \quad R_{22} = -r^2 A \quad \text{and} \quad R_{33} = -r^2 A \sin^2 \theta \quad (11.406)
\]
in which $A = a\ddot{a} + 2\dot{a}^2 + 2k$. The scalar curvature (11.351) is
\[
R = g^{ab} R_{ba} = -\frac{6}{a^2} \left( a\ddot{a} + \dot{a}^2 + k \right). \quad (11.407)
\]

In co-moving coordinates such as those of the Robertson-Walker metric (11.395) $u_i = (1, 0, 0, 0)$, and so the energy-momentum tensor (11.370) is
\[
T_{ij} = \begin{pmatrix}
\rho & 0 & 0 & 0 \\
0 & pg_{11} & 0 & 0 \\
0 & 0 & pg_{22} & 0 \\
0 & 0 & 0 & pg_{33}
\end{pmatrix}. \quad (11.408)
\]
Its trace is
\[
T = g^{ij} T_{ij} = -\rho + 3p. \quad (11.409)
\]
Thus using our formula (11.395) for $g_{00} = -1$, (11.405) for $R_{00}$, (11.408)
density
\[ \rho_c = \frac{3H^2}{8\pi G}. \] (11.415)

The ratio of the energy density \(\rho\) to the critical energy density is called \(\Omega\)
\[ \Omega = \frac{\rho}{\rho_c} = \frac{8\pi G}{3H^2} \rho. \] (11.416)

From (11.414), we see that \(\Omega\) is
\[ \Omega = 1 + \frac{k}{(aH)^2} = 1 + \frac{k}{a^2}. \] (11.417)

Thus \(\Omega = 1\) both in a flat universe \((k = 0)\) and as \(aH \to \infty\). One use of
inflation is to expand \(a\) by \(10^{26}\) so as to force \(\Omega\) almost exactly to unity.

Something like inflation is needed because in a universe in which the energy density is due to matter and/or radiation, the present value of \(\Omega\)
\[ \Omega_0 = 1.000 \pm 0.036 \] (11.418)
is unlikely. To see why, we note that conservation of energy ensures that \(a^3\)
times the matter density \(\rho_m\) is constant. Radiation red-shifts by \(a\), so energy
conservation implies that \(a^4\) times the radiation density \(\rho_r\) is constant. So
with \(n = 3\) for matter and \(4\) for radiation, \(\rho a^n \equiv 3F^2/8\pi G\) is a constant.
In terms of \(F\) and \(n\), Friedmann’s first-order equation (11.412) is
\[ \dot{a}^2 = \frac{8\pi G}{3} \rho a^2 - k = \frac{F^2}{a^{n-2}} - k \] (11.419)

In small-\(a\) limit of the early-universe, we have
\[ \dot{a} = F/a^{(n-2)/2} \quad \text{or} \quad a^{(n-2)/2} da = F dt \] (11.420)
which we integrate to \(a \sim t^{2/n}\) so that \(\dot{a} \sim t^{2/n-1}\). Now (11.417) says that
\[ |\Omega - 1| = \frac{1}{a^2} \propto t^{2-4/n} = \begin{cases} t^{2/3} \quad \text{radiation} \\ t^{2/3} \quad \text{matter} \end{cases} \] (11.421)
Thus, \(\Omega\) deviated from unity faster than \(t^{2/3}\) during the eras of radiation
and matter. At this rate, the inequality \(|\Omega_0 - 1| < 0.036\) could last 13.8
billion years only if \(\Omega\) at \(t = 1\) second had been unity to within six parts in
\(10^{14}\). The only known explanation for such early flatness is inflation.

Manipulating our relation (11.417) between \(\Omega\) and \(aH\), we see that
\[ (aH)^2 = \frac{k}{\Omega - 1}. \] (11.422)
So \(\Omega > 1\) implies \(k = 1\), and \(\Omega < 1\) implies \(k = -1\), and as \(\Omega \to 1\) the
product $aH \to \infty$, which is the essence of flatness since curvature vanishes as the scale factor $a \to \infty$. Imagine blowing up a balloon.

Staying for the moment with a universe without inflation and with an energy density composed of radiation and/or matter, we note that the first-order equation (11.419) in the form $\dot{a}^2 = F^2/a^{n-2} - k$ tells us that for a closed ($k = 1$) universe, in the limit $a \to \infty$ we’d have $\dot{a}^2 \to -1$ which is impossible. Thus a closed universe eventually collapses, which is incompatible with the flatness (11.422) implied by the present value $\Omega_0 = 1.000 \pm 0.036$.

The first-order equation Friedmann (11.412) tells us that $\rho a^2 \geq 3k/8\pi G$. So in a closed universe ($k = 1$), the energy density $\rho$ is positive and increases without limit as $a \to 0$ as in a collapse. In open ($k < 0$) and flat ($k = 0$) universes, the same Friedmann equation (11.412) in the form $\dot{a}^2 = 8\pi G \rho a^2 / 3 - k$ tells us that if $\rho$ is positive, then $\dot{a}^2 > 0$, which means that $\dot{a}$ never vanishes. Hubble told us that $\dot{a} > 0$ now. So if our universe is open or flat, then it always expands.

Due to the expansion of the universe, the wave-length of radiation grows with the scale factor $a(t)$. A photon emitted at time $t$ and scale factor $a(t)$ with wave-length $\lambda(t)$ will be seen now at time $t_0$ and scale factor $a(t_0)$ to have a longer wave-length $\lambda(t_0)$

$$\frac{\lambda(t_0)}{\lambda(t)} = \frac{a(t_0)}{a(t)} = z + 1$$

in which the redshift $z$ is the ratio

$$z = \frac{\lambda(t_0) - \lambda(t)}{\lambda(t)} = \frac{\Delta \lambda}{\lambda}.$$  

(11.424)

Now $H = \dot{a}/a = da/(a dt)$ implies $dt = da/(aH)$, and $z = a_0/a - 1$ implies $dz = -a_0 da/a^2$, so we find

$$dt = -\frac{dz}{(1 + z)H(z)}$$

(11.425)

which relates time intervals to redshift intervals. An on-line calculator is available for macroscopic intervals (Wright, 2006).

### 11.49 Model Cosmologies

The 0-component of the energy-momentum conservation law (11.372) is

$$0 = (T^a_0)_a = \partial_a T^a_0 + \Gamma^a_{ac} T^c_0 - T^a_i \Gamma^i_0 a$$

$$= -\partial_0 T^a_0 - \Gamma^a_{0c} T^c_0 - g^{cc} T_{cc} \Gamma^c_0$$

$$= -\rho - 3\frac{\dot{a}}{a} \rho - 3p \frac{\dot{a}}{a} = -\dot{\rho} - 3\frac{\dot{a}}{a} (\rho + p).$$

(11.426)
or
\[ \frac{d\rho}{da} = -\frac{3}{a} (\rho + p). \quad (11.427) \]

The energy density \( \rho \) is composed of fractions \( \rho_k \) each contributing its own partial pressure \( p_k \) according to its own **equation of state**

\[ p_k = w_k \rho_k \quad (11.428) \]

in which \( w_k \) is a constant. In terms of these components, the energy-momentum conservation law (11.427) is

\[ \sum_k \frac{d\rho_k}{da} = -\frac{3}{a} \sum_k (1 + w_k) \rho_k \quad (11.429) \]

with solution

\[ \rho = \sum_k \rho_k \left( \frac{a}{a_0} \right)^{3(1 + w_k)} = \sum_k \rho_k \left( \frac{a}{a_0} \right)^{3(1 + \frac{p_k}{\rho_k})}. \quad (11.430) \]

Simple cosmological models take the energy density and pressure each to have a single component with \( p = w \rho \), and in this case

\[ \rho = \bar{\rho} \left( \frac{a}{a_0} \right)^{3(1 + w)} = \bar{\rho} \left( \frac{a}{a_0} \right)^{3(1 + \frac{p}{\rho})}. \quad (11.431) \]

**Example 11.25** (\( w = -1/3 \), No Acceleration) If \( w = -1/3 \), then \( p = w \rho = -\rho/3 \) and \( \rho + 3p = 0 \). The second-order Friedmann equation (11.410) then tells us that \( \ddot{a} = 0 \). The scale factor does not accelerate.

To find its constant speed, we use its equation of state (11.431)

\[ \rho = \bar{\rho} \left( \frac{a}{a_0} \right)^{3(1 + w)} = \bar{\rho} \left( \frac{a}{a_0} \right)^2 \quad (11.432) \]

Now all the terms in Friedmann’s first-order equation (11.412) have a common factor of \( 1/a^2 \) which cancels leaving us with the square of the constant speed

\[ \dot{a}^2 = \frac{8\pi G}{3} \bar{\rho} \dot{a}^2 - k. \quad (11.433) \]

Incidentally, \( \bar{\rho} \dot{a}^2 \) must exceed \( 3k/8\pi G \). The scale factor grows linearly with time as

\[ a(t) = \left( \frac{8\pi G}{3} \bar{\rho} \dot{a}^2 - k \right)^{1/2} (t - t_0) + a(t_0). \quad (11.434) \]
Setting $t_0 = 0$ and $a(0) = 0$, we use the definition of the Hubble parameter $H = \dot{a}/a$ to write the constant linear growth $\dot{a}$ as $aH$ and the time as
\[ t = \int_0^a da'/a' H = (1/aH) \int_0^a da' = 1/H. \quad (11.435) \]
So in a universe without acceleration, the age of the universe is the inverse of the Hubble rate. For our universe, the present Hubble time is $1/H_0 = 14.5$ billion years, which isn’t far from the actual age of $13.817 \pm 0.048$ billion years. Presumably, a slower Hubble rate during the era of matter compensates for the higher rate during the era of dark energy.

**Example 11.26 ($w = -1$, Inflation)**  Inflation occurs when the ground state of the theory has a positive and constant energy density $\rho > 0$ that dwarfs the energy densities of the matter and radiation. The internal energy of the universe then is proportional to its volume $U = \rho V$, and the pressure $p$ as given by the thermodynamic relation
\[ p = -\frac{\partial U}{\partial V} = -\rho \quad (11.436) \]
is negative. The equation of state (11.428) tells us that in this case $w = -1$. The second-order Friedmann equation (11.410) becomes
\[ \frac{\dot{a}}{a} = -\frac{4\pi G}{3} (\rho + 3p) = \frac{8\pi G \rho}{3} \equiv g^2 \quad (11.437) \]
By it and the first-order Friedmann equation (11.412) and by choosing $t = 0$ as the time at which the scale factor $a$ is minimal, one may show (exercise 11.37) that in a closed ($k = 1$) universe
\[ a(t) = \frac{\cosh g t}{g}. \quad (11.438) \]
Similarly in an open ($k = -1$) universe with $a(0) = 0$, we have
\[ a(t) = \frac{\sinh g t}{g}. \quad (11.439) \]
Finally in a flat ($k = 0$) expanding universe, the scale factor is
\[ a(t) = a(0) \exp(g t). \quad (11.440) \]
Studies of the cosmic microwave background radiation suggest that inflation did occur in the very early universe—possibly on a time scale as short as $10^{-35}$ s. What is the origin of the vacuum energy density $\rho$ that drove
11.49 Model Cosmologies

inflation? Current theories attribute it to the assumption by at least one scalar field $\phi$ of a mean value $\langle \phi \rangle$ different from the one $\langle 0 | \phi | 0 \rangle$ that minimizes the energy density of the vacuum. When $\langle \phi \rangle$ settled to $\langle 0 | \phi | 0 \rangle$, the vacuum energy was released as radiation and matter in a Big Bang.

Example 11.27 ($w = 1/3$, The Era of Radiation) Until a redshift of $z = 3400$ or 50,000 years after inflation, our universe was dominated by radiation (Frieman et al., 2008). During The First Three Minutes (Weinberg, 1988) of the era of radiation, the quarks and gluons formed hadrons, which decayed into protons and neutrons. As the neutrons decayed ($\tau = 885.7$ s), they and the protons formed the light elements—principally hydrogen, deuterium, and helium in a process called big-bang nucleosynthesis.

We can guess the value of $w$ for radiation by noticing that the energy-momentum tensor of the electromagnetic field (in suitable units)

$$T^{ab} = F^a_c F^{bc} - \frac{1}{4} g^{ab} F_{cd} F^{cd}$$

(11.441)

is traceless

$$T = T^a_a = F^a_c F^c_a - \frac{1}{4} \delta^a_a F_{cd} F^{cd} = 0.$$ 

(11.442)

But by (11.409) its trace must be $T = 3p - \rho$. So for radiation $p = \rho/3$ and $w = 1/3$. The relation (11.431) between the energy density and the scale factor then is

$$\rho = p \left( \frac{\pi}{a} \right)^4.$$ 

(11.443)

The energy drops both with the volume $a^3$ and with the scale factor $a$ due to a redshift; so it drops as $1/a^4$. Thus the quantity

$$f^2 \equiv \frac{8 \pi G \rho a^4}{3}$$

(11.444)

is a constant. The Friedmann equations (11.410 & 11.411) now are

$$\frac{\ddot{a}}{a} = -\frac{4 \pi G}{3} (\rho + 3p) = -\frac{8 \pi G \rho}{3} \quad \text{or} \quad \ddot{a} = -\frac{f^2}{a^3}$$

(11.445)

and

$$a^2 + k = \frac{f^2}{a^2}.$$ 

(11.446)
Tensors and Local Symmetries

With calendars chosen so that \(a(0) = 0\), this last equation (11.446) tells us that for a flat universe \((k = 0)\)

\[
a(t) = (2f t)^{1/2}
\]

while for a closed universe \((k = 1)\)

\[
a(t) = \sqrt{f^2 - (t - f)^2}
\]

and for an open universe \((k = -1)\)

\[
a(t) = \sqrt{(t + f)^2 - f^2}
\]

as we saw in (6.422). The scale factor (11.448) of a closed universe of radiation has a maximum \(a = f\) at \(t = f\) and falls back to zero at \(t = 2f\).  

Example 11.28 \((w = 0, \text{The Era of Matter})\) A universe composed only of dust or non-relativistic collisionless matter has no pressure. Thus \(p = w\rho = 0\) with \(\rho \neq 0\), and so \(w = 0\). Conservation of energy (11.430), or equivalently (11.431), implies that the energy density falls with the volume

\[
\rho = \rho \left( \frac{a}{a_0} \right)^3.
\]

As the scale factor \(a(t)\) increases, the matter energy density, which falls as \(1/a^3\), eventually dominates the radiation energy density, which falls as \(1/a^4\). This happened in our universe about 50,000 years after inflation at a temperature of \(T = 9,400\) K or \(kT = 0.81\) eV. Were baryons most of the matter, the era of radiation dominance would have lasted for a few hundred thousand years. But the kind of matter that we know about, which interacts with photons, is only about 17% of the total; the rest—an unknown substance called dark matter—shortened the era of radiation dominance by nearly 2 million years.

Since \(\rho \propto 1/a^3\), the quantity

\[
m^2 = \frac{4\pi G\rho a^3}{3}
\]

is a constant. For a matter-dominated universe, the Friedmann equations (11.410 & 11.411) then are

\[
\frac{\ddot{a}}{a} = - \frac{4\pi G}{3} (\rho + 3p) = - \frac{4\pi G\rho}{3} \quad \text{or} \quad \ddot{a} = - \frac{m^2}{a^2}
\]
\[ a^2 + k = 2m^2 / a. \] (11.453)

For a flat universe, \( k = 0 \), we get
\[ a(t) = \left[ \frac{3m}{\sqrt{2}} t \right]^{2/3}. \] (11.454)

For a closed universe, \( k = 1 \), we use example 6.46 to integrate
\[ \dot{a} = \sqrt{2m^2 / a - 1} \] (11.455)
to
\[ t - t_0 = -\sqrt{a(2m^2 - a)} - m^2 \arcsin(1 - a/m^2). \] (11.456)

With a suitable calendar and choice of \( t_0 \), one may parametrize this solution in terms of the development angle \( \phi(t) \) as
\[ a(t) = m^2 [1 - \cos \phi(t)] \]
\[ t = m^2 [\phi(t) - \sin \phi(t)]. \] (11.457)

For an open universe, \( k = -1 \), we use example 6.47 to integrate
\[ \dot{a} = \sqrt{2m^2 / a + 1} \] (11.458)
to
\[ t - t_0 = [a(2m^2 + a)]^{1/2} - m^2 \ln \left\{ 2 \left[ a(2m^2 + a) \right]^{1/2} + 2a + 2m^2 \right\}. \] (11.459)

The conventional parametrization is
\[ a(t) = m^2 [\cosh \phi(t) - 1] \]
\[ t = m^2 [\sinh \phi(t) - \phi(t)]. \] (11.460)

**Transparency**: Some 380,000 years after inflation at a redshift of \( z = 1090 \), the universe had cooled to about \( T = 3000 \) K or \( kT = 0.26 \) eV—a temperature at which less than 1% of the hydrogen is ionized. Ordinary matter became a gas of neutral atoms rather than a plasma of ions and electrons, and the universe suddenly became transparent to light. Some scientists call this moment of last scattering or first transparency recombination.
Example 11.29 (w = -1, The Era of Dark Energy)  About 10.3 billion years after inflation at a redshift of z = 0.30, the matter density falling as $1/a^3$ dropped below the very small but positive value of the energy density $\rho_v = (2.23 \text{ meV})^4$ of the vacuum. The present time is 13.817 billion years after inflation. So for the past 3 billion years, this constant energy density, called dark energy, has accelerated the expansion of the universe approximately as (11.439)

$$a(t) = a(t_m) \exp \left( (t - t_m) \sqrt{8\pi G \rho_v/3} \right)$$

in which $t_m = 10.3 \times 10^9$ years.

Observations and measurements on the largest scales indicate that the universe is flat: $k = 0$. So the evolution of the scale factor $a(t)$ is given by the $k = 0$ equations (11.440, 11.447, 11.454, & 11.461) for a flat universe. During the brief era of inflation, the scale factor $a(t)$ grew as (11.440)

$$a(t) = a(0) \exp \left( t \sqrt{8\pi G \rho_i/3} \right)$$

in which $\rho_i$ is the positive energy density that drove inflation.

During the 50,000-year era of radiation, $a(t)$ grew as $\sqrt{t}$ as in (11.447)

$$a(t) = \left( \frac{2}{t_i} \right)^{1/2} + a(t_i)$$

where $t_i$ is the time at the end of inflation, and $t_r'$ is any time during the era of radiation. During this era, the energy of highly relativistic particles dominated the energy density, and $\rho a^4 \propto T^4 a^4$ was approximately constant, so that $T(t) \propto 1/a(t) \propto 1/\sqrt{t}$. When the temperature was in the range $10^{12} > T > 10^{10} \text{ K}$ or $m_{\mu} c^2 > k T > m_e c^2$, where $m_{\mu}$ is the mass of the muon and $m_e$ that of the electron, the radiation was mostly electrons, positrons, photons, and neutrinos, and the relation between the time $t$ and the temperature $T$ was (Weinberg, 2010, ch. 3)

$$t = 0.994 \text{ sec} \times \left[ \frac{10^{10} \text{ K}}{T} \right]^2 + \text{ constant}. \quad (11.464)$$

By $10^9$ K, the positrons had annihilated with electrons, and the neutrinos fallen out of equilibrium. Between $10^9$ K and $10^6$ K, when the energy density of nonrelativistic particles became relevant, the time-temperature relation was (Weinberg, 2010, ch. 3)

$$t = 1.78 \text{ sec} \times \left[ \frac{10^{10} \text{ K}}{T} \right]^2 + \text{ constant}'. \quad (11.465)$$
During the 10.3 billion years of the matter era, \( a(t) \) grew as (11.454)

\[
a(t) = \left[ (t - t_r) \sqrt{3\pi G \rho (t_m')} a(t_m') + a^{3/2}(t_r) \right]^{2/3} + a(t_r) \tag{11.466}
\]

where \( t_r \) is the time at the end of the radiation era, and \( t_m' \) is any time in the matter era. By 380,000 years, the temperature had dropped to 3000 K, the universe had become transparent, and the CMBR had begun to travel freely.

Over the past 3 billion years of the era of vacuum dominance, \( a(t) \) has been growing exponentially (11.461)

\[
a(t) = a(t_m) \exp \left( (t - t_m) \sqrt{8\pi G \rho_v/3} \right) \tag{11.467}
\]

in which \( t_m \) is the time at the end of the matter era, and \( \rho_v \) is the density of dark energy, which while vastly less than the energy density \( \rho_i \) that drove inflation, currently amounts to 68.5\% of the total energy density.

### 11.50 Yang-Mills Theory

The gauge transformation of an abelian gauge theory like electrodynamics multiplies a single charged field by a space-time-dependent phase factor \( \phi'(x) = \exp(iq\theta(x)) \phi(x) \). Yang and Mills generalized this gauge transformation to one that multiplies a vector \( \phi \) of matter fields by a space-time dependent unitary matrix \( U(x) \)

\[
\phi'(x) = \sum_{b=1}^{n} U_{ab}(x) \phi_b(x) \quad \text{or} \quad \phi'(x) = U(x) \phi(x) \tag{11.468}
\]

and showed how to make the action of the theory invariant under such non-abelian gauge transformations. (The fields \( \phi \) are scalars for simplicity.)

Since the matrix \( U \) is unitary, inner products like \( \phi^\dagger(x) \phi(x) \) are automatically invariant

\[
\left( \phi^\dagger(x) \phi(x) \right)' = \phi^\dagger(x) U^\dagger(x) U(x) \phi(x) = \phi^\dagger(x) \phi(x) \tag{11.469}
\]

But inner products of derivatives \( \partial^i \phi^\dagger \partial_i \phi \) are not invariant because the derivative acts on the matrix \( U(x) \) as well as on the field \( \phi(x) \).

Yang and Mills made derivatives \( D_i \phi \) that transform like the fields \( \phi \)

\[
(D_i \phi)' = U D_i \phi. \tag{11.470}
\]

To do so, they introduced gauge-field matrices \( A_i \) that play the role of
11.51 Gauge Theory and Vectors

with \( \phi^\dagger \phi = m^2/\lambda \) so as to minimize their potential energy density \( V(\phi) \). Their kinetic action \((D^i\phi)^\dagger D_i\phi = (\partial^i \phi + A^i \phi)^\dagger (\partial_i \phi + A_i \phi)\) then is in effect \( \phi^\dagger A^i A_i \phi_0 \). The gauge-field matrix \( A_{ab} = i t^{\alpha}_{ab} A_{\alpha}^i \) is a linear combination of the generators \( t^{\alpha} \) of the gauge group. So the action of the scalar fields contains the term \( \phi^\dagger A^i A_i \phi = -M_{\alpha \beta} A_{\alpha}^a A_{\beta}^b \). This **Higgs mechanism** gives masses to those linear combinations \( b^\beta_i A^\beta_i \) of the gauge fields for which \( M_{\alpha \beta} b^\beta_i = m_i^2 \).

The Higgs mechanism also gives masses to the fermions. The mass term \( m \) in the Yang-Mills-Dirac action is replaced by something like \( c \phi \) in which \( c \) is a constant, different for each fermion. In the vacuum and at low temperatures, each fermion acquires as its mass \( c \phi_0 \). On 4 July 2012, physicists at CERN’s Large Hadron Collider announced the discovery of a Higgs-like particle with a mass near 126 GeV/c^2 (Peter Higgs 1929–).

11.51 Gauge Theory and Vectors

This section is optional on a first reading.

We can formulate Yang-Mills theory in terms of vectors as we did relativity. To accomodate noncompact groups, we will generalize the unitary matrices \( U(x) \) of the Yang-Mills gauge group to nonsingular matrices \( V(x) \) that act on \( n \) matter fields \( \psi^a(x) \) as

\[
\psi'^a(x) = \sum_{a=1}^n V^a_b(x) \psi^b(x). \tag{11.480}
\]

The field

\[
\Psi(x) = \sum_{a=1}^n e_a(x) \psi^a(x) \tag{11.481}
\]

will be gauge invariant \( \Psi'(x) = \Psi(x) \) if the vectors \( e_a(x) \) transform as

\[
e'_a(x) = \sum_{b=1}^n e_b(x) V^{-1b}_a(x). \tag{11.482}
\]

In what follows, we will sum over repeated indices from 1 to \( n \) and often will suppress explicit mention of the space-time coordinates. In this compressed notation, the field \( \Psi \) is gauge invariant because

\[
\Psi' = e'_a \psi'^a = e_b V^{-1b}_a V^a_c \psi'^c = e_b \delta^b_c \psi'^c = e_b \psi'^b = \Psi \tag{11.483}
\]

which is \( e'^T \psi' = e^T V^{-1} V \psi = e^T \psi \) in matrix notation.
\( d\rho, d\phi, \) and \( dz, \) and so derive the expressions (11.169) for the orthonormal basis vectors \( \hat{\rho}, \hat{\phi}, \) and \( \hat{z}. \)

11.14 Similarly, derive (11.175) from (11.174).

11.15 Use the definition (11.191) to show that in flat 3-space, the dual of the Hodge dual is the identity: \( **dx^i = dx^i \) and \( **(dx^i \wedge dx^k) = dx^i \wedge dx^k. \)

11.16 Use the definition of the Hodge star (11.202) to derive (a) two of the four identities (11.203) and (b) the other two.

11.17 Show that Levi-Civita’s 4-symbol obeys the identity (11.207).

11.18 Show that 
\[
\varepsilon^{\ell mn} \varepsilon_{\ell mn} = 2 \delta^p_\ell.
\]

11.19 Show that 
\[
\varepsilon_{k \ell mn} \varepsilon^{\ell mn} = 3! \delta^p_k.
\]

11.20 Using the formulas (11.175) for the basis vectors of spherical coordinates in terms of those of rectangular coordinates, compute the derivatives of the unit vectors \( \hat{r}, \hat{\theta}, \) and \( \hat{\phi} \) with respect to the variables \( r, \theta, \) and \( \phi. \) (b) Using the formulas of (a) and our expression (6.28) for the gradient in spherical coordinates, derive the formula (11.297) for the laplacian \( \nabla \cdot \nabla. \)

11.21 Consider the torus with coordinates \( \theta, \phi \) labeling the arbitrary point
\[
p = (\cos \phi (R + r \sin \theta), \sin \phi (R + r \sin \theta), r \cos \theta)
\]
in which \( R > r. \) Both \( \theta \) and \( \phi \) run from 0 to \( 2\pi. \) (a) Find the basis vectors \( e_\theta \) and \( e_\phi. \) (b) Find the metric tensor and its inverse.

11.22 For the same torus, (a) find the dual vectors \( e^\theta \) and \( e^\phi \) and (b) find the nonzero connections \( \Gamma^i_{jk} \) where \( i, j, \) & \( k \) take the values \( \theta \& \phi. \)

11.23 For the same torus, (a) find the two Christoffel matrices \( \Gamma_\theta \) and \( \Gamma_\phi, \)
(b) find their commutator \( [\Gamma_\theta, \Gamma_\phi], \) and (c) find the elements \( R^\theta_{\theta\theta\theta}, R^\phi_{\phi\phi\phi}, R^\theta_{\theta\phi\phi}, \) and \( R^\phi_{\phi\theta\theta} \) of the curvature tensor.

11.24 Find the curvature scalar \( R \) of the torus with points (11.505). **Hint:** In these four problems, you may imitate the corresponding calculation for the sphere in Sec. 11.42.

11.25 By differentiating the identity \( g^{ik} g_{k\ell} = \delta^i_\ell, \) show that \( \delta g^{ik} = -g^{is} g^{kt} \delta g_{st} \) or equivalently that \( dg^{ik} = -g^{is} g^{kt} dg_{st}. \)

11.26 Just to get an idea of the sizes involved in black holes, imagine an isolated sphere of matter of uniform density \( \rho \) that as an initial condition is all at rest within a radius \( r_b. \) Its radius will be less than its Schwarzschild radius if
\[
r_b < \frac{2MG}{c^2} = 2 \left( \frac{4}{3} \pi r_b^3 \rho \right) \frac{G}{c^2}.
\]

If the density \( \rho \) is that of water under standard conditions (1 gram per
cc), for what range of radii \( r_b \) might the sphere be or become a black hole? Same question if \( \rho \) is the density of dark energy.

11.27 For the points (11.392), derive the metric (11.395) with \( k = 1 \). Don’t forget to relate \( d\chi \) to \( dr \).

11.28 For the points (11.393), derive the metric (11.395) with \( k = 0 \).

11.29 For the points (11.394), derive the metric (11.395) with \( k = -1 \). Don’t forget to relate \( d\chi \) to \( dr \).

11.30 Suppose the constant \( k \) in the Roberson-Walker metric (11.391 or 11.395) is some number other than 0 or \( \pm 1 \). Find a coordinate transformation such that in the new coordinates, the Roberson-Walker metric has \( k = k/|k| = \pm 1 \). Hint: You also can change the scale factor \( a \).

11.31 Derive the affine connections in Eq.(11.399).

11.32 Derive the affine connections in Eq.(11.400).

11.33 Derive the affine connections in Eq.(11.401).


11.35 Assume there had been no inflation, no era of radiation, and no dark energy. In this case, the magnitude of the difference \( |\Omega - 1| \) would have increased as \( t^{2/3} \) over the past 13.8 billion years. Show explicitly how close to unity \( \Omega \) would have had to have been at \( t = 1 \) s so as to satisfy the observational constraint \( |\Omega_0 - 1| < 0.036 \) on the present value of \( \Omega \).

11.36 Derive the relation (11.431) between the energy density \( \rho \) and the Robertson-Walker scale factor \( a(t) \) from the conservation law (11.427) and the equation of state \( p = w\rho \).

11.37 Use the Friedmann equations (11.410 & 11.412) for constant \( \rho = -p \) and \( k = 1 \) to derive (11.438) subject to the boundary condition that \( a(t) \) has its minimum at \( t = 0 \).

11.38 Use the Friedmann equations (11.410 & 11.412) with \( w = -1, \rho \) constant, and \( k = -1 \) to derive (11.439) subject to the boundary condition that \( a(0) = 0 \).

11.39 Use the Friedmann equations (11.410 & 11.412) with \( w = -1, \rho \) constant, and \( k = 0 \) to derive (11.440). Show why a linear combination of the two solutions (11.440) does not work.

11.40 Use the conservation equation (11.444) and the Friedmann equations (11.410 & 11.412) with \( w = 1/3, k = 0, \) and \( a(0) = 0 \) to derive (11.447).

11.41 Show that if the matrix \( U(x) \) is nonsingular, then

\[
(\partial_i U) U^{-1} = -U \partial_i U^{-1}. \tag{11.507}
\]

11.42 The gauge-field matrix is a linear combination \( A_k = -ig t^b A^b_k \) of the
There are two ways of thinking about differential forms. The Russian literature views a manifold as embedded in $\mathbb{R}^n$ and so is somewhat more straightforward. We will discuss it first.

**The Russian Way:** Suppose $x(t)$ is a curve with $x(0) = x$ on some manifold $M$, and $f(x(t))$ is a smooth function $f : \mathbb{R}^n \to \mathbb{R}$ that maps points $x(t)$ into numbers. Then the **differential** $df(\dot{x}(t))$ maps $\dot{x}(t)$ at $x$ into

$$df \left( \frac{d}{dt} x(t) \right) = \frac{d}{dt} f(x(t)) = \sum_{j=1}^{n} \dot{x}(t)_j \frac{\partial f(x(t))}{\partial x_j} = \dot{x}(t) \cdot \nabla f(x(t)) \quad (12.18)$$

all at $t = 0$. As physicists, we think of $df$ as a number—the change in the function $f(x)$ when its argument $x$ is changed by $dx$. Russian mathematicians think of $df$ as a linear map of tangent vectors $\dot{x}$ at $x$ into numbers. Since this map is linear, we may multiply the definition (12.18) by $dt$ and arrive at the more familiar formula

$$dt df \left( \frac{d}{dt} x(t) \right) = df \left( dt \frac{d}{dt} x(t) \right) = df (dx(t)) = dx(t) \cdot \nabla f(x(t)) \quad (12.19)$$

all at $t = 0$. So

$$df(dx) = dx \cdot \nabla f. \quad (12.20)$$

is the physicist’s $df$.

Since the differential $df$ is a linear map of vectors $\dot{x}(0)$ into numbers, it is a 1-form; since it is defined on vectors like $\dot{x}(0)$, it is a **differential 1-form**. The term *differential 1-form* underscores the fact that the actual value of the differential $df$ depends upon the vector $\dot{x}(0)$ and the point $x = x(0)$. Mathematicians call the space of vectors $\dot{x}(0)$ at the point $x = x(0)$ the **tangent space** $T_M x$. They say $df$ is a smooth map of the **tangent bundle** $TM$, which is the union of the tangent spaces for all points $x$ in the manifold $M$, to the real line, so $df : TM \to \mathbb{R}$.

In the special case in which $f(x) = x_i(x) = x_i$, the differential $dx_i(\dot{x}(t))$ by (12.18) is

$$dx_i(\dot{x}(t)) = \sum_{j=1}^{n} \dot{x}_j(t) \frac{\partial x_i(x)}{\partial x_j} = \sum_{j=1}^{n} \dot{x}_j(t) \frac{\partial x_i}{\partial x_j} = \sum_{j=1}^{n} \dot{x}_j(t) \delta_{ij} = \dot{x}_i(t). \quad (12.21)$$

These $dx_i$’s are the **basic differentials**. Using $A$ for the vector $\dot{x}(t)$, we find from our definition (12.18) that

$$dx_i(A) = \sum_{j=1}^{n} A_j \frac{\partial x_i}{\partial x_j} = \sum_{j=1}^{n} A_j \delta_{ij} = A_i \quad (12.22)$$
sets of basic differentials. Then by applying the formula (12.24) to the function \( y_k(x) \), we get

\[
dy_k = \sum_{j=1}^{n} \frac{\partial y_k(x)}{\partial x_j} \, dx_j
\]  

(12.32)

which is the familiar rule for changing variables.

The **most general differential 1-form** \( \omega \) on the space \( \mathbb{R}^n \) with coordinates \( x_1 \ldots x_n \) is a linear combination of the basic differentials \( dx_i \) with coefficients \( a_i(x) \) that are smooth functions of \( x = (x_1, \ldots, x_n) \)

\[
\omega = a_1(x) \, dx_1 + \ldots + a_n(x) \, dx_n.
\]  

(12.33)

The **basic differential 2-forms** are \( dx_i \wedge dx_k \) defined as

\[
dx_i \wedge dx_k(A, B) = \begin{vmatrix} dx_i(A) & dx_k(A) \\ dx_i(B) & dx_k(B) \end{vmatrix} = \begin{vmatrix} A_i & A_k \\ B_i & B_k \end{vmatrix} = A_iB_k - A_kB_i.
\]  

(12.34)

So in particular

\[
dx_i \wedge dx_i = 0.
\]  

(12.35)

The **basic differential k-forms** \( dx_1 \wedge \cdots \wedge dx_k \) are defined as

\[
dx_1 \wedge \ldots \wedge dx_k(A_1, \ldots A_k) = \begin{vmatrix} dx_1(A_1) & \ldots & dx_k(A_1) \\ \vdots & \ddots & \vdots \\ dx_1(A_k) & \ldots & dx_k(A_k) \end{vmatrix} = \begin{vmatrix} A_{11} & \ldots & A_{1k} \\ \vdots & \ddots & \vdots \\ A_{k1} & \ldots & A_{kk} \end{vmatrix}.
\]  

(12.36)

**Example 12.4** \( (dx_3 \wedge dr^2) \) If \( r^2 = x_1^2 + x_2^2 + x_3^2 \), then \( dr^2 \) is

\[
dr^2 = 2(x_1 \, dx_1 + x_2 \, dx_2 + x_3 \, dx_3)
\]  

(12.37)

and the differential 2-form \( \omega = dx_3 \wedge dr^2 \) is

\[
\omega = dx_3 \wedge 2(x_1 \, dx_1 + x_2 \, dx_2 + x_3 \, dx_3) = 2x_1 \, dx_3 \wedge dx_1 + 2x_2 \, dx_3 \wedge dx_2
\]  

(12.38)

since in view of (12.35) \( dx_3 \wedge dx_3 = 0 \). So the value of the 2-form \( \omega \) on the vectors \( A = (1, 2, 3) \) and \( B = (2, 1, 1) \) at the point \( x = (3, 0, 3) \) is

\[
\omega(A, B) = 2x_1 \, dx_3 \wedge dx_1(A, B) = 6 \begin{vmatrix} dx_3(A) & dx_1(A) \\ dx_3(B) & dx_1(B) \end{vmatrix} = 6 \begin{vmatrix} 3 & 1 \\ 1 & 2 \end{vmatrix} = 30.
\]  

(12.39)

On the vectors, \( C = (1, 0, 0) \) and \( D = (0, 0, 1) \) at \( x = (2, 3, 4) \), this 2-form has the value \( \omega(C, D) = -4 \).
Forms

\( \Delta q_i \), one may show (exercise 12.10) that this sum of areas remains constant

\[
\frac{d}{dt} \omega^1(\delta p, \delta q; \Delta p, \Delta q) = 0 \tag{12.77}
\]

along the trajectories in phase space (Gutzwiller, 1990, chap. 7).

**Example 12.12** (The Curl) We saw in example 12.7 that the 1-form (12.50) of a vector field \( \mathbf{A} \) is

\[
\omega^A = A_1 \, dx_1 + A_2 \, dx_2 + A_3 \, dx_3
\]

in which the \( h_k \)'s are those that determine (12.44) the squared length \( ds^2 = h_k^2 \, dx_k^2 \) of the triply orthogonal coordinate system with unit vectors \( \hat{e}_1, \hat{e}_2, \hat{e}_3 \). So the exterior derivative of the 1-form \( \omega^A \) is

\[
d\omega^A = \sum_{i,k=1}^3 \partial_k(A_i \, h_i) \, dx_k \wedge dx_i
\]

\[
= \left[ \frac{\partial(A_3 \, h_3)}{\partial x_2} - \frac{\partial(A_2 \, h_2)}{\partial x_3} \right] \, dx_2 \wedge dx_3 \tag{12.78}
\]

\[
+ \left[ \frac{\partial(A_2 \, h_2)}{\partial x_1} - \frac{\partial(A_1 \, h_1)}{\partial x_2} \right] \, dx_1 \wedge dx_2
\]

\[
+ \left[ \frac{\partial(A_1 \, h_1)}{\partial x_3} - \frac{\partial(A_3 \, h_3)}{\partial x_1} \right] \, dx_3 \wedge dx_1 \equiv \omega \nabla \times \mathbf{A}. \tag{12.79}
\]

Comparison with Eq. (12.52) shows that the curl of \( \mathbf{A} \) is

\[\nabla \times \mathbf{A} = \frac{1}{h_2 h_3} \begin{vmatrix} h_1 \hat{e}_1 & h_2 \hat{e}_2 & h_3 \hat{e}_3 \\ \partial_1 & \partial_2 & \partial_3 \\ A_1 h_1 & A_2 h_2 & A_3 h_3 \end{vmatrix} \]

\[
= \frac{1}{h_1 h_2 h_3} \sum_{i,j,k=1}^3 \epsilon_{ijk} h_i \hat{e}_i \frac{\partial(A_k h_k)}{\partial x_j} \tag{12.80}
\]

as we saw in (11.240). This formula gives our earlier expressions for the curl in cylindrical and spherical coordinates (11.241 & 11.242).

**Example 12.13** (The Divergence) We have seen in equations (12.48, 12.49, & 12.52) that the 2-form \( \omega^A(U, V) = A \cdot (U \times V) \) of the vector field \( \mathbf{A} = A_1 \hat{e}_1 + A_2 \hat{e}_2 + A_3 \hat{e}_3 \) is

\[
\omega^2_\mathbf{A} = A_1 \, h_2 \, h_3 \, dx_2 \wedge dx_3 + A_2 \, h_3 \, h_1 \, dx_3 \wedge dx_1 + A_3 \, h_1 \, h_2 \, dx_1 \wedge dx_2. \tag{12.81}
\]
One way of computing \( P_B(n, p, N) \) for large \( N \) is to use Srinivasa Ramanujan’s correction (4.39) to Stirling’s formula \( N! \approx \sqrt{2\pi N} \frac{N}{e} \left(1 + \frac{1}{2N} + \frac{1}{8N^2}\right)^{1/6} \).

When \( N \) and \( N - n \), but not \( n \), are big, one may use (13.51) for \( N! \) and \((N - n)!\) in the formula (13.43) for \( P_B(n, p, N) \) and so may show (exercise 13.11) that

\[
P_B(n, p, N) \approx \frac{(pN)^n}{n!} q^{N-n} R_2(n, N)
\]

in which

\[
R_2(n, N) = \left(1 - \frac{n}{N}\right)^{n-1/2} \left(1 + \frac{1}{2N} + \frac{1}{8N^2}\right)^{1/6} 
\times \left[1 + \frac{1}{2(N - n)} + \frac{1}{8(N - n)^2}\right]^{-1/6}
\]

tends to unity as \( N \to \infty \) for any fixed \( n \).

When all three factorials in \( P_B(n, p, N) \) are huge, one may use Ramanujan’s approximation (13.51) to show (exercise 13.12) that

\[
P_B(n, p, N) \approx \sqrt{\frac{N}{2\pi n(N - n)}} \left(\frac{pN}{n}\right)^n \left(\frac{qN}{N - n}\right)^{N-n} R_3(n, N)
\]

where

\[
R_3(n, N) = \left(1 + \frac{1}{2n} + \frac{1}{8n^2}\right)^{-1/6} \left(1 + \frac{1}{2N} + \frac{1}{8N^2}\right)^{1/6} 
\times \left[1 + \frac{1}{2(N - n)} + \frac{1}{8(N - n)^2}\right]^{-1/6}
\]

tends to unity as \( N \to \infty \), \( N - n \to \infty \), and \( n \to \infty \).

Another way of coping with the unwieldy factorials in the binomial formula \( P_B(n, p, N) \) is to use limiting forms of (13.43) due to Poisson and to Gauss.

### 13.4 The Poisson Distribution

Poisson took the two limits \( N \to \infty \) and \( p = \langle n \rangle / N \to 0 \). So we let \( N \) and \( N - n \), but not \( n \), tend to infinity, and use (13.52) for the binomial distribution (13.43). Since \( R_2(n, N) \to 1 \) as \( N \to \infty \), we get

\[
P_B(n, p, N) \approx \frac{(pN)^n}{n!} q^{N-n} = \frac{\langle n \rangle^n}{n!} q^{N-n}.
\]
13.10 The Einstein-Nernst relation

If a particle of mass \( m \) carries an electric charge \( q \) and is exposed to an electric field \( E \), then in addition to viscosity \( -v/B \) and random buffeting \( f \), the constant force \( qE \) acts on it

\[
m \frac{dv}{dt} = -\frac{v}{B} + qE + f. \tag{13.138}
\]

The mean value of its velocity will then satisfy the differential equation

\[
\left\langle \frac{dv}{dt} \right\rangle = -\frac{\left\langle v \right\rangle}{\tau} + \frac{qE}{m} \tag{13.139}
\]

where \( \tau = mB \). A particular solution of this inhomogeneous equation is

\[
\left\langle v(t) \right\rangle_{pi} = \frac{q\tau E}{m} = qBE. \tag{13.140}
\]

The general solution of its homogeneous version is \( \left\langle v(t) \right\rangle_{gh} = A \exp(-t/\tau) \) in which the constant \( A \) is chosen to give \( \left\langle v(0) \right\rangle \) at \( t = 0 \). So by (6.13), the general solution \( \left\langle v(t) \right\rangle \) to equation (13.139) is (exercise 13.19) the sum of \( \left\langle v(t) \right\rangle_{pi} \) and \( \left\langle v(t) \right\rangle_{gh} \)

\[
\left\langle v(t) \right\rangle = qBE + \left[ \left\langle v(0) \right\rangle - qBE \right] e^{-t/\tau}. \tag{13.141}
\]

By applying the tricks of the previous section (13.9), one may show (exercise 13.20) that the variance of the position \( r \) about its mean \( \left\langle r(t) \right\rangle \) is

\[
\left\langle (r - \left\langle r(t) \right\rangle)^2 \right\rangle = \frac{6kT\tau^2}{m} \left( \frac{t}{\tau} - 1 + e^{-t/\tau} \right) \tag{13.142}
\]

where \( \left\langle r(t) \right\rangle = (q\tau^2E/m) \left( t/\tau - 1 + e^{-t/\tau} \right) \) if \( \left\langle r(0) \right\rangle = \left\langle v(0) \right\rangle = 0 \). So for times \( t \gg \tau \), this variance is

\[
\left\langle (r - \left\langle r(t) \right\rangle)^2 \right\rangle = \frac{6kT\tau t}{m}. \tag{13.143}
\]

Since the diffusion constant \( D \) is defined by (13.134) as

\[
\left\langle (r - \left\langle r(t) \right\rangle)^2 \right\rangle = 6Dt \tag{13.144}
\]

we arrive at the Einstein-Nernst relation

\[
D = BkT = \frac{qB}{q}kT = \frac{\mu}{q}kT \tag{13.145}
\]

in which the electric mobility is \( \mu = qB \).
If we substitute our formula (13.169) for $\langle v^2(t) \rangle$ into the expression (13.123) for the acceleration of $\langle r^2 \rangle$, then we get
\[
\frac{d^2 \langle r^2(t) \rangle}{dt^2} = -\frac{1}{\tau} \frac{d \langle r^2(t) \rangle}{dt} + 2e^{-2t/\tau} \langle v^2(0) \rangle + \frac{6kT}{m} \left( 1 - e^{-2t/\tau} \right). \tag{13.171}
\]
The solution with both $\langle r^2(0) \rangle = 0$ and $d\langle r^2(0) \rangle / dt = 0$ is (exercise 13.21)
\[
\langle r^2(t) \rangle = \langle v^2(0) \rangle \tau^2 \left( 1 - e^{-t/\tau} \right)^2 - \frac{3kT}{m} \tau^2 \left( 1 - e^{-t/\tau} \right) \left( 3 - e^{-t/\tau} \right) + \frac{6kT \tau}{m} t. \tag{13.172}
\]

### 13.12 Characteristic and Moment-Generating Functions

The Fourier transform (3.9) of a probability distribution $P(x)$ is its **characteristic function** $\tilde{P}(k)$ sometimes written as $\chi(k)$
\[
\tilde{P}(k) \equiv \chi(k) \equiv E[e^{ikx}] = \int e^{ikx} P(x) \, dx. \tag{13.173}
\]
The probability distribution $P(x)$ is the inverse Fourier transform (3.9)
\[
P(x) = \int e^{-ikx} \tilde{P}(k) \frac{dk}{2\pi}. \tag{13.174}
\]

**Example 13.10** (Gauss) The characteristic function of the gaussian
\[
P_G(x, \mu, \sigma) = \frac{1}{\sigma \sqrt{2\pi}} \exp \left( -\frac{(x - \mu)^2}{2\sigma^2} \right) \tag{13.175}
\]
is by (3.18)
\[
\tilde{P}_G(k, \mu, \sigma) = \frac{1}{\sigma \sqrt{2\pi}} \int \exp \left( ikx - \frac{(x - \mu)^2}{2\sigma^2} \right) dx = \frac{e^{i\mu k}}{\sigma \sqrt{2\pi}} \int \exp \left( ikx - \frac{x^2}{2\sigma^2} \right) dx = \exp \left( i\mu k - \frac{1}{2} \sigma^2 k^2 \right). \tag{13.176}
\]

For a discrete probability distribution $P_n$ the characteristic function is
\[
\chi(k) \equiv E[e^{ikx}] = \sum_n e^{ikn} P_n. \tag{13.177}
\]
The normalization of both continuous and discrete probability distributions implies that their characteristic functions satisfy $\tilde{P}(0) = \chi(0) = 1.$
Example 13.11 (Poisson) The Poisson distribution (13.58)

\[ P_P(n, \langle n \rangle) = \frac{\langle n \rangle^n}{n!} e^{-\langle n \rangle} \]  

(13.178)

has the characteristic function

\[ \chi(k) = \sum_{n=0}^{\infty} e^{ink} \frac{\langle n \rangle^n}{n!} e^{-\langle n \rangle} = e^{-\langle n \rangle} \sum_{n=0}^{\infty} \frac{(\langle n \rangle e^{ik})^n}{n!} = \exp \left[ \langle n \rangle \left( e^{ik} - 1 \right) \right]. \]  

(13.179)

The moment-generating function is the characteristic function evaluated at an imaginary argument

\[ M(k) \equiv E[e^{kx}] = P(-ik) = \chi(-ik). \]  

(13.180)

For a continuous probability distribution, it is

\[ M(k) = E[e^{kx}] = \int e^{kx} P(x) \, dx \]  

(13.181)

and for a discrete probability distribution, it is

\[ M(k) = E[e^{kx}] = \sum_n e^{kx_n} P_n. \]  

(13.182)

In both cases, the normalization of the probability distribution implies that \( M(0) = 1). \)

Derivatives of the moment-generating function and of the characteristic function give the moments

\[ E[x^n] = \mu_n = \frac{d^n M(k)}{dk^n} \bigg|_{k=0} = (-i)^n \frac{d^n P(k)}{dk^n} \bigg|_{k=0}. \]  

(13.183)

Example 13.12 (Gauss and Poisson) The moment-generating functions for the distributions of Gauss (13.175) and Poisson (13.178) are

\[ M_G(k, \mu, \sigma) = \exp \left( \mu k + \frac{1}{2} \sigma^2 k^2 \right) \]  

and \[ M_P(k, \langle n \rangle) = \exp \left[ \langle n \rangle \left( e^k - 1 \right) \right]. \]  

(13.184)

They give as the first three moments of these distributions

\[ \mu_{G0} = 1, \quad \mu_{G1} = \mu, \quad \mu_{G2} = \mu^2 + \sigma^2 \]  

(13.185)

\[ \mu_{P0} = 1, \quad \mu_{P1} = \langle n \rangle, \quad \mu_{P2} = \langle n \rangle + \langle n \rangle^2 \]  

(13.186)

(exercise 13.22).
Since the characteristic and moment-generating functions have derivatives (13.183) proportional to the moments \( \mu_n \), their Taylor series are

\[
\tilde{P}(k) = E[e^{ikx}] = \sum_{n=0}^{\infty} \frac{(ik)^n}{n!} E[x^n] = \sum_{n=0}^{\infty} \frac{(ik)^n}{n!} \mu_n
\]

and

\[
M(k) = E[e^{kx}] = \sum_{n=0}^{\infty} \frac{k^n}{n!} E[x^n] = \sum_{n=0}^{\infty} \frac{k^n}{n!} \mu_n.
\]

The cumulants \( c_n \) of a probability distribution are the derivatives of the logarithm of its moment-generating function

\[
c_n = \left. \frac{d^n \ln M(k)}{dk^n} \right|_{k=0} = (-i)^n \left. \frac{d^n \ln \tilde{P}(k)}{dk^n} \right|_{k=0}.
\]

One may show (exercise 13.24) that the first five cumulants of an arbitrary probability distribution are

\[
c_0 = 0, \quad c_1 = \mu, \quad c_2 = \sigma^2, \quad c_3 = \nu_3, \quad \text{and} \quad c_4 = \nu_4 - 3\sigma^4
\]

where the \( \nu \)'s are its central moments (13.27). The 3rd and 4th normalized cumulants are the skewness \( \zeta = c_3/\sigma^3 = \nu_3/\sigma^3 \) and the kurtosis \( \kappa = c_4/\sigma^4 = \nu_4/\sigma^4 - 3 \).

**Example 13.13** (Gaussian Cumulants) The logarithm of the moment-generating function (13.184) of Gauss’s distribution is \( \mu k + \sigma^2 k^2/2 \). Thus by (13.189), \( P_G(x, \mu, \sigma) \) has no skewness or kurtosis, its cumulants vanish \( c_{Gn} = 0 \) for \( n > 2 \), and its fourth central moment is \( \nu_4 = 3\sigma^4 \).

### 13.13 Fat Tails

The gaussian probability distribution \( P_G(x, \mu, \sigma) \) falls off for \( |x - \mu| \gg \sigma \) very fast—as \( \exp \left( -\frac{(x - \mu)^2}{2\sigma^2} \right) \). Many other probability distributions fall off more slowly; they have **fat tails**. Rare “black-swan” events—wild fluctuations, market bubbles, and crashes—lurk in their fat tails.

**Gosset’s distribution**, which is known as **Student’s t-distribution** with \( \nu \) degrees of freedom

\[
P_S(x, \nu, a) = \frac{1}{\sqrt{\pi}} \frac{\Gamma((1+\nu)/2)}{\Gamma(\nu/2)} \frac{a^\nu}{(a^2 + x^2)^{(1+\nu)/2}}
\]

(13.191)
13.13 Fat Tails

has **power-law tails**. Its even moments are

\[
\mu_{2n} = (2n - 1)!! \frac{\Gamma(\nu/2 - n)}{\Gamma(\nu/2)} \left( \frac{a^2}{2} \right)^n
\]  

(13.192)

for \(2n < \nu\) and infinite otherwise. For \(\nu = 1\), it coincides with the Breit-Wigner or Cauchy distribution

\[
P_S(x, 1, a) = \frac{1}{\pi} \frac{a}{a^2 + x^2}
\]  

(13.193)

in which \(x = E - E_0\) and \(a = \Gamma/2\) is the half-width at half-maximum.

Two representative cumulative probabilities are (Bouchaud and Potters, 2003, p.15–16)

\[
\Pr(x, \infty) = \int_x^\infty P_S(x', 3, 1) \, dx' = \frac{1}{2} - \frac{1}{\pi} \left[ \arctan x + \frac{x}{1 + x^2} \right] \quad (13.194)
\]

\[
\Pr(x, \infty) = \int_x^\infty P_S(x', 4, \sqrt{2}) \, dx' = \frac{1}{2} - \frac{3}{4} u + \frac{1}{4} u^3 \quad (13.195)
\]

where \(u = x/\sqrt{2} + x^2\) and \(a\) is picked so \(\sigma^2 = 1\). William Gosset (1876–1937), who worked for Guinness, wrote as Student because Guinness didn’t let its employees publish.

The **log-normal** probability distribution on \((0, \infty)\)

\[
P_{ln}(x) = \frac{1}{\sigma x \sqrt{2\pi}} \exp \left[ - \frac{\ln^2(x/x_0)}{2\sigma^2} \right]
\]  

(13.196)

describes distributions of rates of return (Bouchaud and Potters, 2003, p.9). Its moments are (exercise 13.27)

\[
\mu_n = x_0^n e^{n^2\sigma^2/2}.
\]  

(13.197)

The **exponential distribution** on \([0, \infty)\)

\[
P_e(x) = \alpha e^{-\alpha x}
\]  

(13.198)

has (exercise 13.28) mean \(\mu = 1/\alpha\) and variance \(\sigma^2 = 1/\alpha^2\). The sum of \(n\) independent exponentially and identically distributed random variables \(x = x_1 + \cdots + x_n\) is distributed on \([0, \infty)\) as (Feller, 1966, p.10)

\[
P_{n,e}(x) = \alpha (\alpha x)^{n-1} \frac{1}{(n-1)!} e^{-\alpha x}.
\]  

(13.199)

The sum of the squares \(x^2 = x_1^2 + \cdots + x_n^2\) of \(n\) independent normally and
identically distributed random variables of zero mean and variance $\sigma^2$ gives rise to Pearson’s **chi-squared distribution** on $(0, \infty)$

\[
P_{n,G}(x, \sigma)dx = \frac{\sqrt{2}}{\sigma} \frac{1}{\Gamma(n/2)} \left( \frac{x}{\sigma \sqrt{2}} \right)^{n-1} e^{-x^2/(2\sigma^2)}dx
\]

which for $x = \nu$, $n = 3$, and $\sigma^2 = kT/m$ is (exercise 13.29) the Maxwell-Boltzmann distribution (13.100). In terms of $\chi = x/\sigma$, it is

\[
P_n(\chi^2/2) d\chi^2 = \frac{1}{\Gamma(n/2)} \left( \frac{\chi^2}{2} \right)^{n/2-1} e^{-\chi^2/2} d(\chi^2/2).
\]

It has mean and variance

\[
\mu = n \quad \text{and} \quad \sigma^2 = 2n
\]

and is used in the chi-squared test (Pearson, 1900).

Personal income, the amplitudes of catastrophes, the price changes of financial assets, and many other phenomena occur on both small and large scales. **Lévy** distributions describe such multi-scale phenomena. The characteristic function for a symmetric Lévy distribution is for $\nu \leq 2$

\[
\tilde{L}_\nu(k, a) = \exp \left( -a |k|^\nu \right).
\]

Its inverse Fourier transform (13.174) is for $\nu = 1$ (exercise 13.30) the **Cauchy** or **Lorentz** distribution

\[
L_1(x, a_1) = \frac{a_1}{\pi(x^2 + a_1^2)}
\]

and for $\nu = 2$ the gaussian

\[
L_2(x, a_2) = P_G(x, 0, \sqrt{2a_2}) = \frac{1}{2\sqrt{\pi a_2}} \exp \left( -\frac{x^2}{4a_2} \right)
\]

but for other values of $\nu$ no simple expression for $L_\nu(x, a_\nu)$ is available. For $0 < \nu < 2$ and as $x \to \pm \infty$, it falls off as $|x|^{-(1+\nu)}$, and for $\nu > 2$ it assumes negative values, ceasing to be a probability distribution (Bouchaud and Potters, 2003, pp. 10–13).

### 13.14 The Central Limit Theorem and Jarl Lindeberg

We have seen in sections (13.7 & 13.8) that unbiased fluctuations tend to distribute the position and velocity of molecules according to Gauss’s distribution (13.75). Gaussian distributions occur very frequently. The **central limit theorem** suggests why they occur so often.
and the number of points $10^m$ rises further, the probability distributions $\text{Pr}^{(m)}_{e,G,G}(\infty, u)$ converge to the universal cumulative probability distribution $K(u)$ and provide a numerical verification of Kolmogorov’s theorem. Such curves make poor figures, however, because they hide beneath $K(u)$.

The curves labeled $\text{Pr}^{(m)}_{e,S,G}(\infty, u)$ for $m = 2$ and 3 are made from 100 sets of $N = 10^m$ points taken from $P_S(x, 3, 1)$ and tested as to whether they instead come from $P_G(x, 0, 1)$. Note that as $N = 10^m$ increases from 100 to 1000, the cumulative probability distribution $\text{Pr}^{(m)}_{e,S,G}(\infty, u)$ moves farther from Kolmogorov’s universal cumulative probability distribution $K(u)$. In fact, the curve $\text{Pr}^{(4)}_{e,S,G}(\infty, u)$ made from 100 sets of $10^4$ points lies beyond $u > 8$, too far to the right to fit in the figure. Kolmogorov’s test gets more conclusive as the number of points $N \to \infty$.

**Warning, mathematical hazard:** While binned data are ideal for chi-squared fits, they ruin Kolmogorov tests. The reason is that if the data are in bins of width $w$, then the empirical cumulative probability distribution $\text{Pr}^{(N)}_e(\infty, x)$ is a staircase function with steps as wide as the bin-width $w$ even in the limit $N \to \infty$. Thus even if the data come from the theoretical distribution, the limiting value $D_\infty$ of the Kolmogorov distance will be positive. In fact, one may show (exercise 13.21) that when the data do come from the theoretical probability distribution $P_t(x)$ assumed to be continuous, then the value of $D_\infty$ is

$$D_\infty \approx \sup_{-\infty < x < \infty} w \frac{P_t(x)}{2}.$$  

Thus in this case, the quantity $\sqrt{N} D_N$ could diverge as $\sqrt{N} D_\infty$ and lead one to believe that the data had not come from $P_t(x)$.

Suppose we have made some changes in our experimental apparatus and our software, and we want to see whether the new data $x_1', x_2', \ldots, x_N'$ we took after the changes are consistent with the old data $x_1, x_2, \ldots, x_N$ we took before the changes. Then following equations (13.310–13.312), we can make two empirical cumulative probability distributions—one $\text{Pr}^{(N)}_e(\infty, x)$ made from the $N$ old points $x_j$ and the other $\text{Pr}^{(N)}_e(\infty, x)$ made from the $N'$ new points $x_j'$. Next, we compute the distances

$$D_{N,N'}^+ = \sup_{-\infty < x < \infty} \left( \text{Pr}^{(N)}_e(\infty, x) - \text{Pr}^{(N')}_e(\infty, x) \right)$$

$$D_{N,N'} = \sup_{-\infty < x < \infty} \left| \text{Pr}^{(N)}_e(\infty, x) - \text{Pr}^{(N')}_e(\infty, x) \right|$$

which are analogous to (13.313–13.316). Smirnov (Smirnov 1939; Gnedenko
double prob=0, tmpProb=0, fact=0, x=0;
double p[LOOP_ITR];

// probability of no events
p[0] = exp(-AN);
prob = p[0];

// p(k) is the probability of fewer than k+1 events per day
for (k=1; k<=LOOP_ITR; k++)
{
    fact = factorial (k);
    tmpProb = k * exp(-AN) / fact;
    prob += pow(AN, tmpProb);
    p[k] = prob;
}

// Random seed
srand ( time(NULL) );

// Goes through all the histories
for (k=0; k<N; k++)
{
    // Goes through all the days
    for (day=1; day<LOOP_ITR; day++)
    {
        // Generates a random number between 0 and 1
        x = static_cast<double>(rand()) / RAND_MAX;

        // Finds an M with p(M) < X
        for (m=100; m>=0; m--)
        {
            if (x < p[m])
            {
                num = m;
            }
        }

        histories[k][day] = num;
    }
}
Monte Carlo Methods

// Calculates max and sum
numMuons = histories[k].max();
totMuons = histories[k].sum();

// Updates our records
maxEvents[numMuons]++;
totEvents[totMuons]++;
}

// Opens a data file
ofstream fhMaxEvents, fhSumEvents;
fhMaxEvents.open("maxEvents.txt");
fhSumEvents.open("totEvents.txt");

// Sets precision
fhMaxEvents.setf(ios::fixed, ios::floatfield);
fhMaxEvents.precision(7);
fhSumEvents.setf(ios::fixed, ios::floatfield);
fhSumEvents.precision(7);

// Writes the data to a file
for (k=0; k<LOOP_ITR; k++)
{
    fhMaxEvents << k << "  " << maxEvents[k] << endl;
    fhSumEvents << k << "  " << totEvents[k] << endl;
}
}

14.4 Statistical Mechanics

The Metropolis algorithm can generate a sequence of states or configurations of a system distributed according to the Boltzmann probability distribution (1.345). Suppose the state of the system is described by a vector $x$ of many components. For instance, if the system is a protein, the vector $x$ might be the $3N$ spatial coordinates of the $N$ atoms of the protein. A protein composed of 200 amino acids has about 4000 atoms, and so the vector $x$ would have some 12,000 components. Suppose $E(x)$ is the energy
fall within the region $\mathcal{R}$

$$V_\mathcal{R} = \frac{N_\mathcal{R}}{N} L^n. \quad (14.2)$$

The integral formula (14.1) then becomes

$$\int_{\mathcal{R}} f(x) d^n x \approx \frac{L^n}{N} \sum_{k=1}^{N_\mathcal{R}} f(x_k). \quad (14.3)$$

The utility of the Monte Carlo method of numerical integration rises sharply with the dimension $n$ of the hypervolume.

**Example 14.1 (Numerical Integration)** Suppose one wants to integrate the function

$$f(x, y) = \frac{e^{-2x - 3y}}{\sqrt{x^2 + y^2 + 1}} \quad (14.4)$$

over the quarter of the unit disk in which $x$ and $y$ are positive. In this case, $V_\mathcal{R}$ is the area $\pi/4$ of the quarter disk.

To generate fresh random numbers, one must set the seed for the code that computes them. The following program sets the seed by using the subroutine `init_random_seed` defined in a FORTRAN95 program in section 13.16. With some compilers, one can just write “call random_seed().”

```fortran
program integrate
  implicit none ! catches typos
  integer :: k, N
  integer, parameter :: dp = kind(1.0d0)
  real(dp) :: x, y, sum = 0.0d0, f
  real(dp), dimension(2) :: rdn
  real(dp), parameter :: area = atan(1.0d0) ! pi/4
  f(x,y) = exp(-2*x - 3*y)/sqrt(x**2 + y**2 + 1.0d0)
  write(6,*) 'How many points?'
  read(5,*) N
  call init_random_seed() ! set new seed
  do k = 1, N
    call random_number(rdn); x= rdn(1); y = rdn(2)
    if (x**2+y**2 > 1.0d0) then
      go to 10
    end if
    sum = sum + f(x,y)
  end do
  integral = area times mean value < f > of f
```

Once the system is thermalized, one can start measuring properties of the system. One computes a physical quantity every hundred or every thousand sweeps and takes the average of these measurements. That average is the mean value of the physical quantity at temperature $T$.

Why does this work? Consider two configurations $x$ and $x'$ which respectively have energies $E = E(x)$ and $E' = E(x')$ and are occupied with probabilities $P_t(x)$ and $P_t(x')$ as the system is thermalizing. If $E' > E$, then the rate $R(x' \to x)$ of going from $x'$ to $x$ is the rate $v$ of choosing to test $x$ when one is at $x'$ times the probability $P_t(x')$ of being at $x'$, that is, $R(x' \to x) = v P_t(x')$. The reverse rate is $R(x \to x') = v P_t(x) e^{-(E' - E)/kT}$ with the same $v$ since the random walk is symmetric. The net rate from $x' \to x$ then is

$$R(x' \to x) - R(x \to x') = v \left( P_t(x') - P_t(x) e^{-(E' - E)/kT} \right).$$

(14.8)

This net flow of probability from $x' \to x$ is positive if and only if

$$P_t(x')/P_t(x) > e^{-(E' - E)/kT}.$$  

(14.9)

The probability distribution $P_t(x)$ therefore flows with each sweep toward the Boltzmann distribution $\exp(-E(x)/kT)$. The flow slows and stops when the two rates are equal $R(x' \to x) = R(x \to x')$ a condition called detailed balance. At this equilibrium, the distribution $P_t(x)$ satisfies $P_t(x) = P_t(x') e^{-(E' - E)/kT}$ in which $P_t(x') e^{E'/kT}$ is independent of $x$. So the thermalizing distribution $P_t(x)$ approaches the distribution $P(x) = c e^{-E/kT}$ in which $c$ is independent of $x$. Since the sum of these probabilities must be unity, we have

$$\sum_x P(x) = c \sum_x e^{-E/kT} = 1$$

(14.10)

which means that the constant $c$ is the inverse of the partition function

$$Z(T) = \sum_x e^{-E(x)/kT}.$$  

(14.11)

The thermalizing distribution approaches Boltzmann’s distribution (1.345)

$$P_t(x) \to P_B(x) = e^{-E(x)/kT}/Z(T).$$

(14.12)

**Example 14.2 (Z₂ Lattice Gauge Theory)** First, one replaces space-time with a lattice of points in $d$ dimensions. Two nearest neighbor points are separated by the lattice spacing $a$ and joined by a link. Next, one puts an
element $U$ of the gauge group on each link. For the $Z_2$ gauge group (example 10.4), one assigns an action $S_\Box$ to each elementary square or plaquette of the lattice with vertices 1, 2, 3, and 4

$$S_\Box = 1 - U_{1,2} U_{2,3} U_{3,4} U_{4,1}. \tag{14.13}$$

Then, one replaces $E(x)/kT$ with $\beta S$ in which the action $S$ is a sum of all the plaquette actions $S_p$. More details are available at Michael Creutz’s website (latticeguy.net/lattice.html).

Although the generation of configurations distributed according to the Boltzmann probability distribution (1.345) is one of its most useful applications, the Monte Carlo method is much more general. It can generate configurations $x$ distributed according to any probability distribution $P(x)$.

To generate configurations distributed according to $P(x)$, we accept any new configuration $x'$ if $P(x') \geq P(x)$ and also accept $x'$ with probability

$$P(x \to x') = P(x')/P(x) \tag{14.14}$$

if $P(x) > P(x')$.

This works for the same reason that the Boltzmann version works. Consider two configurations $x$ and $x'$. If the system is thermalized, then the probabilities $P_t(x)$ and $P_t(x')$ have reached equilibrium, and so the rate $R(x' \to x)$ from $x \to x'$ must equal that $R(x' \to x)$ from $x' \to x$. If $P(x') < P(x)$, then $R(x' \to x)$ is

$$R(x' \to x) = v P_t(x') \tag{14.15}$$

in which $v$ is the rate of choosing $\delta x = x' - x$, while the rate $R(x \to x')$ is

$$R(x \to x') = v P_t(x) P(x')/P(x) \tag{14.16}$$

with the same $v$ since the random walk is symmetric. Equating the two rates

$$R(x' \to x) = R(x \to x') \tag{14.17}$$

we find that the flow of probability stops when

$$P_t(x) = P(x) P_t(x')/P(x') = c P(x) \tag{14.18}$$

where $c$ is independent of $x'$. Thus $P_t(x) \to P(x)$.

So far we have assumed that the rate of choosing $x \to x'$ is the same as the rate of choosing $x' \to x$. In Smart Monte Carlo schemes, physicists arrange the rates $v_{x \to x'}$ and $v_{x' \to x}$ so as to steer the flow and speed-up thermalization. To compensate for this asymmetry, they change the second
part of the Metropolis step from \( x \rightarrow x' \) when \( E' = E(x') > E = E(x) \) to accept conditionally with probability

\[
P(x \rightarrow x') = P(x') v_{x' \rightarrow x} / [P(x) v_{x \rightarrow x'}]. \tag{14.19}
\]

Now if \( P(x') < P(x) \), then \( R(x' \rightarrow x) \) is

\[
R(x' \rightarrow x) = v_{x' \rightarrow x} P_t(x') \tag{14.20}
\]

while the rate \( R(x \rightarrow x') \) is

\[
R(x \rightarrow x') = v_{x \rightarrow x'} P_t(x) P(x') v_{x' \rightarrow x} / [P(x) v_{x \rightarrow x'}]. \tag{14.21}
\]

Equating the two rates \( R(x' \rightarrow x) = R(x \rightarrow x') \), we find

\[
P_t(x') = P_t(x) P(x') / P(x). \tag{14.22}
\]

That is \( P_t(x) = P(x) P_t(x') / P(x') \) which gives

\[
P_t(x) = N P(x) \tag{14.23}
\]

where \( N \) is a constant of normalization.

### 14.5 Solving Arbitrary Problems

If you know how to generate a suitably large space of trial solutions to a problem, and you also know how to compare the quality of any two of your solutions, then you can use a Monte Carlo method to solve it. The hard parts of this seemingly magical method are characterizing a big enough space of solutions \( s \) and constructing a quality function or functional that assigns a number \( Q(s) \) to every solution in such a way that if \( s \) is a better solution than \( s' \), then

\[
Q(s) > Q(s'). \tag{14.24}
\]

But once one has characterized the space of possible solutions \( s \) and has constructed the quality function \( Q(s) \), then one simply generates zillions of random solutions and selects the one that maximizes the function \( Q(s) \) over the space of all solutions.

If one can characterize the solutions as vectors of a certain dimension, \( s = (x_1, \ldots, x_n) \), then one may use the Monte Carlo method of the previous section (14.4) by replacing \(-E(s)\) with \( Q(s) \) and \( kT \) with a parameter of the same dimension as \( Q(s) \), nominally dimensionless.
evolution more ergodic, which is why most complex modern organisms use sexual reproduction.

Other genomic changes occur when a virus inserts its DNA into that of a cell and when transposable elements (transposons) of DNA move to different sites in a genome.

In evolution, the rest of the Metropolis step is done by the new individual: if he or she survives and multiplies, then the change is accepted; if he or she dies without progeny, then the change is rejected. Evolution is slow, but it has succeeded in turning a soup of simple molecules into humans with brains of 100 billion neurons, each with 1000 connections to other neurons.

John Holland and others have incorporated analogs of these Metropolis steps into Monte Carlo techniques called genetic algorithms for solving wide classes of problems (Holland, 1975; Vose, 1999; Schmitt, 2001).

Evolution also occurs at the cellular level when a cell mutates enough to escape the control imposed on its proliferation by its neighbors and transforms into a cancer cell.

Further Reading

The classic Quarks, Gluons, and Lattices (Creutz, 1983) is a marvelous introduction to the subject; his website (latticeguy.net/lattice.html) is an extraordinary resource, as is Rubinstein’s Simulation and the Monte Carlo Method (Rubinstein and Kroese, 2007).

Exercises

14.1 Go to Michael Creutz’s website (latticeguy.net/lattice.html) and get his C-code for $Z_2$ lattice gauge theory. Compile and run it, and make a graph that exhibits strong hysteresis as you raise and lower $\beta = 1/kT$.

14.2 Modify his code and produce a graph showing the coexistence of two phases at the critical coupling $\beta_t = 0.5 \ln(1 + \sqrt{2})$. Hint: Do a cold start and then 100 updates at $\beta_t$, then do a random start and do 100 updates at $\beta_t$. Plot the values of the action against the update number 1, 2, 3, . . . 100.

14.3 Modify Creutz’s C code for $Z_2$ lattice gauge theory so as to be able to vary the dimension $d$ of space-time. Show that for $d = 2$, there’s no hysteresis loop (there’s no phase transition). For $d = 3$, show that any hysteresis loop is minimal (there’s a second-order phase transition).

14.4 What happens when $d = 5$?
Path Integrals

Linking three of these matrix elements together and using subscripts instead of primes, we have

\[ \langle q_3 | e^{-3\epsilon H} | q_0 \rangle = \int_{-\infty}^{\infty} \langle q_3 | e^{-\epsilon H} | q_2 \rangle \langle q_2 | e^{-\epsilon H} | q_1 \rangle \langle q_1 | e^{-\epsilon H} | q_0 \rangle \, dq_1 \, dq_2 \quad (16.23) \]

\[ = \left( \frac{m}{2\pi\epsilon} \right)^{3/2} \int_{-\infty}^{\infty} \exp \left\{ -\epsilon \sum_{j=0}^{2} \left[ \frac{1}{2} m \dot{q}_j^2 + V(q_j) \right] \right\} \, dq_1 \, dq_2. \]

Boldly passing from 3 to \( n \) and suppressing some integral signs, we get

\[ \langle q_n | e^{-n\epsilon H} | q_0 \rangle = \int_{-\infty}^{\infty} \langle q_n | e^{-\epsilon H} | q_{n-1} \rangle \cdots \langle q_1 | e^{-\epsilon H} | q_0 \rangle \, dq_1 \cdots dq_{n-1} \quad (16.24) \]

\[ = \left( \frac{m}{2\pi\epsilon} \right)^{n/2} \int_{-\infty}^{\infty} \exp \left\{ -\epsilon \sum_{j=0}^{n-1} \left[ \frac{1}{2} m \dot{q}_j^2 + V(q_j) \right] \right\} \, dq_1 \cdots dq_{n-1}. \]

Writing \( dt \) for \( \epsilon \) and taking the limits \( \epsilon \to 0 \) and \( n \equiv \beta/\epsilon \to \infty \), we find that the matrix element \( \langle q_\beta | e^{-\beta H} | q_0 \rangle \) is a path integral of the exponential of the average energy multiplied by \( -\beta \)

\[ \langle q_\beta | e^{-\beta H} | q_0 \rangle = \int \exp \left[ -\int_0^{\beta} \frac{1}{2} m \dot{q}^2(t) + V(q(t)) \, dt \right] Dq \quad (16.25) \]

in which \( Dq \equiv (nm/2\pi\beta)^{n/2} dq_1 dq_2 \cdots dq_{n-1} \) as \( n \to \infty \). We sum over all paths \( q(t) \) that go from \( q(0) = q_0 \) at inverse temperature \( \beta = 0 \) to \( q(\beta) = q_\beta \) at inverse temperature \( \beta \).

In the limit \( \beta \to \infty \), the operator \( \exp(-\beta H) \) becomes proportional to a projection operator (16.1) on the ground state of the theory.

In three-dimensional space, \( q(t) \) replaces \( q(t) \) in equation (16.25)

\[ \langle q_\beta | e^{-\beta H} | q_0 \rangle = \int \exp \left[ -\int_0^{\beta} \frac{1}{2} m \dot{q}^2 + V(q) \, dt \right] Dq. \quad (16.26) \]

Path integrals in imaginary time are called *euclidean* mainly to distinguish them from *Minkowski* path integrals, which represent matrix elements of the time-evolution operator \( \exp(-itH) \) in real time.
16.6 Free Particle in Imaginary Time

a complete set of momentum dyadics $|p\rangle\langle p|$ and doing the resulting Fourier transform.

**Example 16.1** (The Bohm-Aharonov Effect)  From our formula (11.311) for the action of a relativistic particle of mass $m$ and charge $q$, we infer (exercise 16.7) that the action a nonrelativistic particle in an electromagnetic field with no scalar potential is

$$S = \int_{x_1}^{x_2} \left[ \frac{1}{2} mv^2 + q A \right] \cdot dx.$$  

(16.55)

Now imagine that we shoot a beam of such particles past but not through a narrow cylinder in which a magnetic field is confined. The particles can go either way around the cylinder of area $S$ but cannot enter the region of the magnetic field. The difference in the phases of the amplitudes is the loop integral from the source to the detector and back to the source

$$\frac{\Delta S}{\hbar} = \oint \left[ \frac{mv^2}{2} + q A \right] \cdot \frac{dx}{\hbar} = \oint \frac{mv \cdot dx}{2\hbar} + \frac{q}{\hbar} \int S \cdot dS.$$  

(16.56)

in which $\Phi$ is the magnetic flux through the cylinder.

16.6 Free Particle in Imaginary Time

If we mimic the steps of the preceding section (16.5) in which the Hamiltonian is $H = \frac{p^2}{2m}$, set $\beta = it/\hbar = 1/kT$, and use Dirac’s delta function

$$\delta^3(q) = \lim_{t \to 0} \left( \frac{m}{2\pi\hbar t} \right)^{3/2} e^{-mq^2/2\hbar t}$$  

(16.57)

then we get

$$\langle q|e^{-\beta H}|0\rangle = \left( \frac{m}{2\pi\hbar^2\beta} \right)^{3/2} \exp \left[ -\frac{mq^2}{2\hbar^2} \right] = \left( \frac{mkT}{2\pi\hbar^2} \right)^{3/2} e^{-mkTq^2/2\hbar^2}.$$  

(16.58)

To study the ground state of the system, we set $\beta = t/\hbar$ and let $t \to \infty$ in

$$\langle q|e^{-tH/\hbar}|0\rangle = \left( \frac{m}{2\pi\hbar t} \right)^{3/2} \exp \left[ -\frac{m q^2}{2\hbar^2} \right]$$  

(16.59)

which for $D = \hbar/(2m)$ is the solution (3.200 & 13.107) of the diffusion equation.
and the functional delta function
\[ \delta[\nabla \cdot A] = \prod_x \delta(\nabla \cdot A(x)) \] (16.163)

enforces the Coulomb-gauge condition. The term \( \mathcal{L}_m \) is the action density of the matter field \( \psi \).

Tricks are available. We introduce a new field \( A^0(x) \) and consider the

\[
F = \int \exp \left[ i \int \frac{1}{2} (\nabla A^0 + \nabla \triangle^{-1} j^0)^2 \, d^4 x \right] DA^0
\] (16.164)

which is just a number independent of the charge density \( j^0 \) since we can cancel the \( j^0 \) term by shifting \( A^0 \). By \( \triangle^{-1} \), we mean \(-1/4\pi |x - y|\). By integrating by parts, we can write the number \( F \) as (exercise 16.21)

\[
F = \int \exp \left[ i \int \left( \frac{1}{2} (\nabla A^0)^2 - A^0 j^0 - \frac{1}{2} j^0 \triangle^{-1} j^0 \right) \, d^4 x \right] DA^0 + i \int V_C \, dt \] (16.165)

So when we multiply the numerator and denominator of the amplitude (16.161) by \( F \), the awkward Coulomb term cancels, and we get

\[
\langle \Omega | T[\mathcal{O}_1 \ldots \mathcal{O}_n] | \Omega \rangle = \frac{\int \mathcal{O}_1 \ldots \mathcal{O}_n e^{iS'} \delta[\nabla \cdot A] \, DA \, D\psi}{\int e^{iS'} \delta[\nabla \cdot A] \, DA \, D\psi} \] (16.166)

where now \( DA \) includes all four components \( A^\mu \) and

\[
S' = \int \frac{1}{2} \dot{A}^2 - \frac{1}{2} (\nabla \times A)^2 + \frac{1}{2} (\nabla A^0)^2 + A \cdot j - A^0 j^0 + \mathcal{L}_m \, d^4 x. \] (16.167)

Since the delta-function \( \delta[\nabla \cdot A] \) enforces the Coulomb-gauge condition, we can add to the action \( S' \) the term \( (\nabla \cdot \dot{A}) A^0 \) which is \(-\dot{A} \cdot \nabla A^0 \) after we integrate by parts and drop the surface term. This extra term makes the action gauge invariant

\[
S = \int \frac{1}{2} (\dot{A} - \nabla A^0)^2 - \frac{1}{2} (\nabla \times A)^2 + A \cdot j - A^0 j^0 + \mathcal{L}_m \, d^4 x
\] (16.168)
which contributes the terms $-\partial_\mu \omega^*_a \partial^\mu \omega_a$ and

$$-\partial_\mu \omega^*_a g f_{abc} A^\mu_b \omega_c = \partial_\mu \omega^*_a g f_{abc} A^\mu_b \omega_c$$

(16.254) to the action density.

Thus we can do perturbation theory by using the modified action density

$$\mathcal{L}' = -\frac{1}{4} F_{a\mu
u} F^a_{\mu\nu} - \frac{1}{2} (\partial_\mu A^\mu_a)^2 - \partial_\mu \omega^*_a \partial^\mu \omega_a + \partial_\mu \omega^*_a g f_{abc} A^\mu_b \omega_c - \bar{\psi} (\mathcal{D} + m) \psi$$

(16.255) in which $\mathcal{D} \equiv \gamma^\mu D_\mu = \gamma^\mu (\partial_\mu - ig t^a A_\mu)$. The \textit{ghost} field $\omega$ is a mathematical device, not a physical field describing real particles, which would be spinless fermions violating the spin-statistics theorem (example 10.19).

Further Reading


Exercises

16.1 Derive the multiple gaussian integral (16.8) from (5.167).
16.2 Derive the multiple gaussian integral (16.12) from (5.166).
16.3 Show that the vector $\mathbf{Y}$ that makes the argument of the multiple gaussian integral (16.12) stationary is given by (16.13), and that the multiple gaussian integral (16.12) is equal to its exponential evaluated at its stationary point $\mathbf{Y}$ apart from a prefactor involving $\det i S$.
16.4 Repeat the previous exercise for the multiple gaussian integral (16.11).
16.5 Compute the double integral (16.23) for the case $V(q_j) = 0$.
16.6 Insert into the LHS of (16.54) a complete set of momentum dyadics $|p\rangle\langle p|$, use the inner product $\langle q|p\rangle = \exp(i q p \hbar)/\sqrt{2\pi \hbar}$, do the resulting Fourier transform, and so verify the free-particle path integral (16.54).
16.7 By taking the nonrelativistic limit of the formula (11.311) for the action of a relativistic particle of mass $m$ and charge $q$, derive the expression (16.55) for the action a nonrelativistic particle in an electromagnetic field with no scalar potential.
16.8 Show that for the hamiltonian (16.60) of the simple harmonic oscillator the action $S[q_c]$ of the classical path is (16.67).
16.9 Show that the harmonic-oscillator action of the loop (16.68) is (16.69).
16.10 Show that the harmonic-oscillator amplitude (16.72) for $q' = 0$ and $q'' = q$ reduces as $t \to 0$ to the one-dimensional version of the free-particle amplitude (16.54).
16.11 Work out the path-integral formula for the amplitude for a mass \( m \) initially at rest to fall to the ground from height \( h \) in a gravitational field of local acceleration \( g \) to lowest order and then including loops up to an overall constant. Hint: use the technique of section 16.7.

16.12 Show that the action (16.74) of the stationary solution (16.77) is (16.79).


16.14 Derive identity (16.136). Split the time integral at \( t = 0 \) into two halves, use
\[
\epsilon e^{\pm \epsilon t} = \pm \frac{d}{dt} e^{\pm \epsilon t}
\]
and then integrate each half by parts.

16.15 Derive the third term in equation (16.138) from the second term.

16.16 Use (16.143) and the Fourier transform (16.144) of the external current \( j \) to derive the formula (16.145) for the modified action \( S_0[\phi, \epsilon, j] \).

16.17 Derive equation (16.147) from equations (16.145) and (16.146).

16.18 Derive the formula (16.148) for \( Z_0[j] \) from the expression (16.147) for \( S_0[\phi, \epsilon, j] \).


16.20 Derive equation (16.154) from the formula (16.149) for \( Z_0[j] \).

16.21 Show that the time integral of the Coulomb term (16.159) is the term that is quadratic in \( j^0 \) in the number \( F \) defined by (16.164).

16.22 By following steps analogous to those the led to (16.150), derive the formula (16.177) for the photon propagator in Feynman’s gauge.

16.23 Derive expression (16.192) for the inner product \( \langle \zeta | \theta \rangle \).


16.25 Derive the eigenvalue equation (16.200) from the definition (16.198 & 16.199) of the eigenstate \( |\theta\rangle \) and the anticommutation relations (16.196 & 16.197).

16.26 Derive the eigenvalue relation (16.213) for the Fermi field \( \psi_m(x, t) \) from the anticommutation relations (16.209 & 16.210) and the definitions (16.211 & 16.212).

16.27 Derive the formula (16.214) for the inner product \( \langle \chi'| \chi \rangle \) from the definition (16.212) of the ket \( |\chi\rangle \).
# The Renormalization Group

## 17.1 The Renormalization Group in Quantum Field Theory

Most quantum field theories are non-linear with infinitely many degrees of freedom, and because they describe point particles, they are rife with infinities. But short-distance effects, probably the finite sizes of the fundamental constituents of matter, mitigate these infinities so that we can cope with them consistently without knowing what happens at very short distances and very high energies. This procedure is called renormalization.

For instance, in the theory described by the Lagrange density

\[ L = -\frac{1}{2} \partial_v \phi \partial^v \phi - \frac{1}{2} m^2 \phi^2 - \frac{g^2}{24} \phi^4 \]  

(17.1)

we can cut off divergent integrals at some high energy \( \Lambda \). The amplitude for the elastic scattering of two bosons of initial four-momenta \( p_1 \) and \( p_2 \) into two of final momenta \( p_1' \) and \( p_2' \) to one-loop order (Weinberg, 1996, chap. 18) then is proportional to (Zee, 2010, chaps. III & VI)

\[ A = g - \frac{g^2}{32\pi^2} \left[ \ln \left( \frac{\Lambda^6}{stu} \right) - i\pi + 3 \right] \]  

(17.2)

as long as the absolute values of the Mandelstam variables \( s = -(p_1 + p_2)^2 \), \( t = -(p_1 - p_1')^2 \), and \( u = -(p_1 - p_2')^2 \), which satisfy \( stu > 0 \) and \( s + t + u = 4m^2 \) (Stanley Mandelstam, 1928–). We define the physical coupling constant \( g_\mu \), as opposed to the bare one \( g \) that comes with \( L \), to be the real part of the amplitude \( A \) at \( s = -t = -u = \mu^2 \)

\[ g_\mu = g - \frac{3g^2}{32\pi^2} \left[ \ln \left( \frac{\Lambda^2}{\mu^2} \right) + 1 \right]. \]  

(17.3)

Thus the bare coupling constant is \( g = g_\mu + 3g^2 \left[ \ln(\Lambda^2/\mu^2) + 1 \right] \), and using
this formula, we can write our expression (17.2) for the amplitude \( A \) in a form in which the cutoff \( \Lambda \) no longer appears

\[
A = g_\mu - \frac{g^2}{32\pi^2} \left[ \ln \left( \frac{\mu^6}{stu} \right) - i\pi \right]. \tag{17.4}
\]

This is the magic of renormalization.

The physical coupling “constant” \( g_\mu \) is the right coupling at energy \( \mu \) because when all the Mandelstam variables are near the renormalization point \( stu = \mu^6 \), the one-loop correction is tiny, and \( A \approx g_\mu \).

How does the physical coupling \( g_\mu \) depend upon the energy \( \mu \)? The amplitude \( A \) must be independent of the renormalization energy \( \mu \), and so

\[
\frac{dA}{d\mu} = \frac{dg_\mu}{d\mu} - \frac{g^2}{32\pi^2 \mu} = 0 \tag{17.5}
\]

which is a version of the Callan-Symanzik equation.

We assume that when the cutoff \( \Lambda \) is big but finite, the bare and running coupling constants \( g \) and \( g_\mu \) are so tiny that they differ by terms of order \( g^2 \) or \( g_\mu^2 \). Then to lowest order in \( g \) and \( g_\mu \), we can replace \( g^2 \) by \( g_\mu^2 \) in (17.5) and arrive at the simple differential equation

\[
\mu \frac{dg_\mu}{d\mu} = \beta(g_\mu) = \frac{3g_\mu^2}{16\pi^2} \tag{17.6}
\]

which we can integrate

\[
\ln \frac{E}{M} = \int_M^E \frac{d\mu}{\mu} = \int_{g_M}^{g_E} \frac{dg_\mu}{\beta(g_\mu)} = \frac{16\pi^2}{3} \int_{g_M}^{g_E} \frac{dg_\mu}{g_\mu^2} = \frac{16\pi^2}{3} \left( \frac{1}{g_M} - \frac{1}{g_E} \right) \tag{17.7}
\]

to find the running physical coupling constant \( g_\mu \) at energy \( \mu = E \)

\[
g_E = g_M \frac{g_M}{1 - 3g_M \ln(E/M)/16\pi^2}. \tag{17.8}
\]

As the energy \( E = \sqrt{s} \) rises above \( M \), while staying below the singular value \( E = M \exp(16\pi^2/3g_M) \), the running coupling \( g_E \) slowly increases. And so does the scattering amplitude, \( A \approx g_E \).

**Example 17.1** (Quantum Electrodynamics) Vacuum polarization makes the amplitude for the scattering of two electrons proportional to \( \) (Weinberg, 1995, chap. 11)

\[
A(q^2) = e^2 \left[ 1 + \pi(q^2) \right] \tag{17.9}
\]

rather than to \( e^2 \). Here \( e \) is the renormalized charge, \( q = p'_1 - p_1 \) is the
four-momentum transferred to the first electron, and
\[
\pi(q^2) = \frac{e^2}{2\pi^2} \int_0^1 x(1-x) \ln \left[ 1 + \frac{q^2 x(1-x)}{m^2} \right] dx \tag{17.10}
\]
represents the polarization of the vacuum. We define the square of the running coupling constant \(e^2_\mu\) to be the amplitude (17.9) at \(q^2 = \mu^2\)
\[
e^2_\mu = A(\mu^2) = e^2 \left[ 1 + \pi(\mu^2) \right]. \tag{17.11}
\]
For \(\mu^2 \gg m^2\), the vacuum polarization term \(\pi(\mu^2)\) is (exercise 17.1)
\[
\pi(\mu^2) \approx \frac{e^2}{6\pi^2} \left[ \ln \frac{\mu}{m} - \frac{5}{6} \right]. \tag{17.12}
\]
The amplitude (17.9) then is
\[
A(q^2) = e^2_\mu \frac{1 + \pi(q^2)}{1 + \pi(\mu^2)} \tag{17.13}
\]
and since it must be independent of \(\mu\), we have
\[
0 = \frac{d}{d\mu} A(q^2) = \frac{d}{d\mu} \frac{e^2_\mu}{1 + \pi(\mu^2)} \approx \frac{d}{d\mu} \left\{ e^2_\mu \left[ 1 - \pi(\mu^2) \right] \right\}. \tag{17.14}
\]
So we find
\[
0 = 2e_\mu \left[ \frac{de_\mu}{d\mu} \right] \left[ 1 - \pi(\mu^2) \right] - e^2_\mu \frac{d}{d\mu} \pi(\mu^2) = 2e_\mu \left( \frac{de_\mu}{d\mu} \right) \left[ 1 - \pi(\mu^2) \right] - e^2_\mu \frac{e^2}{6\pi^2\mu}. \tag{17.15}
\]
Thus since by (17.10 & 17.11) \(\pi(\mu^2) = O(e^2)\) and \(e^2_\mu = e^2 + O(e^4)\), we find to lowest order in \(e_\mu\)
\[
e_\mu \frac{d}{d\mu} \equiv \beta(e_\mu) = \frac{e^3_\mu}{12\pi^2}. \tag{17.16}
\]
We can integrate this differential equation
\[
\ln \frac{E}{M} = \int_M^E \frac{d\mu}{\mu} = \int_{e_M}^{e_E} \frac{de_\mu}{\beta(e_\mu)} = 12\pi^2 \int_{e_M}^{e_E} \frac{de_\mu}{e^3_\mu} = 6\pi^2 \left( \frac{1 - \alpha}{e_M^2} - \frac{1}{e_E^2} \right) \tag{17.17}
\]
and so get for the running coupling constant the formula
\[
e^2_E = \frac{e^2_M}{1 - e^2_M \ln(E/M)/6\pi^2}. \tag{17.18}
\]
which shows that it slowly increases with the energy \(E\). Thus, the fine-
structure constant \(e^2_\mu/4\pi\) rises from \(\alpha = 1/137.036\) at \(m_e\) to
\[
e^2(45.5\text{GeV})/4\pi = \frac{\alpha}{1 - 2\alpha \ln(45.5/0.00051)/3\pi} = \frac{1}{134.6} \tag{17.19}
\]
where for \(n_f\) flavors of light quarks
\[
\beta_0 = \frac{1}{(4\pi)^2} \left( \frac{11}{3} N - \frac{2}{3} n_f \right)
\]
\[
\beta_1 = \frac{1}{(4\pi)^4} \left( \frac{34}{3} N^2 - \frac{10}{3} N n_f - \frac{N^2 - 1}{N} n_f \right).
\] (17.33)

In quantum chromodynamics, \(N = 3\).

Combining the definition (17.31) of the \(\beta\)-function with its expansion (17.32) for small \(g\), one arrives at the differential equation
\[
\frac{dg}{d\ln a} = \beta_0 g^3 + \beta_1 g^5
\] (17.34)

which one may integrate
\[
\int d\ln a = \ln a - \ln c = \int \frac{dg}{\beta_0 g^3 + \beta_1 g^5} = -\frac{1}{2\beta_0 g^2} + \frac{\beta_1}{2\beta_0^2} \ln \left( \frac{\beta_0 + \beta_1 g^2}{g^2} \right)
\] (17.35)
to find
\[
a(g) = c \left( \frac{\beta_0 + \beta_1 g^2}{g^2} \right)^{\beta_1/2\beta_0^2} e^{-1/2\beta_0 g^2}
\] (17.36)
in which \(c\) is a constant of integration. The term \(\beta_1 g^2\) is of higher order in \(g\), and if one drops it and absorbs a factor of \(\beta_0^2\) into a new constant of integration \(\Lambda\), then one gets
\[
a(g) = \frac{1}{\Lambda} (\beta_0 g^2)^{-\beta_1/2\beta_0^2} e^{-1/2\beta_0 g^2}.
\] (17.37)

As \(g \to 0\), the lattice spacing \(a(g)\) goes to zero very fast (as long as \(n_f < 17\) for \(N = 3\)). The inverse of this relation (17.37) is
\[
g(a) \approx \left[ \beta_0 \ln(a^{-2} \Lambda^{-2}) + (\beta_1/\beta_0) \ln \left( \ln(a^{-2} \Lambda^{-2}) \right) \right]^{-1/2}.
\] (17.38)

It shows that the coupling constant slowly goes to zero with \(a\), which is a lattice version of asymptotic freedom.

### 17.3 The Renormalization Group in Condensed-Matter Physics

The study of condensed matter is concerned mainly with properties that emerge in the bulk, such as the melting point, the boiling point, or the conductivity. So we want to see what happens to the physics when we increase the distance scale many orders of magnitude beyond the size \(a\) of an individual molecule or the distance between nearest neighbors.
The full action of a stretched field is

\[ S(\phi_L) = \int d^d x \left( \frac{1}{2} (\partial \phi)^2 + \sum_n g_{d,n}(L) \phi^n \right) \]  

in which

\[ g_{d,n}(L) = L^d A^n(L) g_n = L^{d-n(d-2)/2} g_{d,n}. \]  

The beta-function

\[ \beta(g_{d,n}) \equiv L \frac{dg_{d,n}(L)}{dL} = d - n(d - 2)/2 \]  

is just the exponent of the coupling “constant” \( g_{d,n}(L) \). If it is positive, then the coupling constant \( g_{d,n}(L) \) gets stronger as \( L \to \infty \); such couplings are called **relevant**. Couplings with vanishing exponents are insensitive to changes in \( L \) and are **marginal**. Those with negative exponents shrink with increasing \( L \); they are **irrelevant**.

The coupling constant \( g_{d,n,p} \) of a term with \( p \) derivatives and \( n \) powers of the field \( \phi \) in a space of \( d \) dimensions varies as

\[ g_{d,n,p}(L) = L^d A^n(L) L^{-p} g_{n,p} = L^{d-n(d-2)/2-p} g_{d,n,p}. \]

**Example 17.3 (QCD)**  In quantum chromodynamics, there is a cubic term \( g f_{abc} A_0^a A_1^b \bar{\psi}_0 \psi_1 c \) which in effect looks like \( g f_{abc} \phi_0 \phi_b \phi_c \). Is it relevant? Well, if we stretch space but not time, then the time derivative has no effect, and \( d = 3 \). So the cubic, \( n = 3 \), grows as \( L^{3/2} \)

\[ g_{3,3,0}(L) = L^{d-n(d-2)/2} g_{3,3,0} = L^{3/2} g_{3,3,0}. \]

Since this cubic term drives asymptotic freedom, its strengthening as space is stretched by the dimensionless factor \( L \) may point to a qualitative explanation of confinement. For if \( g_{3,3,0}(L) \) grows with distance as \( L^{3/2} \), then \( \alpha_s(L) = g_{3,3,0}(L)/4\pi \) grows as \( L^3 \), and so the strength \( \alpha_s(L r)/(L r)^2 \) of the force between two quarks separated by a distance \( L r \) grows linearly with \( L \)

\[ F(L r) = \frac{\alpha_s(L r)}{(L r)^2} = \frac{L^3 \alpha_s(r)}{(L r)^2} = L \frac{\alpha_s(r)}{r^2} \]

which may be enough for quark confinement.

On the other hand, if we stretch both space and time, then the cubic \( g_{4,3,1}(L) \) and quartic \( g_{4,4,0}(L) \) couplings are marginal.