A Course in Mathematical Physics: Groups, Hilbert Spaces and Differential Geometry.

by

Peter Szekeres

Solutions to Problems.

This file produces model answers to all problems in the text (not the exercises, however,which are generally of a simpler nature). These problems vary enormously in difficulty, and any teacher planning to use them in classes should be well aware of this. In many cases a lecturer may prefer to give hints, or break them up in different ways. Some problems may be better placed in a different part of the course after other topics have been developed.

There are certainly places where the reader may know of better solutions than given here. I would be very interested to hear of any such instances, and may have occasion to use such improvements (with appropriate acknowledgements) in subsequent updates of this file.

A number of problems, particularly in the later chapters, have minor annoying mistakes in the printed text which I have corrected here. These have arisen largely because the detailed solutions were not fully worked through until after the proof stage of the manuscript. I must ask for the reader's forbearance in this regard.

Peter Szekeres
Department of Physics and Mathematical Physics
University of Adelaide
South Australia 5005
AUSTRALIA

pszekere@physics.adelaide.edu.au
peterszek@netscape.net

Chapter 1.

Problem 1.1 Show the distributive laws

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C), \qquad A \cup (B \cap C) = (A \cup B) \cap (A \cup C).$$

Solution: Let $x \in A \cap (B \cup C)$. Then $x \in A$ and ($x \in B$ or $x \in C$). Hence ($x \in A$ and $x \in B$) or ($x \in A$ and $x \in C$). Equivalently, ($x \in A \cap B$) or ($x \in A \cap C$), i.e. $x \in (A \cap B) \cup (A \cap C)$. Therefore,

$$A \cap (B \cup C) \subseteq (A \cap B) \cup (A \cap C). \tag{1}$$

Conversely

$$\begin{aligned} x \in (A \cap B) \cup (A \cap C) &\implies (x \in A \cap B) \text{ or } (x \in A \cap C), \\ &\implies (x \in A \text{ and } x \in B) \text{ or } (x \in A \text{ and } x \in C) \\ &\implies x \in A \text{ and } (x \in B \text{ or } x \in C) \\ &\implies x \in A \cap (B \cup C). \end{aligned}$$

Hence

$$(A \cap B) \cup (A \cap C) \subseteq A \cap ((B \cup C). \tag{2}$$

Combining (1) and (2) and using the Axiom of extensionality or the equivalent criterion given in Section 1.2, we have $(A \cap B) \cup (A \cap C) = A \cap (B \cup C)$.

The second part is proved in a similar way

$$\begin{aligned} x \in A \cup (B \cap C) &\implies x \in A \text{ or } (x \in B \text{ and } x \in C) \\ &\implies (x \in A \cup B) \text{ and } (x \in A \cup C) \\ &\implies x \in (A \cup B) \cap (A \cup C) \quad \text{etc.} \end{aligned}$$

Problem 1.2 If $\mathcal{B} = \{B_i \,|\, i \in I\}$ is any family of sets, show that

$$A \cap \bigcup \mathcal{B} = \bigcup \{A \cap B_i \,|\, i \in I\}, \qquad A \cup \bigcap \mathcal{B} = \bigcap \{A \cup B_i \,|\, i \in I\}.$$

Solution: To extend the identities in the previous problem to arbitrary unions and intersections, note the following logical identities: If P is any proposition and $Q_i (i \in I)$ any family of propositions then

$$\begin{aligned} P \text{ and } \exists i \in I(Q_i) &\implies \exists i \in I(P \text{ and } Q_i), \\ P \text{ or } \forall i \in I(Q_i) &\implies \forall i \in I(P \text{ or } Q_i). \end{aligned}$$

Set P to be the propostition $x \in A$ and Q_i the proposition $x \in B_i$, and these logical identities immediately result in the two set theoretical statements of the problem.

Problem 1.3 **Let B be any set. Show that $(A \cap B) \cup C = A \cap (B \cup C)$ if and only if $C \subseteq A$.**

Solution: If $C \subseteq A$ then $A \cup C = A$ and by the second identity of Problem 1.1,

$$(A \cap B) \cup C = C \cup (A \cap B) = (C \cup A) \cap (C \cup B) = A \cap (C \cup B).$$

Conversely, if C is not a subset of A then there exists $y \in C$ such that $y \notin A$. Then, obviously, $y \in (A \cap B) \cup C$ but $y \notin A \cap (B \cup C)$ since $y \notin A$, hence $(A \cap B) \cup C \neq A \cap (B \cup C)$. This proves the *only if* part of the problem.

Problem 1.4 **Show that**

$$A - (B \cup C) = (A - B) \cap (A - C), \qquad A - (B \cap C) = (A - B) \cup (A - C).$$

Solution:

$$\begin{aligned} x \in A - (B \cup C) &\iff x \in A \text{ and } x \notin (B \cup C) \\ &\iff x \in A \text{ and } (x \notin B \text{ and } x \notin C) \\ &\iff (x \in A \text{ and } x \notin B) \text{ and } (x \in A \text{ and } x \notin C) \\ &\iff x \in (A - B) \text{ and } x \in (A - C) \\ &\iff x \in (A - B) \cap (A - C). \end{aligned}$$

The proof of the second identity is totally similar.

Problem 1.5 **If $\mathcal{B} = \{B_i \mid i \in I\}$ is any family of sets, show that**

$$A - \bigcup \mathcal{B} = \bigcup \{A - B_i \mid i \in I\}.$$

Solution: We use the fact that if $Q_i (i \in I)$ is any family of propositions then

$$\text{not } (\forall i \in I(Q_i)) \iff \exists i \in I(\text{not } Q_i).$$

Hence, if P is any other logical proposition then

$$\begin{aligned} P \text{ and not } (\forall i \in I(Q_i)) &\iff P \text{ and } \exists i \in I(\text{not } Q_i) \\ &\iff \exists i \in I(P \text{ and } (\text{not } Q_i)). \end{aligned}$$

Setting P to be $x \in A$ and Q_i to be the proposition $x \in B_i$, we have

$$x \in A \text{ and not } (\forall i \in I(x \in B_i)) \iff \exists i \in I(x \in A \text{ and } (x \notin B_i)),$$

which can be written

$$x \in \big(A - \bigcap_{i \in I} B_i\big) \iff x \in \bigcup_{i \in I}(A - B_i).$$

This implies the required result by the axiom of extensionality.

Problem 1.6 **If E and F are any sets, prove the identities**

$$2^E \cap 2^F = 2^{E \cap F}, \qquad 2^E \cup 2^F \subseteq 2^{E \cup F}.$$

Solution: 1. Let $A \in 2^E \cap 2^F$. Then $A \subseteq E$ and $A \subseteq F$, so that $A \subseteq E \cap F$, or equivalently $A \in 2^{E \cap F}$. Hence

$$2^E \cap 2^F \subseteq 2^{E \cap F}.$$

Conversely, let $A \in 2^{E \cap F}$. Then $A \subseteq E \cap F$, so that $A \subseteq E$ and $A \subseteq F$. Hence $A \in 2^E$ and $A \in 2^F$, or equivalently $A \subseteq 2^E \cap 2^F$. Thus

$$2^{E \cap F} \subseteq 2^E \cap 2^F$$

and by the criterion of Section 1.2, $2^E \cap 2^F = 2^{E \cap F}$.

2. Let $A \in 2^E \cup 2^F$. Then $A \subseteq E$ or $A \subseteq F$, implying that $A \subseteq E \cup F$, or equivalently $A \in 2^{E \cup F}$. Thus

$$2^E \cup 2^F \subseteq 2^{E \cup F}.$$

Equality does not hold in this case, as the converse is not true. For example, if E and F are disjoint non-empy sets ($E \cap F \neq \emptyset$). Then the set $A = E \cup F$ obviously belongs to $2^{E \cup F} = 2^A$, but it is not true that $A \in 2^E$ or $A \in 2^F$.

Problem 1.7 **Show that if $\mathcal{C}$ is any family of sets then**

$$\bigcap_{X \in \mathcal{C}} 2^X = 2^{\bigcap \mathcal{C}}, \qquad \bigcup_{X \in \mathcal{C}} 2^X \subseteq 2^{\bigcup \mathcal{C}}.$$

Solution: 1. Instead of the Axiom of Extensionality we can use a sequence of logical equivalences:

$$\begin{aligned} A \in \bigcap_{X \in \mathcal{C}} 2^X &\iff A \in 2^X \quad \forall X \in \mathcal{C} \\ &\iff A \subseteq X \quad \forall X \in \mathcal{C} \\ &\iff A \in 2^{\bigcap \mathcal{C}}. \end{aligned}$$

This shows the first identity.

2. If $A \in \bigcup_{X\in\mathcal{C}} 2^X$ then there exists $X \in \mathcal{C}$ such that $A \in 2^X$, or equivalently $A \subseteq X$. Hence $A \subset \bigcup \mathcal{C}$, so that $A \in 2^{\bigcup \mathcal{C}}$. This implies the set-theoretical inclusion, $\bigcup_{X\in\mathcal{C}} 2^X \subseteq 2^{\bigcup \mathcal{C}}$.

Problem 1.8 **Show the following identities:**

$$(A \cup B) \times P = (A \times P) \cup (B \times P),$$
$$(A \cap B) \times (P \cap Q) = (A \times P) \cap (B \times P),$$
$$(A - B) \times P = (A \times P) - (B \times P).$$

Solution: The first identity is proved by

$$\begin{aligned}(x,p) \in (A \cup B) \times P \quad &\Longleftrightarrow \quad (x \text{ or } x \in B) \text{ and } p \in P \\ &\Longleftrightarrow \quad (x,p) \in A \times P \text{ or } (x,p) \in B \times P \\ &\Longleftrightarrow \quad (x,p) \in (A \times P) \cup (B \times P).\end{aligned}$$

The second identity is similar, with $\cup$ replaced by $\cap$, and "or" by "and".

For the third identity,

$$\begin{aligned}(x,p) \in (A - B) \times P \quad &\Longleftrightarrow \quad (x \text{ and } x \notin B) \text{ and } p \in P \\ &\Longleftrightarrow \quad (x,p) \in A \times P \text{ and } (x,p) \notin B \times P \\ &\Longleftrightarrow \quad (x,p) \in (A \times P) - (B \times P).\end{aligned}$$

Problem 1.9 **If $\mathcal{A} = \{A_i \mid i \in I\}$ and $\mathcal{B} = \{B_j \mid j \in J\}$ are any two families of sets then**

$$\bigcup \mathcal{A} \times \bigcup \mathcal{B} = \bigcup_{i\in I, j\in J} A_i \times B_j,$$
$$\bigcap \mathcal{A} \times \bigcap \mathcal{B} = \bigcap_{i\in I, j\in J} A_i \times B_j.$$

Solution: For the first identity,

$$\begin{aligned}(x,y) \in \bigcup \mathcal{A} \times \bigcup \mathcal{B} \quad &\Longleftrightarrow \quad \exists i \in I, \exists j \in J \ \text{ s.t. } \ x \in A_i \text{ and } y \in B_j \\ &\Longleftrightarrow \quad \exists i \in I, \exists j \in J \ \text{ s.t. } \ (x,y) \in A_i \times B_j \\ &\Longleftrightarrow \quad (x,y) \in \bigcup_{i\in I, j\in J} A_i \times B_j.\end{aligned}$$

The second identity follows on changing "cups" to "caps" and replacing $\exists$ by $\forall$.

Problem 1.10 Show that both the following two relations

$$(a,b) \leq (x,y) \quad \textbf{iff} \quad a < x \ \textbf{or} \ (a = x \ \textbf{and} \ b \leq y)$$
$$(a,b) \preceq (x,y) \quad \textbf{iff} \quad a \leq x \ \textbf{and} \ b \leq y$$

are partial orders on $\mathbb{R} \times \mathbb{R}$. For any pair of partial orders $\leq$ and $\preceq$ defined on an arbitrary set A, let us say that $\leq$ is *stronger* than $\preceq$ if $a \leq b \ \rightarrow \ a \preceq b$. Is $\leq$ stronger than, weaker than or incomparable with $\preceq$?

Solution: The relation $\leq$ is a kind of "lexicographic" order on ordered pairs of real numbers. It is a partial order since

1. It is reflexive: for all (a,b), we have $(a,b) \leq (a,b)$ on setting $x = a$ and $y = b$, as it is always true the $b \leq b$.

2. Transitive: If $(a,b) \leq (c,d)$ and $(c,d) \leq (e,f)$ then there are four possibilities:

(i) $a < c$ and $c < e$.

(ii) $a < c$ and $(c = e$ and $d \leq f)$.

(iii) $(a = c$ and $b \leq d)$ and $c < e$.

(iv) $(a = c$ and $b \leq d)$ and $(c = e$ and $d \leq f)$.

By the transitivity of $\leq$ on $\mathbb{R}$, we see in the first three cases (i)-(iii) that $c < e$, while in case (iv) $a = e$ and $b \leq f$. In either eventuality $(a,b) \leq (e,f)$.

3. For antisymmetry, if $(a,b) \leq (x,y)$ and $(x,y) \leq (a,b)$ then $a \leq x$ and $x \leq a$, whence $a = x$. We must then have $b \leq y$ and $y \leq b$, so that $b = y$. hence $(a,b) = (x,y)$.

The proof that $\preceq$ is a partial order is similar, but rather easier since it need not be broken down into different cases. For example, to show transitivity: If $(a,b) \preceq (c,d)$ and $(c,d) \preceq (e,f)$ then $a \leq c \leq e$ and $b \leq d \leq f$, whence $a \leq e$ and $b \leq f$ so that $(a,b) \preceq (e,f)$.

If $(a,b) \preceq (x,y)$ then $a \leq x$ and $b \leq y$, so that $a < x$ or $(a = x$ and $b \leq y)$. Hence

$$(a,b) \preceq (x,y) \quad \Longrightarrow \quad (a,b) \leq (x,y),$$

or the partial order $\preceq$ is stronger that $\leq$. Alternatively, we say that $\leq$ is weaker than $\preceq$. It is definitely not stronger, for even if $b > y$ we have $(a,b) \leq (a,y)$, but it is not true that $(a,b) \preceq (a,y)$.

Problem 1.11 There is a technical flaw in the proof of Theorem 1.5, since a decimal number ending in an endless sequence of 1's is identified

with a decimal number ending with a sequence of 0's, for example,

$$.011011111... = .0111000000...$$

Remove this hitch in the proof.

Solution: Decimal numbers in the interval $[0, 1]$ having two representations have the form

$$.a_1 a_2 \dots a_n 10000 \dots$$

They may be set in a sequence

$$\begin{aligned} x_1 &= .1 \\ x_2 &= .01 \\ x_3 &= .11 \\ x_4 &= .001 \\ x_5 &= .011 \\ x_6 &= .101 \\ x_7 &= .111 \\ x_8 &= .0001 \\ &\vdots \end{aligned}$$

Thus these numbers which are "double-counted" form a countable set. The Cantor proof shows that all decimal sequences are uncountable. These consist of all real numbers (doubly-counted numbers given a decimal expansion ending in a sequence of 1's) and the above countable set of decimal expansions ending in 0's. Since the union of two countable sets is always countable the real numbers must be uncountable.

Problem 1.12 **Prove the assertion that the Cantor set is nowhere dense.**

Solution: It is required to show that the Cantor set C is not dense in any open interval (x, y) such that $0 < x < y < 1$. At the n-th stage of creation of the Cantor set, the total length of the closed intervals in which C eventually lies is $(2/3)^n$, while the complement consists of a family of disjoint open intervals of the form (a, b) having total length $1 - (2/3)^n$. Chose n such that $(2/3)^n < |x - y|$. The complement of C must then have non-empty intersection with (x, y), and contains an interval (x', y') such that $x < x' < y' < y$. There is thus no point of $a \in C$ such that $x' < a < y'$, and C is not dense in (x, y). Since (x, y) is an arbitrary open interval, the Cantor set is nowhere dense.

Problem 1.13 **Prove that the set of all real functions $f : \mathbb{R} \to \mathbb{R}$ has a higher cardinality than that of the real numbers by using a Cantor**

diagonal argument to show it cannot be put in one-to-one correspondence with $\mathbb{R}$.

Solution: Suppose the real functions could be put in one-to-one correspondence with the real nubers. Then they could be labelled by real numbers, $f_x : \mathbb{R} \to \mathbb{R}$, one and exactly one function for each real number x. Let $F : \mathbb{R} \to \mathbb{R}$ be the real function defined by $F(x) = f_x(x) + 1$. This function clearly cannot be labelled by any real number, for if $F = f_y$ for some $y \in \mathbb{R}$, then

$$f_y(y) = F(y) = f_y(y) + 1,$$

an obvious contradiction.

Problem 1.14 **If $f : [0,1] \to \mathbb{R}$ is a non-decreasing function such that $f(0) = 0$, $f(1) = 1$, show that the places at which f is not continuous form a countable subset of $[0,1]$.**

Solution: Let f_- and f_+ be the functions defined by

$$f_-(x) = \sup_{y<x} f(y) \qquad \text{and} \qquad f_+(x) = \inf_{y>x} f(y).$$

Since the function is non-decreasing it is clear that $f_-(x) \le f(x)$ and $f_+(x) \ge f(x)$ for all $x \in [0,1]$. The function f is continuous at x iff the limit both from left and right is $f(y) \to f(x)$ as $y \to x$, from which it follows that f is continuous at x if and only if $f_-(x) = f(x) = f_+(x)$. If we set $\Delta f(x) = f_+(x) - f_-(x)$ then $\Delta f(x) \ge 0$ for all $x \in [0,1]$. At the points where f is continuous $\Delta f(x) = 0$, while at points of discontinuity $\Delta f(x) > 0$.

For any finite sequence of points $x_1, x_2, \dots, x_m$ it is evident that $\sum_{i=1}^{m} \Delta f(x_i) \le 1$. Hence, there can be at most one point such that $\frac{1}{2} < \Delta f(x) \le 1$, at most 2 points such that $\frac{1}{4}\Delta f(x) \le \frac{1}{2}$, and in general at most 2^n points such that $2^{-(n+1)} < \Delta f(x) \le 2^{-n}$. Since the points at which $f(x)$ is discontinuous must be such that $\Delta f(x)$ falls in one of these half-open intervals, we can arrange them in a sequence such that the $\Delta f(x)$ lie in intervals of succesively decreasing upper bound 2^{-n}, while arranging points of discontinuity for which Δf lies in the same interval in, say, increasing order. All points of discontinuity of f are thus set out in a sequence and must form a countable set.

Problem 1.15 **Show that in the category of sets a morphism is an epimorphism if and only if it is onto (surjective).**

Solution: 1. Suppose $\varphi : A \to B$ is a surjective (onto) map. We want to show that for any maps $\beta, \beta' : B \to Y$,

$$\beta \circ \varphi = \beta' \circ \varphi \quad \Longrightarrow \quad \beta = \beta'.$$

If $\beta \circ \varphi = \beta' \circ \varphi$ then for all $a \in A$, $\beta(\varphi(a)) = \beta'(\varphi(a))$. Since φ is surjective every $b \in B$ can be written in form $b = \varphi(a)$ for some $a \in A$. Hence $\beta(b) = \beta'(b)$ for all $b \in B$; that is, $\beta = \beta'$ as required for an epimorphism.

2. If φ is an epimorphism, we require to show that it is a surjection. If not, then there exists $b_0 \in B$ such that $b_0 \neq \varphi(a)$ for any $a \in A$. Let Y be any doubleton set $Y = \{x, y\}$ and set $\beta : B \to Y$ to be the map

$$\beta(b_0) = x, \qquad \beta(b) = y \quad \text{for all } b \neq b_0$$

while $\beta' : B \to Y$ is defined by $\beta'(b) = y$ for all $b \in B$. Clearly $\varphi \circ \beta = \varphi \circ \beta'$ since $\varphi(\beta(a)) = \varphi(\beta'(a)) = \varphi(b) = y$ for all $a \in A$. Yet $\beta \neq \beta'$, contradicting the epimorphism property of φ. Hence φ is a surjection.

Problem 1.16 Show that every isomorphism is both a monomorphism and an epimorphism.

Solution: Let $\varphi : B \to A$ be an isomorphism, and $\varphi' : B \to A$ the map such that

$$\varphi' \circ \varphi = \iota_A \qquad \text{and} \qquad \varphi \circ \varphi' = \iota_B.$$

1. To show φ is a monomorphism, let $\alpha, \alpha' : A \to X$ be maps, then

$$\begin{aligned} \varphi \circ \alpha = \varphi \circ \alpha' \quad &\Longrightarrow \quad \varphi' \circ \varphi \circ \alpha = \varphi' \circ \varphi \circ \alpha' \\ &\Longrightarrow \quad \iota_A \circ \alpha = \iota_A \circ \alpha' \\ &\Longrightarrow \quad \iota_A(\alpha(x)) = \iota_A(\alpha'(x)) \quad \forall x \in X \\ &\Longrightarrow \quad \alpha(x) = \alpha'(x) \quad \forall x \in X \\ &\Longrightarrow \quad \alpha = \alpha'. \end{aligned}$$

2. Similarly, φ is an epimorphism since

$$\begin{aligned} \beta \circ \varphi = \beta' \circ \varphi \quad &\Longrightarrow \quad \beta \circ \varphi \circ \varphi' = \beta' \circ \varphi \circ \varphi' \\ &\Longrightarrow \quad \beta \circ \iota_B = \beta' \circ \iota_B \\ &\Longrightarrow \quad \beta(\iota_B(b)) = \beta'(\iota_B(b)) \quad \forall b \in B \\ &\Longrightarrow \quad \beta(b) = \beta'(b) \\ &\Longrightarrow \quad \beta = \beta'. \end{aligned}$$

Chapter 2.

Problem 2.1 Show that the only finite subgroup of the additive reals is the singleton $\{0\}$, while the only finite subgroups of the multiplicative reals are the sets $\{1\}$ and $\{1, -1\}$.

Find all finite subgroups of the multiplicative complex numbers $\dot{\mathbb{C}}$.

Solution: If $a \neq 0$ belongs to a subgroup G of the additive group $\mathbb{R}$, then so do all real numbers

$$na = \underbrace{a + a + \dots a}_{n} \ .$$

Hence G contains an infinite number of elements unless $G = \{0\}$. Similarly if $a \neq \pm 1$ belongs to a multiplicative subgroup G of $\mathbb{R}$, then so do the infinite sequence of real numbers

$$a^n = \underbrace{a.a.\dots.a}_{n} \ .$$

Hence, if G is a finite group it can at most contain the numbers $\{1, -1\}$. This is obviously a subgroup of $\mathbb{R}$, and its only proper subgroup is the trivial singleton group $\{1\}$.

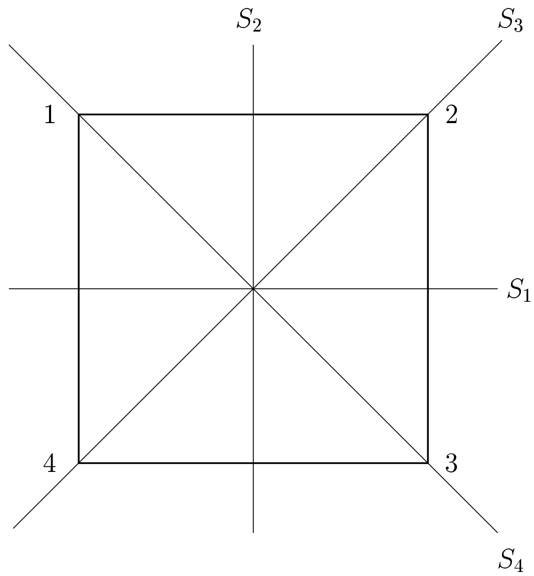
For the multiplicative complex numbers $\mathbb{C}$, if $\lambda \in G$ and $|\lambda| \neq 1$ then the numbers λ^n form a sequence of increasing magnitude if $|\lambda| > 1$, or decreasing magnitude if $|\lambda| < 1$. In either case they form an infinite set. Hence a finite group G can only consist of complex numbers of unit magnitude, $|\lambda| = 1$. Writing $\lambda = \mathrm{e}^{ia}$, if G is a finite group then for some integer n we must have $\lambda^n = \mathrm{e}^{ina} = 1$. Thus $a = 2\pi/n$, and the group G is a cyclic group of the form

$$G = \{\mathrm{e}^{2\pi i/n}, \mathrm{e}^{4\pi i/n}, \dots \mathrm{e}^{2k\pi i/n}, \dots, \mathrm{e}^{2\pi i} = 1\}.$$

Problem 2.2 Write out the complete 8×8 multiplication table for the group of symmetries of the square D_4 described in Example 2.7. Show that R_2 and S_1 generate an abelian subgroup and write out its multiplication table.

Solution: Expressed as permutations in cyclic notation of the corners of the square the elements of D_4 are

$$\begin{aligned} R_0 &= (1)(2)(3)(4) \qquad & S_1 &= (14)(23) \\ R_1 &= (1234) & S_2 &= (12)(34) \\ R_2 &= (13)(24) & S_3 &= (13) \\ R_3 &= (1432) & S_4 &= (24) \end{aligned}$$



Using the multiplication rules for permutations we find the multiplication table (of course this can also be done by following through the geometrical consequences of performing the operations)

R_0	R_1	R_2	R_3	S_1	S_2	S_3	S_4
R_1	R_2	R_3	R_0	S_4	S_3	S_1	S_2
R_2	R_3	R_0	R_1	S_2	S_1	S_4	S_3
R_3	R_0	R_1	R_2	S_3	S_4	S_2	S_1
S_1	S_3	S_2	S_4	R_0	R_2	R_1	R_3
S_2	S_4	S_1	S_3	R_2	R_0	R_3	R_1
S_3	S_2	S_4	S_1	R_3	R_1	R_0	R_2
S_4	S_1	S_3	S_2	R_1	R_3	R_2	R_0

From this table we see that $R_2S_1 = S_1R_2 = S_2$, and

$$S_2R_2 = R_2S_2 = S_1, \qquad S_2S_1 = S_1S_2 = R_2.$$

Hence the elements the elements R_0, R_2, S_1, S_2 are closed with respect to group product. To see they form a subgroup we write the multiplication table,

R_0	R_2	S_1	S_2
R_2	R_0	S_2	S_1
S_1	S_2	R_0	R_2
S_2	S_1	R_2	R_0

These elements form a group since every element is its own inverse, $(R_2)^{-1} = R_2$, $(S_1)^{-1} = S_1$, $(S_2)^{-1} = S_2$, and it is abelian since the multiplication table is symmetric with respect to the main diagonal.

Problem 2.3 **(a) Find the symmetries of the cube, Fig. 2.2(a), which keep the vertex 1 fixed. Write these symmetries as permutations of the vertices in cycle notation.**

(b) Find the group of rotational symmetries of the regular tetrahedron depicted in Fig. 2.2(b).

(c) Do the same for the regular octahedron, Fig. 2.2(c).

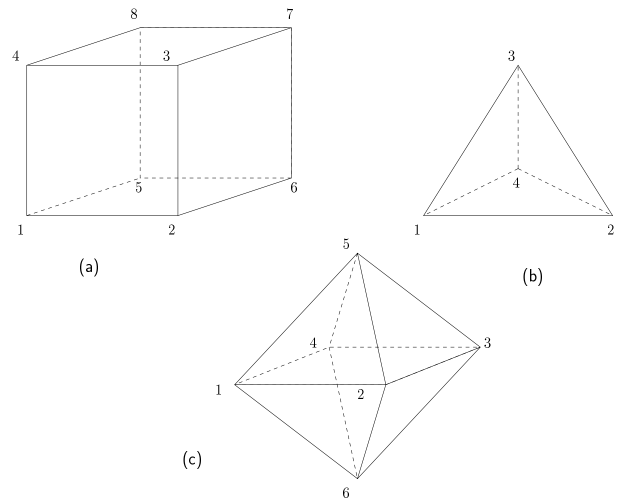


Figure 1: Text Figure 2.2

Solution: (a) There are 3 rotations about the 17 axis: id, (245)(368), (254)(386), and 3 reflections

$$\begin{aligned} \text{with respect to the } 1476-\text{plane}: &\quad (25)(38) \\ 1278-\text{plane}: &\quad (36)(45) \\ 1375-\text{plane}: &\quad (24)(68)\,. \end{aligned}$$

(b) There are 8 rotations about axes from a vertex to the the centre of the opposite face:

$$\begin{aligned} \text{axis 1 to centre of triangle 234}: &\quad (234), \quad (243) \\ \text{2 to centre of triangle 134}: &\quad (134), \quad (143) \\ \text{3 to centre of triangle 124}: &\quad (124), \quad (142) \\ \text{4 to centre of triangle 123}: &\quad (123), \quad (132) \end{aligned}$$

There are 3 further rotations by 180^o about axes connecting midpoints of edges:

$$\begin{aligned} \text{axis: midpoint of 14 to midpoint of 23}: &\quad (14)(23) \\ \text{midpoint of 13 to midpoint of 24}: &\quad (13)(24) \\ \text{midpoint of 12 to midpoint of 34}: &\quad (12)(34) \end{aligned}$$

These 11 rotations and the identity transformation form the 12 element group of rotational symmetries of the tetrahedron. It is straightforward to check group closure: e.g. $(234)(134) = (14)(23)$ etc.

(c) For the octahedron the rotations have axes given on the left:

Axis	:	Rotations		
56	:	(1234),	(13)(24),	(1432)
13	:	(2546),	(24)(56),	(2645)
24	:	(1635),	(13)(56),	(1536)
centre of 145 to centre of 236	:	(154)(236),	(145)(236)	
centre of 125 to centre of 346	:	(125)(346),	(152)(364)	
centre of 162 to centre of 354	:	(162)(354),	(126)(345)	
centre of 146 to centre of 253	:	(146)(253),	(164)(235)	
Midpoint of 12 to midpoint of 34	:	(12)(34)(56)		
Midpoint of 14 to midpoint of 23	:	(14)(23)(56)		
Midpoint of 15 to midpoint of 36	:	(15)(24)(36)		
Midpoint of 16 to midpoint of 35	:	(16)(24)(35)		
Midpoint of 25 to midpoint of 46	:	(13)(25)(46)		
Midpoint of 26 to midpoint of 45	:	(13)(26)(45)		

These 23 rotations and the identity transformation form the 24 element rotational group of symmetries of the octahedron. Closure is straightforward, if tedious, to check: e.g.

$$(2346)(1234) = (154)(236), \qquad (13)(56)(1234) = (12)(34)(56)\text{etc.}$$

Problem 2.4 Show that the multiplicative groups modulo a prime G_7, G_{11}, G_{17}, and G_{23} are cyclic. In each case finding a generator of the group.

Solution: The powers of 2 do not run through G_7, for $2^1 = 2, 2^2 = 4, 2^3 = 8 \equiv 1$ mod 7. However the powers of 3 do run through all non-zero classes modulo 7

$$3^1 = 3,\ 3^2 = 2,\ 3^3 \equiv 2 \cdot 3 = 6,\ 3^4 \equiv 6 \cdot 3 \equiv 4,\ 3^5 \equiv 5,\ 3^6 \equiv 1.$$

Hence G_7 is cyclic with generator 3.

The element 2 is a generator of G_{11} for its successive powers are

$$2,\ 4,\ 8,\ 5,\ 10,\ 9,\ 7,\ 3,\ 6,\ 1\,.$$

G_{17} is not generated by 2 (successive powers 2,4,8,16,15,13,9,1) but it is generated by 3, whose successive powers are

$$3,\ 9,\ 10,\ 13,\ 5,\ 15,\ 11,\ 16,\ 14,\ 8\ 7,\ 4,\ 12,\ 2,\ 6,\ 1\,.$$

G_{23} is not generated by either 2 or 3, as is easily verified, but it is cyclic since it is generated by the element 5 whose successive powers modulo 23 are

$$\begin{gathered}5,\ 2,\ 10,\ 4,\ 20,\ 8,\ 17,\ 16,\ 11,\ 9,\ 22 \equiv -1,\\ 18,\ 21,\ 13,\ 19,\ 3,\ 15,\ 6,\ 7,\ 12,\ 14,\ 1.\end{gathered}$$

Problem 2.5 Show that the order of any cyclic subgroup of S_n is a divisor of $n!$.

Solution: Let $\pi = (a_1 a_2 \dots a_k)(b_1 b_2 \dots b_l)\dots$ be any permutation of $12\dots n$. Every cycle of k elements has order k (see Example 2.9), so that π has order $m =$ l.c.d.$(k, l, \dots)$. It must therefore be shown that

$$m = \text{lcd}\ (k, l, \dots) < n! = 1.2.3.\dots(k + l + \dots)$$

where lcd is short for "lowest common denominator". For two primes p, q it is obviously true that lcd $(p, q) = pq < (p+q)!$ since both $p < p+q$ and $q < p+q$ appear separately in the factorial product (if $p = q$ it is even easier, since lcd $(p, p) = p$). A similar argument applies to prime powers,

$$\text{lcd}\ (p^a, q^b) = p^a q^b < (p^a + q^b)!$$

and extends in a straighforward way to any collection of prime powers

$$\text{lcd}\ (p_1^{a_1}, p_2^{a_2}, \dots, p_r^{a_r}) = p_1^{a_1} p_2^{a_2} \dots p_r^{a_r} < \left(p_1^{a_1} + p_2^{a_2} + \dots + p_r^{a_r}\right)!$$

The general result follows from the fundamental theorem of arithmetic that every natural number k has a unique expansion as a product of prime powers, $k = p_1^{a_1} p_2^{a_2} \dots p_s^{a_s}$.

Problem 2.6 Show that the following sets of matrices form groups with respect to addition of matrices, but that none of them is a group with respect to matrix multiplication: (i) real antisymmetric $n \times n$ matrices ($\mathsf{A}^T = -\mathsf{A}$), (ii) real $n \times n$ matrices having vanishing trace ($\operatorname{tr} \mathsf{A} = \sum_{i=1}^n a_{ii} = 0$), (iii) complex Hermitian $n \times n$ matrices ($\mathsf{H}^\dagger = \mathsf{H}$).

Solution: (i) If $\mathsf{A}^T = -\mathsf{A}$ and $\mathsf{B}^T = -\mathsf{B}$, then we show the closure property by

$$(\mathsf{A} + \mathsf{B})^T = \mathsf{A}^T + \mathsf{B}^T = -\mathsf{A} - \mathsf{B} = -(\mathsf{A} + \mathsf{B}).$$

The additive identity O is obviously antisymmetric, $\mathsf{O}^T = \mathsf{O} = -\mathsf{O}$, and the additive inverse $-\mathsf{A}$ of an antisymmetric matrix is antisymmetric

$$\mathsf{A}^T = -\mathsf{A} \quad \Longrightarrow \quad (-\mathsf{A})^T = -\mathsf{A}^T = \mathsf{A} = -(-\mathsf{A}).$$

Antisymmetric matrices do not form a multiplicative group as, for example, the multiplicative identity I is clearly not antisymmetric.

(ii) The additive closure property is trivial for trace-free matrices,

$$\operatorname{tr}\mathsf{A} = \operatorname{tr}\mathsf{B} = 0 \quad \Longrightarrow \quad \operatorname{tr}(\mathsf{A}+\mathsf{B}) = \operatorname{tr}\mathsf{A} + \operatorname{tr}\mathsf{B} = 0 + 0 = 0.$$

The zero matrix clearly has zero trace, while the negative $-\mathsf{A}$ of a trace-free matrix is also trace-free,

$$\operatorname{tr}\mathsf{A} = 0 \quad \Longrightarrow \quad \operatorname{tr}(-\mathsf{A}) = -\operatorname{tr}\mathsf{A} = 0.$$

These matrices do not form a multiplicative group since $\operatorname{tr}\mathsf{I} = n \neq 0$.

(iii) Additive closure follows from

$$(\mathsf{H}+\mathsf{K})^\dagger = \mathsf{H}^\dagger + \mathsf{K}^\dagger = \mathsf{H} + \mathsf{K}.$$

Additive identity and inverse hold trivially,

$$\mathsf{O}^\dagger = \mathsf{O}, \qquad \mathsf{H}^\dagger = \mathsf{H} \Longrightarrow (-\mathsf{H})^\dagger = -\mathsf{H}^\dagger = -\mathsf{H}.$$

The multiplicative identity I is Hermitian, there need not be inverses. For example the multiplicative inverse of O does not exist. Furthermore, Hermitian matrices are not closed with respect multiplication as the product of two Hermitian matrices need not be Hermitian, since $(\mathsf{HK})^\dagger = \mathsf{K}^\dagger\mathsf{H}^\dagger = \mathsf{KH} \neq \mathsf{HK}$ in general. For example,

$$\begin{pmatrix} 0 & \alpha \\ \overline{\alpha} & 0 \end{pmatrix}\begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix} = \begin{pmatrix} -i\alpha & 0 \\ 0 & i\overline{\alpha} \end{pmatrix} \neq \begin{pmatrix} -i\alpha & 0 \\ 0 & i\overline{\alpha} \end{pmatrix}^\dagger.$$

Problem 2.7 **Find a diagonal complex matrix S such that**

$$\mathsf{I} = \mathsf{S}^T\mathsf{G}_p\mathsf{S}$$

where G_p is defined in Example 2.12. Show

(a) Every complex matrix A satisfying Eq. (2.10) can be written in the form

$$\mathsf{A} = \mathsf{S}\mathsf{B}\mathsf{S}^{-1}$$

where B is a complex orthogonal matrix (i.e. a member of $O(n,\mathbb{C})$).

(b) The complex versions of the pseudo-orthogonal groups, $O(p,q,\mathbb{C})$, are all isomorphic to each other if they have the same dimension,

$$O(p,q,\mathbb{C}) \cong O(n,\mathbb{C}) \qquad \textbf{where} \qquad n = p+q.$$

Solution: The diagonal matrix

$$\mathsf{S} = \begin{pmatrix} \mathsf{I}_p & \mathsf{O} \\ \mathsf{O} & i\mathsf{I}_q \end{pmatrix}$$

where I_p and I_q are the $p \times p$ and $q \times q$ unit matrices respectively, and O refers to rectangular zero matrices of appropriate dimensions, is easily seen to satisfy the equation $\mathsf{I} = \mathsf{S}^T \mathsf{G}_p \mathsf{S}$. Note that, by multiplying from the left by $(\mathsf{S}^{-1})^T$ and on the right by S^{-1} we have

$$(\mathsf{S}^{-1})^T S^{-1} = \mathsf{G}_p \qquad (*)$$

(a) Let A be any matrix satisfying Eq. (2.10), $\mathsf{A}^T \mathsf{G}_p \mathsf{A} = \mathsf{G}_p$, Set $\mathsf{B} = \mathsf{S}^{-1}\mathsf{AS}$. By multiplying from the left by S and the right by S^{-1} we find $\mathsf{A} = \mathsf{SBS}^{-1}$, and

$$\begin{aligned} \mathsf{B}^T\mathsf{B} &= \mathsf{S}^T\mathsf{A}^T(\mathsf{S}^{-1})^T S^{-1}\mathsf{AS} \\ &= \mathsf{S}^T\mathsf{A}^T\mathsf{G}_p\mathsf{AA} \qquad \text{by Eq. } (*) \\ &= \mathsf{S}^T\mathsf{G}_p\mathsf{S} = \mathsf{I}. \end{aligned}$$

That is, B belongs to $O(n, \mathbb{C})$).

(b) Let $\varphi : O(p, q, \mathbb{C}) \to O(n, \mathbb{C})$ be the map $\varphi(\mathsf{A}) = \mathsf{B} = \mathsf{S}^{-1}\mathsf{AS}$. This map is one-to-one, for its inverse is $\varphi^{-1}(\mathsf{B}) = \mathsf{A} = \mathsf{SBS}^{-1}$ and is always an element of $O(p, q, \mathbb{C})$, since

$$\mathsf{A}^T\mathsf{G}_p\mathsf{A} = (\mathsf{S}^{-1})^T\mathsf{B}^T\mathsf{S}^T\mathsf{G}_p\mathsf{SBS}^{-1} = (\mathsf{S}^{-1})^T\mathsf{B}^T\mathsf{IBS}^{-1} = (\mathsf{S}^{-1})^T\mathsf{IS}^{-1} = \mathsf{G}_p.$$

The homomorphism property follows from

$$\varphi(\mathsf{A}_1)\varphi(\mathsf{A}_2) = \mathsf{S}^{-1}\mathsf{A}_1\mathsf{SS}^{-1}\mathsf{A}_2\mathsf{S} = \mathsf{S}^{-1}\mathsf{A}_1\mathsf{A}_2\mathsf{S} = \varphi(\mathsf{A}_1\mathsf{A}_2).$$

Hence φ is an isomorphism.

Problem 2.8 **Show that every element of $SU(2)$ has the form**

$$\mathsf{U} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \qquad \textbf{where} \qquad a = \overline{d} \textbf{ and } b = -\overline{c}.$$

Solution: U is a 2×2 unitary matrix if

$$\mathsf{I} = \mathsf{UU}^\dagger = \begin{pmatrix} a\overline{a} + b\overline{b} & a\overline{c} + b\overline{d} \\ c\overline{a} + d\overline{b} & c\overline{c} + d\overline{d} \end{pmatrix}$$

which gives four equations

$$a\overline{a} + b\overline{b} = 1 \qquad (1)$$

$$c\overline{c} + d\overline{d} = 1 \qquad (2)$$

$$a\overline{c} + b\overline{d} = 0 \qquad (3)$$

$$c\overline{a} + d\overline{b} = 0 \qquad (4)$$

and if $\mathsf{U} \in SU(2)$, we also have unimodularity

$$\det \mathsf{U} = ad - bc = 1 \tag{5}$$

Eq. (4) is the complex conjugate of (3), and therefore adds nothing new. If we multiply Eq. (3) by $\overline{a}$ and use Eqs. (1) and (5) we obtain

$$0 = (1 - b\overline{b})\overline{c} + \overline{a}b\overline{d} = \overline{c} + b(\overline{a}\overline{d} - \overline{b}\overline{c}) = \overline{c} + b.$$

Hence $b = -\overline{c}$. Similarly, multiplying (3) by c gives

$$0 = a + \overline{d}(bc - ad) = a - \overline{d}.$$

Thus $a = \overline{d}$. The only remaining relation, Eq. (1) is equivalent to the unimodularity condition (5).

Problem 2.9 Show that Theorem 2.2 may be extended to infinite groups as well. That is, any group G is isomorphic to a subgroup of Transf(G), the transformation group of the set G.

Solution: The proof of the Cayley's theorem applies essentially unchanged for infinite groups. Let $\varphi : G \to \text{Transf}(G)$ be defined by $\varphi g = L_g$ where $L_g(h) = gh$. The map $L_g : G \to G$ is a transformation of G for it is one-to-one and onto:

$$L_g(h) = L_g(h') \Longrightarrow gh = gh' \Longrightarrow h = h', \qquad h \in G \Longrightarrow h = L_g(g^{-1}h).$$

The proof that φ is a homomorphism follows as in the proof of Theorem 2.2. The map $\varphi : G \to \varphi(G) \subseteq \text{Transf}(G)$ is one-to-one and onto.

Problem 2.10 Find the group multiplication tables for all possible groups on four symbols e, a, b and c, and show that any group of order 4 is either isomorphic to the cyclic group $\mathbb{Z}_4$ or the product group $\mathbb{Z}_2 \times \mathbb{Z}_2$.

Solution: We have $a^2 = e, b$ or c since $a^2 = a \Rightarrow \ a = e$. Firstly consider the case $a^2 = e$. Then $ab = c$, for $ab = a$ would imply that $b = e$; $ab = b$ is only possible if $a = e$; and if $ab = e$ then $b = a$. Similarly $ba = c$ and $ac = ca = b$. Furthermore $b^2 = caac = cec = c^2$. But $b^2 \neq b$ and $c^2 \neq c$ since neither b nor c is equal to the identity e. Hence $b^2 = e$ or $b^2 = a$.
Case (i) $b^2 = c^2 = e$. The only remaining possibility for bc is $bc = cb = a$ and the multiplication table is

e	a	b	c
a	e	c	b
b	c	e	a
c	b	a	e

This is isomorphic with the product group $\mathbb{Z}_2 \times \mathbb{Z}_2$, on making the association $e \longleftrightarrow (0,0)$, $a \longleftrightarrow (0,1)$, $b \longleftrightarrow (1,0)$, $c \longleftrightarrow (1,1)$.
Case(ii) $b^2 = c^2 = a$. Since $be = b$ and $ba = c$ the only remaining possibility for the product of b and c is $bc = cb = e$. The multiplication table is

e	a	b	c
a	e	c	b
b	c	a	e
c	b	e	a

This is the cyclic group of order 4, with b or c as generator.

There remains the case $a^2 = b$ (the case $a^2 = c$ is similar). We cannot have $ab = e$ for then $b = a^{-1}$ and therefore $ba = e$. In this case $ac = b$ and we have $bc = a^2c = ab = e$, so that $c = a$ which is a contradiction to the elements e, a, b, c being all distinct. Hence $ab = a^3 = c$. The group is therefore cyclic with generator a and has multiplication table

e	a	b	c
a	b	c	e
b	c	e	a
c	e	a	b

Problem 2.11 **Show that every cyclic permutation** $(a_1\, a_2 \dots a_n)$ **has the property that for any permutation** π

$$\pi(a_1\, a_2 \dots a_n)\pi^{-1}$$

is also a cycle of length n. **[*Hint*: It is only necessary to show this for transpositions** $\pi = (b_1\, b_2)$ **as every permutation is a product of such transpositions.]**

(a) Show that the conjugacy classes of S_n **consist of those permutations having the same cycle structure, e.g.** $(1\,2\,3)(4\,5)$ **and** $(1\,4\,6)(2\,3)$ **belong to the same conjugacy class.**

(b) Write out all conjugacy classes of S_4 **and calculate the number of elements in each class.**

Solution: For an interchange $\pi = (b_1\, b_2) = \pi^{-1}$ such that $b1, b2 \notin \{a_1\, a_2 \dots a_n\}$, it is obvious that $\pi(a_1\, a_2 \dots a_n)\pi^{-1} = (a_1\, a_2 \dots a_n)$. if $b_1 = a_r$ for some $r = 1, \dots n$, but $b_2 \notin \{a_1\, a_2 \dots a_n\}$, then $\pi(a_1\, a_2 \dots a_n)\pi^{-1} = (a_1\, a_2 \dots a_{r-1} b_2 a_{r+1} \dots a_n)$; i.e. a_r is replaced by b_2 in the cycle. If $b_1 = a_r$ and $b_2 = a_s$ with $r < s$ then the result interchanges a_r and a_s:

$$(a_r a_s)(a_1\, a_2 \dots a_n)(a_r a_s) = (a_1 \dots a_{r-1} a_s a_{r+1} \dots a_{s-1} a_r a_{s+1} \dots a_n).$$

In all cases the result is a cyclic of the same length. For an arbitrary permutation π, write it as a sequence of interchanges, $\pi = \pi_1\pi_2\ldots\pi_k$, and

$$\pi(a_1\ldots a_n)\pi^{-1} = \pi_1\pi_2\ldots\pi_k(a_1\ldots a_n)\pi_k\pi_{k-1}\ldots\pi_1$$

is a cycle of length n.

(a) For any permutation $\sigma = (a_1\ldots a_p)(b_1\ldots b_q)\ldots(c_1\ldots c_r)$ the permutation $\pi\sigma\pi^{-1}$ has the same cycle structure, since

$$\pi\sigma\pi^{-1} = \pi(a_1\ldots a_p)\pi^{-1}\pi(b_1\ldots b_q)\pi^{-1}\pi\ldots\pi^{-1}\pi(c_1\ldots c_r)\pi^{-1}$$

Furthermore, by selected choices of interchanges it is possible to convert any cycle to any other cycle of the same length,

$$(a_1b_1)(a_2b_2)\ldots(a_nb_n)(a_1\,a_2\ldots a_n)(a_1b_1)(a_2b_2)\ldots(a_nb_n) = (b_1b_2\ldots b_n).$$

Hence the conjugacy class $(S_n)_\sigma$ of the permutation σ consists of precisely all permutations having the same cycle structure $(a_1\ldots a_p)(b_1\ldots b_q)\ldots(c_1\ldots c_r)$.

(b) There are 5 possible cycle structures for the elements of S_4, and the number of elements in each is the following:

Representative	No. of elements
(1234)	6
(123)(4)	8
(12)(34)	3
(12)(3)(4)	6
(1)(2)(3)(4)	1

Note the number of elements adds up to 24, which is the order of S_4, as required.

Problem 2.12 Show that the class of groups as objects with homomorphisms between groups as morphisms forms a category—the *category of groups* (see Section 1.7). What are the monomorphisms, epimorphisms and isomorphisms of this category?

Solution: To show that this class of objects and morphisms forms a category we must show:

1. The composition of two homomorphisms $\varphi : H \to K$ and $\psi : G \to H$ is a homomorphism. This follows from

$$\varphi\circ\psi(g_1g_2) = \varphi(\psi(g_1g_2)) = \varphi(\psi(g_1)\psi(g_2)) = \varphi(\psi(g_1))\varphi(\psi(g_2)) = \varphi\circ\psi(g_1)\varphi\circ\psi(g_2)$$

2. The identity map $\mathrm{id}_G : G \to G$ is a homomorphism and for any homomorphisms $\varphi : G \to H$, $\psi : K \to G$ we have $\varphi\circ\mathrm{id}_G = \varphi$ and $\mathrm{id}_G\circ\psi = \psi$. These statements are essentially trivial.

Monomomorphisms in the category of groups are homomorphism which are one-to-one (injective), epimorphism are homomorphisms which are onto (surjective) on the target group, and isomorphisms are group isomorphisms.

Problem 2.13 **(a) Show that if H and K are subgroups of G then their intersection $H \cap K$ is always a subgroup of G.**

(b) Show that the product $HK = \{hk \mid h \in H,\, k \in K\}$ is a subgroup if and only if $HK = KH$.

Solution: (a) If $a, b \in H \cap K$ then $a, b \in H \Rightarrow ab \in H$ and $a, b \in K \Rightarrow ab \in K$. Hence $ab \in H \cap K$. The identity $e \in H$ and $e \in K$, therefore $e \in H \cap K$. If $a \in H \cap K$ then $a \in H \Rightarrow a^{-1} \in H$ and $a \in K \Rightarrow a^{-1} \in K$, so that $a^{-1} \in H \cap K$. Hence $H \cap K$ is a subgroup.

(b) If HK is a subgroup then for all $h_1, h_2 \in H$, $k_1, k_2 \in K$ then by the closure property there exist $h' \in H$, $k' \in K$ such that $h_1 k_1 h_2 k_2 = h'k'$. Hence for arbitrary $k_1 \in K$, $h_2 \in H$

$$k_1 h_2 = (h_1)^{-1} h' k' (k_2)^{-1} \in HK.$$

so that $HK = KH$.

Conversely, if $HK = KH$ then for arbitrary elements $h_1 k_1, h_2 k_2 \in HK$ the product

$$h_1 k_1 h_2 k_2 = h_1 h_2 k_1 k_2 \in HK.$$

Thus HK is closed with respect taking products of elements. The identity $e = ee$ always belongs to the product set HK of two subgroups and if $HK = KH$ the set closed with respect to taking inverses, for

$$\begin{aligned} hk \in HK \quad &\Longrightarrow \quad (hk)^{-1} k^{-1} h^{-1} \in KH = HK \\ &\Longrightarrow \quad (hk)^{-1} = h'k' \in HK \text{ for some } h' \in H, k' \in K. \end{aligned}$$

Problem 2.14 **Find all the normal subgroups of the group of symmetries of the square D_4 described in Example 2.7.**

Solution: A normal subgroup N must contain the entire conjugate class $C_g = \{hgh^{-1} \mid h \in G\}$ of every elemenent $g \in N$. For the group D_4 the conjugacy classes are easily read off from the multiplication table (Problem 2.2),

$$\{R_0\}, \quad \{R_1, R_3\}, \quad \{R_2\}, \quad \{S_1, S_2\}, \quad \{S_3, S_4\}.$$

Apart from the trivial normal subgroups $\{R_0\}$ and the whole group D_4 itself, the only subgroups which contain the entire conjugacy class $\{R_1, R_3\}$ is

$$\{R_0, R_1, R_2, R_3\}$$

for if we try to add any of the elements S_i and the subgroup would include every element of D_4. By similar arguments, the only other normal subgroups are

$$\{R_0, R_2\}, \qquad , \{R_0, R_2, S_1, S_2, \}, \qquad \text{and } \{R_0, R_2, S_3, S_4\}.$$

Problem 2.15 **The *quaternion group* G consists of eight elements denoted**

$$\{1,\ -1,\ i,\ -i,\ j,\ -j,\ k,\ -k\},$$

subject to the following law of composition

$$\begin{aligned} 1g = g1 = g, &\qquad \textbf{for all } g \in Q, \\ -1g = -g, &\qquad \textbf{for } g = i,\ j,\ k, \\ i^2 = j^2 &= k^2 = -1, \\ ij = k,\ & jk = i,\ ki = j. \end{aligned}$$

(a) Write down the full multiplication table for Q, justifying all products not included in the above list.

(b) Find all subgroups of Q and show that all subgroups of Q are normal.

(c) Show that the subgroup consisting of $\{1,\ -1,\ i,\ -i\}$ is the kernel of a homomorphism $Q \to \{1,\ -1\}$.

(d) Find a subgroup H of S_4, the symmetric group of order 4, such that there is a homomorphism $Q \to H$ whose kernel is the subgroup $\{1,\ -1\}$.

Solution: (a) By the associative law, $i(-1) = ii^2 = i^2 i = -1i = -i$. Similarly

$$j(-1) - j, \qquad k(-1) - k.$$

Now

$$ji = jjk = j^2 k = -1k = -k$$

and similarly $ik = -j$, $kj = -i$. Other products can be evaluated in the following way:

$$\begin{aligned} &(-1)(-i) = (-1)kj = (-k)j = k(-1)j = kiij = jk = i, \\ &(-1)(-j) = j, \quad (-1)(-k) = k. \end{aligned}$$

Since $-i = (-1)i = (-1)(-1)(-i)$ we must have $(-1)^2 = (-1)(-1) = 1$. Hence

$$(-i)(-1) = i(-1)(-1) = i.1 = i, \qquad (-j)(-1) = j, \quad (-k)(-1) = k,$$

and

$$\begin{aligned} (-i)i = (-1)ii = (-1)(-1) = 1, \quad i(-i) = ii(-1) = (-1)(-1) = 1, \\ (-j)j = j(-j) = (-k)k = k(-k) = 1. \end{aligned}$$

This essentially completes the multiplication table:

1	−1	i	$-i$	j	$-j$	k	$-k$
−1	1	$-i$	i	$-j$	j	$-k$	k
i	$-i$	−1	1	k	$-k$	$-j$	j
$-i$	i	1	−1	$-k$	k	j	$-j$
j	$-j$	$-k$	k	−1	1	i	$-i$
$-j$	j	k	$-k$	1	−1	$-i$	i
k	$-k$	j	$-j$	$-i$	i	−1	1
$-k$	k	$-j$	j	i	$-i$	1	−1

(b) Every subgroup H contains 1. For example, $\{1, -1\}$ is a subgroup. If $i \in H$ then $-1 = ii \in H$ and also $-i = (-1)i \in H$. If H contains j as well then it also contains $-j = (-1)j$, $k = ij$ and $-k = (-1)k$—that is, H is the entire quaternion group Q. Thus the only proper subgroups are

$$\{1\}, \qquad \{1, -1\}, \qquad \{1, -1, i, -i\}, \qquad \{1, -1, j, -j\}, \qquad \{1, -1, k, -k\}.$$

The subgroup $H = \{1, -1\}$ is normal since $iHi^{-1} = \{ii^{-1}, (-1)ii^{-1}\} = \{1, -1\} = H$ etc.
$H = \{1, -1, i, -i\}$ is normal since

$$iHi^{-1} = iH(-i) = \{1, -1, i, -i\} = H,$$

$$jHj^{-1} = jH(-j) = \{1, -1, -jij, jij\} = \{1, -1, -i, i\}$$

as $jij = jk = i$, and similarly $kHk^{-1} = H$. In the same way, the other 2 subgroups may be shown to be normal, so all subgroups of Q are normal.

(c) Let $\phi : Q \to \{1,\ -1\}$ be the map defined by

$$\phi(1) = \phi(-1) = \phi(i) = \phi(-i) = 1, \qquad \phi(j) = \phi(-j) = \phi(k) = \phi(-k) = -1.$$

This is a homomorphism since $\phi(jk) = \phi(i) = 1 = \phi(j)\phi(k) = (-1)(-1)$ etc. The kernel of ϕ is clearly $\{1, -1, i, -i\}$.

(d) Set ψ to be the map defined by $\psi(\pm 1) = (1)(2)(3)(4) = \mathrm{id}$ and

$$\psi(\pm i) = (12)(34), \qquad \psi(\pm j) = (13)(24), \qquad \psi(\pm k) = (14)(23).$$

This is a homomorphism since $\psi(\pm i)\psi(\pm j) = \psi(\pm ij) = \psi(\pm k) = (14)(23)$ etc. The kernel of the homomorphism is clearly $H_1 = \{1, -1\}$.

Problem 2.16 **A *Möbius transformation* is a complex map,**

$$z \mapsto z' = \frac{az + b}{cz + d} \qquad \textbf{where} \qquad a, b, c, d \in \mathbb{C}, \quad ad - bc = 1.$$

(a) Show that these are one-to-one and onto transformations of the extended complex plane which includes the point $z = \infty$, and write out the

composition of an arbitrary pair of transformations given by constants (a, b, c, d) and (a', b', c', d').

(b) Show that they form a group, called the *Möbius group.*

(c) Show that the map μ from $SL(2,\mathbb{C})$ to the Möbius group, which takes the unimodular matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ to the above Möbius transformation, is a homomorphism, and that the kernel of this homomorphism is $\{\mathsf{I}, -\mathsf{I}\}$; i.e. the Möbius group is isomorphic to $SL(2,\mathbb{C})/\mathbb{Z}_2$.

Solution: (a) Since

$$z = \frac{dz' - b}{-cz' + a} \tag{1}$$

the map is invertible and therefore one-to-one. The point $z = -d/c$ is sent to $z' = \infty$ and $z' = a/c$ corresponds to $z = \infty$. The composition of

$$z' = \alpha(z)\frac{az+b}{cz+d} \qquad \text{and} \qquad z'' = \alpha'(z') = \frac{a'z'+b'}{c'z'+d'}$$

is

$$z'' = \alpha' \circ \alpha(z) = \frac{(a'a + b'c)z + (a'b + b'd)}{(c'a + d'c)z + (c'b + d'd)}. \tag{2}$$

(b) The Möbius transformations form a group since they are closed under functional composition by (2), the identity transformation corresponds to $b = c = 0$, $a = d = 1$, and the inverse of a Möbius transformation is given by (1).

(c) Let $\mu : SL(2,\mathbb{C}) \to \mathcal{M}$, where $\mathcal{M}$ is the set of Möbius transformations, be the map

$$\mu : \begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto \alpha(z) = \frac{az+b}{cz+d} \qquad \text{where} \qquad ad - bc = 1$$

This map is a homomorphism since, by Eq. (2)

$$\mu\left(\begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix}\right) = \mu\left(\begin{pmatrix} a'a + b'c & a'b + b'd \\ c'a + d'c & c'b + d'd \end{pmatrix}\right) = \alpha' \circ \alpha.$$

Two unimodular matrices are mapped to the same Möbius transformation α if and only if they are proportional. Since $\det \mathsf{A} = 1$ the factor of proportionality must be ± 1, so the kernel of the homomorphism is $\{\mathsf{I}, -\mathsf{I}\}$. Thus the Möbius group is isomorphic to the factor group $SL(2,\mathbb{C})/\{\mathsf{I}, -\mathsf{I}\} \cong SL(2,\mathbb{C})/\mathbb{Z}_2$.

Problem 2.17 **Assuming the identification of G with (G, e_H) and H with (e_G, H), show that $G \cong (G \times H)/H$ and $H \cong (G \times H)/G$.**

Solution: The elements $h \in H$ are identified with elements of $G \times H$ by setting $h \equiv (e_G, h)$. Elements of $(G \times H)/H$ are then cosets $(g, h)H = H(g, h)$. Two cosets are identical, $(g, h)H = (g', h')H$ if and only if $g = g'$ and $h' = hh_1$ for some $h_1 \in H$.

Every element of the group $(G \times H)/H$ can thus be written uniquely in the form $(g, e_H)H$.
Let $\phi : G \to (G \times H)/H$ be defined by $\phi(g) = (g, e_H)H$. This map is one-to-one and onto, and is a homomorphism since

$$\phi(gg') = (gg', e_H)H = (g, e_H)H(g', e_H)H = \phi(g)\phi(g').$$

ϕ is therefore an isomorphism.
The isomorphism between H and $(G \times H)/G$ is completely analogous.

Problem 2.18 Show that the conjugacy classes of the direct product $G \times H$ of two groups G and H consist precisely of products of conjugacy classes from the groups

$$(C_i, D_j) = \{(g_i, h_j) \,|\, g_i \in C_i,\, h_j \in D_j\}$$

where C_i is a conjugacy class of G and D_j a conjugacy class of H.

Solution: Every conjugacy class of G may be written

$$C_i = \{g g_i g^{-1} \,|\, g_i \in G\}$$

and the conjugacy classes of H are

$$D_j = \{h h_j h^{-1} \,|\, h_j \in H\}.$$

Similarly, the conjugacy classes of $G \times H$ are of the form

$$\begin{aligned} C_{ij} &= \{(g,h)(g_i,h_j)(g,h)^{-1} \,|\, g_i \in G,\, h_j \in H, g \in G, h \in H\} \\ &= \{(g,h)(g_i,h_j)(g^{-1},h^{-1}) \,|\, g_i \in G,\, h_j \in H, g \in G, h \in H\} \\ &= \{(g g_i g^{-1}, h h_j h^{-1}) \,|\, g_i \in G,\, h_j \in H, g \in G, h \in H\} \\ &= C_i \times D_j. \end{aligned}$$

Problem 2.19 If H is any subgroup of a group G define the action of G on the set of left cosets G/H by $g : g'H \mapsto gg'H$.

(a) Show that this is always a transitive action of H on G.

(b) Let G have a transitive left action on a set X, and set $H = G_x$ to be the isotropy group of any point x. Show that the map $i : G/H \to X$ defined by $i(gH) = gx$ is well-defined, one-to-one and onto.

(c) Show that the left action of G on X can be identified with the action of G on G/H defined in (a).

(d) Show that the group of proper orthogonal transformations $SO(3)$ acts transitively on the 2-sphere S^2,

$$S^2 = \{(x, y, z) \,|\, r^2 = x^2 + y^2 + z^2 = 1\} = \{\mathbf{r} \,|\, r^2 = \mathbf{r}^T\mathbf{r} = 1\}$$

where r is a column vector having real components x, y, z. Show that the isotropy group of any point r is isomorphic to $SO(2)$, and find a bijective correspondence between the factor space $SO(3)/SO(2)$ and the 2-sphere S^2 such that $SO(3)$ has identical left action on these two spaces.

Solution: (a) For any g' the orbit of $g'H$ is $Gg'H = \{gg'H \,|\, g \in G\} = G/H$. Hence the action of G on G/H is transitive.

(b) The action i is well-defined for if $gH = g'H$ then $g' = gh$ for some $h \in H = G_x$, so that $g'x = ghx = gx$. Thus $i(gH) = i(g'H)$. The map i is one-to-one, for if $gx = g'x$ then $g^{-1}g' \in G_x = H$, or equivalently $g' \in gH$, i.e. $g'H = gH$. Finally, it is onto since G's action on X is transitive; thus, for all $y \in X$ there exists $g \in G$ such that $gx = y$, so that $i(gH) = y$.

(c)The action of G on X can be identified with the left action of G on G/H since $g(g'H) = (gg')H$. Hence, if $y = g'x$ then $y = i(g'H)$ and

$$gi(g'H) = gy = gg'x = i(gg'H).$$

Another way of expressing this equivalence of actions is $g \circ i = i \circ g$, expressed by the commutative diagram.

(d) The action of $SO(3)$ is transitive on S^2 since every unit vector $\mathbf{x}$ can be rotated to any other unit vector $\mathbf{y}$: for example, if $\mathbf{e}$ and $\mathbf{f}$ are any two unit vectors orthogonal to $\mathbf{x}$ and orthogonal to each other then the matrix A_1 whose columns are $(\mathbf{efx})$ is an orthogonal matrix and has the property

$$\mathsf{A}_1\mathbf{k} = \begin{pmatrix} \mathbf{e} & \mathbf{f} & x \\ \downarrow & \downarrow & y \\ & & z \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$

If we choose A_2 to be an orthogonal matrix such that $\mathsf{A}_2\mathbf{k} = \mathbf{y}$ then the orthogonal matrix $\mathsf{A} = \mathsf{A}_2(\mathsf{A}_1)^{-1}$ has the required property $\mathsf{A}\mathbf{x} = \mathbf{y}$.
The isotropy group of any point $\mathbf{n} \in S^2$ is the set of orthogonal transformations leaving the direction $\mathbf{n}$ invariant. It is the set of rotations having $\mathbf{n}$ as axis. For example, if $\mathbf{n} = \mathbf{k} = (0, 0, 1)^T$ then the isotropy group of $\mathbf{r}$ consists of orthogonal transformations of the form

$$\mathsf{A} = \begin{pmatrix} a_{11} & a_{12} & 0 \\ a_{21} & a_{22} & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} \cos\theta & \pm\sin\theta & 0 \\ \mp\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

for some θ. These are rotations about the axis $\mathbf{k}$ by an angle θ, forming the group $SO(2)$.

Every rotation can be represented by an *axis of rotation* $\mathbf{n}$ and an angle of rotation θ about this axis. The possible axes or rotation are in one-to-one correspondence with points on the unit sphere, while the angular freedom charactetrizes an element of $SO(2)$. Thus cosets $SO(3)/SO(2)$ are essentially the possible axes of rotation—i.e., points of S^2.

A more algebraic way of establishing this correspondence is the use of Euler angles (see, for example Goldstein *Classical Mechanics*). With a slight variant on Goldstein's notation, every rotation can be shown to be a product of three successive rotations about the axes z, x and z again, $\mathsf{A} = \mathsf{DCB}$ where

$$\mathsf{B} = \begin{pmatrix} \cos\psi & \sin\psi & 0 \\ -\sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{pmatrix}, \ \mathsf{C} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & \sin\theta \\ 0 & -\sin\theta & \cos\theta \end{pmatrix}, \ \mathsf{D} = \begin{pmatrix} \cos\phi & \sin\phi & 0 \\ -\sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

The rotation B belongs to $SO(2)$, while the product DB is determined by two angles θ and ϕ which can be set in one-to-one correspondence with points of the sphere expressed in polar coordinates ($0 \le \theta \le \pi$, $0 \le \phi < 2\pi$, with all points having $\theta = 0$ identified as are points having $\theta = \pi$).

Problem 2.20 **The *projective transformations* of the line are defined by**

$$x' = \frac{ax+b}{cx+d}, \qquad \textbf{where } ad - bc = 1.$$

Show that projective transformations preserve the cross-ratio

$$\frac{(x_1 - x_2)(x_3 - x_4)}{(x_3 - x_2)(x_1 - x_4)}$$

between any four points x_1, x_2, x_3 and x_4. Is every analytic transformation that preserves the cross-ratio between any four points on the line necessarily a projective transformation? Do the projective transformations form a group?

Solution: For any two points x_1 and x_2

$$x_1' - x_2' = \frac{(ax_1+b)(cx_2+d) - (ax_2+b)(cx_1+d)}{(cx_1+d)(cx_2+d)} = \frac{x_1 - ax_2}{(cx_1+d)(cx_2+d)},$$

whence, by simple cancellations

$$\frac{(x_1'-x_2')(x_3'-x_4')}{(x_3'-x_2')(x_1'-x_4')} = \frac{(x_1-x_2)(x_3-x_4)}{(x_3-x_2)(x_1-x_4)}.$$

If a transformation $x \to x'$ preserves all cross-ratios, then fixing x_2, x_3 and x_4, gives

$$\frac{x_1' - \text{const.}}{x_1' - \text{const.}} = \text{const.}\frac{x_1 - \text{const.}}{x_1 - \text{const.}}$$

from which there exist constants a', b', c', d' such that

$$x_1' = \frac{a'x_1 + b'}{c'x_1 + d'}$$

We may assume $a'd' - b'c' \neq 0$ else $c'/d' = a'/b'$ and the transformation degenerates to $x' = \text{const}$. By rescaling to $a' = \lambda a$, $b' = \lambda b$, $c' = \lambda c$, $d' = \lambda d$ we still have

$$x_1' = \frac{ax_1 + b}{cx_1 + d}$$

and we may choose λ such that $ad - bc = 1$.
The projective transformationsform a group in the same way as the Möbius transformations (Problem 2.16).

Problem 2.21 Show that a matrix U is unitary, satisfying Eq. (2.12), if and only if it preserves the 'norm'

$$\|\mathbf{z}\|^2 = \sum_{i=1}^{n} z_i \bar{z}_i$$

defined on column vectors $(z_1, z_2, \ldots, z_n)^T$ in $\mathbb{C}^n$. Verify that the set of $n \times n$ complex unitary matrices $U(n)$ forms a group.

Solution: We may write

$$\|\mathbf{z}\|^2 = \mathbf{z}^T \mathbf{z} = \mathbf{z}^\dagger \mathbf{z}$$

where $\mathbf{z}$ is the column vector having components

$$\mathbf{z} = \begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{pmatrix}$$

Setting $\mathbf{z}' = \mathsf{U}\mathbf{z}$,

$$\|\mathbf{z}'\|^2 = \mathbf{z}'^\dagger \mathbf{z}' = \mathbf{z}^\dagger \mathsf{U}^\dagger \mathsf{U} \mathbf{z} = \mathbf{z}^\dagger \mathbf{z} = \|\mathbf{z}\|^2$$

for arbitrary $\mathbf{z}$ if and only if $\mathsf{U}^\dagger \mathsf{U} = \mathsf{I}$, i.e. $\mathsf{U}^\dagger = \mathsf{U}^{-1}$. Thus this norm preserving condition holds iff U is a unitary matrix, $\mathsf{U}\mathsf{U}^\dagger = \mathsf{I}$.

The set of unitary matrices is a group as it satisfies:

Closure If $\mathsf{U}_1 \mathsf{U}_1^\dagger = \mathsf{I}$, $\mathsf{U}_2 \mathsf{U}_2^\dagger = \mathsf{I}$, then $\mathsf{U}_1 \mathsf{U}_2$ is unitary, for

$$\mathsf{U}_1 \mathsf{U}_2 (\mathsf{U}_1 \mathsf{U}_2)^\dagger = \mathsf{U}_1 \mathsf{U}_2 \mathsf{U}_2^\dagger \mathsf{U}_1^\dagger = \mathsf{U}_1 \mathsf{I} \mathsf{U}_1^\dagger = \mathsf{I}.$$

Identity The identity matrix is unitary, $\mathsf{I}\mathsf{I}^\dagger = \mathsf{I}^2 = \mathsf{I}$.

Inverse If $\mathsf{UU}^\dagger = \mathsf{I}$ then μ^{-1} is unitary, for

$$\mathsf{U}^{-1}(\mathsf{U}^\dagger)^{-1} = \mathsf{U}^{-1}\mathsf{U}\mathsf{U}^\dagger(\mathsf{U}^\dagger)^{-1} = \mathsf{I}\mathsf{I} = \mathsf{I}.$$

Problem 2.22 Show that two rotations belong to the same conjugacy classes of the rotation group $SO(3)$ if and only if they have the same magnitude; that is, they have the same angle of rotation but possibly a different axis of rotation.

Solution: A and A' belong to the same conjugacy class iff there exists a rotation $\mathsf{C} \in SO(3)$ such that $\mathsf{A}' = \mathsf{CAC}^{-1}$. The axis of rotation defined by $\mathsf{A} \in SO(3)$ is the unit vector $\mathbf{n}$ left invariant by the action of A, i.e. $\mathsf{A}\mathbf{n} = \mathbf{n}$ (see Problem 2.19 (d)). Then $\mathbf{n}' = \mathsf{C}\mathbf{n}$ is the axis of rotation of A', since

$$\mathsf{A}'\mathbf{n}' = \mathsf{A}'\mathsf{C}\mathbf{n} = \mathsf{CAC}^{-1}\mathsf{C}\mathbf{n} = \mathsf{CA}\mathbf{n} = \mathsf{C}\mathbf{n} = \mathbf{n}'.$$

If $\mathbf{e}$ is any vector orthogonal to $\mathbf{n}$ then the angle of rotation θ defined by A is determined by the equation

$$\mathbf{e} \cdot (\mathsf{A}\mathbf{e}) \equiv \mathbf{e}^T\mathsf{A}\mathbf{e} = \cos\theta.$$

Now $\mathbf{e}' = \mathsf{C}\mathbf{e}$ is orthogonal to $\mathbf{n}'$, for

$$\mathbf{e}' \cdot \mathbf{n}' = \mathsf{C}\mathbf{e} \cdot C\mathbf{n} = \mathbf{e}^T\mathsf{C}^T\mathsf{C}\mathbf{n} = \mathbf{e}^T\mathsf{I}\mathbf{n} = \mathbf{e} \cdot \mathbf{n} = 0$$

and the angle of rotation is unchanged under the action of conjugation, for

$$\begin{aligned}\mathbf{e}' \cdot \mathsf{A}'\mathbf{e}' &= \mathbf{e}^T\mathsf{C}^T\mathsf{A}'\mathsf{C}\mathbf{e} \\ &= \mathbf{e}^T\mathsf{C}^{-1}\mathsf{A}'\mathsf{C}\mathbf{e} \\ &= \mathbf{e}^T\mathsf{A}\mathbf{e} = \cos\theta.\end{aligned}$$

Hence all rotations belonging to the same conjugacy class have the the same angle of rotation.

Conversely, let A and A' be rotations having the same angle of rotation; i.e. if $\mathbf{n}$ and $\mathbf{n}'$ are respective axes of rotation, $\mathsf{A}\mathbf{n} = \mathbf{n}$ and $\mathsf{A}'\mathbf{n}' = \mathbf{n}'$, then for all vectors $\mathbf{e} \perp \mathbf{n}$ and $\mathbf{e}' \perp \mathbf{n}'$ we have

$$\mathbf{e}^T\mathsf{A}\mathbf{e} = \mathbf{e}'^T\mathsf{A}'\mathbf{e}' = \cos\theta.$$

There always exists a rotation $\mathbf{C}$ such that $\mathbf{n}' = \mathsf{C}\mathbf{n}$ (see Problem 2.19 (d)). Then

$$\mathsf{A}'\mathbf{n}' = \mathbf{n}' \implies \mathsf{A}'\mathsf{C}\mathbf{n} = \mathsf{C}\mathbf{n} \implies \mathsf{C}^{-1}\mathsf{A}'\mathsf{C}\mathbf{n} = \mathbf{n} = \mathsf{A}\mathbf{n}.$$

Now, as shown above, if $\mathbf{e} \perp \mathbf{n}$ and then $\mathbf{e}' = \mathsf{C}\mathbf{e} \perp \mathbf{n}'$ and

$$\mathsf{A}\mathbf{e} = \cos\theta \Longrightarrow \mathsf{A}'\mathbf{e}' = \cos\theta\,\mathbf{e}'.$$

Hence

$$\mathsf{A}'\mathsf{C}\,\mathbf{e} = \cos\theta\,\mathsf{C}\,\mathbf{e}$$

from which we have

$$\mathsf{C}^{-1}\mathsf{A}'\mathsf{C}\,\mathbf{e} = \cos\theta\,\mathbf{e} = \mathsf{A}\mathbf{e}.$$

As any vector $\mathbf{r}$ can be decomposed into components parallel and orthogonal to $\mathbf{n}$,

$$\mathbf{r} = (\mathbf{r}\cdot\mathbf{n})\mathbf{n} + \mathbf{r}^{\perp}$$

it follows that

$$\mathsf{C}^{-1}\mathsf{A}'\mathsf{C}\,\mathbf{r} = \mathsf{A}\mathbf{r},$$

so that

$$\mathsf{C}^{-1}\mathsf{A}'\mathsf{C} = \mathsf{A},$$

i.e. A and A' belong to the same conjugacy class, as $\mathsf{A}' = \mathsf{C}\mathsf{A}\mathsf{C}^{-1}$.

Problem 2.23 The general Galilean transformation

$$t' = t + a, \qquad \mathbf{r}' = \mathsf{A}\mathbf{r} - \mathbf{v}t + \mathbf{b} \qquad \textbf{where} \qquad \mathsf{A}^T\mathsf{A} = \mathsf{I}$$

may be denoted by the abstract symbol $(a, \mathbf{v}, \mathbf{b}, \mathsf{A})$. Show that the result of performing two Galilean transformations

$$G_1 = (a_1,\, \mathbf{v}_1,\, \mathbf{b}_1,\, \mathsf{A}_1) \qquad \textbf{and} \qquad G_2 = (a_2,\, \mathbf{v}_2,\, \mathbf{b}_2,\, \mathsf{A}_2)$$

in succession is

$$G = G_2G_1 = (a,\, \mathbf{v},\, \mathbf{b},\, \mathsf{A})$$

where

$$a = a_1 + a_2, \quad \mathbf{v} = \mathsf{A}_2\mathbf{v}_1 + \mathbf{v}_2, \quad \mathbf{b} = \mathbf{b}_2 - a_1\mathbf{v}_2 + \mathsf{A}_2\mathbf{b}_1 \quad \textbf{and} \;\; \mathsf{A} = \mathsf{A}_2\mathsf{A}_1.$$

Show from this rule of composition that the Galilean transformations form a group. In particular verify explicitly that the associative law holds.

Solution: The transformation G_1 is

$$t' = t + a, \qquad \mathbf{r}' = \mathsf{A}_1\mathbf{r} - \mathbf{v}_1 t + \mathbf{b}_1$$

whence G_2G_1 is

$$\begin{aligned} t'' &= t' + a_2 = t + (a_1 + a_2) \\ \mathbf{r}'' &= \mathsf{A}_2\mathbf{r}' - \mathbf{v}_2 t' + \mathbf{b}_2 \\ &= \mathsf{A}_2(\mathsf{A}_1\mathbf{r} - \mathbf{v}_1 t + \mathbf{b}_1) - \mathbf{v}_2(t + a_1) + \mathbf{b}_2 \\ &= \mathsf{A}_2\mathsf{A}_1\mathbf{r} - (\mathsf{A}_2\mathbf{v}_1 + \mathbf{v}_2) + \mathsf{A}_2\mathbf{b}_1 - \mathbf{v}_2 a_1 + \mathbf{b}_2 \end{aligned}$$

Thus

$$G = G_2 G_1 = (a,\, \mathbf{v},\, \mathbf{b},\, \mathsf{A})$$

where

$$a = a_1 + a_2, \quad \mathbf{v} = \mathsf{A}_2\mathbf{v}_1 + \mathbf{v}_2, \quad \mathbf{b} = \mathbf{b}_2 - a_1\mathbf{v}_2 + \mathsf{A}_2\mathbf{b}_1 \quad \text{and} \quad \mathsf{A} = \mathsf{A}_2\mathsf{A}_1.$$

This is a Galilean transformation since $\mathsf{A}_2\mathsf{A}_1$ is orthogonal (see Example 2.11). Hence Galilean transformations are closed with respect to this product law.

The *associative law* follows from

$$G_3(G_2G_1) = (a_1 + a_2 + a_3,\, \mathsf{A}_3(\mathsf{A}_2\mathbf{v}_1 + \mathbf{v}_2) + \mathbf{v}_3,\\ \mathsf{A}_3(\mathsf{A}_2\mathbf{b}_1 - \mathbf{v}_2 a_1 + \mathbf{b}_2) - \mathbf{v}_3(a_1 + a_2) + \mathbf{b}_3,\, \mathsf{A}_3\mathsf{A}_2\mathsf{A}_1)$$

which agrees with

$$(G_3G_2)G_1 = (a_1 + a_2 + a_3,\, \mathsf{A}_3\mathsf{A}_2\mathbf{v}_1 + \mathsf{A}_3\mathbf{v}_2 + \mathbf{v}_3,\\ \mathsf{A}_3\mathsf{A}_2\mathbf{b}_1 - (\mathsf{A}_3\mathbf{v}_2 + \mathbf{v}_3)a_1 + \mathsf{A}_3\mathbf{b}_2 - \mathbf{v}_3 a_2 + \mathbf{b}_3,\, \mathsf{A}_3\mathsf{A}_2\mathsf{A}_1).$$

The reason for requiring a proof of the associative law is that we are proving the group property from the law of composition of 4-tuples $(a,\, \mathbf{v},\, \mathbf{b},\, \mathsf{A})$, not from the composition of Galilean transformations as maps.

The identity transform $t' = t$, $\mathbf{r}' = \mathbf{r}$ corresponds to the Galilean transformation with $a = 0$, $\mathbf{b} = \mathbf{v} = \mathbf{0}$, $\mathsf{A} = \mathsf{I}$, i.e. $G = (o,\, \mathbf{0},\, \mathbf{0},\, \mathsf{I})$.

If $G = (a,\, \mathbf{v},\, \mathbf{b},\, \mathsf{A})$ is a Galilean transformation we may find its inverse by solving $G'G = (0,\, \mathbf{0},\, \mathbf{0},\, \mathsf{I})$, i.e.

$$a' + a = 0$$

$$\mathsf{A}'\mathbf{v} + \mathbf{v}' = \mathbf{0}$$

$$\mathsf{A}'\mathbf{b} - \mathbf{v}'a + \mathbf{b}' = \mathbf{0}$$

$$\mathsf{A}'\mathsf{A} = \mathsf{I}.$$

Hence

$$a' = -a, \qquad \mathsf{A}' = \mathsf{A}^{-1}, \qquad \mathbf{v}' = -\mathsf{A}^{-1}\mathbf{v}, \qquad \mathbf{b}' = -\mathsf{A}^{-1}(\mathbf{b} + a\mathbf{v})$$

so that

$$(a,\, \mathbf{v},\, \mathbf{b},\, \mathsf{A})^{-1}) = \big(-a,\, -\mathsf{A}^{-1}\mathbf{v},\, -\mathsf{A}^{-1}(\mathbf{b} + a\mathbf{v}), \mathsf{A}^{-1}\big).$$

Problem 2.24 **(a) From the matrix relation defining a Lorentz transformation L,**

$$\mathsf{G} = \mathsf{L}^T\mathsf{G}\mathsf{L},$$

where G is the 4×4 diagonal matrix whose diagonal components are $(1, 1, 1, -1)$; show that Lorentz transformations form a group.

(b) Denote the Poincaré transformation

$$\mathbf{x}' = \mathsf{L}\mathbf{x} + \mathbf{a}$$

by $(\mathsf{L}, \mathbf{a})$, and show that two Poincaré transformations $(\mathsf{L}_1, \mathbf{a})$ and $(\mathsf{L}_2, \mathbf{b})$ performed in succession is equivalent to the Poincaré transformation

$$(\mathsf{L}_2\mathsf{L}_1, \mathbf{b} + \mathsf{L}_2\mathbf{a}).$$

(c) From this law of composition show that the Poincaré transformations form a group. As in the previous problem the associative law should be shown explicitly.

Solution: (a) *Closure* follows from

$$(\mathsf{L}_1\mathsf{L}_2)^T\mathsf{G}\mathsf{L}_1\mathsf{L}_2 = \mathsf{L}_2^T\mathsf{L}_1^T\mathsf{G}\mathsf{L}_1\mathsf{L}_2 = \mathsf{L}_2^T\mathsf{G}\mathsf{L}_2 = \mathsf{G}.$$

The identity transformation $\mathsf{L} = \mathsf{I}$ is trivially a Lorentz tranformation, since $\mathsf{I}^T\mathsf{G}\mathsf{I} = \mathsf{G}$.
The inverse of any Lorentz transformation is a Lorentz transformation, for on multiplying the defining equation from the left by $(\mathsf{L}^T)^{-1}$ and the right by L^{-1} we have

$$\mathsf{G} = \mathsf{L}^T\mathsf{G}\mathsf{L} \quad \Longrightarrow \quad (\mathsf{L}^T)^{-1}\mathsf{G}\mathsf{L}^{-1} = \mathsf{I}\mathsf{G}\mathsf{I} = \mathsf{G}.$$

(b) If $\mathbf{x}' = \mathsf{L}\mathbf{x} + \mathbf{a}$ and $\mathbf{x}'' = \mathsf{L}_2\mathbf{x}' + \mathbf{b}$ then

$$\mathbf{x}'' = \mathsf{L}_2(\mathsf{L}_1\mathbf{x} + \mathbf{a}) + \mathbf{b} = \mathsf{L}_2\mathsf{L}_1\mathbf{x} + \mathsf{L}_2\mathbf{a} + \mathbf{b}$$

so that the law of composition of Poincaré transformations is

$$(\mathsf{L}_2, \mathbf{b})(\mathsf{L}_1, \mathbf{a}) = (\mathsf{L}_2\mathsf{L}_1, \mathbf{b} + \mathsf{L}_2\mathbf{a}).$$

(c) This product law shows closure. The associative law follows from

$$\begin{aligned}(\mathsf{L}_3, \mathbf{c})\big((\mathsf{L}_2, \mathbf{b})(\mathsf{L}_1, \mathbf{a})\big) &= (\mathsf{L}_3, \mathbf{c})(\mathsf{L}_2\mathsf{L}_1, \mathbf{b} + \mathsf{L}_2\mathbf{a}) \\ &= (\mathsf{L}_3\mathsf{L2L1}, \mathbf{c} + \mathsf{L}_3\mathbf{b} + \mathsf{L}_3\mathsf{L}_2\mathbf{a}) \\ &= (\mathsf{L}_3\mathsf{L}_2, \mathbf{c} + \mathsf{L}_3\mathbf{b})(\mathsf{L}_1, \mathbf{a}) \\ &= \big((\mathsf{L}_3, \mathbf{c})\big((\mathsf{L}_2, \mathbf{b})\big)(\mathsf{L}_1, \mathbf{a})\end{aligned}$$

The identity Poincaré transformation is $(\mathsf{I}, \mathbf{0})$, and the inverse of $(\mathsf{L}, \mathbf{a})$ is

$$(\mathsf{L}, \mathbf{a})^{-1} = (\mathsf{L}^{-1}, -\mathsf{L}^{-1}\mathbf{a}).$$

Problem 2.25 **Let V be an abelian group with law of composition $+$, and G any group with a left action on V, denoted as usual by $g : v \mapsto gv$. Assume further that this action is a homomorphism of V,**

$$g(v + w) = gv + gw.$$

(a) Show that $G \times V$ is a group with respect to the law of composition

$$(g, v)(g', v') = (gg', v + gv').$$

This group is known as the semi-direct product of G and V, and is denoted $G \circledS V$.

(b) Show that the elements of type $(g, 0)$ form a subgroup of $G \circledS V$ that is isomorphic with G and that V is isomorphic with the subgroup (e, V). Show that the latter is a normal subgroup.

(c) Show that every element of $G \circledS V$ has a unique decomposition of the form vg, where $g \equiv (g, 0) \in G$ and $v \equiv (e, v) \in V$.

Solution: (a) To show the group properties:
Closure: This is obvious from the law of composition.
Associative law:

$$\begin{aligned}(g_1, v_1)\big((g_2, v_2)(g_3, v_3)\big) &= (g_1, v_1)(g_2 g_3, v_2 + g_2 v_3) \\ &= \big(g_1(g_2 g_3), v_1 + g_1(v_2 + g_2 v_3)\big) \\ &= \big((g_1 g_2) g_3, (v_1 + g_1 v_2) + g_2 v_3\big) \\ &= \big((g_1, v_1)(g_2, v_2)\big)(g_3, v_3).\end{aligned}$$

Identity: The identity element is $(e, 0)$, where e is the identity of the group G, since

$$(e, 0)(g, v) = (eg, 0 + ev) = (g, v) \qquad \text{for all } g \in G, v \in V.$$

[Note: $ev = v$ for all $v \in V$ since a left action ϕ of a group G on a space X is a homomomorphism into the group of transformations of the space X, so that $\phi(e) = \mathrm{id}_X$.]
Inverse: The inverse of an element (g, v) is an element (g', v') such that $(g, v)(g', v') = (e, 0)$. That is, $gg' = e$ and $v + gv' = 0 \Rightarrow \; v' = -g^{-1}v$, so that

$$(g, v)^{-1} = (g^{-1}, -g^{-1}v).$$

(b) The map $G \to G \circledS V$ defined by $g \mapsto (g,\, 0)$ is obviously one-to-one and is a homomorphism, since $gh \mapsto (gh,\, 0) = (g,\, 0)(h,\, 0)$.
The map $V \to G \circledS V$ defined by $v \mapsto (e,\, v)$ is also an isomorphism since $v + w \mapsto (e,\, v + w) = (e,\, v)(e,\, w)$ and it is clearly one-to-one. The subgroup $V \equiv \{(e, v)\}$ is a normal subgroup since for all $g \in G$, $w \in V$

$$\begin{aligned}(g, w)V(g^{-1}, -g^{-1}w) &= \{(g, w)(e, v) - g^{-1}w)\} \\ &= \{(g, w + gv)(g^{-1}, -g^{-1}w)\} \\ &= \{(e, w + gv - gg^{-1}w)\} = \{(e, gv)\} \subseteq V.\end{aligned}$$

(c) For all $g \in G$, $v \in V$ we have

$$vg \equiv (e, v)(g, 0) = (eg, v + e0) = (g, v).$$

The decomposition is unique for if $vg = v'g'$ then $(g, v) = (g', v')$ so that $g = g'$, $v = v'$.

Problem 2.26 The following provide examples of the concept of semi-direct product defined in Problem 2.25:

(a) Show that the Euclidean group is the semi-direct product of the rotation group $SO(3, \mathbb{R})$ and $\mathbb{R}^3$, the space of column 3-vectors.

(b) Show that the Poincaré group is the semi-direct product of the Lorentz group $O(3, 1)$ and the abelian group of four-dimensional vectors $\mathbb{R}^4$ under vector addition (see Problem 2.24).

(c) Display the Galilean group as the semi-direct product of two groups.

Solution: (a) The Euclidean group consists of transformations $(\mathsf{A}, \mathbf{a}) : \mathbb{R}^3\mathbb{R}^3$ of the form

$$\mathbf{r}' = \mathsf{A}\mathbf{r} + \mathbf{a} \qquad \mathsf{A}^T\mathsf{A} = \mathsf{I}, \det \mathsf{A} = 1.$$

Since

$$(\mathsf{A}, \mathbf{a})(\mathsf{B}, \mathbf{b})\mathbf{r} = (\mathsf{A}, \mathbf{a})(\mathsf{B}\mathbf{r} + \mathbf{b}) = \mathsf{A}(\mathsf{B}\mathbf{r} + \mathbf{b}) + \mathbf{a} = (\mathsf{AB}, \mathbf{a} + \mathsf{A}\mathbf{b})\mathbf{r}$$

the law of composition is $(\mathsf{A}, \mathbf{a})(\mathsf{B}, \mathbf{b}) = (\mathsf{AB}, \mathbf{a} + \mathsf{A}\mathbf{b})$. Hence, setting $G = SO(3)$ and $V = \mathbb{R}^3$ in Problem 2.24, the Euclidean group is the semidirect product $SO(3)Ⓢ\mathbb{R}^3$.

(b) From Problem 2.24 the composition of two Poincaré transformations is

$$(\mathsf{L}_2, \mathbf{b})(\mathsf{L}_1, \mathbf{a}) = (\mathsf{L}_2\mathsf{L}_1, \mathbf{b} + \mathsf{L}_2\mathbf{a}) \qquad \mathsf{L}_1, \mathsf{L}_2 \in O(3, 1), \quad \mathbf{a}, \mathbf{b} \in \mathbb{R}^4$$

Thus the Poincaré group is the semi-direct product $O(3, 1)Ⓢ\mathbb{R}^4$.

(c) Galilean transformations $t' = t + a$, $\mathbf{r}' = \mathsf{A}\mathbf{r} - \mathbf{v}t + \mathbf{b}$ where $\mathsf{A}^T\mathsf{A} = \mathsf{I}$ may be written as transformations on $\mathbb{R}^4$:

$$\begin{pmatrix} \mathbf{r}' \\ t' \end{pmatrix} = \begin{pmatrix} \mathsf{A} & -\mathbf{v} \\ \mathbf{0}^T & 1 \end{pmatrix} \begin{pmatrix} \mathbf{r} \\ t \end{pmatrix} + \begin{pmatrix} \mathbf{b} \\ a \end{pmatrix}$$

The composition of 4×4 matrices of the type

$$\begin{pmatrix} \mathsf{A} & -\mathbf{v} \\ \mathbf{0}^T & 1 \end{pmatrix}$$

form a group G since they are closed with respect to matrix products,

$$\begin{pmatrix} \mathsf{A} & -\mathbf{v} \\ \mathbf{0}^T & 1 \end{pmatrix} \begin{pmatrix} \mathsf{B} & -\mathbf{w} \\ \mathbf{0}^T & 1 \end{pmatrix} = \begin{pmatrix} \mathsf{AB} & -\mathsf{A}\mathbf{w} - \mathbf{v} \\ \mathbf{0}^T & 1 \end{pmatrix}.$$

The 4×4 identity matrix obiously belongs to this class of matrices, and the inverse of $(\mathsf{A}, \mathbf{v})$ is found by setting $\mathsf{B} = \mathsf{A}^T$, $\mathbf{w} = -\mathsf{A}^T\mathbf{v}$. If we set $V = \mathbb{R}^4$, then the Galilean group can be written $G \circledS \mathbb{R}^4$ for if we write a Galilean transformation as

$$\begin{pmatrix} \mathbf{r}' \\ t' \end{pmatrix} = \left(\begin{pmatrix} \mathsf{A} & -\mathbf{v} \\ \mathbf{0}^T & 1 \end{pmatrix}, \begin{pmatrix} \mathbf{b} \\ a \end{pmatrix} \right) \begin{pmatrix} \mathbf{r} \\ t \end{pmatrix},$$

the composition law given in Problem 2.23 reduces to the semidirect product law

$$\left(\begin{pmatrix} \mathsf{A}_2 & -\mathbf{v}_2 \\ \mathbf{0}^T & 1 \end{pmatrix}, \begin{pmatrix} \mathbf{b}_2 \\ a_2 \end{pmatrix} \right) \left(\begin{pmatrix} \mathsf{A}_1 & -\mathbf{v}_1 \\ \mathbf{0}^T & 1 \end{pmatrix}, \begin{pmatrix} \mathbf{b}_1 \\ a_1 \end{pmatrix} \right)$$
$$= \left(\begin{pmatrix} \mathsf{A}_2\mathsf{A}_1 & -\mathsf{A}_2\mathbf{v}_1 - \mathbf{v}_2 \\ \mathbf{0}^T & 1 \end{pmatrix}, \begin{pmatrix} \mathbf{b}_2 - a_1\mathbf{v}_2 + \mathsf{A}_2\mathbf{b}_1 \\ a_1 + a_2 \end{pmatrix} \right).$$

Problem 2.27 The group A of affine transformations of the line consists of transformations of the form

$$x' = ax + b, \qquad a \neq 0.$$

Show that these form a semi-direct product on $\dot{\mathbb{R}} \times \mathbb{R}$. Although the multiplicative group of reals $\dot{\mathbb{R}}$ and the additive group $\mathbb{R}$ are both abelian, demonstrate that their semi-direct product is not.

Solution: $Ify = ax' + b$ and $x' = a'x' + b')$ then $y = aa'x + ab' + b$. Thus writing affine transformations as elements (a, b) of $\dot{\mathbb{R}} \times \mathbb{R}$, where $a \neq 0$, the composition law is

$$(a, b)(a'b') = (aa', b + ab')$$

which is precisely the law of composition of a semidirect product (see Problem 2.25). The multiplicative group of reals $\dot{\mathbb{R}}$ is here assumed to act on the additive group of reals $\mathbb{R}$ by taking a product ab.

The semidirect product is not abelian since

$$(a', b')(a, b) = (a'a, b' + a'b) \neq (aa', b + ab') \qquad \text{in general.}$$

Chapter 3

Problem 3.1 **Show that the integers modulo a prime number p form a finite field.**

Solution: Since the rules for adding and multiplying modulo p imply $[a]+[b]=[a+b]$ and $[a][b]=[ab]$, the conditions for a ring, axioms (R1)-(R7) follow immediately. To show that they form a field it is only necessary to show that the non-zero elements all have inverses. The element $[1]$ is clearly an identity, $[1][a]=[a]$ for all integers a. Let $[a]\neq[0]$, i.e. $a\neq np$ for any integer n. Since a and p are relatively prime there exist integers k and l such that $ka+lp=1$, i.e. $ka\equiv 1 \bmod p$. Hence $[k][a]=[1]$ and $k=[a]^{-1}$. Thus the ring of integers modulo p forms a field.

Problem 3.2 **Show that the set of all real numbers of the form $a+b\sqrt{2}$, where a and b are rational numbers, is a field. If a and b are restricted to the integers show that this set is a ring, but is not a field.**

Solution: The set of numbers of the form $a+b\sqrt{2}$ where a and b are rational, is closed with respect to addition and multiplication, for

$$(a+b\sqrt{2})+(a'+b'\sqrt{2})=(a+a')+(b+b')\sqrt{2},$$

$$(a+b\sqrt{2})(a'+b'\sqrt{2})=aa'+2bb'+(ab'+ba')\sqrt{2}.$$

Axioms (R1), (R2), (R5), R(6), R(7) all follow from addition and multiplication of real numbers. The number 0 is clearly of this form, as are additive inverses (negatives)

$$0=0+0\sqrt{2},\qquad -(a+b\sqrt{2})=-a-b\sqrt{2}.$$

The identity $1=1+0\sqrt{2}$ is of this type, and the inverse of any non-zero number of this form is also of this form for

$$\frac{1}{a+b\sqrt{2}}=\frac{a-b\sqrt{2}}{(a+b\sqrt{2})(a-b\sqrt{2})}=\frac{a-b\sqrt{2}}{a^2-2b^2}.$$

The right hand side numerator $a^2-2b^2\neq 0$, since $a\neq b\sqrt{2}$ for any rational numbers a and b since $\sqrt{2}$ is an irrational number.

If a and b are restricted to be integers, all the above arguments hold for the ring axioms (R1)-(R7), but inverses are clearly not in this class (e.g. $(2+0\sqrt{2})^{-1}=\frac{1}{2}$).

Problem 3.3 **Show that the infinite dimensional vector space $\mathbb{R}^\infty$ is isomorphic with a proper subspace of itself.**

Solution: Let V be the subspace consisting of all sequences having first element 0,

$$V=\{(0,v_2,v_3,v_4,\dots)\}$$

This subspace is isomorphic to $\mathbb{R}^\infty$, by setting $T : \mathbb{R}^\infty \to V$ to be the linear map

$$T(u_1, u_2, u_3, \dots) = (0, u_1, u_2, u_3, \dots).$$

This map is linear

$$\begin{aligned} T((u_1, u_2, \dots) + a(v_1, v_2, \dots)) = T(u_1 + av_1, u_2 + av_2, \dots) = (0, u_1 + av_1, u_2 + av_2, \dots) \\ (0, u_1, u_2, \dots) + a(0, v_1, v_2, \dots) = T(u_1, u_2, \dots) + aT(v_1, v_2, \dots). \end{aligned}$$

It is clearly one-to-one, $Tu = Tu' \Rightarrow u = u'$ and onto as every point of V is the image of a point of $\mathbb{R}^\infty$. The map T is therefore an isomorphism.

Problem 3.4 On the vector space $\mathcal{P}(x)$ of polynomials with real coefficients over a variable x, let x be the operation of multiplying by the polynomial x, and let D be the operation of differentiation,

$$x : f(x) \mapsto xf(x) \qquad D : f(x) \mapsto \frac{\mathrm{d}f(x)}{\mathrm{d}x}.$$

Show that both of these are linear operators over $\mathcal{P}(x)$ and that $Dx - xD = I$, where I is the identity operator.

Solution: The operators x and D are

$$x(a_0 + a_1x + a_2x^2 + \cdots + a_nx^n) = a_0x + a_1x^2 + a_2x^3 + \cdots + a_nx^{n+1},$$

$$D(a_0 + a_1x + a_2x^2 + \cdots + a_nx^n) = a_1 + a_2x + \cdots + na_nx^{n-1}.$$

Both these operators are clearly linear,

$$x(f(x) + ag(x)) = xf(x) + axg(x),$$

$$D(f(x) + ag(x)) = \frac{\mathrm{d}}{\mathrm{d}x}(f(x) + ag(x)) = \frac{\mathrm{d}f(x)}{\mathrm{d}x} + a\frac{\mathrm{d}g(x)}{\mathrm{d}x} = Df(x) + aDg(x).$$

If $f(x) = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n$ then

$$\begin{aligned} Dxf(x) &= D(a_0x + a_1x^2 + a_2x^3 + \cdots + a_nx^{n+1} \\ &= a_0 + 2a_1x + 3a_2x^2 + \cdots + (n+1)a_nx^n \end{aligned}$$

and

$$\begin{aligned} xDf(x) &= x(a_1 + 2a_2x + \cdots + na_nx^{n-1}) \\ &= a_1x + 2a_2x^2 + \cdots + na_nx^n, \end{aligned}$$

so that

$$(Dx - xD)f(x) = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n = f(x).$$

Hence $Dx - xD = I$, the identity map on $\mathcal{P}(x)$.

Problem 3.5 **If L, M and N are vector subspaces of V show that**

$$L \cap (M + (L \cap N)) = L \cap M + L \cap N$$

but it is not true in general that

$$L \cap (M + N) = L \cap M + L \cap N.$$

Solution: Let $u \in L \cap (M + (L \cap N))$. Then $u \in L$ and there exist vectors $m \in M$ and $n \in L \cap N$ such that $u = m + n$. Since $m = u - n$ and both u and n belong to L, we must have $m \in L$. Hence $m \in L \cap M$ and thus $u \in L \cap M + L \cap N$. Hence $L \cap (M + (L \cap N)) \subseteq L \cap M + L \cap N$. For the converse, let $u \in L \cap M + L \cap N \subseteq M + (L \cap N)$. Then $u = m + n$ where $m \in L \cap M$ and $m \in L \cap N$, and since both m and n belong to L, we must have $u \in L$. Hence $u \in L \cap (M + (L \cap N))$, and therefore $L \cap M + L \cap N \subseteq L \cap (M + (L \cap N))$. The set theoretical equality follows from the standard criterion in Chapter 1 (the Axiom of Extensionality).

Let $M = \{(x, 0) \mid x \in \mathbb{R}\} \subset \mathbb{R}^2$ and $N = \{(0, y) \mid y \in \mathbb{R}\} \subset \mathbb{R}^2$. If $L = \{(a, a) \mid a \in \mathbb{R}\}$ then $L \cap M = L \cap N = \{0\}$, but since $M + N = \mathbb{R}^2$ it follows that $L \cap (M + N) = L$. Thus $(L \cap M) + (L \cap N) = \{0\} \neq L \cap (M + N)$.

Problem 3.6 **Let $V = U \oplus W$, and let $v = u + w$ be the unique decomposition of a vector v into a sum of vectors from $u \in U$ and $w \in W$. Define the *projection operators* $P_U : V \to U$ and $P_W : V \to W$ by**

$$P_U(v) = u, \qquad P_W(v) = w.$$

Show that

(a) $P_U^2 = P_U$ and $P_W^2 = P_W$.

(b) Show that if $P : V \to V$ is an operator satisfying $P^2 = P$, said to be an *idempotent* operator, then there exists a subspace U such that $P = P_U$. [*Hint*: Set $U = \{u \mid Pu = u\}$ and $W = \{w \mid Pw = 0\}$ and show that these are complementary subspaces such that $P = P_U$ and $P_W = \mathrm{id}_V - P$.]

Solution: (a) For any vector $v \in V$, $P_U^2 v = P_U(P_U v) = P_U u = u$ since $u = u + 0$ is the unique decomposition of the vector $u \in U$. Hence $P_U^2 v = P_U v$ for all $v \in V$. Similarly $P_W^2 = P_W$.

(b) If $P^2 = P$ we set $U = \{u \mid Pu = u\}$ and $W = \{w \mid Pw = 0\}$. These are vector subspaces of V, for if $u, u' \in U$ then so is any linear combination $u + au'$, then by linearity of the operator P

$$Pu = u,\ Pu' = u' \implies P(u + au') = Pu + aPu' = u + au'$$

and

$$w, w' \in W \implies Pw = 0,\ Pw' = 0 \implies P(w + aw') = Pw + aPw' = 0 + a0 = 0.$$

The intersection of these subspaces is $U \cap W = \{0\}$, for if $x \in U \cap W$ then $Px = x = 0$. For any vector $v \in V$, set $u = Pv$ and $w = v - u$. Then $u \in U$, for $Pu = P^2 v = Pv = u$, and $w \in W$ for $Pw = Pv - Pu = u - u = 0$, and the decomposition of any vector $v = u + w$ is as required for a direct sum. The uniqueness of the decomposition follows from $U \cap W = \{0\}$, for if $v = u + w = u' + w'$ then $u - u' = w - w'$ is a vector belonging both to U and W, so that $u - u' = w - w' = 0$. It is immediate from the construction that $P = P_U$ and $P_W = \mathrm{id}_V - P$, for $Pv = u$ and $(\mathrm{id}_V - P)v = v - u = w$.

Problem 3.7 **Show that the vectors $(1, x)$ and $(1, y)$ in $\mathbb{R}^2$ are linearly dependent iff $x = y$. In $\mathbb{R}^3$, show that the vectors $(1, x, x^2)$, $(1, y, y^2)$ and $(1, z, z^2)$ are linearly dependent iff $x = y$ or $y = z$ or $x = z$.**

Generalize these statements to $(n + 1)$ dimensions.

Solution: If there exists a non-trivial linear relation $a(1, x) + b(1, y) = 0$ then

$$a + b = ax + by = 0$$

which holds iff $a = -b \neq 0$, and $x - y = 0$. Hence $x = y$.
In $\mathbb{R}^3$ the vectors $(1, x, x^2)$, $(1, y, y^2)$ and $(1, z, z^2)$ are linearly dependent iff there exist three scalars a,b, c, not all zero, such that the matrix equation

$$\begin{pmatrix} a & b & c \end{pmatrix} \begin{pmatrix} 1 & x & x^2 \\ 1 & y & y^2 \\ 1 & z & z^2 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \end{pmatrix}$$

has a non-trivial solution for a, b, c. This is so iff

$$\det \begin{vmatrix} 1 & x & x^2 \\ 1 & y & y^2 \\ 1 & z & z^2 \end{vmatrix} = 0,$$

which on expanding by the first row becomes the polynomial equation

$$\begin{aligned} \det &= yz^2 - y^2 z - x(z^2 - y^2) + x^2(z - y) \\ &= -(x - y)(x - z)(y - z) \end{aligned}$$

This determinant only vanishes if $x = y$ or $y = z$ or $x = z$.
In $\mathbb{R}^{n+1}$ the vectors $(1, x_1, x_1^2, \dots, x_1^n)$, $(1, x_2, x_2^2, \dots, x_2^n), \dots, (1, x_{n+1}, x_{n+1}^2, \dots, x_{n+1}^n)$, are linearly dependent iff

$$\det \begin{vmatrix} 1 & x_1 & x_1^2 & \dots & x_1^n \\ 1 & x_2 & x_2^2 & \dots & x_2^n \\ \dots & \dots & \dots & \dots & \dots \\ 1 & x_{n+1} & x_{n+1}^2 & \dots & x_{n+1}^n \end{vmatrix} = (-1)^{n(n+1)/2} \prod_{i<j} (x_i - x_j) = 0.$$

The evaluation of the determinant is from the fact that it is a polynomial of degree $n(n+1)/2$ which vanishes whenever $x_i = x_j$ for any pair of indices $i \neq j$, and the coefficient is found by evaluating contribution from the diagonal term. Thus the vectors are linearly dependent if and only if some pair of vectors are equal, $x_i = x_j$.

Problem 3.8 **Let V and W be any vector spaces, which are possibly infinite dimensional, and $T : V \to W$ a linear map. Show that if M is a l.i. subset of W, then $T^{-1}(M) = \{v \in V \mid Tv \in M\}$ is a linearly independent subset of V.**

Solution: Assume that M is an l.i. subset of W. If $v_1, v_2, \ldots, v_k$ are any vectors belonging to $T^{-1}(M)$ such that

$$a^1 v_1 + a^2 v_2 + \cdots + a^k v_k = 0,$$

then

$$T(a^1 v_1 + \cdots + a^k v_k) = a^1 T(v_1) + \cdots + a^k T(v_k) = 0.$$

Since $T(v_k) \in M$ and M is l.i. all coefficients $a^1, \ldots, a^k$ must vanish. Hence $T^{-1}(M)$ is a linearly independent subset of V.

Problem 3.9 **Let V and W be finite dimensional vector spaces of dimensions n and m respectively, and $T : V \to W$ a linear map. Given a basis $\{e_1, e_2, \ldots, e_n\}$ of V and a basis $\{f_1, f_2, \ldots, f_m\}$ of W, show that the equations**

$$Te_k = \sum_{a=1}^{m} T^a_k f_a \qquad (k = 1, 2, \ldots, n)$$

serve to uniquely define the $m \times n$ matrix of components $\mathsf{T} = [T^a_k]$ of the linear map T with respect to these bases.

If $v = \sum_{k=1}^{n} v^k e_k$ is an arbitrary vector of V show that the components of its image vector $w = Tv$ are given by

$$w^a = (Tv)^a = \textstyle\sum_{k=1}^{n} T^a_k v^k \ .$$

Write this as a matrix equation.

Solution: For each $k = 1, \ldots, n$ the vector $Te_k \in W$ and therefore has a unique expansion in the basis of f_a of W, with coefficients T^a_k $(a = 1, \ldots, m)$. Hence the matrix of coefficients $\mathsf{T} = [T^a_k]$ is uniquely defined by the given bases e_1 and f_a.
For any vector $v = \sum_{k=1}^{n} v^k e_k \in V$

$$w = Tv = \sum_{k=1}^{n} v^k T e_k = \sum_{k=1}^{n} \sum_{a=1}^{m} T^a_k v^k f_a.$$

Expanding w in the basis f_a, namely $w = \sum_{a=1}^{m} w^a f_a$ we see that

$$w^a = \sum_{k=1}^{n} T^a_k v^k \ .$$

Writing $\mathbf{w}$ as the $m \times 1$ column matrix having components w^a and $\mathbf{v}$ as the $n \times 1$ column vector with components v^k, this equation may be written

$$\mathbf{w} = \mathsf{T}\mathbf{v}.$$

Problem 3.10 **Let V be a four-dimensional vector space and $T : V \to V$ a linear operator whose effect on a basis $e_1, \dots, e_4$ is**

$$\begin{aligned} Te_1 &= 2e_1 - e_4 \\ Te_2 &= -2e_1 + e_4 \\ Te_3 &= -2e_1 + e_4 \\ Te_4 &= e_1. \end{aligned}$$

Find a basis for $\ker T$ and $\operatorname{im} T$ and calculate the rank and nullity of T.

Solution: $\ker T = \{u \in V \,|\, Tu = 0\}$, hence $u = ae_1 + be_2 + ce_3 + de_4 \in \ker T$ iff

$$(2a - 2b - 2c + d)e_1 + (-a + b + c)e_4 = 0.$$

Setting the coefficients of e_1 and e_4 to zero, we have $a = b + c$ and $d = 0$. Thus

$$u = b(e_1 + e_2) + c(e_1 + e_3)$$

where b and c are arbitrary scalars. The vectors $f_1 = e_1 + e_2$ and $f_2 = e_1 + e_3$ therefore form a basis of $\ker T$.
Since $Tu = (a - b - c)(2e_1 - e_4) + de_1$, and a, b, c, d are arbitrary scalars, the space $\operatorname{im} T$ is spanned by $g_1 = 2e_1 - e_4$ and $g_2 = e_1$.
The rank of T is $\dim(\operatorname{im} T) = 2$, and its nullity is $\dim(\ker T) = 2$.

Problem 3.11 **Let $\{e_1, e_2, e_3\}$ be a basis of a three-dimensional vector space V. Show that the vectors $\{e'_1, e'_2, e'_3\}$ defined by**

$$\begin{aligned} e'_1 &= e_1 + e_3 \\ e'_2 &= 2e_1 + e_2 \\ e'_3 &= 3e_2 + e_3 \end{aligned}$$

also form a basis of V.

What are the elements of the matrix $\mathsf{A} = [A^j{}_i]$ in Eq. (3.16) ? Calculate the components of the vector

$$v = e_1 - e_2 + e_3$$

with respect to the basis $\{e'_1, e'_2, e'_3\}$, and verify the column vector transformation $\mathbf{v}' = \mathsf{A}\mathbf{v}$.

Solution: The vectors e'_i must be shown to be linearly independent. One method is to express the original basis e_i explicitly in terms of the e'_j, by solving the three linear equations:

$$\begin{aligned} e_1 &= \tfrac{1}{7}(e'_1 + 3e'_2 - e'_3) \\ e_2 &= \tfrac{1}{7}(-2e'_1 + e'_2 + 2e'_3) \\ e_3 &= \tfrac{1}{7}(6e'_1 - 3e'_2 + e'_3) \end{aligned}$$

The matrix A in the equation $e_i = A^j{}_i e'_j$ has components

$$\mathsf{A} = [A^j{}_i] = \begin{pmatrix} \frac{1}{7} & \frac{-2}{7} & \frac{6}{7} \\ \frac{3}{7} & \frac{1}{7} & \frac{-3}{7} \\ \frac{-1}{7} & \frac{2}{7} & \frac{1}{7} \end{pmatrix}.$$

Expanding the vector $v = e_1 - e_2 + e_3$ in the new basis by substituting for e_1, e_2, e_3 from the above, gives

$$\begin{aligned} v &= \frac{1}{7}(e'_1 + 3e'_2 - e'_3 + 2e'_1 - e'_2 - 2e'_3 + 6e'_1 - 3e'_2 + e'_3) \\ &= \frac{1}{7}(9e'_1 - e'_2 - 2e'_3) \end{aligned}$$

In matrices,

$$\mathbf{v}' = \begin{pmatrix} \frac{9}{7} & \frac{-1}{7} & \frac{-2}{7} \end{pmatrix} = \mathsf{A}\mathbf{v} = \begin{pmatrix} \frac{1}{7} & \frac{-2}{7} & \frac{6}{7} \\ \frac{3}{7} & \frac{1}{7} & \frac{-3}{7} \\ \frac{-1}{7} & \frac{2}{7} & \frac{1}{7} \end{pmatrix} \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}.$$

Problem 3.12 **Let $T : V \to W$ be a linear map between vector spaces V and W. If $\{e_i \,|\, i = 1, \dots, n\}$ is a basis of V and $\{f_a \,|\, a = 1, \dots, m\}$ a basis of W, how does the matrix T, defined in Problem 3.9, transform under a transformation of bases**

$$e_i = A^j{}_i e'_j\,, \qquad f_a = B^b{}_a f'_b\,?$$

Express your answer both in component and in matrix notation.

Solution:
$$Te_i = T^a_i f_a \quad \Longrightarrow \quad \begin{aligned} Te'_j &= T(A'^k_{\ j} e_k) \qquad \text{by Eq. (3.20)} \\ &= A'^k_{\ j} Te_k \\ &= A'^k_{\ j} T^a_k f_a \\ &= A'^k_{\ j} T^a_k B^b_{\ a} f'_b \\ &= T'^b_{\ j} f'_b \end{aligned}$$

where
$$T'^b_{\ j} = A'^k_{\ j} T^a_k B^b_{\ a} = B^b_{\ a} T^a_k (A^{-1})^k_j.$$

In matrices this equation reads
$$\mathsf{T}' = \mathsf{B}\mathsf{T}\mathsf{A}^{-1}.$$

Problem 3.13 **Let e_1, e_2, e_3 be a basis for a three-dimensional vector space and e'_1, e'_2, e'_3 a second basis given by**
$$\begin{aligned} e'_1 &= e_3, \\ e'_2 &= e_2 + 2e_3, \\ e'_3 &= e_1 + 2e_2 + 3e_3. \end{aligned}$$

(a) Express the e_i in terms of the e'_j, and write out the transformation matrices $\mathsf{A} = [A^i_{\ j}]$ and $\mathsf{A}' = \mathsf{A}^{-1} = [A'^i_{\ j}]$.

(b) If $u = e_1 + e_2 + e_3$, compute its components in the e'_i basis.

(c) Let T be the linear transformation defined by
$$Te_1 = e_2 + e_3, \quad Te_2 = e_3 + e_1, \quad Te_3 = e_1 + e_2.$$

What is the matrix of components of T with respect to the basis e_i?

(d) By evaluating Te'_1, etc. in terms of the e'_j, write out the matrix of components of T with respect to the e'_j basis and verify the similarity transformation $\mathsf{T}' = \mathsf{A}\mathsf{T}\mathsf{A}^{-1}$.

Solution: (a) Solving for e_1, e_2, e_3 we find
$$\begin{aligned} e_1 &= e'_1 - 2e'_2 + e'_3 \\ e_2 &= -2e'_1 + e'_2 \\ e_3 &= e'_1 \end{aligned}$$

and from the equations $e_i = A^j_{\ i} e'_j$, $e'_j = A'^i_{\ j} e_i$ we read off the matrix coefficients
$$\mathsf{A} = [A^j_{\ i}] = \begin{pmatrix} 1 & -2 & 1 \\ -2 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \qquad \mathsf{A} = [A'^i_{\ j}] = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 2 \\ 1 & 2 & 3 \end{pmatrix}.$$

(b) Substtuting for the e_i we have $u = e_1 + e_2 + e_3 = -e'_1 + e'_3$, whence

$$u'^1 = 0, \qquad u'^2 = -1, \qquad u'^3 = 1.$$

(c) Reading off the coefficents of the matrix T from $Te_i = T^j_i e_j$,

$$\mathsf{T} = [T^j_i] = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}.$$

(d)

$$\begin{aligned} Te'_1 &= Te_3 = e_1 + e_2 = -e'_1 - e'_2 + e'_3 \\ Te'_2 &= T(e_2 + 2e_3) = 3e_1 + 2e_2 + e_3 = -4e'_2 + 3e'_3 \\ Te'_3 &= T(e_1 + 2e_2 + 3e_3) = 5e_1 + 4e_2 + 3e_3 = -6e'_2 + 5e'_3 \end{aligned}$$

so that

$$\mathsf{T}' = \begin{pmatrix} -1 & 0 & 0 \\ -1 & -4 & -6 \\ 1 & 3 & 5 \end{pmatrix} = \mathsf{A}(\mathsf{T}\mathsf{A}^{-1}) = \begin{pmatrix} 1 & -2 & 1 \\ -2 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 3 & 5 \\ 1 & 2 & 4 \\ 0 & 1 & 3 \end{pmatrix}.$$

Problem 3.14 Find the dual basis to the basis of $\mathbb{R}^3$ having column vector representation

$$\mathbf{e}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad \mathbf{e}_2 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \quad \mathbf{e}_3 = \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}.$$

Solution: Let $\varepsilon^1 = (a_1, b_1, c_1)$, $\varepsilon^2 = (a_2, b_2, c_2)$, $\varepsilon^3 = (a_3, b_3, c_3)$. Then the linear equations

$$\varepsilon^1(e_1) = a_1 + b_1 + c_1 = 1, \quad \varepsilon^1(e_2) = a_1 - c_1 = 0, \quad \varepsilon^1(e_3) = -b_1 + c_1 = 0$$

are readily solved to give $a_1 = b_1 = c_1 = \frac{1}{3}$. Similarly, equations $\varepsilon^2(e_i) = \delta^2_i$ and $\varepsilon^3(e_i) = \delta^3_i$ result in

$$a_2 + b_2 + c_2 = 0, \quad a_2 - c_2 = 1, \quad -b_2 + c_2 = 0 \qquad \Longrightarrow \qquad a_2 = \tfrac{2}{3}, \quad b_2 = c_2 = -\tfrac{1}{3},$$

and

$$a_3 + b_3 + c_3 = 0, \quad a_3 - c_3 = 0, \quad -b_3 + c_3 = 1 \qquad \Longrightarrow \qquad a_3 = c_3 = \tfrac{1}{3}, \quad b_3 = -\tfrac{2}{3}.$$

Hence

$$\varepsilon^1 = (\tfrac{1}{3}, \tfrac{1}{3}, \tfrac{1}{3}), \qquad \varepsilon^2 = (\tfrac{2}{3}, -\tfrac{1}{3}, -\tfrac{1}{3}), \qquad \varepsilon^3 = (\tfrac{1}{3}, -\tfrac{2}{3}, \tfrac{1}{3}).$$

Problem 3.15 **Let $\mathcal{P}(x)$ be the vector space of real polynomials $f(x) = a_0 + a_1x + \dots + a_nx^n$. If $(b^0, b^1, b^2, \dots)$ is any sequence of real numbers, show that the map $\beta : \mathcal{P}(x) \to \mathbb{R}$ given by**

$$\beta(f(x)) = \sum_{i=0}^{n} b^i a_i$$

is a linear functional on $\mathcal{P}(x)$.

Show that *every* linear functional β on $\mathcal{P}(x)$ can be obtained in this way from such a sequence and hence that $(\hat{\mathbb{R}}^\infty)^* = \mathbb{R}^\infty$.

Solution: If $f(x) = \sum_{i=0}^{n} a_i x^i$ and $f'(x) = \sum_{i=0}^{m} a'_i x^i$ we may write

$$f(x) = \sum_{i=0}^{\infty} a_i x^i, \qquad f'(x) = \sum_{i=0}^{\infty} a'_i x^i$$

where we set $a_i = 0$ for all $i > n$ and $a'_j = 0$ for all $j > m$. Then

$$f(x) + af'(x) = \sum_{i=0}^{\infty} (a_i + aa'_i)x^i, \qquad \text{where} \qquad a_i + aa'_i = 0 \text{ for all } i > \max(n, m),$$

and

$$\beta\big(f(x) + af'(x)\big) = \sum_{i=0}^{\infty} b^i(a_i + aa'_i) = \sum_{i=0}^{\infty} b^i a_i + a\sum_{i=0}^{\infty} b^i a'_i = \beta\big(f(x)\big) + a\beta\big(f'(x)\big).$$

Hence β is a linear functional on $\mathcal{P}(x)$.

Conversely if β is a linear functional on $\mathcal{P}(x)$, set $b^0 = \beta(1)$, $b^1 = \beta(x)$, $b^2 = \beta(x^2), \dots, b^n = \beta(x^n), \dots$, and it follows from linearity that $\beta\big(f(x)\big) = \sum_{i=0}^{\infty} b^i a_i$. Since all such sequences $(b^0, b^1, b^2, \dots)$ form the vector space $\mathbb{R}^\infty$, and $\mathcal{P}(x) \cong \hat{\mathbb{R}}^\infty$ (see Example 3.10), it follows that $(\hat{\mathbb{R}}^\infty)^* = \mathbb{R}^\infty$.

Problem 3.16 **Define the annihilator $S^\perp$ of a subset $S \subseteq V$ as the set of all linear functionals which vanish on S,**

$$S^\perp = \{\omega \in V^* \mid \omega(u) = 0 \ \forall u \in S\}.$$

(a) Show that for any subset S, $S^\perp$ is a vector subspace of V^*.

(b) If $T \subseteq S$, show that $S^\perp \subseteq T^\perp$.

(c) If U is a vector subspace of V show that $(V/U)^* \cong U^\perp$ [*Hint*: For each

$\omega \in U^\perp$ define the element $\bar\omega \in (V/U)^*$ by $\bar\omega(v+U) = \omega(v)$.]

(d) Show that $U^* \cong V^*/U^\perp$.

(e) If V is finite dimensional with $\dim V = n$ and W is any subspace of V with $\dim W = m$, show that $\dim W^\perp = n - m$. [*Hint*: Use a basis adapted to the subspace W by Theorem 3.7 and consider its dual basis in V^*.]

(f) Adopting the natural identification of V and V^{**}, show that $(W^\perp)^\perp = W$.

Solution: (a) If ω and ρ are any pair of linear functionals belonging to $S^\perp$. Then, for any $a \in \mathbb{R}$,

$$(\omega + a\rho)(u) = \omega(u) + a\rho(u) = 0 \qquad \text{for all } u \in S.$$

Hence $\omega + a\rho \in S^\perp$, showing that $S^\perp$ is a vector subspace of V^*.

(b) If $\omega \in S^\perp$ then $\omega(u) = 0$ for all $u \in S$. In particular, since $T \subseteq S$, $\omega(u) = 0$ for all $u \in T$; in other words, $\omega \in T^\perp$. Hence $S^\perp \subseteq T^\perp$.

(c) For any $\omega \in U^\perp$ define $\bar\omega \in (V/U)^*$ by $\bar\omega(v+U) = \omega(v)$. This definition is independent of the coset representative v, for if $v' = v + u$ where $u \in U$ then $\omega(v') = \omega(v)$. The correspondence $\omega \to \bar\omega$ is one-to-one for if $\bar\omega(v+U) = \bar\rho(v+U)$ for all cosets $v + U$ then $\omega(v) = \rho(v)$ for all $v \in V$. Furthermore the correspondence is a vector space homomorphism, i.e. linear, since

$$\overline{\omega + a\rho}(v+U) = (\omega + a\rho)(v) = \omega(v) + a\rho(v) = \bar\omega(v+U) + a\bar\rho(v+U).$$

(d) Every element of $V^*/U^\perp$ can be written as a coset $[\omega] = \omega + U^\perp$ where $\omega \in V^*$. This defines a unique element of U^* by setting $[\omega](u) = \omega(u)$, which definition is independent of the choice of coset representative ω, for if $[\omega] = [\omega']$ then $\omega' = \omega + \rho$ for some $\rho \in U^\perp$ so that

$$[\omega'](u) = (\omega + \rho)(u) = \omega(u) + \rho(u) = \omega(u) = [\omega](u).$$

The correspondence $V^*/U^\perp \to U^*$ is one-to-one, for

$$\begin{aligned}
[\omega] = [\omega'] \quad &\Longrightarrow \quad [\omega](u) = [\omega'](u) \qquad \forall u \in U \\
&\Longrightarrow \quad (\omega - \omega')(u) = 0 \qquad \forall u \in U \\
&\Longrightarrow \quad \omega - \omega' \in U^\perp \\
&\Longrightarrow \quad \omega + U^\perp = \omega' + U^\perp.
\end{aligned}$$

Finally, the correspondence is *onto*. For, let W be any complementary subspace to U (see Theorem 3.2), such that $V = U \oplus W$. Given any $\alpha \in U^*$ we can extend it to a unique linear functional on V by setting $\alpha(W) = \{0\}$, and using the fact that every $v \in V$ has a unique decomposition $v = u + w$ where $u \in U$, $w \in W$. Then $\alpha = [\alpha]$, since $\alpha(v) = \alpha(u) = [\alpha](u)$. We thus have a vector space isomorphism

$U^* \cong V^*/U^\perp$.

(e) Let $\{e_1, \dots, e_n\}$ be a basis of V such that $\{e_1, \dots, e_m\}$ is a basis of the subspace W (see Theorem 3.7). If $\{\varepsilon^1, \dots, \varepsilon^n\}$ is the dual basis then a linear functional $\omega = w_i \varepsilon^i$ belongs to $W^\perp$ if and only if $w_i = 0$ for all $i = 1, \dots m$. Since $w_{m+1}, \dots, w_n$ may be arbitrary, the annihilator subspace $W^\perp$ is the subspace of V^* spanned by the linearly independent covectors $\varepsilon^{m+1}, \dots, \varepsilon^n$. Hence $\dim W^\perp = n - m$.

(f) Since $W^\perp = \{\omega \in V^* \,|\, \omega(u) = 0 \quad \forall u \in W\}$, we have

$$\begin{aligned}(W^\perp)^\perp &= \{v \in V^{**} = V \,|\, \langle v, \omega\rangle = 0 \; \forall \omega \in W^\perp\} \\ &= \{v \in V \,|\, \omega(v) = 0 \; \forall \omega \in W^\perp\} \supseteq W\end{aligned}$$

If $W^{\perp\perp} \neq W$ then there exists a vector in $W^{\perp\perp}$ linearly indepdendent of any basis of W, so $\dim W^{\perp\perp} > m$. However, by part (e), $\dim W^\perp = n - m$, so that $\dim W^{\perp\perp} = n - (n - m) = m = \dim W$. Hence we must have $W^{\perp\perp} = W$.

Problem 3.17 **Let u be a vector in the vector space V of dimension n.**

(a) If ω is a linear functional on V such that $a = \omega(u) \neq 0$, show that a basis $e_1, \dots, e_n$ can be chosen such that

$$u = e_1 \qquad \textbf{and} \qquad \omega = a\varepsilon^1$$

where $\{\varepsilon^1, \dots, \varepsilon^n\}$ is the dual basis. [*Hint*: Apply Theorem 3.7 to the vector u and try a further basis transformation of the form $e'_1 = e_1$, $e'_2 = e_2 + a_2 e_1, \dots,$ $e'_n = e_n + a_n e_1$.]

(b) If $a = 0$, show that the basis may be chosen such that

$$u = e_1 \qquad \textbf{and} \qquad \omega = \varepsilon^2.$$

Solution: (a) Let

$$e'_1 = e_1 = u, \qquad e'_2 = e_2 + a_2 e_1, \quad \dots \quad , \quad e'_n = e_n + a_n e_1.$$

If $\{\varepsilon^1, \dots, \varepsilon^n\}$ and $\{\varepsilon'^1, \dots, \varepsilon'^n\}$ are dual bases to the bases $\{e_i\}$ and $\{e'_i\}$ resp., write $\omega = w_i \varepsilon^i = w'_i \varepsilon'^i$ where $w_i = \omega(e_i)$ and

$$w'_i = \omega(e'_i) = \omega(e_i + a_i e_i) = w_i + a_i a \qquad (i = 2, \dots n)$$

Pick $a_i = -w_i/a$ and we have $w'_i = 0$ for $i = 2, \dots, n$. Thus

$$\omega = w'_1 \varepsilon'^1 = \omega(e'_1)\varepsilon'^1 = \omega(u)\varepsilon'^1 = a\varepsilon'^1,$$

as required.

(b) If $a = 0$, then in any basis such that $u = e_1$ we have $w_1 = a = 0$. We may assume without loss of generality that $w_2 \neq 0$, for if $w_i = 0$ for all $i \neq 1$ then $\omega = 0$ (since $w_1 = 0$). Set $e'_1 = e_1 = u$ and

$$\begin{aligned} e'_2 &= \frac{1}{w_2}\, e_2 \\ e'_i &= e_1 + a_i e_i \qquad (i = 3, \dots, n). \end{aligned}$$

We then have

$$w'_2 = \omega(e'_2) = \frac{1}{w_2}\omega(e_2) = \frac{w_2}{w_2} = 1,$$

while for $i > 2$

$$w'_i = \omega(e'_i) = w_i + a_i w_2 = 0$$

if we set $a_i = -w_i/w_2$. The primed basis then has the required properties, $u = e'_1$, $\omega = \varepsilon'^2$.

Problem 3.18 **For the three-dimensional basis transformation of Problem 3.13 evaluate the ε'^j dual to e'_i in terms of the dual basis ε^j. What are the components of the linear functional $\omega = \varepsilon^1 + \varepsilon^2 + \varepsilon^3$ with respect to the new dual basis?**

Solution: Using $\varepsilon'^i = A'^i{}_j \varepsilon^j$ gives

$$\begin{aligned} \varepsilon'^1 &= \varepsilon^1 - 2\varepsilon^2 + \varepsilon^3 \\ \varepsilon'^2 &= -2\varepsilon^1 + \varepsilon^2 \\ \varepsilon'^3 &= \varepsilon^1 \end{aligned}$$

Either by solving for the ε^i or by invoking $\varepsilon^i = A^i{}_j \varepsilon'^j$, we find

$$\begin{aligned} \varepsilon^1 &= \varepsilon'^3 \\ \varepsilon^2 &= \varepsilon'^2 + 2\varepsilon'^3 \\ \varepsilon^3 &= \varepsilon'^1 + 2\varepsilon'^2 + 3\varepsilon'^3 \end{aligned}$$

Substituting these in ω we find

$$\omega = \varepsilon^1 + \varepsilon^2 + \varepsilon^3 = \varepsilon'^1 + 3\varepsilon'^2 + 6\varepsilon'^3.$$

Thus the components of ω with respect to the new basis are $(1, 3, 6)$.

Problem 3.19 **If $A : V \to V$ is a linear operator, define its transpose to be the linear map $A' : V^* \to V^*$ such that**

$$A'\omega(u) = \omega(Au) \qquad \forall\, u \in V,\ \omega \in V^*.$$

Show that this relation uniquely defines the linear operator A' and that

$$O' = O, \quad (\mathbf{id}_V)' = \mathbf{id}_{V^*}, \quad (aA + bB)' = aA' + bB' \quad \forall a, b \in \mathbb{K}.$$

(a) Show that $(BA)' = A'B'$.

(b) If A is an invertible operator then show that $(A')^{-1} = (A^{-1})'$

(c) If V is finite dimensional show that $A'' = A$, if we make the natural identification of V^{**} and V.

(d) Show that the matrix of components of the transpose map A' with respect to the dual basis is the transpose of the matrix of A, $\mathsf{A}' = \mathsf{A}^T$.

(e) Using Problem 3.16 show that $\ker A' = (\operatorname{im} A)^\perp$.

(f) Use (3.10) to show that the rank of A' equals the rank of A.

Solution: The function $A'\omega : V \to V$ is clearly well defined since $A'\omega(u) = \omega(Au)$ is unique for each $u \in V$. It is a linear functional on V since

$$A'\omega(au+bv) = \omega(A(au+bv)) = \omega(aAu+bAv) = a\omega(Au)+b\omega(Av) = aA'\omega(u)+bA'\omega(v).$$

Furthermore the operator $A' : V^* \to V^*$ is linear, since

$$A'(a\omega + b\rho)(u) = (a\omega + b\rho)(Au) = (aA'\omega + bA'\rho)(u)$$

for all $u \in V$, so that $A'(a\omega + b\rho) = aA'\omega + bA'\rho$. The relations $O' = O, \quad (\mathrm{id}_V)' = \mathrm{id}_{V^*}$ and $(aA+bB)' = aA'+bB'$ are true of any linear map $' : L(V,V) \to L(V^*, V^*)$.

(a) For all $u \in V$

$$A'B'\omega(u) = B'\omega(Au) = \omega(BAu) = (BA)'\omega(u).$$

Hence $A'B' = (BA)'$.

(b) If A is invertible then

$$(A^{-1})'A' = (AA^{-1})' = (\mathrm{id}_V)' = \mathrm{id}_{V^*},$$

whence $(A')^{-1} = (A^{-1})'$

(c) For all $u \in V$ and all $\omega \in V^*$

$$\omega(A''u) = \langle \omega, A''u \rangle = \langle A'\omega, u \rangle = \langle \omega, Au \rangle = \omega(Au)$$

whence $A'' = A$.

(d) For any $u \in V$

$$A'\varepsilon^k(u) = \varepsilon^k(Au) = \varepsilon^k(A^j{}_l u^l e_j) = \delta^k_j A^j{}_l u^l = A^k{}_l u^l = A^k{}_l \varepsilon^l(u).$$

Hence

$$A'\varepsilon^k = A^k{}_l \varepsilon^l,$$

and the matrix components of A' with respect to the dual basis are given by $(A')_l{}^k = A^k{}_l$. Thus $\mathsf{A}' = \mathsf{A}^T$. Equivalently, if $\omega = w_k\varepsilon^k$ the components of $A'\omega$ are $w_k A^k{}_l$, since

$$A' w_k \varepsilon^k = A^k_l w_k \varepsilon^l.$$

(e)

$$\begin{aligned}\ker A' &= \{\omega \mid A'\omega = 0\} \\ &= \{\omega \mid A'\omega(u) = 0 \quad \forall u \in V\} \\ &= \{\omega \mid \omega(Au) = 0 \quad \forall u \in V\} \\ &= (\operatorname{im} A)^\perp\end{aligned}$$

since $\operatorname{im} A = \{Au \mid u \in V\}$.

(f) Using Eq. (3.10)

$$\begin{aligned}\rho(A') &= n - \nu(A') \\ &= n - \dim\ker A' \\ &= n - \dim(\operatorname{im} A)^\perp \\ &= n - (n - \dim\operatorname{im} A) \qquad \text{by Problem 3.16 (e)} \\ &= \dim\operatorname{im} A = \rho(A).\end{aligned}$$

Problem 3.20 The *row rank* of a matrix is defined as the maximum number of linearly independent rows, while its *column rank* is the maximum number of linearly independent columns.

(a) Show that the rank of a linear operator A on a finite dimensional vector space V is equal to the column rank of its matrix A with respect to any basis of V.

(b) Use parts (d) and (e) of Problem 3.19 to show that the row rank of a square matrix is equal to its column rank.

Solution: (a) Let $\{e_1, \dots, e_n\}$ be any basis of V and set $\mathbf{u}_1, \dots, \mathbf{u}_n$ to be the columns of the matrix $\mathsf{A} = [A^i{}_j]$ of A with respect to this basis. That is, $\mathbf{u}_j$ is the column vector having components $A^i{}_j$ $(i = 1, \dots, n)$. Let u_j be the vector whose components with respect to the basis $\{e_1, \dots, e_n\}$ are the components of the column vector $\mathbf{u}_j$, i.e. $u_j = A^i{}_j e_i = Ae_j$. Since the vectors Ae_j span $\operatorname{im} A$, it follows that the rank of A, $\rho(A) = \dim\operatorname{im} A$, is the maximum number of vectors $u_1, \dots, u_n$ which are linearly independent. Hence this is the maximum number of l.i. columns of the matrix A, and is equal to the column rank of A.

(b) Since $\mathsf{A}' = \mathsf{A}^T$ (see Problem 3.19 (d)) the columns of A' are precisely the rows of A. Hence, since $\rho(A') = \rho(A)$ (see Problem 3.19 (f))

$$\text{row rank of } \mathsf{A} = \text{col. rank of } \mathsf{A}' = \rho(A') = \rho(A) = \text{col. rank of } \mathsf{A}.$$

Problem 3.21 **Let S be a linear operator on a vector space V.**

(a) Show that the rank of S is one, $\rho(S) = 1$, if and only if there exists a non-zero vector u and a non-zero linear functional α such that

$$S(v) = u\alpha(v).$$

(b) With respect to any basis $\{e_i\}$ of V and its dual basis $\{\varepsilon^j\}$, show that

$$S^i{}_j = u^i a_j \qquad \textbf{where} \qquad u = u^i e_i, \quad \alpha = a_j \varepsilon^j.$$

(c) Show that every linear operator A of rank r can be written as a sum of r linear operators of rank one.

(d) Show that the last statement is equivalent to the assertion that for every matrix A of rank r there exist column vectors $\mathbf{u}_i$ and $\mathbf{a}_i$ ($i = 1, \dots, r$) such that

$$\mathsf{A} = \sum_{i=1}^{r} \mathbf{u}_i \mathbf{a}_i^T.$$

Solution: (a) $\rho(S) = 1$ if and only if $\dim \operatorname{im} S = 1$, which is true iff $\operatorname{im} S$ is spanned by a single vector u. Thus a necessary and sufficient condition for the operator S to be of rank 1 is that there exists a real-valued function $\alpha : V \to \mathbb{R}$ such that

$$S(v) = \alpha(v)u.$$

The function α is linear, for $S(av + bw) = \alpha(av + bw)u$, and by the linearity of the operator S

$$S(av + bw) = aS(v) + bS(w) = a\alpha(v)u + b\alpha(w)u = (a\alpha(v) + b\alpha(w))u$$

so that

$$\alpha(av + bw) = a\alpha(v) + b\alpha(w).$$

The linear functional α cannot vanish everywhere on V, else S would have rank 0.

(b) Setting $v = e_j$ in $Sv = \alpha(v)u$ we have

$$Se_j = \alpha(e_j)u^i e_i = a_j u^i e_i = S^i{}_j e_i\,.$$

Hence $S^i{}_j = u^i a_j$, as required.

(c) The vector subspace $\operatorname{im} A$ is spanned by $r = \rho(A) = \dim \operatorname{im} A$ l.i. vectors $u_1, \dots, u_r$. Hence every vector $Av \in \operatorname{im} A$ can be uniquely written in the form

$$Av = \alpha_1(v)u_1 + \alpha_2(v)u_2 + \dots + \alpha_r(v)u_r$$

where the real numbers $\alpha_i(v)$ are uniquely determined by v, since the u_i are linearly independent. By an identical argument to that given in (a) it follows that the functions $\alpha_i : V \to \mathbb{R}$ are linear functionals, and we can write

$$A = A_1 + A_2 + \cdots + A_r \qquad \text{where} \qquad A_i(v) = \alpha_i(v)u_i.$$

That is, A is a sum of r linear operators of rank one.

(d) In $\mathbb{R}^n$ every linear functional α can be written in matrix form $\alpha(\mathbf{v}) = \alpha^T\mathbf{v}$. Thus every linear operator represented by a matrix A of rank r can be expressed in the form

$$\mathsf{A}\mathbf{v} = \sum_{i=1}^{r} \mathbf{u}_i \alpha_i^T \mathbf{v}.$$

That is,

$$\mathsf{A} = \sum_{i=1}^{r} \mathbf{u}_i \alpha_i^T.$$

Chapter 4

Problem 4.1 **The trace of an $n \times n$ matrix $\mathsf{T} = [T^i_{\ j}]$ is defined as the sum of its diagonal elements,**

$$\operatorname{tr} \mathsf{T} = T^i_{\ i} = T^1_{\ 1} + T^2_{\ 2} + \cdots + T^n_{\ n}.$$

Show that

(a) $\operatorname{tr}(\mathsf{ST}) = \operatorname{tr}(\mathsf{TS})$**.**

(b) $\operatorname{tr}(\mathsf{ATA}^{-1}) = \operatorname{tr} T$**.**

(c) If $T : V \to V$ is any operator define its trace to be the trace of its matrix with respect to a basis $\{e_i\}$. Show that this definition is independent of the choice of basis, so that there is no ambiguity in writing $\operatorname{tr} T$.

(d) If $f(z) = a_0 + a_1 z + a_2 z^2 + \cdots + (-1)^n z^n$ is the characteristic polynomial of the operator T, show that $\operatorname{tr} T = (-1)^{n-1} a_{n-1}$.

(e) If T has eigenvalues $\lambda_1, \dots, \lambda_m$ with multiplicities $p_1, \dots, p_m$, show that

$$\operatorname{tr} T = \sum_{i=1}^{m} p_i \lambda_i \ .$$

Solution: (a) $\operatorname{tr}(\mathsf{ST}) = (ST)^i_{\ i} = S^i_{\ k} T^k_{\ i} = T^k_{\ i} S^i_{\ k} = (TS)^k_{\ k} = \operatorname{tr}(\mathsf{TS})$.

(b) $\operatorname{tr}(\mathsf{ATA}^{-1}) = \operatorname{tr}(\mathsf{A}^{-1}\mathsf{AT}) = \operatorname{tr}(\mathsf{IT}) = \operatorname{tr} T$.

(c) The matrix $\mathsf{T} = [T^i_{\ j}]$ of the operator T with respect to a basis $\{e_i\}$ is defined by $Te_j = T^i_{\ j} e_i$, and the trace of T is set to be $\operatorname{tr} \mathsf{T} = T^i_{\ i}$. If $\{e'_j\}$ is a second basis, where $e_j = A^i_{\ j} e'_j$, the matrix of T is (see Eq. (4.5))

$$\mathsf{T}' = \mathsf{ATA}^{-1},$$

whence $\operatorname{tr} \mathsf{T}' = \mathsf{T}$.

(d) The characteristic polynomial $f(z) = a_0 + a_1 z + a_2 z^2 \cdots + \cdots + (-1)^n z^n$ is given by

$$f(z) = \det \begin{vmatrix} T^1_{\ 1} - z & T^1_{\ 2} & \dots & T^1_{\ n} \\ & T^2_{\ 2} - z & & \\ & & \ddots & \\ & & & T^n_{\ n} - z \end{vmatrix}.$$

Hence the coefficient of z^{n-1} is $a_{n-1} = (-1)^{n-1}(T^1_{\ 1} + T^2_{\ 2} + \cdots + T^n_{\ n})$; i.e. $\operatorname{tr} T = (-1)^{n-1} a_{n-1}$.

(e) By the fundamental theorem of algebra we have n complex roots of $f(z)$,

$$f(z) = (z_1 - z)(z_2 - z) \dots (z_n - z).$$

The coefficient of z^{n-1} is $a_{n-1} = (-1)^{-n-1}(z_1 + z_2 + \cdots + z_n$. If the eigenvalues of T are $\lambda_1, \dots \lambda_k$, with multiplicities $p_1, \dots p_k$, these values run through the roots $z_1, z_2, \dots, z_n$, where the eigenvlaue λ_i is repeated p_i times. Hence

$$\sum_{i=1}^{m} p_i \lambda_i = \sum_{j=1}^{n} z_j = (-1)^{n-1} a_{n-1} = \operatorname{tr} T$$

by part (d).

Problem 4.2 On a complex vector space V let S and T be two commuting operators, $ST = TS$.

(a) Show that if v is an eigenvector of T then so is Sv.

(b) Show that a basis for V can be found such that the matrices of both S and T with respect to this basis are in upper triangular form.
[NOTE: vector space must be complex]

Solution: (a) If $Tv = \lambda v$ and $ST = TS$ then

$$T(Sv) = TSv = STv = S(\lambda v) = \lambda Sv.$$

(b) Let λ_1 be an eigenvalue of T and V_1 the corresponding eigenspace consisting of all vectors v such that

$$Tv = \lambda_1 v.$$

The subspace V_1 is invariant subspace of T and also invariant under the action of S by part (a), i.e. $SV_1 \subseteq V_1$. If μ_1 is an eigenvalue of $S\big|_{V_1}$, there exists a corresponding eigenvector $u_1 \in V_1$ of S,

$$Su_1 = \mu_1 u_1, \qquad Tu_1 = \lambda_1 u_1.$$

The operators T and S both have a natural action on the quotient space $V/L(u_1)$ (where $L(u_1) = \{au_1 \mid a \in \mathbb{C}\}$), defined by

$$T\big(v + L(u_1)\big) = Tv + L(u_1)$$
$$S\big(v + L(u_1)\big) = Sv + L(u_1)$$

since $T\big(L(u_1)\big) \subseteq L(u_1)$ and $S\big(L(u_1)\big) \subseteq L(u_1)$. Thus, there exists a common eigenvector

$$T\big(u_2 + L(u_1)\big) = \lambda_2\big(u_2 + L(u_1)\big)$$
$$S\big(u_2 + L(u_1)\big) = \mu_2\big(u_2 + L(u_1)\big)$$

i.e.

$$Tu_2 = \lambda_2 u_2 + T^1_2 u_1$$
$$Su_2 = \mu_2 u_2 + S^1_2 u_1$$

It follows that $L(u_1, u_2)$ is a subspace which is invariant under both operators T and S: $TL(u_1, u_2) \subseteq L(u_1, u_2)$ and $SL(u_1, u_2) \subseteq L(u_1, u_2)$. Continue inductively. At the kth step we consider the action of T and S on $V/L(u_1, \dots, u_{k-1})$. There is a common eigenvector

$$T\big(u_k + L(u_1, \dots, u_{k-1})\big) = \lambda_k\big(u_k + L(u_1, \dots, u_{k-1})\big)$$
$$S\big(u_k + L(u_1, \dots, u_{k-1})\big) = \mu_k\big(u_k + L(u_1, \dots, u_{k-1})\big)$$

so that

$$Tu_k = \sum_{i=1}^{k} T^i_k u_i, \qquad Su_k = \sum_{i=1}^{k} S^i_k u_i.$$

This says that $T^i_k = S^i_k = 0$ for all $k < i \le n$; i.e. in the basis $u_1, \dots, u_n$ the matrices T and S are both upper triangular.

Problem 4.3 **For the operator $T : V \to V$ on a four-dimensional vector space given in Problem 3.10, show that no basis exists such that the matrix of T is diagonal. Find a basis in which the matrix of T has the Jordan form**

$$\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & \lambda & 1 \\ 0 & 0 & 0 & \lambda \end{pmatrix}$$

for some λ, and calculate the value of λ.

Solution: The operator T in Problem 3.10 is given by

$$\begin{aligned} Te_1 &= 2e_1 - e_4 \\ Te_2 &= -2e_1 + e_4 \\ Te_3 &= -2e_1 + e_4 \\ Te_4 &= e_1 \end{aligned}$$

and has matrix

$$\mathsf{T} = \begin{pmatrix} 2 & -2 & -2 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 1 & 1 & 0 \end{pmatrix}.$$

$u = ae_1 + be_2 + ce_3 + de_4$ is an eignevector , $Tu = \lambda u$, if

$$\begin{aligned} 2a - 2b - 2c + d &= \lambda a \\ 0 &= \lambda b \\ 0 &= \lambda c \\ -a + b + c &= \lambda d \end{aligned}$$

If $\lambda = 0$ then $d = 0$ and $a = b + c$. As shown in Problem 3.10 there are are two independent eigenvectors (spanning $\ker T$),

$$f_1 = e_1 + e_2, \qquad f_2 = e_1 + e_3.$$

For $\lambda \neq 0$ we must have $b = c = 0$ and

$$\begin{aligned} 2a + d &= \lambda a \\ -a &= \lambda d \end{aligned}$$

giving rise to the eigenvalue euqation

$$d(\lambda^2 - 2\lambda + 1) = 0$$

so that $\lambda = 1$ with unique independent eigenvector $f_3 = e_1 - e_4$. All eigenvectors of T are therefore of the form $af_1 + bf_2$ or cf_3, which span at most a 3-dimensional subspace of T. Hence the four-dimensional space V cannot be spanned by eigenvectors T, which cannot therefore be diagonalized (since all basis vectors in a diagonal representation are eigenvectors). Since $\lambda = 1$ has multiplicity 2, we seek a vector f_4 such that

$$Tf_4 = \lambda f_4 + f_3 = f_4 + f_3.$$

Setting $f_4 = ae_1 + be_2 + ce_3 + de_4$ we have

$$\begin{aligned} 2a - 2b - 2c + d &= a + 1 \\ 0 &= b \\ 0 &= c \\ -a + b + c &= d - 1, \end{aligned}$$

giving $b = c = 0$, $a = 1 - d$. We may therefore set $f_4 = e_4$ (taking, for example, $a = 0$) and in the basis f_1, f_2, f_3, f_4 the matrix of T has the Jordan form

$$\mathsf{T} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Problem 4.4 **Let S be the matrix**

$$\mathsf{S} = \begin{pmatrix} i-1 & 1 & 0 & 0 \\ -1 & 1+i & 0 & 0 \\ -1-2i & 2i & -i & 1 \\ 2i-1 & 1 & 0 & -i \end{pmatrix}.$$

Find the minimal annihilating polynomial and the characteristic polynomial of this matrix, its eigenvalues and eigenvectors, and find a basis that reduces it to its Jordan canonical form.

Solution: The characteristic polynomial is

$$f(z) = \det \begin{vmatrix} -1+i-z & 1 & 0 & 0 \\ -1 & 1+i-z & 0 & 0 \\ -1-2i & 2i & -i-z & 1 \\ -1+2i & 1 & 0 & -i-z \end{vmatrix} = z^4 + 2z^2 + 1.$$

Factorization gives $f(z) = (z-i)^2(z+i)^2$, and the eigenvalues are $z = \pm 1$, both with multiplicity 2. To find the minimal annihilating polynomial we compute powers of S,

$$\mathsf{S} = \begin{pmatrix} -1+i & 1 & 0 & 0 \\ -1 & 1+i & 0 & 0 \\ -1-2i & 2i & -i & 1 \\ -1+2i & 1 & 0 & -i \end{pmatrix} \qquad \mathsf{S}^2 = \begin{pmatrix} -1-2i & 2i & 0 & 0 \\ -2i & -1+2i & 0 & 0 \\ 2i & 0 & -1 & -2i \\ -2i & 2i & 0 & -1 \end{pmatrix}$$

$$\mathsf{S}^3 = \begin{pmatrix} 3-i & -3 & 0 & 0 \\ 3 & -3-i & 0 & 0 \\ 3+2i & -2i & i & -3 \\ 3-2i & -3 & 0 & i \end{pmatrix} \qquad \mathsf{S}^4 = \begin{pmatrix} 1+4i & -4i & 0 & 0 \\ 4i & 1-4i & 0 & 0 \\ -4i & 0 & 1 & 4i \\ 4i & -4i & 0 & 1 \end{pmatrix}$$

Component by component, a polynomial equation of the form

$$\Delta(\mathsf{A}) = a\mathsf{I} + b\mathsf{S} + c\mathsf{S}^2 + d\mathsf{S}^3 + e\mathsf{S}^4 = \mathsf{O}$$

gives, after some manipulation, the equations

$$c = a + e, \qquad b = d, \qquad b = i(c - 2e), \qquad b = -i(c - 2e)$$

whence $b = d = c - 2e = 0$ and $c = a + e$. Setting $a = 1$ gives the minimal annihilating polynomial $\Delta(z) = 1 + 2z^2 + z^4 = f(z)$. Since the minimal annihilating polynomial and characteristic polynomial are identical the multiple eigenvalues $\lambda = \pm i$ can only have one independent eigenvector each. Solving for eigenvectors gives

$$\mathsf{S}\mathbf{f}_1 = i\mathbf{f}_1, \qquad \mathsf{S}\mathbf{f}_3 = -i\mathbf{f}_3.$$

where

$$\mathbf{f}_1 = \begin{pmatrix} 1 & 1 & 0 & 1 \end{pmatrix}, \qquad \mathbf{f}_3 = \begin{pmatrix} 0 & 0 & 1 & 0 \end{pmatrix}$$

and solving for $\mathsf{S}\mathbf{f}_2 = i\mathbf{f}_2 + \mathbf{f}_1$ and $\mathsf{S}\mathbf{f}_4 = i\mathbf{f}_4 + \mathbf{f}_3$ gives

$$\mathbf{f}_2 = \begin{pmatrix} 0 & 1 & 1 & 0 \end{pmatrix}, \qquad \mathbf{f}_4 = \begin{pmatrix} 0 & 0 & 0 & 1 \end{pmatrix}.$$

In the basis $\mathbf{f}_1$, $\mathbf{f}_2$, $\mathbf{f}_3$, $\mathbf{f}_4$ we have the Jordan form

$$\mathsf{S}' = \mathsf{A}\mathsf{S}\mathsf{A}^{-1} = \begin{pmatrix} i & 1 & 0 & 0 \\ 0 & i & 0 & 0 \\ 0 & 0 & -i & 1 \\ 0 & 0 & 0 & -i \end{pmatrix}.$$

Problem 4.5 **Verify that (4.30), (4.31) is a solution of $\dot{\mathbf{x}}_r = \mathsf{A}\mathbf{x}_r(t)$ provided**

$$\begin{aligned}\mathsf{A}\mathbf{h}_1 &= \lambda_i \mathbf{h}_1\\ \mathsf{A}\mathbf{h}_2 &= \lambda_i \mathbf{h}_2 + \mathbf{h}_1\\ &\vdots\\ \mathsf{A}\mathbf{h}_r &= \lambda_i \mathbf{h}_r + \mathbf{h}_{r-1}\end{aligned}$$

where λ_i is an eigenvalue of A.

Solution: Let

$$\mathbf{x}_r(t) = \mathbf{w}_r(t)\mathrm{e}^{\lambda_i t}$$

where

$$\mathbf{w}_r(t) = \frac{t^{r-1}}{(r-1)!}\mathbf{h}_1 + \frac{t^{r-2}}{(r-2)!}\mathbf{h}_2 + \cdots + t\mathbf{h}_{r-1} + \mathbf{h}_r .$$

Then

$$\dot{\mathbf{x}}_r(t) = \mathrm{e}^{\lambda_i t}\big(\lambda_i \mathbf{w}_r(t) + \dot{\mathbf{w}}_r(t)\big)$$

and the equation $\dot{\mathbf{x}}_r = \mathsf{A}\mathbf{x}_r(t)$ reads

$$\begin{aligned}&\mathsf{A}\Big(\frac{t^{r-1}}{(r-1)!}\mathbf{h}_1 + \frac{t^{r-2a}}{(r-2)!}\mathbf{h}_2 + \cdots + t\mathbf{h}_{r-1} + \mathbf{h}_r\Big)\\ &\quad = \lambda_i\Big(\frac{t^{r-1}}{(r-1)!}\mathbf{h}_1 + \frac{t^{r-2}}{(r-2)!}\mathbf{h}_2 + \cdots + t\mathbf{h}_{r-1} + \mathbf{h}_r\Big)\\ &\qquad + \frac{t^{r-2}}{(r-2)!}\mathbf{h}_1 + \frac{t^{r-3}}{(r-3)!}\mathbf{h}_2 + \cdots + \mathbf{h}_{r-1}\end{aligned}$$

The coefficients of t^{r-1}, t^{r-2}, $\ldots$, $t^{r-k}, \ldots$, t^0 give respectively

$$\begin{aligned}\mathsf{A}\mathbf{h}_1 &= \lambda_i \mathbf{h}_1\\ \mathsf{A}\mathbf{h}_2 &= \lambda_i \mathbf{h}_2 + \mathbf{h}_1\\ &\vdots\\ \mathsf{A}\mathbf{h}_k &= \lambda_i \mathbf{h}_k + \mathbf{h}_{k-1}\\ &\vdots\\ \mathsf{A}\mathbf{h}_r &= \lambda_i \mathbf{h}_r + \mathbf{h}_{r-1} \ .\end{aligned}$$

Problem 4.6 **Discuss the remaining cases for two-dimensional autonomous systems: (a) $\lambda_1 = \lambda_2 = \lambda \neq 0$ and (i) two distinct eigenvectors**

$\mathbf{h}_1$ and $\mathbf{h}_2$, (ii) only one eigenvector $\mathbf{h}_1$; (b) A a singular matrix. Sketch the solutions in all instances.

Solution: (a) $\lambda_1 = \lambda_2 = \lambda$. Case (i): There are two distinct eigenvectors $\mathbf{h}_1$ and $\mathbf{h}_2$. The general solution is

$$\mathbf{x} = c_1 e^{\lambda t}\mathbf{h}_1 + c_2 e^{\lambda t}\mathbf{h}_2,$$

representing straight lines through the origin. If $\lambda < 0$ the solution approaches the critical point at the origin, while if $\lambda > 0$ the solution expands away from the origin.

Case (ii): $\lambda_1 = \lambda_2 = \lambda \neq 0$, but there exists only one eigenvector $\mathbf{h}_1$. There then exists a vector $\mathbf{h}_2$ such that

$$\mathsf{A}\mathbf{h}_2 = \lambda\mathbf{h}_2 + \mathbf{h}_1$$

and the general solution is (see Eqs. (4.30), (4.31) and the previous problem)

$$\begin{aligned}\mathbf{x} &= c_1 e^{\lambda t}\mathbf{h}_1 + c_2 e^{\lambda t}(t\mathbf{h}_1 + \mathbf{h}_2).\\ &= (c_1 + c_2 t)e^{\lambda t}\mathbf{h}_1 + c_2 e^{\lambda t}\mathbf{h}_2.\end{aligned}$$

If $\lambda < 0$ the solutions approach the origin tangentially to the $\mathbf{h}_1$-direction, while if $\lambda > 0$ they depart from the origin asymptotically to this direction.

(b) If A is a singular matrix then one eigenvalue vanishes, say $\lambda_2 = 0$.
Case(i): $\lambda_1 \neq 0$. The solution has the form

$$\mathbf{x} = c_1 e^{\lambda_1 t}\mathbf{h}_1 + c_2\mathbf{h}_2.$$

If $\lambda_1 < 0$ the line $\mathbf{x} = c\mathbf{h}_2$ is a line of stable critical points (solutions approach this line as $t \to \infty$), while if $\lambda_1 > 0$ it is a line of unstable critical points (solutions diverge from this line).
Case(ii): $\lambda_1 = \lambda_2 = 0$. If there are two independent eignvectors then the matrix $\mathsf{A} = \mathsf{O}$ and the equation is trivial (all solutions are constant and all points are critical points). Assume therefore that there is a single eigenvector $\mathbf{h}_1$. There then exists a vector $\mathbf{h}_2$ such that

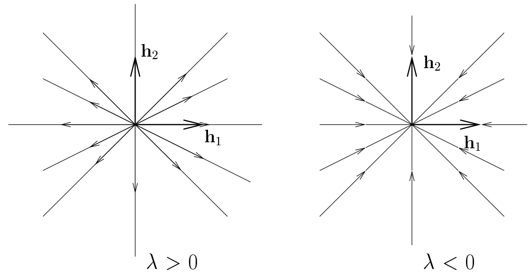
$$\mathsf{A}\mathbf{h}_1 = 0, \qquad \mathsf{A}\mathbf{h}_2 = \mathbf{h}_1$$

and the solution of the differential equation is

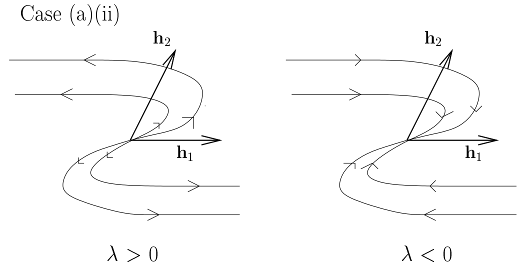
$$\mathbf{x} = c_1\mathbf{h}_1 + c_2(t\mathbf{h}_1 + \mathbf{h}_2)$$

All points along the $\mathbf{h}_1$-axis are critical points, and the general solution is a straight line parallel to this axis.

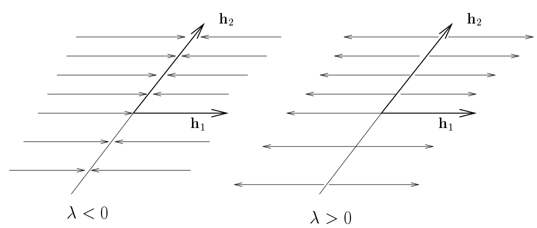
Case (a)(i)



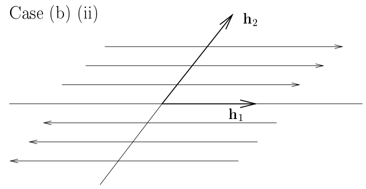
Case (a)(ii)



Case (b) (i)



Case (b) (ii)



Problem 4.7 Classify all three-dimensional autonomous systems of linear differential equations having constant coefficients.

Solution: Case (A): The matrix A is non-singular. In this instance the only critical point is $\mathbf{x}_0 = \mathbf{0}$ and all eigenvalues are $\neq 0$.
(1) Suppose all eigenvalues are real and unequal, $\lambda_1 < \lambda_2 < \lambda_3$: The eigenvectors $\mathbf{h}_1$, $\mathbf{h}_2$ and $\mathbf{h}_3$ form a basis of $\mathbb{R}^3$ and the general solution is

$$\mathbf{x} = c_1 e^{\lambda_1 t}\mathbf{h}_1 + c_2 e^{\lambda_2 t}\mathbf{h}_2 + c_3 e^{\lambda_3 t}\mathbf{h}_3.$$

We may assume all $c_i \neq 0$ else it reduces to the two-dimensional case in the text.
Case (1a) $\lambda_1 < \lambda_2 < \lambda_3 < 0$: The critical point $\mathbf{x} = \mathbf{0}$ is a stable node. All solutions approach it as $t \to \infty$ asymptotically to the $\mathbf{h}_3$-axis.
Case (1b) $\lambda_1 < 0$, $\lambda_3 > 0$: $\mathbf{x} = \mathbf{0}$ is a saddle point, with solutions asymptoting to the $\mathbf{h}_1$-axis as $t \to -\infty$ and to the $\mathbf{h}_1$-axis as $t \to \infty$.
Case (1c) $0 < \lambda_1 < \lambda_2 < \lambda_3$: The critical point $\mathbf{x} = \mathbf{0}$ is an unstable node asymptoting to the $\mathbf{h}_1$-axis as $t \to -\infty$.

(2) λ_3 real and $\lambda_1 = \lambda$, $\lambda_2 = \overline{\lambda}$ where λ is complex: As in the text the arbitrary real solution is can be written in the form

$$\mathbf{x} = Re^{\mu t}\big(\cos(\nu t + \alpha)\mathbf{h}_1 + \sin(\nu t + \alpha)\mathbf{h}_2\big) + c_3 e^{\lambda_3 t}\mathbf{h}_3.$$

If $\lambda_3 < 0$ the solutions spiral towards the 2-dimensional cases in the plane $x^3 = 0$. If $\lambda_3 > 0$ the same solutions are stretched in spirals around the $\mathbf{h}_3$-axis towards $x^3 \to \pm\infty$.

(3) $\lambda_1 = \lambda_2 = \lambda \neq \lambda_3$: If the eigenvalue λ has two distinct eigenvectors the solutions can be written in the form

$$\mathbf{x} = e^{\lambda t}(c_1\mathbf{h}_1 + c_2\mathbf{h}_2) + c_3 e^{\lambda_3 t}\mathbf{h}_3.$$

and are essentially two-dimensional in character, lying in the plane spanned by vectors $c_1\mathbf{h}_1 + c_2\mathbf{h}_2$ and $\mathbf{h}_3$.
If the eigenvalue λ has only one eigenvector the solutions can be written in the form

$$\mathbf{x} = e^{\lambda t}\big((c_1 + c_2 t)\mathbf{h}_1 + c_2\mathbf{h}_2\big) + c_3 e^{\lambda_3 t}\mathbf{h}_3.$$

Stability and instabilty is determined by the relative signs and magnitudes of λ and λ_3.

(4) $\lambda_1 = \lambda_2 = \lambda_3 = \lambda \neq 0$: The case of 3, 2 and 1 independent eigenvectors are given repsectively by

$$\mathbf{x} = e^{\lambda t}(c_1\mathbf{h}_1 + c_2\mathbf{h}_2 + c_3\mathbf{h}_3)$$
$$\mathbf{x} = e^{\lambda t}\big((c_1 + c_2 t)\mathbf{h}_1 + c_2\mathbf{h}_2 + c_3\mathbf{h}_3\big)$$
$$\mathbf{x} = e^{\lambda t}\Big((c_1 + c_2 t + c_3\frac{t^2}{2})\mathbf{h}_1 + (c_2 c_3 t)\mathbf{h}_2 + c_3\mathbf{h}_3\Big)$$

The first solution is a straight line in the direction $c_1\mathbf{h}_1 + c_2\mathbf{h}_2 + c_3\mathbf{h}_3$.

Case (B): If A is singular then at least one eigenvalue is zero, say $\lambda_3 = 0$.
(1) If $\lambda_1 \neq \lambda_2 \neq 0$ or $\lambda_1 = \lambda$, $\lambda_2 = \overline{\lambda}$, the treatment is almost identical with cases A(1) and A(2) above, with λ_3 set to zero. The only adjustment is that any asymptotic behaviour in the $\mathbf{h}_3$ direction is replaced by the fact that all solutions lie in a plane $x^3 = c_3 = \text{const}$.
(2) $\lambda_2 = 0$, $\lambda_1 \neq 0$: there are two cases, depending on whether there are two or one eigenvector having eigenvalue $\lambda = 0$. The case $\lambda_1 = 0$, $\lambda_2 \neq 0$ is similar.
(3) $\lambda_1 = \lambda_2 = \lambda_3 = 0$: The case of 3 eigenvectors is trivial, while 2 and 1 eigenvector is essentially given in A(4) with $\lambda = 0$.

Chapter 5

Problem 5.1 **Let $(V, \cdot)$ be a real Euclidean inner product space and denote the length of a vector $x \in V$ by $|x| = \sqrt{x \cdot x}$. Show that two vectors u and v are orthogonal iff $|u + v|^2 = |u|^2 + |v|^2$.**

Solution:
$$|u + v|^2 = (u + v) \cdot (u + v) = u \cdot u + v \cdot v + 2u \cdot v = |u|^2 + |v|^2$$
if and only if $u \cdot v = 0$.

Problem 5.2 **Let**

$$\mathsf{G} = [g_{ij}] = [u_i \cdot u_j] = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & -1 \\ 0 & -1 & 1 \end{pmatrix}$$

be the components of a real inner product with respect to a basis u_1, u_2, u_3. Use Gram-Schmidt orthogonalization to find an orthonormal basis e_1, e_2, e_3, expressed in terms of the vectors u_i, and find the index of this inner product.

Solution: Both u_1 and u_2 are null vectors, therefore the process must be started off with a non-null combination such as $f_1 = u_1 + u_2$. Then $f_1 \cdot f_1 = 2u_1 \cdot u_2 = 2$, and normalizing gives

$$e_1 = \frac{1}{\sqrt{2}} f_1 = \frac{1}{\sqrt{2}}(u_1 + u_2), \qquad e_1 \cdot e_1 = 1.$$

A vector which is easily seen to be orthogonal to f_1 is $f_2 = u_1 - u_2$, having magnitude $f_2 \cdot f_2 = -2$. Again, normalizing gives

$$e_2 = \frac{1}{\sqrt{2}} f_2 \quad \Longrightarrow \quad e_2 \cdot e_2 = -1, \quad e_1 \cdot e_2 = 0.$$

A third vector orthogonal to e_1 and e_2 must have a component in the u_3 direction, so try a vector of the form $f_3 = u_3 + au_1 + bu_2$. Then

$$f_3 \cdot u_1 = b, \qquad f_3 \cdot u_2 = a - 1$$

and

$$f_3 \cdot e_1 = f_3 \cdot e_2 = 0 \quad \Longrightarrow \quad f_3 \cdot u_1 = f_3 \cdot u_2 = 0 \quad \Longrightarrow \quad b = 0, \quad a = 1.$$

Hence $f_3 = u_1 + u_3$, and $f_3 \cdot f_3 = 1$. Thus we may take $e_3 = f_3$ and an orthonormal basis is

$$\begin{aligned} e_1 &= \frac{1}{\sqrt{2}}(u_1 + u_2) \\ e_2 &= \frac{1}{\sqrt{2}}(u_1 - u_2) \\ e_3 &= u_1 + u_3, \end{aligned}$$

giving

$$G' = [e_i \cdot e_j] = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

The index is $2 - 1 = 1$.

Problem 5.3 **Let G be the symmetric matrix of components of a real inner product with respect to a basis u_1, u_2, u_3**

$$\mathsf{G} = [g_{ij}] = [u_i \cdot u_j] = \begin{pmatrix} 1 & 0 & 1 \\ 0 & -2 & 1 \\ 1 & 1 & 0 \end{pmatrix}.$$

Using Gram–Schmidt orthogonalization, find an orthonormal bases e_1, e_2, e_3 expressed in terms of the vectors u_i.

Solution: We may set $e_1 = u_1$ since

$$e_1 \cdot e_1 = u_1 \cdot u_1 = g_{11} = 1.$$

Since $u_2 \cdot u_2 = -2$ and $u_2 \cdot u_1 = 0$, we may set

$$e_2 = \frac{1}{\sqrt{2}} u_2 \quad \Longrightarrow \quad e_2 \cdot e_2 = -1, \quad e_2 \cdot e_1 = 0.$$

For the third vector, try $v = u_3 + ae_1 + be_2$, so that

$$v \cdot e_1 = a + u_3 \cdot e_1 = a + 1$$
$$v \cdot e_2 = -b + u_3 \cdot e_2 = -b + \frac{1}{\sqrt{2}}$$

Hence, set $a = -1$, $b = 1/\sqrt{2}$, so that $v = u_3 - u_1 + \frac{1}{2}u_2$. Then

$$v \cdot v = u_3 \cdot u_3 + u_1 \cdot u_1 + \tfrac{1}{4}u_2 \cdot u_2 - 2u_3 \cdot u_1 + u_3 \cdot u_2 - u_1 \cdot u_2$$
$$= 0 + 1 - \tfrac{1}{2} - 2 + 1 - 0$$
$$= -\frac{1}{2}$$

Hence, on normalizing,

$$e_3 = \sqrt{2}v = \sqrt{2}(u_3 - u_1 + \frac{1}{2}u_2), \qquad e_3 \cdot e_3 = -1.$$

and $e_3 \cdot e_1 = e_3 \cdot e_2 = 0$. With respect to the $\{e_i\}$ basis

$$G' = [e_i \cdot e_j] = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

The index is $1 - 2 = -1$.

Problem 5.4 **Define the concept of a 'symmetric operator' $S : V \to V$ as one that satisfies**

$$(Su) \cdot v = u \cdot (Sv) \qquad \textbf{for all} \qquad u, v \in V.$$

Show that this results in the component equation

$$S^k_i g_{kj} = g_{ik} S^k_j,$$

equivalent to the matrix equation

$$\mathsf{S}^T \mathsf{G} = \mathsf{G}\mathsf{S}.$$

Show that for an orthonormal basis in a Euclidean space this results in the usual notion of symmetry, but fails for pseudo-Euclidean spaces.

Solution: Setting $u = e_i$, $v = e_j$ in the definition of a symmetric operator gives

$$(Se_i) \cdot e_j = e_i \cdot (Se_j).$$

The definition of operator components $Se_i = S^k_i e_k$ gives

$$S^k_i e_k \cdot e_j = e_i \cdot S^k_j e_k$$

and substituting $g_{kj} = e_k \cdot e_j$ results in the required

$$S^k_i g_{kj} = g_{ik} S^k_j.$$

In terms of the matrices $\mathsf{S} = [S^k_i]$ and $\mathsf{G} = [g_{ij}]$ the standard rule for matrix multiplication gives

$$\mathsf{S}^T \mathsf{G} = \mathsf{G}\mathsf{S}.$$

In an orthonormal basis for a Euclidean inner product we have $g_{ij} = e_i \cdot e_j = \delta_{ij}$, or equivalently $\mathsf{G} = \mathsf{I}$. The matrix equation then reads $\mathsf{S}^T = \mathsf{S}$, which is the usual concept of a symmetric matrix.
For a pseudo-Euclidean space $g_{ij} = \eta_i \delta_{ij}$ where $\eta_i = \pm 1$, and $S^j{}_i \eta_j = S^i{}_j \eta_i$ (summation convention temporarily suspended here). Hence in this case, $S^j_i = \pm S^i{}_j$, depending on the sign of $\eta_i \eta_j$.

Problem 5.5 **Let V be a Minkowskian vector space of dimension n with index $n - 2$ and let $k \neq 0$ be a null vector ($k \cdot k = 0$) in V.**

(a) Show that there is an orthonormal basis $e_1, \dots, e_n$ such that

$$k = e_1 - e_n.$$

(b) Show that if u is a 'time-like' vector, defined as a vector with negative magnitude $u \cdot u < 0$, then u is not orthogonal to k.

(c) Show that if v is a null vector such that $v \cdot k = 0$, then $v \propto k$.

(d) If $n \geq 4$ which of these statements generalize to a space of index $n-4$?

Solution: (a) If $f_1, \dots f_n$ is any o.n. basis, such that $f_i \cdot f_j = \delta_{ij}$ for $i, j = 1, \dots, n-1$ and $f_n \cdot f_n = -1$ then let $k = k^i f_i$. The null vector requirement is

$$k \cdot k = 0 = \sum_{i=1}^{n-1} (k^i)^2 \; - \; (k^n)^2.$$

By Gram-Schmidt orthonormalization in a Euclidean space, a rotation $e_i = A^j{}_i f_j$ $(i, j, = 1, \dots, n-1)$ of the first $n-1$ vectors can always be performed such that $e_1 \propto \sum_{i=1}^{n-1} k^i f_i$. Then there exists a real number α such that $k = \alpha(e_1 - e_n)$. To set $\alpha = 1$ we try a further transformation of the form

$$e_1 = a e'_1 + b e'_n$$
$$e_2 = c e'_1 + d e'_n$$

so that

$$k = \alpha\big((a - c)e'_1 + (b - d)e'_n\big).$$

Since it is required that $e_1 \cdot e_1 = 1$, $e_n \cdot e_n = -1$, and $e_1 \cdot e_n = 0$, we have

$$a^2 - b^2 = 1, \qquad c^2 - d^2 = -1, \qquad ac - bd = 0.$$

Further since we want $k = e'_1 - e'_n$ we must have

$$a - c = d - b = \frac{1}{\alpha}, \quad \Longrightarrow \quad c = a - \frac{1}{\alpha}, \quad d = b + \frac{1}{\alpha}$$

and combining these equations gives, after a little algebra,

$$a = \frac{1}{2}\big(\alpha + \frac{1}{\alpha}\big), \quad b = \frac{1}{2}\big(\alpha - \frac{1}{\alpha}\big), \qquad c = b, \quad d = a.$$

We now have the desired result $k = e_1 - e_n$.

(b) In a basis derived in (a) where $k = e_1 - e_n$, a vector u is orthogonal to k, $u \cdot k = 0$, if and only if $u^1 + u^n = 0$. Hence, as $u^1 = -u^n$,

$$u \cdot u = \sum_{i=1}^{n-1} (u^i)^2 \; - \; (u^n)^2 = \sum_{i=2}^{n-1} (u^i)^2 \geq 0.$$

Thus u cannot be timelike—it cannot hold that $u \cdot u < 0$.

(c) Similarly in a basis if v is a null vector orthogonal to k, then in the special basis of part (a), $v^1 = -v^n$, and the magnitude of v is

$$v \cdot v = \sum_{i=2}^{n-1} (v^i)^2 = 0$$

iff $v^2 = v^3 = \cdots = v^{n-1} = 0$. Thus v is a null vector orthogonal to k iff $v = v^1(e_1 - e_n) \propto k$.

(d) If the metric has index $n-4$ then there exists an orthonormal basis such that $e_i \cdot e_j = \delta_{ij}$ for $i, j = 1, \ldots, n-2$, and $e_{n-1} \cdot e_{n-1} = e_n \cdot e_n = -1$, $e_{n-1} \cdot e_n = 0$:

$$\mathsf{G} = [e_i \cdot e_j] = \begin{pmatrix} 1 & 0 & 0 & \ldots \\ 0 & 1 & 0 & \ldots \\ \ldots & \ldots & \ldots & \ldots \\ \ldots & 0 & -1 & 0 \\ \ldots & 0 & 0 & -1 \end{pmatrix}.$$

Part (a) still holds, but we may need an extra rotation in the $e_{n-1} - e_n$ plane.
Part(b) is no longer true. For example, in $n = 4$, the null vector $k = e_1 - e_4$ is orthogonal to the time-like vector $u = e_3$.
Part (c) also does not hold as $k = e_1 - e_4$ and $v = e_2 - e_3$ are both null vectors and orthogonal to each other, $v \cdot k = 0$.

Problem 5.6 Show that the norm defined by an inner product satisfies the parallelogram law.

$$\|u+v\|^2 + \|u-v\|^2 = 2\|u\|^2 + 2\|v\|^2.$$

Solution:

$$\begin{aligned} \|u+v\|^2 + \|u-v\|^2 &= \langle u+v \,|\, u+v\rangle + \langle u-v \,|\, u-v\rangle \\ &= \langle u \,|\, u\rangle + \langle v \,|\, u\rangle + \langle u \,|\, v\rangle + \langle v \,|\, v\rangle \\ &\quad + \langle u \,|\, u\rangle - \langle v \,|\, u\rangle - \langle u \,|\, v\rangle + \langle v \,|\, v\rangle \\ &= 2\|u\|^2 + 2\|v\|^2. \end{aligned}$$

Problem 5.7 On an inner product space show that

$$4\langle u \,|\, v\rangle = \|u+v\|^2 - \|u-v\|^2 - i\|u+iv\|^2 + i\|u-iv\|^2.$$

Hence show that a linear transformation $U : V \to V$ is unitary iff it is norm preserving,

$$\langle Uu \,|\, Uv\rangle = \langle u \,|\, v\rangle, \quad \forall\, u, v \in V \quad \iff \quad \|Uv\| = \|v\|, \quad \forall\, v \in V.$$

Solution: Using a calculation similar to that in Problem 5.6 we have

$$\begin{aligned}
\|u+v\|^2 &= \|u\|^2 + \|v\|^2 + \langle u \,|\, v\rangle + \langle v \,|\, u\rangle \\
\|u-v\|^2 &= \|u\|^2 + \|v\|^2 - \langle u \,|\, v\rangle - \langle v \,|\, u\rangle \\
\|u+iv\|^2 &= \|u\|^2 + \|v\|^2 + \langle u \,|\, iv\rangle + \langle iv \,|\, u\rangle \\
&= \|u\|^2 + \|v\|^2 + i\langle u \,|\, v\rangle - i\langle v \,|\, u\rangle \\
\|u-iv\|^2 &= \|u\|^2 + \|v\|^2 - i\langle u \,|\, v\rangle + i\langle v \,|\, u\rangle
\end{aligned}$$

Hence

$$\begin{aligned}
&\|u+v\|^2 - \|u-v\|^2 - i\|u+iv\|^2 + i\|u-iv\|^2 \\
&\qquad = (1-1-i+i)(\|u\|^2 + \|v\|^2) + (1+1+1+1)\langle u \,|\, v\rangle + (1+1-1-1)\langle v \,|\, u\rangle \\
&\qquad = 4\langle u \,|\, v\rangle.
\end{aligned}$$

From Eq. (5.19) a linear transformation U is unitary if $\langle Uu \,|\, Uv\rangle = \langle u \,|\, v\rangle$ for all $u, v \in V$. Setting $u = v$ this implies it is norm preserving,

$$\|Uv\| = \sqrt{\langle Uv \,|\, Uv\rangle} = \sqrt{\langle v \,|\, v\rangle} = \|v\|.$$

Conversely, if it is norm preserving, then using the above identity and linearity of U

$$\begin{aligned}
4\langle Uu \,|\, Uv\rangle &= \|U(u+v)\|^2 - \|U(u-v)\|^2 - i\|U(u+iv)\|^2 + i\|U(u-iv)\|^2 \\
&= \|u+v\|^2 - \|u-v\|^2 - i\|u+iv\|^2 + i\|u-iv\|^2 \\
&= 4\langle u \,|\, v\rangle.
\end{aligned}$$

Problem 5.8 Show that a pair of vectors u and v in a complex inner product space are orthogonal iff

$$\|\alpha u + \beta v\|^2 = \|\alpha u\|^2 + \|\beta v\|^2, \qquad \forall\, \alpha, \beta \in \mathbb{C}.$$

Find a non-orthogonal pair of vectors u and v in a complex inner product space such that $\|u+v\|^2 = \|u\|^2 + \|v\|^2$.

Solution:

$$\begin{aligned}
\|\alpha u + \beta v\|^2 &= \langle \alpha u + \beta v \,|\, \alpha u + \beta v\rangle \\
&= \langle \alpha u \,|\, \alpha u\rangle + \langle \beta v \,|\, \beta v\rangle + \langle \alpha u \,|\, \beta v\rangle + \langle \beta v \,|\, \alpha u\rangle \\
&= \|\alpha u\|^2 + \|\beta v\|^2 + \overline{\alpha}\beta\langle u \,|\, v\rangle + \overline{\beta}\alpha\langle v \,|\, u\rangle.
\end{aligned}$$

Hence, if $\langle u \,|\, v\rangle = 0$ then $\langle v \,|\, u\rangle = 0$ and for all $\alpha, \beta \in \mathbb{C}$

$$\|\alpha u + \beta v\|^2 = \|\alpha u\|^2 + \|\beta v\|^2.$$

Conversely, set $\alpha = \beta = 1$ in the identity gives

$$\|u+v\|^2 = \|u\|^2 + \|v\|^2 + \langle u \,|\, v\rangle + \langle v \,|\, u\rangle = \|u\|^2 + \|v\|^2$$

while setting $\alpha = 1$, $\beta = i$ gives

$$\|u + iv\|^2 = \|u\|^2 + \|v\|^2 + i\langle u \,|\, v\rangle - i\langle v \,|\, u\rangle = \|u\|^2 + \|v\|^2.$$

Hence

$$\langle u \,|\, v\rangle + \langle v \,|\, u\rangle = \langle u \,|\, v\rangle - \langle v \,|\, u\rangle = 0$$

from which it follows at once that $\langle u \,|\, v\rangle = 0$.

In $\mathbb{C}^2$ with the standard inner product, let $u = (1,0)$ and $v = (i,0)$. Then

$$\|u + v\|^2 = \|u\|^2 + \|v\|^2 = 1 + 1 = 2$$

but $\langle u \,|\, v\rangle = i \neq 0$.

Problem 5.9 **Show that the formula**

$$\langle \mathsf{A} \,|\, \mathsf{B}\rangle = \mathrm{tr}(\mathsf{B}\,\mathsf{A}^\dagger)$$

defines an inner product on the vector space of $m \times n$ complex matrices $M(m,n)$.

(a) Calculate $\|\mathsf{I}_n\|$ where I_n is the $n \times n$ identity matrix.

(b) What characterizes matrices orthogonal to I_n?

(c) Show that all unitary $n \times n$ matrices U have the same norm with respect to this inner product.

Solution:

$$\langle \mathsf{A} \,|\, \mathsf{B}\rangle = \mathrm{tr}(\mathsf{B}\,\mathsf{A}^\dagger) = \sum_{i=1}^{m}\sum_{j=1}^{n} b_{ij}\overline{a_{ij}}$$

We must verify the three criteria (IP1)-(IP3) of an inner product:

(IP1)

$$\overline{\langle \mathsf{B} \,|\, \mathsf{A}\rangle} = \overline{\sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}\overline{b_{ij}}} = \sum_{i=1}^{m}\sum_{j=1}^{n} b_{ij}\overline{a_{ij}} = \langle \mathsf{A} \,|\, \mathsf{B}\rangle.$$

(IP2)

$$\langle \mathsf{A} \,|\, \alpha\mathsf{B} + \beta\mathbb{C}\rangle = \mathrm{tr}((\alpha\mathsf{B} + \beta\mathbb{C})\mathsf{A}^\dagger) = \alpha\,\mathrm{tr}(\mathsf{B}\mathsf{A}^\dagger) + \beta\,\mathrm{tr}(\mathsf{C}\mathsf{A}^\dagger) = \alpha\langle \mathsf{A} \,|\, \mathsf{B}\rangle + \beta\langle \mathsf{A} \,|\, \mathsf{C}\rangle.$$

(IP3)

$$\langle \mathsf{A} \,|\, \mathsf{A}\rangle = \sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}\overline{a_{ij}} = \sum_{i=1}^{m}\sum_{j=1}^{n} |a_{ij}|^2 \geq 0.$$

(a)

$$\|\mathsf{I}_n\| = \sqrt{\langle \mathsf{I}_n \,|\, \mathsf{I}_n\rangle} = \sqrt{\mathrm{tr}(\mathsf{I}_n)^2} = \sqrt{\mathrm{tr}\,\mathsf{I}_n} = \sqrt{n}.$$

(b) The matrix A is orthogonal to I_n, i.e. $\langle \mathsf{A} \,|\, \mathsf{I}_n\rangle = 0$ if and only if

$$0 = \sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}\overline{\delta_{ij}} = \sum_{i=1}^{m} a_{ii}.$$

That is, iff $\operatorname{tr} \mathsf{A} = 0$.

(c) For an $n \times n$ unitary matrix $\mathsf{U}\mathsf{U}^\dagger = \mathsf{I}_n$,

$$\|\mathsf{U}\| = \sqrt{\langle \mathsf{U} \,|\, \mathsf{U} \rangle} = \sqrt{\operatorname{tr}(\mathsf{U}\mathsf{U}^\dagger)} = \sqrt{\operatorname{tr} \mathsf{I}_n} = \sqrt{n}.$$

Problem 5.10 **Let S and T be complex inner product spaces and let $U : S \to T$ be a linear map such that $\|Ux\| = \|x\|$. Prove that**

$$\langle Ux \,|\, Uy \rangle = \langle x \,|\, y \rangle \qquad \textbf{for all } x, y \in S.$$

Solution: The identity $\|U(x+y)\|^2 = \|x+y\|^2$ can be written

$$\langle U(x+y) \,|\, U(x+y) \rangle = \langle Ux + Uy) \,|\, Ux + Uy) \rangle == \langle x+y \,|\, x+y \rangle.$$

Expanding gives

$$\begin{aligned} \langle Ux \,|\, Ux \rangle + \langle Uy \,|\, Uy \rangle + \langle Ux \,|\, Uy \rangle + \langle Uy \,|\, Ux \rangle \\ = \langle x \,|\, x \rangle + \langle y \,|\, y \rangle + \langle x \,|\, y \rangle + \langle y \,|\, x \rangle \end{aligned}$$

and using $\|Ux\|^2 = \|x\|^2$, $\|Uy\|^2 = \|y\|^2$, we find

$$\langle Ux \,|\, Uy \rangle + \langle Uy \,|\, Ux \rangle = \langle x \,|\, y \rangle + \langle y \,|\, x \rangle \,.$$

A similar treatment of the equation $\|U(x+iy)\|^2 = \|x+iy\|^2$ gives

$$i\langle Ux \,|\, Uy \rangle - i\langle Uy \,|\, Ux \rangle = i\langle x \,|\, y \rangle - i\langle y \,|\, x \rangle \,.$$

Removing the factor i and adding the last two displayed equations gives

$$\langle Ux \,|\, Uy \rangle = \langle x \,|\, y \rangle$$

as required.

Problem 5.11 **Let V be a complex vector space with an 'indefinite inner product', defined as an inner product which satisfies (IP1), (IP2) but with (IP3) replaced by the non-singularity condition**

(IP3′) $\langle u \,|\, v \rangle = 0$ **for all** $v \in V$ **implies that** $u = 0$.

(a) Show that similar results to Theorems 5.2 and 5.3 can be proved for such an indefinite inner product.

(b) If there are p $+1$'s along the diagonal and q -1's, find the defining

relations for the group of transformations $U(p,q)$ between orthonormal basis.

Solution: (a) The proofs of these theorems follow on lines very similar to those given in the text. To start of the orthonormalization procedure, note that there can exist "null vectors" $u \neq 0$ such that $\langle u \mid u\rangle = 0$, but not *every* vector can be null else. For, if $\langle u \,|\, u\rangle = 0 \forall u \in V$ then for all $u, v \in V$

$$0 = \langle u+v \,|\, u+v\rangle - i\langle u+iv \,|\, u+iv\rangle = 2\langle u \,|\, v\rangle$$

contradicting condition (IP3$'$). Pick a vector $u \in V$ such that $\langle u \mid u\rangle \neq 0$, and set $e_1 = u/\sqrt{|\langle u \,|\, u\rangle|}$, so that $\langle e_1 \,|\, e_1\rangle = \pm 1$.
Set $V_1 = \{w \in V \mid \langle w \mid e_1\rangle = 0\}$ to be the vector subspace orthogonal to e_1. The demonstration that every $v \in V$ has a unique deomposition $v = ae_1 + v'$ where $v' \in V_1$ follows exactly as in the proof of Theorem 5.2, and the rest of the proof is essentially identical.

The analogue of Sylvester's theorem 5.3 follows on similar lines to that given in the text. Assume there are two orthonormal bases $\{e_i\}$ and $\{f_j\}$ such that

$$\langle e_i \,|\, e_i\rangle = \begin{cases} 1 \text{ if } i \leq r \\ -1 \text{ if } i > r \end{cases} , \qquad \langle f_j \,|\, f_j\rangle = \begin{cases} 1 \text{ if } j \leq s \\ -1 \text{ if } j > s \end{cases}$$

where $s > r$. Then there exists a non-zero vector

$$u = \sum_{j=1}^{s} a^j f_j = -\sum_{i=1}^{n-r} b^i e_{r+i}$$

which has the contradictory properties $\langle u \,|\, u\rangle > 0$ and $\langle u \,|\, u\rangle < 0$.
(b) If $\{e_i\}$ and $\{e'_i\}$ are orthonormal bases, related by $e_i = U^j_{\ i} e'_j$ then

$$g_{ij} = \langle e_i \,|\, e_j\rangle = \langle U^k_{\ i} e'_k \,|\, U^l_{\ j} e'_l\rangle = \overline{U^k_{\ i}} g_{kl} U^l_{\ j}$$

where $\mathsf{G} = [g_{ij}]$ is the $(p+q)\times(p+q)$ diagonal matrix whose first p diagonal elements are $+1$ and the last q are -1. In matrices, this equation reads

$$\mathsf{G} = \mathsf{U}^\dagger \mathsf{G} \mathsf{U}.$$

Problem 5.12 **If V is an inner product space, an operator $K : V \to V$ is called self-adjoint if**

$$\langle u \,|\, Kv\rangle = \langle Ku \,|\, v\rangle$$

for any pair of vectors $u, v \in V$. Let $\{e_i\}$ be an arbitrary basis, having $\langle e_i \,|\, e_j\rangle = h_{ij}$, and set $Ke_k = K^j_{\ k} e_j$. Show that if $\mathsf{H} = [h_{ij}]$ and $\mathsf{K} = [K^k_{\ j}]$ then

$$\mathsf{HK} = \mathsf{K}^\dagger \mathsf{H} = (\mathsf{HK})^\dagger.$$

If $\{e_i\}$ is an orthonormal basis, show that K is a hermitian matrix.

Solution: Setting $u = e_i$, $v = e_j$ we have

$$\langle e_i \,|\, K e_j \rangle = \langle K e_i \,|\, e_j \rangle$$

i.e.

$$\langle e_i \,|\, K^k_j e_k \rangle = \langle K^k_i e_k \,|\, e_j \rangle$$

whence

$$h_{ik} K^k_j = \overline{K^k_i} h_{kj}$$

which are the components of the matrix equation

$$\mathsf{HK} = \mathsf{K}^\dagger \mathsf{H}.$$

Since $h_{ij} = \langle e_i \,|\, e_j \rangle = \overline{\langle e_j \,|\, e_i \rangle} = \overline{h_{ji}}$ it follows that the matrix H is Hermitian, $\mathsf{H}^\dagger = \mathsf{H}$, so that

$$\mathsf{K}^\dagger \mathsf{H} = \mathsf{K}^\dagger \mathsf{H}_\dagger = (\mathsf{HK})^\dagger \,.$$

For an o.n. basis $\mathsf{H} = \mathsf{I}$ and the above equation reduces to $\mathsf{K} = \mathsf{K}^\dagger$, the matrix K is Hermitian.

Problem 5.13 **For a function $\phi : G \to \mathbb{C}$, if we set $g\phi$ to be the function $(g\phi)(a) = \phi(g^{-1}a)$ show that $(gg')\phi = g(g'\phi)$.**
Show that the inner product (5.28) is G-invariant, $(g\phi, g\psi) = (\phi, \psi)$ for all $g \in G$.

Solution:

$$(gg')\phi(a) = \phi((gg')^{-1}a) = \phi((g')^{-1}g^{-1}a) = g'\phi(g^{-1}a) = gg'\phi(a).$$

For the second part of the question, by Eq. (5.28)

$$\begin{aligned}(g\phi, g\psi) &= \frac{1}{|G|} \sum_{a \in G} \overline{g\phi(a)} g\psi(a) \\ &= \frac{1}{|G|} \sum_{a \in G} \overline{\phi(g^{-1}a)} \psi(g^{-1}a).\end{aligned}$$

As a ranges over G, the elements $b = g^{-1}a$ alos range over all elements of G, whence

$$(g\phi, g\psi) = \frac{1}{|G|} \sum_{b \in G} \overline{\phi(b)} \psi(b) = (\phi, \psi).$$

Problem 5.14 **Let the character of a representation T of a group G on a vector space V be the function $\chi : G \to \mathbb{C}$ defined by**

$$\chi(g) = \operatorname{tr} T(g) = T^i_i(g).$$

(a) Show that the character is independent of the choice of basis and is a member of $\mathcal{F}(G)$, and that characters of equivalent representations are identical. Show that $\chi(e) = \dim V$.

(b) Any complex-valued function on G that is constant on conjugacy classes (see Section 2.4) is called a central function. Show that characters are central functions.

(c) Show that with respect to the inner product (5.28), characters of any pair of inequivalent irreducible representations $T_1 \nsim T_2$, are orthogonal to each other, $(\chi_1, \chi_2) = 0$, while the character of any irreducible representation T has unit norm $(\chi, \chi) = 1$.

(d) From Theorem 5.8 and Theorem 5.7 every unitary representation T can be decomposed into a direct sum of inequivalent irreducible unitary representations $T_\mu : G \to GL(V_\mu)$

$$T \sim m_1 T_1 \oplus m_2 T_2 \oplus \cdots \oplus m_N T_N \qquad (m_\mu \geq 0).$$

Show that the multiplicities m_μ of the representations T_μ are given by

$$m_\mu = (\chi, \chi_\mu) = \frac{1}{|G|} \sum_{g \in G} \overline{\chi(g)} \chi_\mu(g)$$

and T is irreducible if and only if its character has unit magnitude, $(\chi, \chi) = 1$. Show that T and T' have no irreducible representations in common in their decompositions if and only if their characters are orthogonal.

Solution: (a) Let T be a representation of a finite group G on a vector space V of dimension n. As in problem 4.1(b), the trace of an operator is basis-independent since its matrix is unchanged by similarity transformations, using $\mathrm{tr}(\mathsf{AB}) = \mathrm{tr}(\mathsf{BA})$:

$$\mathrm{tr}(\mathsf{STS}^{-1}) = \mathrm{tr}(\mathsf{S}^{-1}\mathsf{ST}) = \mathrm{tr}\,\mathsf{T}.$$

The character χ is therefore a member of $\mathcal{F}(G)$ which is entirely determined by the representation. Characters of equivalent representations are identical, for if $T_1 \sim T_2$ there exists a vector space isomorphism $A : V_1 \to V_2$ such that

$$T_2(g) = AT_1(g)A^{-1},$$

and

$$\chi_2(g) = \mathrm{tr}\, T_2(g) = \mathrm{tr}(AT_1(g)A^{-1}) = \mathrm{tr}\, T_1(g) = \chi_1(g).$$

(b) From Section 2.4 two elements g and g' are said to belong to the same conjugacy class C if and only if there exists $h \in G$ such that $g' = hgh^{-1}$. Then if χ is the character of a representation T then

$$\chi(g') = \mathrm{tr}\, T(g') = \mathrm{tr}\, T(hgh^{-1}) = \mathrm{tr}\, T(h)T(g)T(h)^{-1} = \mathrm{tr}\, T(g) = \chi(g).$$

Hence χ is a central function.

(c) From Eq. (5.29) it follows that for any pair of inequivalent irreducible representation $T_1 \nsim T_2$, their characters are orthogonal to each other,

$$(\chi_1, \chi_2) = \sum_{i=1}^{n_1} \sum_{a=1}^{n_2} \big(T_{(1)ii}, T_{(2)aa}\big) = 0,$$

while from (5.30) the character any irreducible representation T has unit norm

$$\begin{aligned}(\chi, \chi) &= \sum_{i=1}^{n} \sum_{k=1}^{n} (T_{ii}, T_{kk}) \\ &= \frac{1}{n} \sum_{i=1}^{n} \sum_{k=1}^{n} \delta_{ik}\delta_{ik} \\ &= \frac{1}{n} \sum_{i=1}^{n} \delta_{ii} = \frac{n}{n} = 1.\end{aligned}$$

(d) Setting

$$T \sim m_1 T_1 \oplus m_2 T_2 \oplus \cdots \oplus m_N T_N \qquad (m_\mu \geq 0)$$

where $T_\mu : G \to GL(V_\mu)$ are the inequivalent irreducible unitary representations of G. The character of T is clearly an additive function,

$$\chi(g) = \sum_{\mu=1}^{N} m_\mu \chi_\mu(g)$$

where χ_μ is the character of the μth irreducible representation T_μ. Using the orthonormality of characters of inequivalent reps shown in (c),

$$(\chi_\mu, \chi_\nu) = \delta_{\mu\nu},$$

we have

$$m_\mu = (\chi, \chi_\mu) = \frac{1}{|G|} \sum_{g \in G} \overline{\chi(g)} \chi_\mu(g).$$

T is irreducible iff it is equivalent to T_μ for some μ, i.e. $m_\nu = \delta_{\mu\nu}$. Hence a necessary and sufficient condition for T to irreducible is that

$$(\chi, \chi) = (\chi_\mu, \chi_\mu) = 1.$$

From

$$(\chi, \chi') = \sum_{\mu=1}^{N} m_\mu m'_\mu \geq 0$$

it follows that the character χ and χ' are orthogonal, $(\chi, \chi') = 0$, iff every product $m_\mu m'_\mu = 0$, i.e. iff T and T' have no irreducible representations in common in their decompositions.

Chapter 6

Problem 6.1 **The following is an alternative method of defining the algebra of complex numbers. Let $\mathcal{P}$ be the associative algebra consisting of real polynomials on the variable x, defined in Example 6.3. Set $\mathcal{C}$ to be the ideal of $\mathcal{P}$ generated by x^2+1; i.e., the set of all polynomials of the form $f(x)(x^2+1)g(x)$. Show that the linear map $\phi : \mathbb{C} \to \mathcal{P}/\mathcal{C}$ defined by**

$$\phi(i) = [x] = x + \mathcal{C}, \qquad \phi(1) = [1] = 1 + \mathcal{C},$$

is an algebra isomorphism.

Which complex number is identified with the polynomial class $[1 + x + 3x^2 + 5x^3] \in \mathcal{P}/\mathcal{C}$?

Solution: Firstly, since ϕ is a linear map $\phi(a + bi) = [a + bx]$. Hence

$$\phi\big((a+ib)(c+id)\big) = \phi\big(ac - bd + i(bc + ad)\big) = [ac - bd + (bc + ad)x].$$

Using the rule for multiplying elements of the factor algebra, $[u][v] = [uv]$, (see proof of Theorem 6.1), we have

$$\begin{aligned}\phi(a+ib)\phi(c+id) &= [a + bx]\,[c + dx]\\ &= [ac + bcx + abx + bdx^2]\\ &= [ac - bd + (bc + ad)x]\end{aligned}$$

since $[x^2] = -[1]$ as follows from the fact that $[x^2+1] = 0$ (equivalently, $x^2 + 1 \in \mathcal{C}$). Hence ϕ is an algebra homomorphism,

$$\phi\big((a+ib)(c+id)\big) = \phi(a+ib)\phi(c+id).$$

All higher powers $[x^n]$ can be equated with $\pm[1]$ or $\pm[x]$, e.g. $[x^2] = -1,\ [x^3] = -[x]$, $[x^4] = [1]$ etc. it follows that every polynomial $[p(x)]$ can be reduced to the form $[a + bx]$. Hence ϕ is an isomorphism as it is one-to-one and onto.

$$[1 + x + 3x^2 + 5x^3] = [1 + x - 3 - 5x] = [-2 - 4x]$$

which corresponds to the complex number $-2 - 4i$.

Problem 6.2 **Let J be a complex structure on a real vector space V, and set**

$$V(J) = \{v = u - iJu \,|\, u \in V\} \subseteq V^C, \qquad \bar{V}(J) = \{v = u + iJu \,|\, u \in V\}.$$

(a) Show that $V(J)$ and $\bar{V}(J)$ are complex vector subspaces of V^C.

(b) Show that $v \in V(J) \Rightarrow Jv = iv$ and $v \in \bar{V}(J) \Rightarrow Jv = -iv$.

(c) Prove that the complexification of V is the direct sum of $V(J)$ and $\bar{V}(J)$,

$$V^C = V(J) \oplus \bar{V}(J).$$

Solution: (a) If $v = u - iJu$, $v' = u' - iJu' \in V(J)$, where u, $u' \in V$, then

$$v + v' = (u + u') - iJ(u + u') \in V(J)$$

and

$$(a + ib)(u - iJu) = au - iaJu + ibu + bJu = (au + bJu) - iJ(au + bJu) \in V(J).$$

Hence $V(J)$ is closed with respect to vector addition and complex scalar multiplication, and is therefore a vector subspace of V^C. Similarly, $(a + ib)(u + iJu) = (au + bJu) + iJ(au + bJu)$ shows that $\bar{V}(J)$ is a vector subspace of V^C.

(b) If $v = u - iJu \in V(J)$ then

$$Jv = Ju - iJ^2u = Ju + iu = i(u - iJu) = iv.$$

Similarly, if $v = u + iJu \in \bar{V}(J)$ then $Jv = i(u + iJu) = -iu$.

(c) If $u + iv$ is any vector in V^C where $u, v \in V$, try for a decomposition

$$u + iv = u' - iJu' + v' + iJv'.$$

The real and and imaginary parts of this equation give

$$u' + v' = u, \qquad -Ju' + Jv' = v.$$

Multiplying the second equation of this pair by J on the left and right gives

$$u' - v' = Jv$$

and solving for u' and v' we find

$$u' = \frac{1}{2}(u + Jv), \qquad v' = \frac{1}{2}(u - Jv).$$

This provides a unique decomposition of $u + iv$ into a sum of two vectors, one from $V(J)$, athe other from $\bar{V}(J)$. The uniqueness guarantees that V^C is a direct sum, $V^C = V(J) \oplus \bar{V}(J)$, for it implies that $V(J) \cap \bar{V}(J) = \{0\}$ since the zero vector has the unique decomposition $0 = 0 + 0$.

Problem 6.3 **If V is a real vector space and U and $\bar{U}$ are complex conjugate subspaces of V^C such that $V^C = U \oplus \bar{U}$, show that there exists a complex structure J for V such that $U = V(J)$ and $\bar{U} = \bar{V}(J)$, where $V(J)$ and $\bar{V}(J)$ are defined in the previous problem.**

Solution: Every $u \in V^C$ has a unique decomposition $u = z + \overline{z'}$ where $z, z' \in U$. In particular the vectors $u \in V$ are characterized by $u = \overline{u}$, so that

$$z + \overline{z'} = \bar{z} + z'$$

and, since $U \cap \bar{U} = \{0\}$,

$$z - z' = \overline{z - z'} = 0.$$

Hence $z = z'$ and every $u \in V$ has a unique decomposition

$$u = z + \bar{z} \qquad (z \in U).$$

Conversely, of course, every vector of this form belongs to V since $\overline{z + \bar{z}} = z + \bar{z}$. Now set

$$v = J(u) = i(z - \bar{z}) = iz + \overline{iz} \in V.$$

The operator is obviously real linear on V and is a complex structure, $J^2 = -\mathrm{id}_V$, since

$$J^2 u = Jv = i(iz - \overline{iz}) = i(iz + i\bar{z}) = -(z + \bar{z}) = -u.$$

Furthermore every element $u - iJu \in V(J)$ belongs to U, since

$$u - iJu = (z + \bar{z}) - ii(z - \bar{z}) = 2z \in U,$$

while conversely every $z \in U$ can be written $z = v - iJv$ where $v = \frac{1}{2}z + \frac{1}{2}\bar{z} \in V$. Hence $U = V(J)$, and similarly $\bar{U} = \bar{V}(J)$.

Problem 6.4 **Let J be a complex structure on a real vector space V of dimension $n = 2m$. Let $u_1, u_2, \ldots, u_m$ be a basis of the subspace $V(J)$ defined in Problem 6.2, and set**

$$u_a = e_a - ie_{m+a} \qquad \textbf{where} \qquad e_a,\, e_{m+a} \in V \quad (a = 1, \ldots, m).$$

Show that the matrix $\mathsf{J}_0 = [J^j_i]$ of the complex structure, defined by $Je_i = J^j_i e_j$ where $i = 1, 2, \ldots, n = 2m$, has the form

$$\mathsf{J}_0 = \begin{pmatrix} \mathsf{O} & \mathsf{I} \\ -\mathsf{I} & \mathsf{O} \end{pmatrix}.$$

Show that the matrix of any complex structure with respect to an arbitrary basis has the form

$$\mathsf{J} = \mathsf{A}\mathsf{J}_0\mathsf{A}^{-1}.$$

Solution: The e_i are linearly independent for if there exist scalars A^a, B^a such that

$$\sum_{a=1}^{m} A^a e_a + B^a e_{m+a} = 0$$

then, using

$$e_a = \frac{1}{2}(u_a + \overline{u_a}), \qquad e_{m+a} = \frac{i}{2}(u_a - \overline{u_a}),$$

we have

$$\frac{1}{2}\sum_{a=1}^{m}(A^a + iB^a)u_a + (A^a - iB^a)\overline{u_a} = 0.$$

Since the $u_a, \overline{u_a}$ are l.i. it follows that $A^a + iB^a = A^a - iB^a = 0$ for all $a = 1, \dots m$, whence all A^a and B^a vanish , proving that the e_i are l.i. and form a basis of V since they are n in number.

Since $Ju_a = iu_a$ (see Problem 6.2 (b)), we have

$$Je_a - iJe_{m+a} = ie_a + e_{m+a}$$

and equating real and imaginary parts,

$$Je_a = e_{m+a}, \qquad Je_{m+a} = -e_a$$

so that $J^a{}_{m+b} = \delta^a_b$, $J^{m+a}{}_b = \delta^a_b$ while all other components vanish. Every other basis e'_j is related to such a special basis by $e_j = A^i{}_j e'_i$, where $\mathsf{A} = [A^i{}_j]$ is a non-singular matrix, so that as in Chapter 3

$$Je'_i = J'^k_i e'_k \qquad \text{where} \qquad J'^k_i = A^k_j J^j_l (A^{-1})^l{}_i,$$

i.e.

$$\mathsf{J} = [J'^k_i] = \mathsf{A}\mathsf{J}_0\mathsf{A}^{-1}.$$

Problem 6.5 **Show the 'anticommutation law' of conjugation,**

$$\overline{PQ} = \overline{Q}\,\overline{P}.$$

Hence prove

$$|PQ| = |P|\,|Q|.$$

Solution: Set $P = p_0 + \mathbf{p}$, $Q = q_0 + \mathbf{q}$ and we have, from Eq. (6.6)

$$PQ = p_0q_0 - \mathbf{p}\cdot\mathbf{q} + p_0\mathbf{q} + q_0\mathbf{p} + \mathbf{p}\times\mathbf{q},$$

and taking the conjugate quaternion

$$\overline{PQ} = p_0q_0 - \mathbf{p}\cdot\mathbf{q} - p_0\mathbf{q} - q_0\mathbf{p} - \mathbf{p}\times\mathbf{q}.$$

On the other hand,

$$\begin{aligned}\overline{Q}\,\overline{P} &= (q_0 - \mathbf{q})(p_0 - \mathbf{p}) \\ &= q_0p_0 - q_0\mathbf{p} - p_0\mathbf{q} + \mathbf{qp} \\ &= p_0q_0 - p_0\mathbf{q} - q_0\mathbf{p} - \mathbf{q}\cdot\mathbf{p} + \mathbf{q}\times\mathbf{p} \qquad \text{by Eq. (6.5)} \\ &= p_0q_0 - \mathbf{q}\cdot\mathbf{p} - p_0\mathbf{q} - q_0\mathbf{p} - \mathbf{p}\times\mathbf{q} = \overline{PQ}.\end{aligned}$$

Hence

$$|PQ|^2 = PQ\overline{PQ} = PQ\overline{Q}\,\overline{P} = P|Q|^2\overline{P} = |Q|^2P\overline{P} = |Q|^2|P|^2,$$

and taking the positive square root gives $|PQ| = |P|\,|Q|$.

Problem 6.6 **Show that the set of 2×2 matrices of the form**

$$\begin{pmatrix} z & w \\ -\bar{w} & \bar{z} \end{pmatrix},$$

where z and w are complex numbers, forms an algebra of dimension 4 over the real numbers.

(a) Show that this algebra is isomorphic to the algebra of quaternions by using the bijection

$$Q = a + bi + cj + dk \longleftrightarrow \begin{pmatrix} a+ib & c+id \\ -c+id & a-ib \end{pmatrix}.$$

(b) Using this matrix representation prove the identities given in Problem 6.5.

Solution: Matrices of the form

$$\mathsf{Q} = \begin{pmatrix} z & w \\ -\bar{w} & \bar{z} \end{pmatrix}, \mathsf{P} = \begin{pmatrix} u & v \\ -\bar{v} & \bar{u} \end{pmatrix}$$

are closed with respect to formation of real linear combinations

$$a\mathsf{Q} + b\mathsf{P} = \begin{pmatrix} z+u & w+v \\ -\overline{w+v} & \overline{z+u} \end{pmatrix},$$

and matrix products

$$\mathsf{QP} = \begin{pmatrix} zu - w\bar{v} & zv + w\bar{u} \\ -\bar{w}u - \bar{z}\bar{v} & -\bar{w}v + \bar{z}\bar{u} \end{pmatrix} = \begin{pmatrix} zu - w\bar{v} & zv + w\bar{u} \\ -\overline{zv - w\bar{u}} & \overline{zu - w\bar{v}} \end{pmatrix}.$$

Hence they form an algebra over the real numbers.

(a) Setting

$$\mathsf{I} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \mathsf{X}_1 = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}, \quad \mathsf{X}_2 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad \mathsf{X}_3 = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}$$

it is straightforward to verify that

$$\mathsf{I}^2 = \mathsf{I}, \qquad (\mathsf{X}_1)^2 = (\mathsf{X}_2)^2 = (\mathsf{X}_3)^2 = -\mathsf{I}$$

$$\begin{aligned} \mathsf{X}_1\mathsf{X}_2 &= -\mathsf{X}_2\mathsf{X}_1 = \mathsf{X}_3 \\ \mathsf{X}_2\mathsf{X}_3 &= -\mathsf{X}_3\mathsf{X}_2 = \mathsf{X}_1 \\ \mathsf{X}_3\mathsf{X}_1 &= -\mathsf{X}_1\mathsf{X}_3 = \mathsf{X}_2 \end{aligned}$$

Hence the algebra generated by these matrices,

$$a\mathsf{I} + b\mathsf{X}_1 + c\mathsf{X}_2 + d\mathsf{X}_3 = \begin{pmatrix} a+ib & c+id \\ -c+id & a-ib \end{pmatrix}$$

is isomorphic to the algebra of quaternions, by the correspondence

$$1 \to \mathsf{I}, \quad i \to \mathsf{X}_1, \quad j \to \mathsf{X}_2, \quad k \to \mathsf{X}_3.$$

(b) If

$$Q \longleftrightarrow \begin{pmatrix} z & w \\ -\bar{w} & \bar{z} \end{pmatrix}, \qquad \text{then} \qquad \overline{Q} \longleftrightarrow \begin{pmatrix} \bar{z} & -w \\ \bar{w} & z \end{pmatrix},$$

i.e. the action of conjugation in quaternions is equivalent to the replacements

$$z \longrightarrow \overline{z}, \qquad w \longrightarrow -w$$

in the above matrices. Hence if P is as above then

$$\overline{\mathsf{P}}\,\overline{\mathsf{Q}} = \begin{pmatrix} \overline{uz} - v\bar{w} & -\bar{u}w - vz \\ \overline{vz} + u\bar{w} & -\bar{v}w + uz \end{pmatrix} = \overline{\mathsf{QP}}.$$

The identity $|PQ| = |P|\,|Q|$ follows by the same reasoning as in Problem 6.5.

Problem 6.7 **Find a quaternion Q such that**

$$Q^{-1}iQ = j, \qquad Q^{-1}jQ = k.$$

[*Hint*: Write the first equation as $iQ = Qj$.] For this Q calculate $Q^{-1}kQ$.

Solution: Set $Q = a_0 + a_1 i + a_2 j + a_3 k$. Then $iQ = Qj$ implies

$$a_0 i - a_1 + a_2 k - a_3 j = a_0 j + a_1 k - a_2 - a_3 i.$$

Equating coefficents of $1, i, j, k$ separately we have

$$-a_1 = -a_2, \quad a_0 = -a_3, \quad -a_3 = a_0, \quad a_2 = a_1,$$

i.e. $a_1 = a_2$ and $a_0 = -a_3$. Similarly $jQ = Qk$ leads to $a_0 = -a_1$, $a_2 = a_3$. Hence $a_0 = -a_1 = -a_2 = -a_3$, and Q is the quaternion

$$Q = a_0(1 - i - j - k).$$

The inverse quaternion is

$$Q^{-1} = \frac{\overline{Q}}{Q\overline{Q}} = \frac{1}{4a_0}(1+i+j+k)$$

and

$$\begin{aligned}
Q^{-1}kQ &= \frac{1}{4}(1+i+j+k)k(1-i-j-k) \\
&= \frac{1}{4}(1+i+j+k)(k-j+i=1) \\
&= \frac{1}{4}(k-j+i+1-j-k-1+i+1+1-k+j-1+i+j+k) \\
&= i\,.
\end{aligned}$$

Problem 6.8 **Let e_A and e_B where $A,\, B \subseteq \{1,2,\ldots,n\}$ be two basis elements of the Clifford algebra associated with the Euclidean inner product space having $g_{ij} = \delta_{ij}$. Show that $e_Ae_B = \pm e_C$ where $C = A \cup B - A \cap B$. Show that a plus sign appears in this rule if the number of pairs**

$$\{(i_r, j_s)\,|\, i_r \in A,\ j_s \in B,\ i_r > j_s\}$$

is even, while a minus sign occurs if this number of pairs is odd.

Solution: Let

$$\begin{aligned}
e_A &= e_{i_1}e_{i_2}\ldots e_{i_a} \qquad \text{where} \qquad i_1 < i_2 < \cdots < i_a \in A, \\
e_B &= e_{j_1}e_{j_2}\ldots e_{j_b} \qquad \text{where} \qquad j_1 < j_2 < \cdots < j_b \in B.
\end{aligned}$$

Then

$$e_Ae_B = e_{i_1}\ldots e_{i_a}e_{j_1}e_{j_2}\ldots e_{j_b}.$$

If $e_{j_1} < e_{i_a}$ move it to the left in this product by a succession of interchanges until $j_1 \geq i_k$. In each such interchanges the sign of the product is changed since $e_ie_j = -e_je_i$ for $i \neq j$. If $i_k = j_1$ we simply eliminate the pair $e_{i_k}e_{j_1}$, since $e_ie_j = 1$ for $i = j$. Repeat the process successively for $e_{j_2}\ldots e_{j_b}$ until an element $e_C = e_{k_1}\ldots e_{k_c}$ with $k_1 < k_2 < \cdots < k_c$ is arrived at. The indices k_l all belong to the set A or B, but indices which belong to both A and B areexcluded, i.e. $C = A \cup B - A \cap B$.
In the above process a sign change arises in any interchange of elements $e_{i_r}e_{j_s}$ whenever $i_r > j_s$. Hence the sign of e_C in the product e_Ae_B is equal to the number of pairs (i_r, j_s) such that $i_r \in A,\ j_s \in B$ and $i_r > j_s$.

Problem 6.9 **Let $\{e_1,\, e_2, \ldots,\, e_n\}$ be a basis of a vector space V of dimension $n \geq 5$. By calculating their wedge product, decide whether the following vectors are linearly dependent or independent:**

$$e_1 + e_2 + e_3, \quad e_2 + e_3 + e_4, \quad e_3 + e_4 + e_5, \quad e_1 + e_3 + e_5.$$

Can you find a linear relation among them?

Solution:

$$
\begin{aligned}
&(e_1+e_2+e_3)\wedge(e_2+e_3+e_4)\wedge(e_3+e_4+e_5)\wedge(e_1+e_3+e_5)\\
&=(e_1\wedge e_2+e_1\wedge e_3+e_1\wedge e_4+e_2\wedge e_4+e_3\wedge e_4)\\
&\qquad\wedge(e_3\wedge e_1+e_4\wedge e_1+e_4\wedge e_3+e_4\wedge e_5+e_5\wedge e_1)\\
&=e_1\wedge e_2\wedge e_4\wedge e_3+e_1\wedge e_2\wedge e_4\wedge e_5+e_1\wedge e_3\wedge e_4\wedge e_5\\
&\qquad+e_2\wedge e_4\wedge e_3\wedge e_1+e_2\wedge e_4\wedge e_5\wedge e_1+e_3\wedge e_4\wedge e_5\wedge e_1)\\
&=0
\end{aligned}
$$

since the first and fourth, second and fifth, third and sixth terms concel in the penultimate expression (check that the vectors in the wedge products are odd permutations of each other). Hence these vecotrs are linearly dependent.

To find a linear relation among them try

$$e_1+e_3+e_5=a(e_1+e_2+e_3)+b(e_2+e_3+e_4)+c(e_3+e_4+e_5).$$

Then, equating coefficients of the different e_i gives

$$a=1,\qquad 0=a+b,\qquad 1=a+b+c,\qquad b+c=0,\qquad 1=c$$

so that $a=1$, $b=-1$, $c=1$.

Problem 6.10 **Let W be a vector space of dimension 4 and $\{e_1, e_2, e_3, e_4\}$ a basis. Let A be the 2-vector on W,**

$$A=e_2\wedge e_1+ae_1\wedge e_3+e_2\wedge e_3+ce_1\wedge e_4+be_2\wedge e_4.$$

Write out explicitly the equations $A\wedge u=0$ where $u=u^1e_1+u^2e_2+u^3e_3+u^4e_4$ and show that they have a non-trivial solution if and only if $c=ab$. In this case find two vectors u and v such that $A=u\wedge v$.

Solution:

$$
\begin{aligned}
A\wedge u=&(u^1-au^2-u^3)e_1\wedge e_2\wedge e_3+(bu^1-cu^2-u^4)e_1\wedge e_2\wedge e_4\\
&+(-cu^3+au^4)e_1\wedge e_3\wedge e_4+(u^4-bu^3)e_2\wedge e_3\wedge e_4
\end{aligned}
$$

so that $A\wedge u=0$ if

$$
\begin{aligned}
u^1-au^2-u^3&=0\\
bu^1-cu^2-u^4&=0\\
-cu^3+au^4&=0\\
u^4-bu^3&=0
\end{aligned}
$$

. This set of linear equations has a non-trivial solution iff

$$\det \begin{vmatrix} 1 & -a & -1 & 0 \\ b & -c & 0 & -1 \\ 0 & 0 & -c & a \\ 0 & 0 & -b & 1 \end{vmatrix} = 0.$$

The determinant evaluates to $(ab - c)^2$, so that $A \wedge u = 0$ iff $c = ab$.
The general solution of the linear set of equations is easily found to be

$$u^1 = ax + y, \qquad u^2 = x, \qquad u^3 = y, \qquad u^4 = by$$

where x nd y are arbitrary real numbers. Hence, setting $x = 1, y = 0$ and $x = 0, y = 1$ gives the two independent solutions

$$u = ae_1 + e_2, \qquad v = e_1 + e_3 + be_4.$$

A straightforward computation gives $A = u \wedge v$.

Problem 6.11 **Let U be a subspace of V spanned by linearly independent vectors $\{u_1, u_2, \dots, u_p\}$.**

(a) Show that the p-vector $E_U = u_1 \wedge u_2 \wedge \dots \wedge u_p$ is defined uniquely up to a factor by the subspace U in the sense that if $\{u'_1, u'_2, \dots, u'_p\}$ is any other linearly independent set spanning U then the p-vector $E'_U \equiv u'_1 \wedge \dots \wedge u'_p$ is proportional to E_U; i.e., $E'_U = cE_U$ for some scalar c.

(b) Let W be a q-dimensional subspace of V, with corresponding q-vector E_W. Show that $U \subseteq W$ if and only if there exists a simple $(q-p)$-vector F such that $E_W = E_U \wedge F$.

(c) Show that if $p > 0$ and $q > 0$ then $U \cap W = \{0\}$ if and only if $E_U \wedge E_W \neq 0$.

Solution: (a) Let $u'_i = \sum_{j=1}^p B^j_{\ i} u_j$ be a second basis of U, then

$$u'_1 \wedge \dots \wedge u'_p = \sum_{j_1=1}^p \dots \sum_{j_p=1}^p B^{j_1}_{\ 1} B^{j_2}_{\ 2} \dots B^{j_p}_{\ p} u_{j_1} \wedge u_{j_2} \wedge \dots \wedge u_{j_p}.$$

Since every summand on the RHS is either 0 if any pair of vectors is equal, or $\pm u_1 \wedge u_2 \wedge \dots \wedge u_p$ whenever all u_i are unequal, we have

$$E'_U = u'_1 \wedge \dots \wedge u'_p = cu_1 \wedge u_2 \wedge \dots \wedge u_p = cE_U.$$

(b) If $U \subseteq W$, we may use the basis extension theorem, Theorem 3.7, to extend any basis $u_1, \dots, u_p$ of U to a basis $u_1, \dots, u_p, u_{p+1}, \dots u_q$ of W. Then

$$E_W = u_1 \wedge \dots \wedge u_p \wedge u_{p+1} \wedge \dots \wedge u_q = E_U \wedge F$$

where $F = u_{p+1} \wedge \cdots \wedge u_q$. Conversely, if there exists a simple vector $F = v_1 \wedge v_2 \wedge \cdots \wedge v_{q-p} \neq 0$ such that

$$E_W = E_U \wedge F = u_1 \wedge \cdots \wedge u_p \wedge v_1 \wedge v_2 \wedge \cdots \wedge v_{q-p} \neq 0$$

then the vectors $u_1, \ldots, u_p, v_p, \ldots v_{q-p}$ are linearly independent and q in number. They therefor form a basis of W and the subspace U spanned by $u_1, \ldots, u_p$ is clearly a subspace.

(c) Let $u_1, \ldots, u_p$ be a basis of U and $w_1 \ldots w_q$ a basis of W. Their intersection $U \cap W = \{0\}$ if and only if these vectors are all linearly independent (else there is a linear combination of u''s equal to a linear combination of w's which would be a non-zero vector in $U \cap W$). By Theorem 6.2 linear independence holds iff

$$E_U \cap E_W = u_1 \wedge \cdots \wedge u_p \wedge w_1 \wedge \cdots \wedge w_q \neq 0.$$

Problem 6.12 As in Example 6.12, $n \times n$ unitary matrices satisfy $\mathsf{U}\mathsf{U}^\dagger = \mathsf{I}$ and those near the identity have the form

$$\mathsf{U} = \mathsf{I} + \epsilon\mathsf{A} \qquad (\epsilon \ll 1)$$

where A is anti-hermitian, $\mathsf{A} = -\mathsf{A}^\dagger$.

(a) Show that the set of anti-hermitian matrices form a Lie algebra with respect to the commutator $[\mathsf{A}, \mathsf{B}] = \mathsf{AB} - \mathsf{BA}$ as bracket product.

(b) The four *Pauli matrices* σ_μ $(\mu = 0, 1, 2, 3)$ are defined by

$$\sigma_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Show that $\mathsf{Y}_\mu = \frac{1}{2} i\, \sigma_\mu$ form a basis of the Lie algebra of $U(2)$ and calculate the structure constants.

(c) Show that the one-parameter subgroup generated by Y_1 consists of matrices of the form

$$\mathrm{e}^{t\,\mathsf{Y}_1} = \begin{pmatrix} \cos\frac{1}{2}t & i\sin\frac{1}{2}t \\ i\sin\frac{1}{2}t & \cos\frac{1}{2}t \end{pmatrix}.$$

Calculate the one-parameter subgroups generated by Y_2, Y_3 and Y_0.

Solution: (a) We need to show that anti-hermitian matrices are closed with respect to commutator products,

$$\begin{aligned}
[\mathsf{A}, \mathsf{B}]^\dagger &= (\mathsf{AB} - \mathsf{BA})^\dagger \\
&= \mathsf{B} \dagger \mathsf{A}^\dagger - \mathsf{A}^\dagger \mathsf{B}^\dagger \\
&= (-\mathsf{B})(-\mathsf{A}) - (-\mathsf{A})(-\mathsf{B}) \\
&= \mathsf{BA} - \mathsf{AB} = -[\mathsf{A}, \mathsf{B}]
\end{aligned}$$

(b) A general 2×2 anti-hermitian matrix A has the form

$$\mathsf{A} = \begin{pmatrix} ia & c+id \\ -c+id & ib \end{pmatrix} = (a+b)\mathsf{Y}_0 + (a-b)\mathsf{Y}_3 + 2d\mathsf{Y}_1 + 2c\mathsf{Y}_2$$

where

$$\mathsf{Y}_0 = \tfrac{1}{2}i\sigma_0 = \begin{pmatrix} i & 0 \\ 0 & i \end{pmatrix}, \qquad \mathsf{Y}_1 = \tfrac{1}{2}i\sigma_1 = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix},$$

$$\mathsf{Y}_2 = \tfrac{1}{2}i\sigma_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \mathsf{Y}_3 = \tfrac{1}{2}i\sigma_3 = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}.$$

Straightforward matrix calculations result in

$$[\mathsf{Y}_0, \mathsf{Y}_i] = \mathsf{Y}_0\mathsf{Y}_i - \mathsf{Y}_i\mathsf{Y}_0 \propto \mathsf{Y}_i - \mathsf{Y}_i = \mathsf{O},$$

$$[\mathsf{Y}_1, \mathsf{Y}_2] = -\mathsf{Y}_3, \qquad [\mathsf{Y}_2, \mathsf{Y}_3] = -\mathsf{Y}_1, \qquad [\mathsf{Y}_3, \mathsf{Y}_1] = -\mathsf{Y}_2.$$

and the structure constants are

$$C^0_{ij} = 0, \qquad C^i_{jk} = \epsilon_{ijk} = \begin{cases} 1 \text{ if } ijk \text{ is an even perm of} 123 \\ -1 \text{ if } ijk \text{ is an odd perm of} 123 \\ 0 \text{ if any pair of } ijk \text{ are equal.} \end{cases}$$

(c) Using $(\mathsf{Y}_1)^2 = -\frac{1}{2}\mathsf{I}$, $(\mathsf{Y}_1)^3 = -\frac{1}{4}\mathsf{Y}_1 \dots$, we have

$$\begin{aligned} e^{t\mathsf{Y}_1} &= \mathsf{I} + t\mathsf{Y}_1 + \frac{1}{2!}t^2(\mathsf{Y}_1)^2 + \frac{1}{3!}t^3(\mathsf{Y}_1)^3 + \dots \\ &= \mathsf{I}\Big(1 - \frac{1}{2!}\frac{t^2}{2^2} + \frac{1}{4!}\frac{t^4}{2^4} - \dots\Big) + 2\mathsf{Y}_1\Big(\frac{t}{2} - \frac{1}{3!}\frac{t^3}{2^3} + \dots\Big) \\ &= \mathsf{I}\cos(t/2) + 2\mathsf{Y}_1\sin(t/2) \\ &= \begin{pmatrix} \cos\frac{1}{2}t & i\sin\frac{1}{2}t \\ i\sin\frac{1}{2}t & \cos\frac{1}{2}t \end{pmatrix} \end{aligned}$$

Similarly

$$e^{t\mathsf{Y}_2} = \begin{pmatrix} \cos\frac{1}{2}t & \sin\frac{1}{2}t \\ -\sin\frac{1}{2}t & \cos\frac{1}{2}t \end{pmatrix}, \qquad e^{t\mathsf{Y}_3} = \begin{pmatrix} e^{\frac{1}{2}it} & 0 \\ 0 & e^{\frac{1}{2}it} \end{pmatrix}$$

and

$$e^{t\mathsf{Y}_0} = e^{\frac{1}{2}it}\mathsf{I}.$$

Problem 6.13 Let u be an $n \times 1$ column vector. A non-singular matrix A is said to *stretch* u if it is an eigenvector of A,

$$\mathsf{A}\mathbf{u} = \lambda\mathbf{u}.$$

Show that the set of all non-singular matrices that stretch u forms a group with respect to matrix multiplication, called the *stretch group of* u.

(a) Show that the 2×2 matrices of the form

$$\begin{pmatrix} a & a+c \\ b+c & b \end{pmatrix} \qquad (c \neq 0,\ a+b+c \neq 0)$$

form the stretch group of the 2×1 column vector $\mathbf{u} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$.

(b) Show that the Lie algebra of this group is spanned by the matrices

$$\mathsf{X}_1 = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}, \quad \mathsf{X}_2 = \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix}, \quad \mathsf{X}_3 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Calculate the structure constants for this basis.

(c) Write down the matrices that form the one-parameter subgroups $e^{t\mathsf{X}_1}$ and $e^{t\mathsf{X}_3}$.

Solution: The set of non-singular matrices A which stretch $\mathbf{u}$ is a group, for if

$$\mathsf{A}\mathbf{u} = \lambda\mathbf{u}, \qquad \mathsf{B}\mathbf{u} = \mu\mathbf{u}$$

then $\mathsf{AB}\mathbf{u} = \mu\lambda\mathbf{u}$; hence this set of matrices is closed with respect to matrix multiplication. The identity matrix I clearly stretches $\mathbf{u}$, since $\mathsf{I}\mathbf{u} = \mathbf{u} = 1.\mathbf{u}$, and the inverse of any stretching matrix also stretches $\mathbf{u}$:

$$\mathsf{A}\mathbf{u} = \lambda\mathbf{u} \quad \Longrightarrow \quad \mathsf{A}^{-1}\mathbf{u} = \lambda^{-1}\mathbf{u}.$$

Hence these matrices form a group.

(a) If

$$\begin{pmatrix} x & y \\ u & v \end{pmatrix} \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \lambda \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

then

$$x - y = \lambda, \qquad u - v = -\lambda.$$

Setting $x = a$, $v = b$, $\lambda = -c$ gives the desired form

$$\mathsf{A} = \begin{pmatrix} a & a+c \\ b+c & b \end{pmatrix}$$

where $c \neq 0$ and $a + b + c \neq 0$, else $\det \mathsf{A} = c(a+b+c) = 0$ and A is a singular matrix if either $c = 0$ or $a + b + c = 0$. Conversely, every matrix of this form is a stretch matrix of the given vector $\mathbf{u} = (1, -1)^T$.

(b) Near the identity, $\mathsf{A} = \mathsf{I} + \epsilon\mathsf{X}$ where $\epsilon \ll 1$, we may define real numbers α, β and γ such that

$$a = 1 + \epsilon\alpha, \qquad b = 1 + \epsilon\beta, \qquad c = -1 + \epsilon\gamma.$$

Thus

$$\mathsf{X} = \begin{pmatrix} \alpha & \alpha+\gamma \\ \beta+\gamma & \beta \end{pmatrix} = \alpha\mathsf{X}_1 + \beta\mathsf{X}_2 + \gamma\mathsf{X}_3,$$

where

$$\mathsf{X}_1 = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}, \quad \mathsf{X}_2 = \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix}, \quad \mathsf{X}_3 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

These matrices span the Lie algebra of the stretch group of **u**. Their commutators are calculated by matrix multiplication, $[\mathsf{X},\mathsf{Y}] = \mathsf{XY} - \mathsf{YX}$,

$$\begin{aligned}
&[\mathsf{X}_1, \mathsf{X}_2] = \mathsf{X}_1 - \mathsf{X}_2 \;: && C^1_{12} = -C^1_{21} = 1, \;\; C^2_{12} = -C^2_{21} = -1 \\
&[\mathsf{X}_1, \mathsf{X}_3] = \mathsf{X}_1 - \mathsf{X}_2 \;: && C^1_{13} = -C^1_{31} = 1, \;\; C^2_{13} = -C^2_{31} = -1 \\
&[\mathsf{X}_2, \mathsf{X}_3] = -\mathsf{X}_1 + \mathsf{X}_2 : && C^1_{23} = -C^1_{32} = -1, C^2_{23} = -C^2_{32} = 1
\end{aligned}$$

(c) Since $\mathsf{X}^2 = \mathsf{X}$, we have $(\mathsf{X}_1)^p = \mathsf{X}_1$ for all $p \geq 1$, whence

$$\mathrm{e}^{t\mathsf{X}_1} = \mathsf{I} + (\mathrm{e}^t - 1)\mathsf{X}_1 = \begin{pmatrix} \mathrm{e}^t & \mathrm{e}^t - 1 \\ 0 & 1 \end{pmatrix}.$$

For X_3 alternate powers are equal, $(\mathsf{X}_3)^2 = \mathsf{I}$, $(\mathsf{X}_3)^3 = \mathsf{X}_3$ etc. so that

$$\mathrm{e}^{t\mathsf{X}_3} = \begin{pmatrix} \cosh t & \sinh t \\ \sinh t & \cosh t \end{pmatrix}.$$

Problem 6.14 **Show that 2×2 trace-free matrices, having $\operatorname{tr} \mathsf{A} = A_{11} + A_{22} = 0$, form a Lie algebra with respect to bracket product $[\mathsf{A},\mathsf{B}] = \mathsf{AB} - \mathsf{BA}$.**

(a) Show that the following matrices form a basis of this Lie algebra

$$\mathsf{X}_1 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \qquad \mathsf{X}_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \qquad \mathsf{X}_3 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$$

and compute the structure constants for this basis.

(b) Compute the one-parameter subgroups $\mathrm{e}^{t\mathsf{X}_1}$, $\mathrm{e}^{t\mathsf{X}_2}$ and $\mathrm{e}^{t\mathsf{X}_3}$.

Solution: It is only necessary to show closure with respect bracket product,

$$\operatorname{tr}[\mathsf{A},\mathsf{B}] = \operatorname{tr}(\mathsf{AB} - \mathsf{BA}) = \sum_{i=1}^{2}\sum_{j=1}^{2} a_{ij}b_{ji} - b_{ij}a_{ji} = 0.$$

(a) The general matrix of this Lie algebra has the form

$$\mathsf{A} \begin{pmatrix} a & b \\ c & -a \end{pmatrix} = a\mathsf{X}_1 + b\mathsf{X}_2 + c\mathsf{X}_3$$

where

$$\mathsf{X}_1 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \qquad \mathsf{X}_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \qquad \mathsf{X}_3 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}.$$

These three matrices form a basis of the Lie algebra since they are l.i. Commutators are

$$\begin{aligned} [\mathsf{X}_1, \mathsf{X}_2] &= 2\mathsf{X}_2 \ : \qquad C^2_{12} = -C^2_{21} = 2 \\ [\mathsf{X}_1, \mathsf{X}_3] &= -2\mathsf{X}_3 : \qquad C^3_{13} = -C^3_{31} = -2 \\ [\mathsf{X}_2, \mathsf{X}_3] &= \mathsf{X}_1 \quad : \qquad C^1_{23} = -C^1_{32} = 1. \end{aligned}$$

(b) Since $(\mathsf{X}_1)^2 = \mathsf{I}$, $(\mathsf{X}_1)^3 = \mathsf{X}_1$ etc., we have

$$e^{t\mathsf{X}_1} = \mathsf{I} + t\mathsf{X}_1 + \frac{t^2}{2!}\mathsf{I} + \frac{t^3}{3!}\mathsf{X}_1 + \cdots = \begin{pmatrix} e^t & 0 \\ 0 & e^{-t} \end{pmatrix}.$$

Higher powers of X_2 and X_3 all vanish, $(\mathsf{X}_2)^2 = (\mathsf{X}_3)^2 = \mathsf{O}$ etc., so

$$\begin{aligned} e^{t\mathsf{X}_2} &= \mathsf{I} + t\mathsf{X}_2 = \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix} \\ e^{t\mathsf{X}_3} &= \mathsf{I} + t\mathsf{X}_3 = \begin{pmatrix} 1 & 0 \\ -t & 1 \end{pmatrix} \end{aligned}$$

Problem 6.15 Let $\mathcal{L}$ be the Lie algebra spanned by the three matrices

$$\mathsf{X}_1 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad \mathsf{X}_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \qquad \mathsf{X}_3 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Write out the structure constants for this basis, with respect to the usual matrix commutator bracket product.
Write out the three one-parameter subgroups $e^{t\mathsf{X}_i}$ generated by these basis elements, and verify in each case that they do in fact form a one-parameter group of matrices.

Solution:

$$[\mathsf{X}_1, \mathsf{X}_2] = \mathsf{X}_1\mathsf{X}_2 - \mathsf{X}_2\mathsf{X}_1 = \mathsf{X}_3, \qquad [\mathsf{X}_1, \mathsf{X}_3] = \mathsf{O}, \qquad [\mathsf{X}_2, \mathsf{X}_3] = \mathsf{O}$$

so that $C^3_{12} = -C^3_{21} = 1$ and all other $C^i_{jk} = 0$.
Since $(\mathsf{X}_1)^2 = (\mathsf{X}_2)^2 = (\mathsf{X}_3)^2 = \mathsf{O}$, it follows that

$$e^{t\mathsf{X}_1} = \mathsf{I} + t\mathsf{X}_1 = \begin{pmatrix} 1 & t & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad e^{t\mathsf{X}_2} = \mathsf{I} + t\mathsf{X}_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & t \\ 0 & 0 & 1 \end{pmatrix},$$

$$e^{t\mathsf{X}_3} = \mathsf{I} + t\mathsf{X}_3 = \begin{pmatrix} 1 & 0 & t \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

The one-parameter group properties follow from

$$e^{t\mathsf{X}_1} e^{s\mathsf{X}_1} = \mathsf{I} + t\mathsf{X}_1 + s\mathsf{X}_1 + ts(\mathsf{X}_1)^2 = \mathsf{I} + (t+s)\mathsf{X}_1 = e^{(t+s)\mathsf{X}_1},$$

and similarly

$$e^{t\mathsf{X}_2} e^{s\mathsf{X}_2} = e^{(t+s)\mathsf{X}_2}, \qquad e^{t\mathsf{X}_3} e^{s\mathsf{X}_3} = e^{(t+s)\mathsf{X}_3}.$$

Chapter 7.

Problem 7.1 **Show that the direct sum $V \oplus W$ of two vector spaces can be defined from the free vector space as $F(V \times W)/U$ where U is a subspace generated by all linear combinations of the form**

$$(av + bv', aw + bw') - a(v, w) - b(v', w').$$

Solution: The direct sum $V \oplus W$ is the vector space $V \times W$ with the operations of vector addition and scalar multiplication defined by

$$a(v, w) + b(v', w') = (av + bv', aw + bw'),$$

(see Section 3.4). Let U be the subspace of $F(V \times W)$,

$$U = \{(av + bv', aw + bw') - a(v, w) - b(v', w') \mid v, v' \in V,\ w, w' \in W\}.$$

For any pair of vectors $v \in V$, $w \in W$ let $[(v, w)]$ be the coset of (v, w) in $F(V \times W)/U$,

$$[(v, w)] = (v, w) + U.$$

Then

$$\begin{aligned} a[(v, w)] + b[(v', w')] &= a(v, w) + b(v', w') + U \\ &= a(v, w) + b(v', w') + (av + bv', aw + bw') \\ &\quad - a(v, w) - b(v', w') + U \\ &= (av + bv', aw + bw') + U \\ &= [(av + bv', aw + bw')] \end{aligned}$$

That is, the rule for addition and scalar multiplication of cosets is identical to the rule for representatives in $V \oplus W$.
To show that the correspondence $(v, w) \longrightarrow [(v, w])$ is one-to-one, note that if

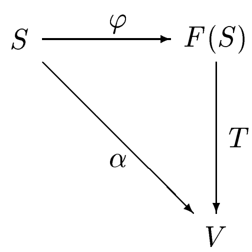
$$[(v, w)] = [(v', w')]$$

i.e. $(v, w) - (v', w') \in U$, then

$$(v - v', w - w') = (v, w) - (v', w') + \big((v - v', w - w') - (v, w) - (v', w')\big) \in U.$$

Since every non-trivial element of U is a sum of more than one linearly independent terms from $F(V \times W)$ the only term of the form (x, y) $(x \in V, y \in W)$ which belongs U is $(0, 0)$. Hence $v - v' = w - w' = 0$, or equivalently $v = v'$, $w = w'$.

Problem 7.2 **Prove the so-called *universal property* of free vector spaces. Let $\varphi : S \to F(S)$ be the map that assigns to any element $s \in S$ its characteristic function $\chi_s \in F(S)$. If V is any vector space and $\alpha : S \to V$ any map from S to V, show that there exists a unique linear map $T : F(S) \to V$ such that $\alpha = T \circ \varphi$, as depicted by the *commutative diagram***



Show that this process is reversible and may be used to define the free vector space on S as being the unique vector space $F(S)$ for which the bove commutative diagram holds.

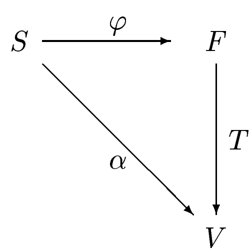
Solution: If $\alpha(s) = v \in V$ then

$$T(\chi_s) = T(\varphi(s)) = T \circ \varphi(s) = \alpha(s) = v.$$

This defines $T : F(S) \to V$ uniquely by the requirement of linearity:

$$T\Big(\sum_{i=1}^{k} a_i \chi_{s_i}\Big) = \sum_{i=1}^{k} a_i T(\chi_{s_i}) = \sum_{i=1}^{k} a_i \alpha(s_i).$$

We can define a free vector space over S as a vector space F and map $\varphi : S \to F$ such that for any map $\alpha : S \to V$ where V is a vector space there exists a unique linear map $T : F \to V$ such that $\alpha = T \circ \varphi$,



Let V be the vector space $F(S) = \{\chi_s \mid s \in S\}$ and $\alpha : S \to F(S)$ defined by $\alpha(s) = \chi_s$. Then T restricted to domain $\varphi(S) \subset F$ is one-to-one, for if $T(\varphi(s)) = T(\varphi(s'))$ then $\chi_s = \chi_{s'}$. Hence $s = s'$, so that $\varphi(s) = \varphi(s')$. Furthermore $L(\varphi(S)) = F$ since the map T is unique, for if $f \notin L(\varphi(S))$ then we can define $T(f) \in V$ arbitrarily and still maintain the commutative diagram. Thus T is a vector space isomorphism and $\varphi(s) = T^{-1}(\chi_s)$. This could therefore be used as a definition of the free vector space which provides essentially the same structure as $F(S)$ and the map α.

Problem 7.3 **Let $\mathcal{F}(V)$ be the free associative algebra over a vector space V.**

(a) Show that there exists a linear map $I : V \to \mathcal{F}(V)$ such that if $\mathcal{A}$ is any associative algebra over the same field $\mathbb{K}$ and $S : V \to \mathcal{A}$ a linear map, then there exists a unique algebra homomorphism $\alpha : \mathcal{F}(V) \to \mathcal{A}$ such that $S = \alpha \circ I$.

(b) Depict this property by a commutative diagram.

(c) Show the converse: any algebra $\mathcal{F}$ for which there is a map $I : V \to \mathcal{F}$ such that the commutative diagram holds for an arbitrary linear map S is isomorphic with the free associative algebra over V.

Solution: (a) The free associative algebra $\mathcal{F}(V)$ is defined as the infinite direct sum

$$\mathcal{F}(V) = V^{(0)} \oplus V^{(1)} \oplus V^{(2)} \oplus V^{(3)} \oplus \dots .$$

where $V^{(r)} = V \otimes V \otimes \cdots \otimes V$. The product rule on $\mathcal{F}(V)$ is defined by

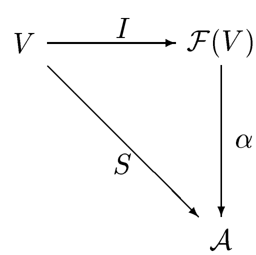
$$\underbrace{u_1 \otimes u_2 \otimes \cdots \otimes u_r}_{\in V^{(r)}} \underbrace{v_1 \otimes v_2 \otimes \cdots \otimes v_s}_{\in V^{(s)}} = \underbrace{u_1 \otimes \cdots \otimes u_r \otimes v_1 \otimes \cdots \otimes v_s}_{\in V^{(r+s)}} .$$

and extending to all of $\mathcal{F}(V)$ by linearity. Set $I(u) = u \in V^{(1)}$. Then we must have $\alpha(u) = \alpha(I(u)) = \alpha \circ I(u) = S(u)$. For products,

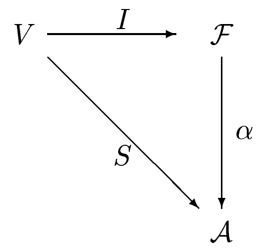
$$\alpha(u \otimes v) = \alpha(uv) = \alpha(u)\alpha(v) \qquad \text{etc.}$$

Continuing in this way α is uniquely defined on all elements of the form $u \otimes v \otimes w \otimes \cdots \otimes x$. Since every element of $\mathcal{F}(V)$ is a linear sum of elements of this form, α is uniquely defined on all of $\mathcal{V}$ by linearity.

(b)



(c) Let $\mathcal{F}$ be an associative and $I : V \to \mathcal{F}$ a linear map such that for an arbitrary linear map $S : V \to \mathcal{A}$ where $\mathcal{A}$ is any associative algebra the commutative diagram



holds a unique algebra homomorphism α. Set $\mathcal{A} = \mathcal{F}(V) = V^{(0)} \oplus V^{(1)} \oplus V^{(2)} \oplus \dots$ and $S(u) = u$ for all $u \in V$. An essentially identical argument to that in the second part of Problem 7.2 implies that α is an algebra isomorphism $\mathcal{F} \cong \mathcal{F}(V)$ and $I(u) = \alpha^{-1}(u)$.

Problem 7.4 Give a definition of quaternions as a factor algebra of the free algebra on a three-dimensional vector space.

Solution: Let V be a vector space generated by three vectors denoted i, j and k, and let $\mathcal{F}(V)$ be the free associative algebra over V. Set $\mathcal{N}$ to be the two-sided ideal generated by the elements to be thought of as zero under the rules of quaternion multiplication, namely

$$i \otimes i + 1, \quad j \otimes j + 1, \quad k \otimes k + 1, \quad i \otimes j - k, \quad j \otimes k - i, \quad k \otimes i - j.$$

The quaternion algebra can then be defined as $\mathcal{Q} = \mathcal{F}(V)/\mathcal{N}$.

Problem 7.5 The Clifford algebra $\mathcal{C}_g$ associated with an inner product space V with scalar product $g(u,v) \equiv u \cdot v$ can be defined in the following way. Let $\mathcal{F}(V)$ be the free associative algebra on V and $\mathcal{C}$ the two-sided ideal generated by all elements of the form

$$A \otimes (u \otimes v + v \otimes u - 2g(u,v)1) \otimes B \qquad (A, B \in \mathcal{F}(V)).$$

The Clifford algebra in question is now defined as the factor space $\mathcal{F}(V)/\mathcal{C}$. Verify that this algebra is isomorphic with the Clifford algebra as defined in Section 6.3, and could serve as a basis-independent definition for the Clifford algebra associated with a real inner product space.

Solution: Let $[T] = T + \mathcal{C}$ be the coset in $\mathcal{F}(V)/\mathcal{C}$ of the tensor T in $\mathcal{F}(V)$. We then have

$$\begin{aligned} [T][S] &= (T + \mathcal{C})(S + \mathcal{C}) \\ &= T \otimes S + \mathcal{C} \otimes S + T \otimes \mathcal{C} + \mathcal{C} \otimes \mathcal{C} \\ &= T \otimes S + \mathcal{C} = [T \otimes S]. \end{aligned}$$

In particular

$$[u][v]+[v][u]=[u\otimes v+v\otimes u]=2g(u,v)[1],$$

and for any basis $e_1,\dots e_n$ of V

$$[e_i][e_j]+[e_j][e_i]=2g(e_i,e_j)[1]=2g_{ij}[1].$$

Let $\varphi:\mathcal{F}(V)\to\mathcal{C}_g$ be the map defined by

$$\varphi(T)=T^{i_1\dots i_r}e_{i_1}\dots e_{i_r}\qquad\text{where}\qquad T=T^{i_1\dots i_r}e_{i_1}\otimes\dots\otimes e_{i_r}$$

and extend by linearity to all direct sums of tensors of different orders in $\mathcal{F}(V)$. This map is an algebra homomorphism,

$$\varphi(TS)=\varphi(T\otimes S)=\varphi(T)\varphi(S)$$

and passes to a map $\varphi:\mathcal{F}(V)/\mathcal{C}\to\mathcal{C}_g$ by setting $\varphi([T])=\varphi(T)$. This definition is independent of the choice of representative T in the coset $[T]=T+\mathcal{C}$ for if $S\in\mathcal{C}$ then S has the form $A\otimes(u\otimes v+v\otimes u-2g(u,v)1)\otimes B$ and

$$\varphi(S)=\varphi(A)\varphi(u\otimes v+v\otimes u-2g(u,v)1)\varphi(B)=0$$

for

$$\varphi(u\otimes v+v\otimes u-2g(u,v)1)=\varphi(u)\varphi(v)+\varphi(v)\varphi(u)-2g(u,v)1=0.$$

The map φ is one-to-one, for if $\varphi([T])=0$ the only way in which $\varphi(T)$ can vanish is if it is a sum of terms of the form $P(uv+vu-2g(u,v))Q$, as $uv+vu=2g(u,v)$ is the only identity assumed to hold in Clifford algebras. In other words T must be a sum of terms of the form $A\otimes(u\otimes v+v\otimes u-2g(u,v)1)\otimes B$ for tensors A and B; i.e. $T\in\mathcal{C}$.

Problem 7.6 **Show that E_2 is a basis of $\Lambda^2(V)$. In outline: define the maps $\varepsilon^{kl}:V^2\to\mathbb{R}$ by**

$$\varepsilon^{kl}(u,v)=\varepsilon^k(u)\varepsilon^l(v)-\varepsilon^l(u)\varepsilon^k(v)=u^kv^l-u^lv^k.$$

Extend by linearity to the tensor space $V^{(2,0)}$ and show there is a natural passage to the factor space, $\hat{\varepsilon}^{kl}:\Lambda^2(V)=V^{(2,0)}/\mathcal{S}^2\to\mathbb{R}$. If a linear combination from E_2 were to vanish,

$$\sum_{i<j}A^{ij}e_i\wedge e_j=0,$$

apply the map $\hat{\varepsilon}^{kl}$ to this equation, to show that all coefficients A^{ij} must vanish separately.

Indicate how the argument may be extended to show that if $r\le n=\dim V$ then E_r is a basis of $\Lambda^r(V)$.

Solution: Any bilinar function $f : V^2 \to \mathbb{R}$ passes naturally to a linear function $f : \mathcal{F}(V^2) \to \mathbb{R}$, which vanishes by linearity on the subspace U of $F(V^2)$ generated by elements of the form $(av + bv',\, w) - a(v, w) - b(v', w)$ and $(v,\, aw + bw') - a(v, w) - b(v, w')$. Hence we may define a map $\bar{f} : V^{(2,0)} \to \mathbb{R}$ by setting

$$\bar{f}(\sum_r a^r u_r \otimes v_r) = \sum_r a^r f((u_r, v_r)),$$

which is independent of the representative $\sum_r a^r(u_r, v_r) \in \mathcal{F}(V^2)$. The linear functions $\varepsilon^{kl} : V^2 \to \mathbb{R}$ defined by

$$\varepsilon^{kl}(u \otimes v) \equiv \varepsilon^{kl}(u, v) = \varepsilon^k(u)\varepsilon^l(v) - \varepsilon^l(u)\varepsilon^k(v) = u^k v^l - u^l v^k$$

vanish on the subspace $\mathcal{S}^2$ generated by tensors of the form $u \otimes v + v \otimes u$, since

$$\varepsilon^{kl}(u \otimes v) = \varepsilon^k(u)\varepsilon^l(v) - \varepsilon^l(u)\varepsilon^k(v) + \varepsilon^k(v)\varepsilon^l(u) - \varepsilon^l(v)\varepsilon^k(u) = 0.$$

Hence they pass to linear functions $\hat{\varepsilon}^{kl} : \Lambda^2(V) = V^{(2,0)}/\mathcal{S}^2 \to \mathbb{R}$, by setting

$$\hat{\varepsilon}^{kl}([A]) = \varepsilon^{kl}(A) \qquad \text{where } A \in V^{(2,0)} \text{ and } [A] = A + \mathcal{C}^2.$$

In particular

$$\hat{\varepsilon}^{kl}(e_i \wedge e_j) = \hat{\varepsilon}^{kl}([e_i \otimes e_j]) = \varepsilon^{kl}(e_i \otimes e_j) = \delta^k_i \delta^l_j - \delta^l_i \delta^k_j.$$

If

$$\sum_{i<j} A^{ij} e_i \wedge e_j = 0,$$

then applying the map $\hat{\varepsilon}^{kl}$ with $k < l$ to this equation gives

$$0 = \sum_{i<j} A^{ij}(\delta^k_i \delta^l_j - \delta^l_i \delta^k_j) = \sum_{i<j} A^{ij} \delta^k_i \delta^l_j = A^{kl}.$$

Hence the tensors $E_2 = \{e_{kl} = e_k \wedge e_l \mid k < l\}$ are linearly independent are form a basis of $\Lambda^2(V)$.

The argument that $E_r = \{e_{i_1 i_2 \dots i_r} \mid i_1 < i_2 < \dots < i_r\}$ is a basis of $\Lambda^r(V)$ is essentially identical. In this case we define the alternating maps $\hat{\varepsilon}^{k_1 k_2 \dots k_r} : \Lambda^r(V) \to \mathbb{R}$ from

$$\varepsilon^{k_1 k_2 \dots k_r}(u_{i_1} \otimes u_{i_2} \otimes \dots \otimes u_{i_r}) = \sum_\pi (-1)^\pi \varepsilon^{k_{\pi(1)}}(u_{i_1}) \varepsilon^{k_{\pi(2)}}(u_{i_2}) \dots \varepsilon^{k_{\pi(r)}}(u_{i_r})$$

where the sum is over all permutations π of $12 \dots r$. If we have an equation

$$\sum_{i_1 < \dots i_r} A^{i_1 i_2 \dots i_r} e_{i_1} \wedge e_{i_2} \wedge \dots \wedge e_{i_r} = 0$$

then application of the linear function $\hat{\varepsilon}^{k_1 k_2 \dots k_r}$ where $k_1 < k_2 \dots < k_r$ results in all coefficients $A^{i_1 i_2 \dots i_r} = 0$.

Problem 7.7 **Let $\bar{T}$ be the linear map defined by a covariant tensor T of degree 2 as in Example 7.6. If $\{e_i\}$ is a basis of V and $\{\varepsilon^j\}$ the dual basis, define the matrix of components of $\bar{T}$ with respect to these bases as $[\bar{T}_{ji}]$ where**

$$\bar{T}(e_i) = \bar{T}_{ji}\varepsilon^j .$$

Show that the components of the tensor T in this basis are identical with the components as a map, $T_{ij} = \bar{T}_{ij}$.

Similarly if S is a contravariant tensor of degree 2 and $\bar{S}$ the linear map defined in Example 7.7, show that the components $\bar{S}^{ij}$ are identical with the tensor components S^{ij}.

Solution: Set $u = e_i$, $v = e_j$ in Eq. (7.10) gives

$$\langle \bar{T}e_j, e_i\rangle = T(e_i, e_j) = T_{ij}$$

and substituting $\bar{T}(e_j) = \bar{T}_{kj}\varepsilon^k$,

$$\langle \bar{T}e_j, e_i\rangle = \langle \bar{T}_{kj}\varepsilon^k, e_i\rangle = \bar{T}_{kj}\langle \varepsilon^k, e_i\rangle = \bar{T}_{kj}\delta^k_i = \bar{T}_{ij}.$$

Hence $T_{ij} = \bar{T}_{ij}$, as required.

Similarly if $\bar{S}V^* \to V$ is defined by $\langle \bar{S}\rho, \omega\rangle = S(\omega,\, \rho)$, having components $\bar{S}\varepsilon^j = \bar{S}^{kj}e_k$, then

$$S^{ij} = S(\varepsilon^i,\, \varepsilon^j) = \langle \bar{S}\varepsilon^j, \varepsilon^i\rangle = \langle \bar{S}^{kj}e_k, \varepsilon^i\rangle = \bar{S}^{kj}\delta^i_{\ k} = \bar{S}^{ij}.$$

Problem 7.8 **Show that every tensor R of type $(1,1)$ defines a map $\tilde{R} : V^* \to V^*$ by**

$$\langle \tilde{R}\omega, u\rangle = R(\omega, u)$$

and show that for a natural definition of components of this map, $\tilde{R}^k_{\ i} = R^k_{\ i}$.

Solution: The map $\tilde{R}\omega : V \to \mathbb{R}$ defined by $\tilde{R}\omega(u) = R(\omega, u)$ is a linear functional since

$$\tilde{R}\omega(au + bv) = R(\omega, au + bv) = aR(\omega, u) + bR(\omega, v) = a\tilde{R}\omega(u) + b\tilde{R}\omega(v).$$

Hence $\tilde{R}$ defines a map $\tilde{R} : V^* \to V^*$, and we may write $\langle \tilde{R}\omega, u\rangle \equiv \tilde{R}\omega(u)$. For a basis $\{e_i\}$ and dual basis $\{\varepsilon^j\}$, define the components of $\tilde{R}$ by

$$\tilde{R}\varepsilon^k = \tilde{R}^k_{\ j}\varepsilon^j .$$

Then

$$\tilde{R}^k_{\ i} = R(\varepsilon^k,\, e_i) = \langle \tilde{R}\varepsilon^k, e_i\rangle = \tilde{R}^k_{\ j}\langle \varepsilon^j, e_i\rangle = \tilde{R}^k_{\ j}\delta^j_{\ \ i} = \tilde{R}^k_{\ i}.$$

Problem 7.9 **Show that the definition of tensor product of two vectors $u \times v$ given in Eq.(7.11) agrees with that given in Section 7.1 after relating the two concepts of tensor by isomorphism.**

Solution: Using the dual representation of tensor spaces, and setting $W = V$ in Eq. (7.3), we have

$$V^{(2,0)} = V \otimes V \cong (V^*, V^*)^*,$$

where isomorphism is defined by

$$A(\rho, \varphi) = \sum_r a^r \rho(v_r)\varphi(w_r).$$

for $A = \sum_r a^r v_r \otimes w_r$. Setting $r = 1$, $a^1 = 1$, $v_1 = v$, $w_1 = w$ we have

$$v \otimes w(\rho, \varphi) = \rho(v)\varphi(w).$$

which agrees with the definition of tensor product in Eq. (7.11).

Problem 7.10 **Let e_1, e_2 and e_3 be a basis of a vector space V and $e'_{i'}$ a second basis given by**

$$\begin{aligned} e'_1 &= e_1 - e_2, \\ e'_2 &= e_3, \\ e'_3 &= e_1 + e_2. \end{aligned}$$

(a) Display the transformation matrix $\mathsf{A}' = [A'^{i}_{\ i'}]$.

(b) Express the original basis e_i in terms of the $e'_{i'}$ and write out the transformation matrix $\mathsf{A} = [A^{j'}_{\ j}]$.

(c) Write the old dual basis ε^i in terms of the new dual basis $\varepsilon'^{i'}$ and conversely.

(d) What are the components of the tensors $T = e_1 \otimes e_2 + e_2 \otimes e_1 + e_3 \otimes e_3$ and $S = e_1 \otimes \varepsilon^1 + 3e_1 \otimes \varepsilon^3 - 2e_2 \otimes \varepsilon^3 - e_3 \otimes \varepsilon^1 + 4e_3 \otimes \varepsilon^2$ in terms of the basis e_i and its dual basis?

(e) What are the components of these tensors in terms of the basis $e'_{i'}$ and its dual basis?

Solution: (a) The components of the matrix A' are read off from the definition Eq. (7.26), $e'_{j'} = A'^{j}_{\ j'} e_j$.

$$\mathsf{A}' = [A'^{i}_{\ i'}] = \begin{pmatrix} 1 & 0 & 1 \\ -1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}.$$

(b) Solving the linear equations for e_1, e_2 and e_3,

$$\begin{aligned} e_1 &= \tfrac{1}{2}(e'_1 + e'_3) \\ e_2 &= \tfrac{1}{2}(-e'_1 + e'_3) \\ e_3 &= e'_2, \end{aligned}$$

and, again reading off from Eq. (7.26),

$$\mathsf{A} = [A^{j'}_{j}] = \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} & 0 \\ 0 & 0 & 1 \\ \frac{1}{2} & \frac{1}{2} & 0 \end{pmatrix}.$$

(c) From Eq. (7.28) we can read off

$$\begin{aligned} \varepsilon^i = A'^{i}{}_{j'}\varepsilon'^{j'} &\implies \varepsilon^1 = \varepsilon'^1 + \varepsilon'^3, \quad \varepsilon^2 = -\varepsilon'^1 + \varepsilon'^3, \quad \varepsilon^3 = \varepsilon'^2, \\ \varepsilon'^{i'} = A^{i'}_{j}\varepsilon^j &\implies \varepsilon'^1 = \tfrac{1}{2}(\varepsilon^1 - \varepsilon^2), \quad \varepsilon'^2 = \varepsilon^3, \quad \varepsilon'^3 = \tfrac{1}{2}(\varepsilon^1 + \varepsilon^2). \end{aligned}$$

(d) The components of the tensors T and S are read off as the coefficients of the various $e_i \otimes e_j$ etc.,

$$[T^{ij}] = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad [S_{ij}] = \begin{pmatrix} 1 & 0 & 3 \\ 0 & 0 & -2 \\ -1 & 4 & 0 \end{pmatrix}.$$

(e) Substituting for e_i from (b) we find

$$T = -\frac{1}{2}e'_1 \otimes e'_1 + \frac{1}{2}e'_3 \otimes e'_3 + e'_2 \otimes e'_2,$$

and the components are

$$[T'_{ij}] = \begin{pmatrix} -\frac{1}{2} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \frac{1}{2} \end{pmatrix}.$$

Similarly, substituting (b) and (c) into S, a rather longer computation gives

$$\begin{aligned} S &= \frac{1}{2}(e'_1 + e'_3) \otimes (\varepsilon'^1 + \varepsilon'^3) + \dots \\ &= \frac{1}{2}\left(e'_1 \otimes \varepsilon'^1 + e'_3 \otimes \varepsilon'^1 + e'_1 \otimes \varepsilon'^3 + e'_3 \otimes \varepsilon'^3\right) + \dots \\ &= \frac{1}{2}e'_1 \otimes \varepsilon'^1 + \frac{5}{2}e'_1 \otimes \varepsilon'^2 + \frac{1}{2}e'_1 \otimes \varepsilon'^3 \\ &\quad - 5e'_2 \otimes \varepsilon'^1 + 3e'_2 \otimes \varepsilon'^3 \\ &\quad \frac{1}{2}e'_3 \otimes \varepsilon'^1 + \frac{1}{2}e'_2 \otimes \varepsilon'^2 + \frac{1}{2}e'_3 \otimes \varepsilon'^3 \end{aligned}$$

and the components of S in the primed basis are

$$[S'^{i}{}_{j}] = \begin{pmatrix} \frac{1}{2} & \frac{5}{2} & \frac{1}{2} \\ -5 & 0 & 3 \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \end{pmatrix}.$$

Problem 7.11 **Let V be a vector space of dimension 3, with basis e_1, e_2, e_3. Let T be the contravariant tensor of rank 2 whose components in this basis are $T^{ij} = \delta^{ij}$, and let S be the covariant tensor of rank 2 whose components are given by $S_{ij} = \delta_{ij}$ in this basis. In a new basis e'_1, e'_2, e'_3 defined by**

$$\begin{aligned} e'_1 &= e_1 + e_3 \\ e'_2 &= 2e_1 + e_2 \\ e'_3 &= 3e_2 + e_3 \end{aligned}$$

calculate the components $T'^{i'j'}$ and $S'_{i'j'}$.

Solution: Solving for e_1, e_2 and e_3, we find

$$\begin{aligned} e_1 &= \frac{1}{7}\big(e'_1 + 3e'_2 - e'_3\big) \\ e_2 &= \frac{1}{7}\big(-2e'_1 + e'_2 + 2e'_3\big) \\ e_3 &= \frac{1}{7}\big(6e'_1 - 3e'_2 + e'_3\big) \end{aligned}$$

so that

$$[A^{i'}_{i}] = \frac{1}{7}\begin{pmatrix} 1 & -2 & 6 \\ 3 & 1 & -3 \\ -1 & 2 & 1 \end{pmatrix}.$$

Using Eq. (7.28) we can read off the transformed dual basis

$$\begin{aligned} \varepsilon'^1 &= \frac{1}{7}\big(\varepsilon^1 - 2\varepsilon^2 + 6\varepsilon^3\big) \\ \varepsilon'^2 &= \frac{1}{7}\big(3\varepsilon^1 + \varepsilon^2 - 3\varepsilon^3\big) \\ \varepsilon'^1 &= \frac{1}{7}\big(-\varepsilon^1 + 2\varepsilon^2 + \varepsilon^3\big) \end{aligned}$$

The transformed components of T are $T'^{ij} = T(\varepsilon'^i, \varepsilon'^j)$, so that $T'^{11} = \frac{1}{49}(1+4+36)$ etc.,

$$[T'^{ij}] = \frac{1}{49}\begin{pmatrix} 41 & -17 & 1 \\ -17 & 19 & -4 \\ 1 & -4 & 6 \end{pmatrix}.$$

Similarly,

$$[S'_{ij}] = [S(e'_i,\, e'_j)] = \begin{pmatrix} 2 & 2 & 1 \\ 2 & 5 & 3 \\ 1 & 3 & 10 \end{pmatrix}.$$

Problem 7.12 **Let $T : V \to V$ be a linear operator on a vector space V. Show that its components T^i_j given by Eq. (3.6) are those of the tensor $\hat{T}$ defined by**

$$\hat{T}(\omega, v) = \langle \omega, Tv \rangle.$$

Prove that they are also the components with respect to the dual basis of a linear operator $T^* : V^* \to V^*$ defined by

$$\langle T^*\omega, v \rangle = \langle \omega, Tv \rangle.$$

Show that tensors of type (r, s) are in one-to-one correspondence with linear maps from $V^{(s,0)}$ to $V^{(r,0)}$, or equivalently from $V^{(0,r)}$ to $V^{(0,s)}$.

Solution: The components of the operator T are defined by Eq. (3.6), $Te_j = T^k_j e_k$. Thus the components of the tensor $\hat{T}$ are given by (see Problem 7.7 for a similar discussion),

$$\hat{T}^i_j = \hat{T}(\varepsilon^i, e_j) = \langle \varepsilon^i, Te_j \rangle = \langle \varepsilon^i, T^k_j e_k \rangle = T^k_j \delta^i_k = T^i_j.$$

The dual operator T^* has components defined by $T^*\varepsilon^i = T^{*i}{}_k \varepsilon^k$ and we have

$$\langle T^*\varepsilon^i, e_j \rangle = \langle T^{*i}{}_k \varepsilon^k, e_j \rangle = T^{*i}_k \delta^k_j = T^{*i}_j,$$

while also, from the above

$$\langle T^*\varepsilon^i, e_j \rangle = \langle \varepsilon^i, Te_j \rangle = T^i_j = \hat{T}^i_j.$$

If $\bar{T} : V^{(s,0)} \to V^{(r,0)}$ is a linear map then for any vectors $u_1, u_2, \dots u_r \in V$ the image $\bar{T}(u_1 \otimes u_2 \otimes \cdots \otimes u_s)$ is a tensor of type $(r, 0)$, and we can a tensor T of type (r, s) by

$$T(\omega_1, \dots, \omega_r, u_1, u_2, \dots, u_s) = \bar{T}(u_1 \otimes u_2 \otimes \cdots \otimes u_s)\,(\omega_1, \dots, \omega_r).$$

It is only necessary to check that T is multilinear in all arguments — trivial for all ω_i and very simple for any u_j since

$$u_1 \otimes \cdots \otimes (au_j + bv_j) \cdots \otimes u_s = au_1 \otimes \cdots \otimes u_j \cdots \otimes u_s + bu_1 \otimes \cdots \otimes v_j \cdots \otimes u_s$$

and $\bar{T}$ is a linear map.
The map $\bar{T} \to T$ is one-to-one for if $\bar{T} = 0$ then $\bar{T}(u_1 \otimes u_2 \otimes \cdots \otimes u_s)$ vanishes for all vectors $u_1, \dots, u_s$. Hence the tensor $T = 0$ for

$$T(\omega_1, \dots, \omega_r, u_1, u_2, \dots, u_s) = 0$$

for all $\omega_i \in V^*$ and all $u_j \in V$. Finally the map is onto, for if $e_1, \dots, e_n$ is any basis of V then define the map $\bar{T}$ by setting

$$\bar{T}(e_{i_1} \otimes e_{i_2} \otimes \cdots \otimes e_{i_s})\,(\omega_1, \dots, \omega_r) = T(\omega_1, \dots, \omega_r, e_{i_1} \otimes e_{i_2} \otimes \cdots \otimes e_{i_s})$$

and extend to all tensor $A \in V^{(0,s)}$ by linearity,

$$\bar{T}(A^{i_1 i_2 \dots i_s} e_{i_1} \otimes e_{i_2} \otimes \cdots \otimes e_{i_s}) = A^{i_1 i_2 \dots i_s} e_{i_1} \bar{T}(e_{i_1} \otimes e_{i_2} \otimes \cdots \otimes e_{i_s}).$$

Thus the map $\bar{T} \to T$ is a one-to-one correspondence.
Similarly a linear map $\tilde{T} : V^{(0,r)} \to V^{(0,s)}$ defines a tensor $T \in V^{(r,s)}$ by

$$T(\omega_1, \dots, \omega_r, u_1, u_2, \dots, u_s) = \tilde{T}(\omega_1 \otimes \omega \cdots \otimes \omega_r)\,(u_1, u_2, \dots, u_s),$$

and an essentially identical argument to that above shows that this is a one-to-one correspondence.

Problem 7.13 **Let $T : V \to V$ be a linear operator on a vector space V. Show that its components T^i_j defined through Eq. (3.6) transform as those of a tensor of rank (1,1) under an arbitrary basis transformation.**

Solution: The components of the operator T are defined by $Te_j = T^i_j e_i$. Hence, substituting from the transformation equation (7.26),

$$\begin{aligned} Te'_{j'} &= T(A'^j_{\ j'} e_j) \\ &= A'^j_{\ j'} T e_j \\ &= A'^j_{\ j'} T^i_j e_i \\ &= A'^j_{\ j'} T^i_j A^{i'}_i e'_{i'} \\ &= T'^{i'}_{\ j'} e'_{i'} \end{aligned}$$

where

$$T'^{i'}_{\ j'} = T^i_j A^{i'}_i A'^j_{\ j'}$$

which is precisely the transformation law of a tensor of type $(1,1)$.

Problem 7.14 **Show directly from Eq. (7.14) and the transformation law of components g_{ij}**

$$g'_{j'k'} = g_{jk} A'^j_{\ j'} A'^k_{\ k'},$$

that the components of an inverse metric tensor g^{ij} transform as a contravariant tensor of degree 2,

$$g'^{i'k'} = A^{i'}_l g^{lk} A^{k'}_k.$$

Solution: From Eq. (7.14), $g^{kj} g_{ji} = \delta^k_i$, in the primed coordinates we must have

$$g'^{i'm'} g'_{m'k'} = \delta^{i'}_{\ k'}$$

Hence

$$g'^{i'm'} A'^{m}_{m'} A'^{k}_{k'} g_{mk} = \delta^{i'}_{k'}$$

and multiplying through by $A^{k'}_{a}$ gives

$$g'^{i'm'} A'^{m}_{m'} A'^{k}_{k'} A^{k'}_{a} g_{mk} = A^{i'}_{a}.$$

Hence

$$g'^{i'm'} A'^{m}_{m'} \delta^{k}_{a} g_{mk} = g'^{i'm'} A'^{m}_{m'} g_{ma} = A^{i'}_{a}.$$

Multiplying equation on the right by g^{ab},

$$g'^{i'm'} A'^{m}_{m'} \delta^{b}_{m} = g'^{i'm'} A'^{b}_{m'} = A^{i'}_{a} g^{ab}.$$

Finally, multiplying through the equation on the right by $A^{j'}_{b}$ gives

$$g'^{i'm'} \delta^{j'}_{m'} = A^{i'}_{a} g^{ab} A^{j'}_{b},$$

i.e.

$$g'^{i'j'} = A^{i'}_{i} A^{j'}_{j} g^{ij}.$$

Problem 7.15 **Let g_{ij} be the components of an inner product with respect to a basis u_1, u_2, u_3**

$$g_{ij} = u_i \cdot u_j = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$

(a) Find an orthonormal basis of the form $e_1 = u_1$, $e_2 = u_2$, $e_3 = au_1 + bu_2 + cu_3$ such that $a > 0$, and find the index of this inner product.

(b) If $v = u_1 + \frac{1}{2}u_3$ find its lowered components v_i.

(c) Express v in terms of the orthonormal basis found above, and write out its lowered components with respect to that basis.

Solution: (a) $e_1 \cdot e_1 = e_2 \cdot e_2 = 1$ and $e_1 \cdot e_2 = 0$. If $e_3 = au_1 + bu_2 + cu_3$ then

$$\begin{aligned} e_3 \cdot e_1 &= a + c = 0 \\ e_3 \cdot e_2 &= b = 0 \\ e_3 \cdot e_3 &= a^2 + b^2 + 2ac = a^2 + 2ac = \pm 1 \end{aligned}$$

Hence $c = -a$ and $-a^2 = \pm 1$. The only possibility is $a = 1$, $c = -1$, so that $e_3 = u_1 - u_3$, and $e_3 \cdot e_3 = -1$. The index is $2 - 1 = 1$.

(b) If $v = u_1 + \frac{1}{2}u_3$ then $v^1 = 1$, $v^2 = 0$ and $v^3 = \frac{1}{2}$. Hence lowered components $v_i = g_{ij}v^j$ are

$$v_1 = g_{1i}v^i = v^1 + v^3 = \frac{3}{2}, \qquad v_2 = v^2 = 0, \qquad v_3 = v^1 = 1.$$

(c) Since $u_3 = e_1 - e_3$, we have $v = e_1 + \frac{1}{2}(e_1 - e_3) = \frac{3}{2}e_1 - \frac{1}{2}e_3$ and lowered components are

$$v_1 = \frac{3}{2}, \qquad v_2 = 0, \qquad v_3 = \frac{1}{2}.$$

Problem 7.16 **Let g be a metric tensor on a vector space V and define T to be the tensor**

$$T = ag^{-1} \otimes g + \delta \otimes u \otimes \omega$$

where u is a non-zero vector in V and ω is a covector.

(a) Write out the components $T^{ij}{}_{kl}$ of the tensor T.

(b) Evaluate the components of the following four contractions

$$A = C_1^1 T, \quad B = C_2^1 T, \quad C = C_1^2 T, \quad D = C_2^2 T,$$

and show that $B = C$.

(c) Show that $D = 0$ iff $\omega(u) = -a$. Hence show that if $T^{ij}_{kl} u^l u_j = 0$, then $D = 0$.

(d) Show that if $n = \dim V > 1$ then $T^{ij}_{kl} u^l u_j = 0$ if and only if $a = \omega(u) = 0$ or $u_i u^i = 0$.

Solution: (a) If $u = u^i e_i$ and $\omega = w_j \varepsilon^j$ then

$$T^{ij}{}_{kl} \equiv T^{ij}_{kl} = ag^{ij} g_{kl} + \delta^i{}_k u^j w_l.$$

(b)

$$\begin{aligned}
A = C_1^1 T &\implies A^j_l = T^{ij}_{il} = a\delta^j_l + nu^j w_l \\
B = C_2^1 T &\implies B^j_k = T^{ij}_{ki} = a\delta^j_k + u^j w_k \\
C = C_1^2 T &\implies C^i_l = T^{ij}_{jl} = a\delta^i_l + u^i w_l \\
D = C_2^2 T &\implies D^i_k = T^{ij}_{kj} = a\delta^i_k + \delta^i_k u^j w_j.
\end{aligned}$$

(c) $D = \delta^i_k(a + u^j w_j) = 0$ iff $a = -u^j w_j = -\omega(u)$.
Hence

$$\begin{aligned}
T^{ij}_{kl} u^l u_i = au_k u^j + u^l w_l u_k u^j &= u_k u^j (a + u^l w_l) = 0 \\
&\implies a + u^l w_l = 0 \implies D = 0.
\end{aligned}$$

(d) Since

$$T^{ij}_{kl} u^l u_j = au^i u_k + u^j w_j u^l u_l \delta^i_k$$

it is immediate that $a = u^j w_j = 0$ implies that $T^{ij}_{kl} u^l u_j = 0$. Conversely if $T^{ij}_{kl} u^l u_j = 0$ then contracting the equation over the free indices ik,

$$au^i u_i + nu^j w_j u^l u_l = u^i u_i (a + nu^j w_j) = 0.$$

On the other hand multiplying the equation $T^{ij}_{kl}u^l u_j = 0$ by u_i gives

$$u^i u_i u_k (a + u^j w_j) = 0.$$

Comparing these two equations we have either $a = w_j u^j = 0$ or $u^i u_i$.

Problem 7.17 On a vector space V of dimension n let T be a tensor of rank $(1,1)$, S a symmetric tensor of rank $(0,2)$ and δ the usual 'invariant tensor' of rank $(1,1)$. Write out the components $R^{ij}{}_{klmr}$ of the tensor

$$R = T \otimes S \otimes \delta + S \otimes \delta \otimes T + \delta \otimes T \otimes S.$$

Perform the contraction of this tensor over i and k, using any available contraction properties of δ^i_j. Perform a further contraction over the indices j and r.

Solution:

$$\begin{aligned}
R^{ij}{}_{klmr} &= T^i_k S_{lm} \delta^j_r + S_{kl} \delta^i_m T^j_r + \delta^i_k T^j_l S_{mr}. \\
(C^1_1 R)^j{}_{lmr} &= T^i_i S_{lm} \delta^j_r + S_{lm} T^j_r + n T^j_l S_{mr}. \\
(C^1_3 C^1_1 R)_{lm} &= T^i_i S_{lm} n + S_{lm} T^j_j + n T^j_l S_{mj} \\
&= (n+1) T^i_i S_{lm} + n T^j_l S_{mj}.
\end{aligned}$$

Problem 7.18 Show that covariant symmetric tensors of rank 2, satisfying $T_{ij} = T_{ji}$, over a vector space V of dimension n form a vector space of dimension $n(n+1)/2$.

(a) A tensor S of type $(0,r)$ is called *totally symmetric* if $S_{i_1 i_2 \dots i_r}$ is left unaltered by any interchange of indices. What is the dimension of the vector space spanned by the totally symmetric tensors on V?

(b) Find the dimension of the vector space of covariant tensors of rank 3 having the cyclic symmetry

$$T(u,v,w) + T(v,w,u) + T(w,u,v) = 0.$$

Solution: If S and T are covariant symmetric tensors of rank 2 then so is $S + aT$, for

$$S_{ij} + aT_{ij} = S_{ji} + aT_{ji}.$$

Hence covariant symmetric tensors of rank 2 form a vector subspace of $V^{(0,2)}$. Given a basis $e_1, \dots, e_n$ of V, dual basis $\varepsilon^1, \dots \varepsilon^n$ of V^*, every symmetric tensor may be written in the form

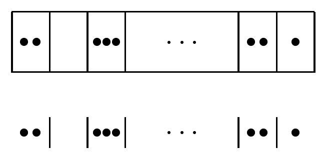
$$S = \sum_{i<j}^{n} S_{ij}(\varepsilon^i \otimes \varepsilon^j + \varepsilon^j \otimes \varepsilon^i) + \sum_{i=1}^{n} S_{ii}\varepsilon^i \otimes \varepsilon^i.$$

The tensors $\varepsilon^i \otimes \varepsilon^j + \varepsilon^j \otimes \varepsilon^i$ $(i < j)$ and $\varepsilon^i \otimes \varepsilon^i$ are linearly independent and thus form a basis of this space. Its dimension is therefore $n + n(n-1)/2 = n(n+1)/2$.

(a) Similarly, the space of all totally symmetric tensors on V has a basis

$$\varepsilon^{i_1} \otimes \varepsilon^{i_2} \otimes \cdots \otimes \varepsilon^{i_r} + \varepsilon^{i_2} \otimes \varepsilon^{i_1} \otimes \cdots \otimes \varepsilon^{i_r} + \dots$$

where the sum is taken over all permutations of indices $i_1 i_2 \dots i_r$. In each such basis tensor we can set the first term to have increasing order of indices, $i_1 \le i_2 \le \cdots \le i_r$. Hence the dimension of the space of symmetric tensors is equal to the number of ascending sequences of this type from the numbers $1, \dots n$. This is equal to the number of ways of distributing r balls in n boxes, with no limit placed on the number of balls in any one box. An equivalent problem is to find the number of ways of distributing r balls and $(n-1)$ box partitions in $n + r - 1$ places:



The dimension of the space of symmetric tensors is therefore

$$\binom{n+r-1}{r} = \frac{n(n+1)\dots(n+r-1)}{r!}.$$

(c) Let U be the subspace of tensor of type $(0,3)$ spanned by elements of the form

$$E^{ijk} = \varepsilon^i \otimes \varepsilon^j \otimes \varepsilon^k + \varepsilon^j \otimes \varepsilon^k \otimes \varepsilon^i + \varepsilon^k \otimes \varepsilon^i \otimes \varepsilon^j.$$

The space of tensors for which $T(u,v,w) + T(v,w,u) + T(w,u,v) = 0$ is equivalent to those whose components have the cyclic symmetry $T_{ijk} + T_{jki} + T_{kij} = 0$, and are in one-to-one correspondence with elements of the factor space $V^{(0,3)}/U$. The basis of U has

$$\begin{array}{lll} n & \text{elements of the form } E^{iii} & (i = 1, \dots n) \\ 2\binom{n}{2} & \text{elements of the form } E^{iij} & (i \ne j,\ i,j = 1, \dots n) \\ 2\binom{n}{3} & \text{elements of the form } E^{ijk} & (i \ne j \ne k,\ i,j,k = 1, \dots n) \end{array}$$

In the second line the factor 2 appears because we may have $i < j$ or $j < i$, while in the third line the factor 2 appears because $E^{123} \neq E^{213}$ but all other permutations of 123 are equivalent to one of these, e.g. $E^{231} = E^{123}$ etc. Hence the dimension of the space of tensors with the cyclic symmetry is

$$n^3 - n - 2\binom{n}{2} - 2\binom{n}{3} = n^3 - \frac{n(n^2+2)}{3} = \frac{2n(n^2-1)}{3},$$

Chapter 8.

Problem 8.1 **Express components of the exterior product of two 2-vectors $A = A^{ij}e_{ij}$ and $B = B^{kl}e_{kl}$ as a sum of six terms,**

$$(A \wedge B)^{ijkl} = \tfrac{1}{6}(A^{ij}B^{kl} + A^{ik}B^{lj} + \dots)$$

How many terms would be needed for a product of a 2-vector and a 4-vector? Show that in general the components of the exterior product of an r-vector and an s-vector can be expressed as a sum of $\dfrac{(r+s)!}{r!s!}$ terms.

Solution:

$$\begin{aligned}(A \wedge B)^{ijkl} &= \frac{1}{24}(A^{ij}B^{kl} - A^{ji}B^{kl} - A^{ij}B^{lk} + A^{ji}B^{lk} \\ &\quad + A^{ik}B^{lj} - A^{ki}B^{lj} - \dots \\ &\quad + A^{il}B^{jk} - \dots) \\ &= \frac{1}{6}(A^{ij}B^{kl} + A^{ik}B^{lj} + A^{il}B^{jk} \\ &\quad A^{kl}B^{ij} + A^{lj}B^{ik} + A^{jk}B^{il})\end{aligned}$$

For a product of a 2-vector A^{ij} and 4-vector B^{lkmn} we find

$$(A \wedge B)^{ijklmn} = \frac{1}{720}(A^{ij}B^{klmn} - A^{ji}B^{klmn} + \dots)$$

and the $6! = 720$ permutations are broken into groups of permutations on two indices over A and 4 indices over B, which give rise to $2!4! = 48$ identical terms. The result is a sum of $720/48 = 15$ independent terms.

For an r-vector A and an s-vector B, the resulting $r+s$-tensor $A \wedge B$ is a sum of $(r+s)!$ terms arising from all possibly permutations on the $(r+s)$ indices, of which the $r!s!$ terms which permute the r indices on A and the s indices on B are all identical. The resulting wedge product is therefore a sum of $\dfrac{(r+s)!}{r!s!}$ terms.

Problem 8.2 **Let V be a four-dimensional vector space with basis $\{e_1, e_2, e_3, e_4\}$, and A a 2-vector on V.**

(a) Show that a vector u satisfies the equation

$$A \wedge u = 0$$

if and only if there exists a vector v such that

$$A = u \wedge v.$$

[***Hint***: **Pick a basis such that $e_1 = u$.**]

(b) If

$$A = e_2 \wedge e_1 + ae_1 \wedge e_3 + e_2 \wedge e_3 + ce_1 \wedge e_4 + be_2 \wedge e_4$$

write out explicitly the equations $A \wedge u = 0$ where $u = u^1e_1 + u^2e_2 + u^3e_3 + u^4e_4$ and show that they have a solution if and only if $c = ab$. In this case find two vectors u and v such that $A = u \wedge v$.

(c) In general, show that a 4-vector $A \wedge A = 8\alpha e_1 \wedge e_2 \wedge e_3 \wedge e_4$ where

$$\alpha = A^{12}A^{34} + A^{23}A^{14} + A^{31}A^{24},$$

and

$$\det[A^{ij}] = \alpha^2.$$

(d) Show that A is the wedge product of two vectors $A = u \wedge v$ if and only if $A \wedge A = 0$.

Solution: (a) Setting a basis such that $u = e_1$ and writing $A = A^{ij}e_i \wedge e_j$ we have

$$A \wedge u = A^{ij}e_i \wedge e_j \wedge e_1 = 0$$

iff $A^{ij} = 0$ whenever both $i > 1$ and $j > 1$. That is,

$$A = A^{1j}e_1 \wedge e_j + A^{i1}e_i \wedge e_1 = 2A^{1j}e_1 \wedge e_j = u \wedge v$$

where $v = 2A^{1j}e_j$.

(b)

$$\begin{aligned} A \wedge u &= e_{213}u^3 + e_{214}u^4 + ae_{132}u^2 + ae_{134}u^4 + e_{231}u^1 \\ &\quad + e_{234}u^4 + ce_{142}u^2 + ce_{143}u^3 + be_{241}u^1 + be_{243}u^3 \\ &= e_{123}(u^1 - au^2 - u^3) + e_{124}(bu^1 - cu^2 - u^4) \\ &\quad + e_{134}(-cu^3 + au^4) + e_{234}(-bu^3 + u^4). \end{aligned}$$

Hence $A \wedge u$ has a non-trivial solution iff the matrix of coefficients of the linear set of equations for u^i vanishes,

$$\det \begin{vmatrix} 1 & -a & -1 & 0 \\ b & -c & 0 & -1 \\ 0 & 0 & -c & a \\ 0 & 0 & -b & 1 \end{vmatrix} = (ab - c)^2 = 0.$$

Hence the necessary and sufficient condition for $A = v \wedge u$ is $ab = c$.
The general solution for u^i is $u^1 = au^2 + u^3$, $u^4 = bu^3$, so that

$$u = (ax + y)e_1 + xe_2 + ye_3 + bye_3.$$

Set $x = 1$, $y = 0$ we have $u = ae_1 + e_2$, and solving

$$A = u \wedge v = (ae_1 + e_2) \wedge (v^1e_1 + v^2e_2 + v^3e_3 + v^4e_4)$$

results in a set of equations in the coefficients of e_{ij} for v^i:

$$v^1 = 1 + av^2, \qquad v^3 = 1, \qquad v^4 = b.$$

Since v^2 is arbitrary we may set $v^2 = 0$, so that $v = e_1 + e_3 + be_4$. It is straightforward to verify that

$$A = u \wedge v = (ae_1 + e_2) \wedge (e_1 + e_3 + be_4).$$

(c)

$$\begin{aligned} A \wedge A &= (A^{ij} e_i \wedge e_j) \wedge (A^{kl} e_k \wedge e_l) \\ &= A^{ij} A^{kl} e_i \wedge e_j \wedge e_k \wedge e_l \\ &= A^{ij} A^{kl} \epsilon_{ijkl} e_1 \wedge e_2 \wedge e_3 \wedge e_4 \end{aligned}$$

Now

$$A^{ij} A^{kl} \epsilon_{ijkl} = A^{12} A^{34} - A^{21} A^{34} - A^{12} A^{43} + A^{34} A^{12} + A^{34} A^{12} - \dots$$

The expansion on the RHS consists of 24 terms which breaks up into three groups of 8 equal terms,

$$A^{ij} A^{kl} \epsilon_{ijkl} = 8(A^{12} A^{34} + A^{13} A^{42} + A^{14} A^{23}) = 8\alpha$$

as required.
The determinant is straightforward to expand,

$$\det A^{ij} = \det \begin{vmatrix} 0 & A^{12} & -A^{31} & A^{14} \\ -A^{12} & 0 & A^{23} & A^{24} \\ A^{31} & -A^{23} & 0 & A^{34} \\ -A^{14} & -A^{24} & -A^{34} & 0 \end{vmatrix} = (A^{12} A^{34} + A^{13} A^{42} + A^{14} A^{23})^2.$$

(d) By (a) the condition for A to be a simple bivector, $A = u \wedge v$, is that there exists a vector u such that $A \wedge u = 0$. Explicitly writing out the components of this 3-vector equation,

$$\begin{aligned} u^1 A^{23} - u^2 A^{13} + u^3 A^{12} &= 0 \\ u^1 A^{24} - u^2 A^{14} + u^4 A^{12} &= 0 \\ u^1 A^{23} - u^3 A^{14} + u^4 A^{13} &= 0 \\ u^2 A^{34} - u^3 A^{24} + u^4 A^{23} &= 0 \end{aligned}$$

which has a non-trivial solution iff

$$\det \begin{vmatrix} A^23 & -A^{13} & -A^{12} & 0 & \\ A^{24} & -A^{14} & 0 & A^{12} & \\ A^{34} & 0 & -A^{14} & A^{13} & \\ 0 & A^{34} & -A^{24} & A^{23} & 0 \end{vmatrix} = 0.$$

Interchange the 2nd and 3rd columns, first and 4th columns, multiply first and 3rd columns by -1, and this determinant results in the same determinant as $\det[A^{ij}]$

with A^{23} and A^{14} interchanged. Hence the determinant is again α^2. Thus $A \wedge u = 0$ iff $\alpha = 0$, or equivalently $A \wedge A = 0$ by part (c).

Problem 8.3 **Prove Cartan's lemma, that if $u_1, u_2, \dots, u_p$ are linearly independent vectors and $v_1, \dots, v_p$ are vectors such that**

$$u_1 \wedge v_1 + u_2 \wedge v_2 + \cdots + u_p \wedge v_p = 0,$$

then there exists a symmetric set of coefficients $A_{ij} = A_{ji}$ such that

$$v_i = \sum_{j=1}^{p} A_i j u_j.$$

[*Hint*: Extend the u_i to a basis for the whole vector space V.]

Solution: By the basis extension theorem, Theorem 3.7, there exists a basis $e_1, \dots e_n$ of V such $e_i = u_i$ for $i = 1, 2, \dots p$. In this basis we may set

$$v_i = \sum_{j=1}^{p} A_{ij} u_j + \sum_{\alpha=p+1}^{n} A_{i\alpha} e_\alpha.$$

The equation

$$\sum_{i=1}^{p} u_i \wedge v_i = 0$$

reads, in this basis,

$$\sum_{i=1}^{p} \sum_{j=1}^{p} A_{ij} e_i \wedge e_j + \sum_{i=1}^{p} \sum_{\alpha=p+1}^{n} A_{i\alpha} e_i \wedge e_\alpha = 0.$$

Treating this as a tensor equation, with $e_i \wedge e_j = \frac{1}{2}(e_i \otimes e_j - e_j \otimes e_i)$, we see from the coefficients of $e_i \otimes e_\alpha$ that $A_{i\alpha} = 0$. The coefficient of $e_i \otimes e_j$ with $i < j \leq p$, is $A_{ij} - A_{ji}$. As this must vanish, we have the symmetry condition $A_{ij} = A_{ji}$.

Problem 8.4 **If V is an n-dimensional vector space and A a 2-vector, show that there exists a basis $e_1, e_2, \dots, e_n$ of V such that**

$$A = e_1 \wedge e_2 + e_3 \wedge e_4 + \dots e_{2r-1} \wedge e_{2r},$$

for some number $2r$, called the *rank* of A.

(a) Show that the rank only depends on the 2-vector A, not on the choice of basis, by showing that $A^r \neq 0$ and $A^{r+1} = 0$ where

$$A^p = \underbrace{A \wedge A \wedge \cdots \wedge A}_{p}.$$

(b) If $f_1, f_2, \dots, f_n$ is any basis of V and $A = A^{ij} f_i \otimes f_j$ where $A^{ij} = -A^{ji}$, show that the rank of the matrix of components $\mathsf{A} = [A^{ij}]$ coincides with the rank as defined above.

Solution: If f_i is any basis of V and $A = A^{ij} f_i \wedge f_j$ we may assume, possibly after a rearrangement of basis vectors, that $A^{12} \neq 0$ (else $A = 0$ and we are done, with $r = 0$). Consider a change of basis $\{f_1, f, \dots f_n\} \to \{e_1, e_2, f_3, \dots, f_n\}$ of the form

$$f_1 = e_1 + a^3 f_3 + \dots + a^n f_n$$
$$f_2 = e_2 + b^3 f_3 + \dots + b^n f_n.$$

Then

$$A = A^{12}(e_1 + a^3 f_3 + \dots + a^n f_n) \wedge (e_2 + b^3 f_3 + \dots + b^n f_n) + A^{13} e_1 \wedge f_3 + \dots + A^{23} e_2 \wedge f_3 + \dots$$

If we pick

$$a^i = \frac{A^{2i}}{A^{12}}, \qquad b^i = -\frac{A^{1i}}{A^{12}} \qquad \text{for} \qquad i = 3, 4, \dots n$$

then

$$A = A^{12} e_1 \wedge e_2 + A'$$

where A' is a 2-vector entirely spanned by $f_i \wedge f_j$ $(i, j = 3, 4, \dots n)$. On rescaling $e_1 \to e_1 / A^{12}$ we have $A = e_1 \wedge e_2 + A'$. If $A' = 0$, the procedure has come to an end with $r = 1$, else repeat the process on A', using only the basis vectors $f_3, \dots, f_n$. Continue until A has the required form $A = e_1 \wedge e_2 + e_3 \wedge e_4 + \dots e_{2r-1} \wedge e_{2r}$.

(a) $$A^2 = A \wedge A = 2 e_1 \wedge e_2 \wedge e_3 \wedge e_4 + 2 e_1 \wedge e_2 \wedge e_5 \wedge e_6 + \dots$$

which is clearly non-zero if $r > 1$. Continuing, we find

$$A^r = r! e_1 \wedge e_2 \wedge \dots \wedge e_{2r-1} \wedge e_{2r}.$$

and $A^{r+1} = A^r \wedge A = 0$ since all terms in this product involve a repeated e_i.

(b) Let the components of the special basis vectors e_i with respect to the basis f_j be denoted e_i^j. Then

$$A^{ij} = \frac{1}{2}(e_1^i e_2^j - e_2^i e_1^j + e_3^i e_4^j - \dots - e_{2r}^i e_{2r-1}^j),$$

and for an arbitrary vector $\mathbf{x}$ in $\mathbb{R}^n$ having components x_i we have

$$A^{ij} x_j = \frac{1}{2}(e_1^i (e_2^j x_j) - e_2^i (e_1^j x_j) + e_3^i (e_4^j x_j) - \dots - e_{2r}^i (e_{2r-1}^j x_j)).$$

The RHS are the components of an arbitrary vector spanned by $e_i, e_2, \dots, e_{2r}$, and is therefore an arbitrary vector of the subspace V_A. Since this is also the entire range of the matrix operator $\mathsf{A} = [A^{ij}]$, we see that $\dim V_A = \rho(\mathsf{A})$, the rank of the operator A. By Problem 3.20 (a), this implies that the column rank of $\mathsf{A} = \rho(\mathsf{A}) = 2r$.

Problem 8.5 **Let V be an n-dimensional space and A an arbitrary $(n-1)$-vector.**

(a) Show that the subspace V_A of vectors u such that $u \wedge A = 0$ has dimension $n-1$.

(b) Show that every $(n-1)$-vector A is decomposable, $A = v_1 \wedge v_2 \wedge \cdots \wedge v_{n-1}$ for some vectors $v_1, \ldots, v_{n-1} \in V$. [*Hint*: Take a basis for $e_1, \ldots, e_n$ of V such that the first $n-1$ vectors span the subspace V_A, which is always possible by Theorem 3.7, and expand A in terms of this basis.]

Solution: (a) The set V_A of vectors such that $u \wedge A = 0$ is clearly a vector subspace, for

$$u \wedge A = 0 \text{ and } v \wedge A = 0 \quad \Longrightarrow \quad (u + av) \wedge A = 0.$$

In any basis e_i, an $(n-1)$-vector A can be written

$$A = a_1 e_{23\ldots n} - a_2 e_{134\ldots n} + a_3 e_{124\ldots n} - \cdots + (-1)^{n-1} a_n e_{123\ldots n-1}$$

where $e_{23\ldots n} = e_2 \wedge e_3 \wedge \cdots \wedge e_n$ etc. If $u = u^i e_i$ the equation $u \wedge A = 0$ reads

$$(u^1 a_1 + u^2 a_2 + \ldots u^n a_n) e_{123\ldots n} = 0.$$

We may assume, without loss of generality, that $a_1 \neq 0$, so that

$$u^1 = -u^2 \frac{a_2}{a_1} - u^3 \frac{a_3}{a_1} - \cdots - u^n \frac{a_n}{a_1}$$

where $u^2, \ldots u^n$ are arbitrary real numbers. Hence V_A is spanned by the $n-1$ linearly independent vectors

$$e_i - \frac{a_i}{a_1} e_1 \qquad (i = 2, 3, \ldots n)$$

and therefore has dimension $n-1$.

(b) Choose a basis $e_1, \ldots e_{n-1}, e_n$ of V such that the first $n-1$ vectors span V_A. Then for all $u \in V_A$, $u = u^1 e_1 + u^2 e_2 + \cdots + u^{n-1} e_{n-1}$. Writing

$$A = a e_{12\ldots n-1} + \sum_{i=1}^{n-1} b_i e_{12\ldots i-1\, i+1 \ldots n}$$

we see that all b_i must vanish $(i = 1, \ldots, n-1)$, for if $b_i \neq 0$ then

$$e_i \wedge A = a.0 + b_i (-1)^{i-1} e_{12\ldots n} \neq 0.$$

contradicting $e_i \in V_A$. Hence

$$A = a e_{12\ldots n-1} = v_1 \wedge v_2 \wedge \cdots \wedge v_{n-1}$$

if we set $v_1 = a e_1$, and $v_i = e_i$ for $i = 2, \ldots, n-1$.

Problem 8.6 **Show that the quantity $\langle A, \beta\rangle$ defined in Eq. (8.39) vanishes for all p-vectors A if and only if $\beta = 0$. Hence show that the correspondence between linear functionals on $\Lambda(V)$ and p-forms is one-to-one,**

$$(\Lambda^p(V))^* \cong \Lambda^{*p}(V).$$

Solution: If

$$\langle A, \beta\rangle = p! A^{i_1 i_2 \dots i_p} \beta_{i_1 i_2 \dots i_p} = 0$$

for all p-vectors A, then we may set $A = e_{j_1 j_2 \dots j_p}$, so that

$$A^{i_1 i_2 \dots i_p} = (e_{j_1 j_2 \dots j_p})^{i_1 i_2 \dots i_p} = \frac{1}{p!} \delta^{i_1 i_2 \dots i_p}_{j_1 j_2 \dots j_p}.$$

Then

$$\begin{aligned} 0 = \langle A, \beta\rangle &= \frac{1}{p!} \delta^{i_1 i_2 \dots i_p}_{j_1 j_2 \dots j_p} \beta_{i_1 i_2 \dots i_p} \\ &= \frac{1}{p!} \sum_{\pi} (-1)^{\pi} \beta_{j_{\pi(1)} j_{\pi(2)} \dots j_{\pi(p)}} \\ &= \beta_{j_1 j_2 \dots j_p} \end{aligned}$$

so that the p-form $\beta = 0$.

Define the map $\varphi : \Lambda^{*p}(V) \to (\Lambda^p(V))^*$ by

$$\varphi(\beta)(A) = \langle A, \beta\rangle.$$

$\varphi(\beta)$ is a linear functional on $(\Lambda^p(V))$, for

$$\varphi(\beta)(aA + bB) = \langle aA + bB, \beta\rangle = a\langle A, \beta\rangle + b\langle B, \beta\rangle = a\varphi(\beta)(A) + b\varphi(\beta)(B).$$

Furthermore the map φ is obviously linear

$$\varphi(a\alpha + b\beta)(A) = \langle A, a\alpha + b\beta\rangle = a\langle A, \alpha\rangle + b\langle A, \beta\rangle = a\varphi(\alpha)(A) + b\varphi(\beta)(A).$$

It is one-to-one for, by the above,

$$\begin{aligned} \varphi(\alpha) = \varphi(\beta) \quad &\Longrightarrow \quad \langle A, \alpha\rangle = \langle A, \beta\rangle \quad \forall A \in \Lambda^p(V) \\ &\Longrightarrow \quad \langle A, \beta - \alpha\rangle = 0 \quad \forall A \in \Lambda^p(V) \\ &\Longrightarrow \quad \beta = \alpha \end{aligned}$$

Finally, the map is onto, for every liner functional f on $\Lambda^p(V)$ is uniquely defined by its values on the basis elements,

$$f(e_{i_1 i_2 \dots i_p}) = a_{i_1 i_2 \dots i_p}.$$

Let $A = a_{j_1 j_2 \dots j_p} \varepsilon^{j_1 j_2 \dots j_p}$, then

$$\langle e_{i_1 i_2 \dots i_p}, a_{j_1 j_2 \dots j_p} \varepsilon^{j_1 j_2 \dots j_p} \rangle = a_{j_1 j_2 \dots j_p} \frac{1}{p!} \delta^{j_1 j_2 \dots j_p}_{i_1 i_2 \dots i_p} = a_{i_1 i_2 \dots i_p}.$$

Hence $\varphi(A) = f$, so that φ is onto, one-to-one and therefore an isomorphism.

Problem 8.7 Show that the interior product between basis vectors e_i and $\varepsilon^{i_1 i_2 \dots i_r}$ is given by

$$\mathrm{i}_{e_i} \varepsilon^{i_1 \dots i_r} = \begin{cases} 0 & \textbf{if } i \notin \{i_1, \dots i_r\}, \\ (-1)^{a-1} \varepsilon^{i_1 \dots i_{a-1} i_{a+1} \dots i_r} & \textbf{if } i = i_a. \end{cases}$$

Solution: Using $\varepsilon^{i_1 \dots i_r} = \varepsilon^{i_1} \wedge \varepsilon^{i_2 \dots i_r}$ and Eq. (8.20), we have

$$\mathrm{i}_{e_i} \varepsilon^{i_1 \dots i_r} = \langle e_i, \varepsilon^{i_1} \rangle \varepsilon^{i_2 \dots i_r} - \varepsilon^{i_1} \wedge \mathrm{i}_{e_i} \varepsilon^{i_2 \dots i_r}$$

and repeating on the last term on the RHS, we find by induction

$$\begin{aligned} \mathrm{i}_{e_i} \varepsilon^{i_1 \dots i_r} &= \sum_{a=1}^{r} (-1)^{a-1} \langle e_i, \varepsilon^{i_a} \rangle \varepsilon^{i_1} \wedge \dots \wedge \varepsilon^{i_{a-1}} \wedge \varepsilon^{i_{a+1}} \wedge \dots \wedge \varepsilon^{i_r} \\ &= \sum_{a=1}^{r} (-1)^{a-1} \delta^{i_a}_{i} \varepsilon^{i_1 \dots i_{a-1} i_{a+1} \dots i_r} \end{aligned}$$

which immediately gives the desired result.

Problem 8.8 Prove Eq. (8.52).

Solution: The components of Eq. (8.47), $A \wedge B = (*A, B)E$ for arbitrary $B \in \Lambda^{(n-p)}(V)$, in an orthonormal frame are

$$(A \wedge B)^{i_1 \dots i_p j_1 \dots j_{n-p}} = (*A, B) E^{i_1 \dots i_p j_1 \dots j_{n-p}}$$

and using Eqs. (8.15) and (8.30) with $g = \pm 1$ we have

$$A^{[i_1 \dots i_p} B^{j_1 \dots j_{n-p}]} = (*A, B) \frac{1}{n!} \epsilon^{i_1 \dots i_p j_1 \dots j_{n-p}}.$$

Multiplying (contracting) throughout by $E_{i_1 \dots i_p j_1 \dots j_{n-p}} = ((-1)^s / n!) \epsilon_{i_1 \dots j_{n-p}}$ gives

$$A^{i_1 \dots i_p} B^{j_1 \dots j_{n-p}} E_{i_1 \dots i_p j_1 \dots j_{n-p}} = (*A, B) \frac{1}{n!} (-1)^s.$$

Hence, from Eqs. (8.42), (8.39)

$$\begin{aligned} A_{i_1 \dots i_p} B_{j_1 \dots j_{n-p}} E^{i_1 \dots i_p j_1 \dots j_{n-p}} &= (*A, B) \frac{1}{n!} (-1)^s \\ &= (-1)^s \frac{(n-p)!}{n!} (*A)^{j_1 \dots j_{n-p}} B_{j_1 \dots j_{n-p}}. \end{aligned}$$

Since $B_{j_1\dots j_{n-p}}$ are arbitrary we have, on setting $E^{i_1\dots i_p j_1\dots j_{n-p}} = (1/n!)\epsilon^{i_1\dots i_p j_1\dots j_{n-p}}$

$$A^{j_1\dots j_{n-p}} = \frac{(-1)^s}{(n-p)!}\epsilon^{i_1\dots i_p j_1\dots j_{n-p}} A_{i_1\dots i_p}.$$

Problem 8.9 **Every p-form α can be regarded as a linear functional on $\Lambda^p(V)$ through the action $\alpha(A) = \langle A, \alpha \rangle$. Show that the basis $\varepsilon^{\mathbf{i}}$ is dual to the basis $e_{\mathbf{j}}$ where $\mathbf{i} = i_1 < i_2 < \dots < i_p$, $\mathbf{j} = j_1 < j_2 < \dots < j_p$,**

$$\langle e_{\mathbf{j}}, \varepsilon^{\mathbf{i}} \rangle = \delta^{\mathbf{i}}_{\mathbf{j}} \equiv \delta^{i_1}_{j_1}\delta^{i_2}_{j_2}\dots\delta^{i_p}_{j_p}.$$

Verify that

$$\sum_{i_1<i_2<\dots<i_p} \langle e_{i_1 i_2 \dots i_p}, \varepsilon^{i_1 i_2 \dots i_p} \rangle = \dim \Lambda^p(V).$$

Solution: Writing out $e_{\mathbf{j}}$ and $\varepsilon^{\mathbf{i}}$ as tensor products,

$$e_{\mathbf{j}} = e_{j_1} \wedge \dots \wedge e_{j_p} = \frac{1}{p!}\sum_{\sigma}(-1)^{\sigma} e_{j_{\sigma(1)}} \otimes \dots \otimes e_{j_{\sigma(r)}},$$
$$\varepsilon^{\mathbf{i}} = \varepsilon^{i_1} \wedge \dots \wedge \varepsilon^{i_p} = \frac{1}{p!}\sum_{\pi}(-1)^{\pi} \varepsilon^{i_{\pi(1)}} \otimes \dots \otimes \varepsilon^{i_{\pi(r)}}.$$

Hence, by Eq. (8.39)

$$\begin{aligned}\langle e_{\mathbf{j}}, \varepsilon^{\mathbf{i}} \rangle &= p! C^1_1 C^2_2 \dots C^p_p e_{j_1\dots j_p} \otimes \varepsilon^{i_1\dots i_p} \\ &= \frac{1}{p!}\sum_{\sigma}\sum_{\pi}(-1)^{\sigma}(-1)^{\pi}\langle e_{j_{\sigma(1)}}, \varepsilon^{i_{\pi(1)}}\rangle \dots \langle e_{j_{\sigma(p)}}, \varepsilon^{i_{\pi(p)}}\rangle\end{aligned}$$

Non-zero contributions from these summands arise only if $i_1 = j_1, i_2 = j_2, \dots, i_p = j_p$ since both sequence are increasing and must consist of the same integers, else one of the products in the summand will vanish. The product of $\langle \cdot, \cdot \rangle$ products will be 1 if and only if the permutations are identical, $\pi = \sigma$. As there are $p!$ permutations, we have

$$\langle e_{\mathbf{j}}, \varepsilon^{\mathbf{i}} \rangle = \begin{cases} 0 & \text{if } \mathbf{i} \neq \mathbf{j} \\ (p!/p!) = 1 & \text{if } \mathbf{i} = \mathbf{j} \end{cases}.$$

This is the required result.
In the implied sum

$$\langle e_{i_1<i_2<\dots<i_p}, \varepsilon^{i_1<i_2<\dots<i_p} \rangle \equiv \sum_{i_1<i_2<\dots<i_p} \langle e_{\mathbf{i}}, \varepsilon^{\mathbf{j}} \rangle$$

each term contributes 1, and the number of terms is the number of ways a sequence $i_1 < i_2 < \cdots < i_p$ can be selected from $1, 2, \ldots n$; that is

$$\langle e_{i_1<i_2<\cdots<i_p}, \varepsilon^{i_1<i_2<\cdots<i_p} \rangle = \binom{n}{p} = \dim \Lambda^p(V).$$

Problem 8.10 Show that if u, v and w are vectors in an n-dimensional real inner product space then

(a) $(u \wedge v, u \wedge v) = (u \cdot u)(v \cdot v) - (u \cdot v)^2$.

(b) $u \wedge *(v \wedge w) = (u \cdot w) * v - (u \cdot v) * w$.

(c) Which identities do these equations reduce to in three-dimensional cartesian vectors?

Solution: (a) From Eqs. (8.39) and (8.42) we have for any p-vectors A and B,

$$(A, B) = p! A^{i_1 i_2 \ldots i_p} g_{i_1 j_1} g_{i_1 j_2} \ldots g_{i_p j_p} B^{j_1 j_2 \ldots j_p},$$

and setting $A = B = u \wedge v = \frac{1}{2}(u \otimes v - v \otimes u)$,

$$\begin{aligned}(u \wedge v, u \wedge v) &= 2! \frac{1}{2}(u^i v^j - v^i u^j) g_{ik} g_{jl} \frac{1}{2}(u^k v^l - v^k u^l) \\ &= \frac{1}{2}(u^i u_i v^j v_j - v^i u_i u^j v_j - u^i v_i v^j u_j + v^i v_i u^j u_j) \\ &= (u \cdot u)(v \cdot v) - (v \cdot u)^2.\end{aligned}$$

(b) Since $** A = \pm A$ any p-vector equation is equivalent to its Hodge dual. The Hodge dual of the equation in question is, using Eq. (8.49)

$$*(u \wedge *(v \wedge w)) = (u \cdot w) * *v - (u \cdot v) * *w = (u \cdot w)(-1)^{s+n-1} v - (u \cdot v)(-1)^{s+n-1} w.$$

Calculating components of the LHS in an o.n. basis

$$(u \wedge *(v \wedge w))^{i_1 \ldots i_{n-1}} = \frac{(-1)^s}{(n-2)!} u^{[i_1} \epsilon^{i_2 \ldots i_{n-1}] j_1 j_2} v_{j_1} w_{j_2}$$

and lowering the indices, removes the determinant of g sign $(-1)^s$:

$$(u \wedge *(v \wedge w))_{i_1 \ldots i_{n-1}} = \frac{1}{(n-2)!} u_{[i_1} \epsilon_{i_2 \ldots i_{n-1}] j_1 j_2} v^{j_1} w^{j_2}.$$

Hence

$$\begin{aligned}(*(u\wedge *(v\wedge w))^j &= (-1)^s\epsilon^{i_1\dots i_{n-1}j}(u\wedge *(v\wedge w))_{i_1\dots i_{n-1}}\\ &= (-1)^s\epsilon^{i_1\dots i_{n-1}j}\frac{1}{(n-2)!}(-1)^{n-2}u_{[i_{n-1}}\epsilon_{i_1 i_2\dots i_{n-2}]j_1j_2}v^{j_1}w^{j_2}\end{aligned}$$

after a cyclic permutation of indices

$$\begin{aligned}&= (-1)^{s+n}\frac{(n-2)!}{(n-2)!}u_{i_{n-1}}(\delta^{i_{n-1}}_{j_1}\delta^j{}_{j_2} - \delta^{i_{n-1}}_{j_2}\delta^j{}_{j_1})v^{j_1}w^{j_2}\\ &= (-1)^{s+n}((u\cdot v)w^j - (u\cdot w)v^j)\\ &= (-1)^{s+n-1}((u\cdot w)v^j - (u\cdot v)w^j)\\ &= (*((u\cdot w)*v - (u\cdot v)*w))^j,\end{aligned}$$

as required.

(c) In 3-dimensional Cartesian tesnors property (a) is equivalent to the vector identity

$$(\mathbf{u}\times\mathbf{v})^2 = \mathbf{u}^2\mathbf{v}^2 - (\mathbf{u}.\mathbf{v})^2$$

and (b) is equivalent to

$$\mathbf{u}\times(\mathbf{v}\times\mathbf{w}) = (\mathbf{u}\cdot\mathbf{w})\mathbf{v} - (\mathbf{u}\cdot\mathbf{v})\mathbf{w}.$$

Problem 8.11 Let g_{ij} be the Minkowski metric on a four-dimensional space, having index 2 (so that there are three + signs and one − sign).

(a) By calculating the inner products $(e_{i_1i_2}, e_{j_1,j_2})$, using (8.44) show that there are three +1's and three -1's in these inner products, and the index of the inner product defined on the six-dimensional space of bivectors $\Lambda^2(V)$ is therefore 0.

(b) What is the index of the inner product on $\Lambda^2(V)$ if V is n-dimensional and g_{ij} has index t? [*Ans.*: $\frac{1}{2}(t^2-n)$.]

Solution: (a) The 6-dimensional space $\Lambda^2(V)$ has basis e_{12}, e_{13}, e_{23}, e_{14}, e_{24}, e_{34}. Assuming $\eta_1=\eta_2=\eta_3=1$, $\eta_4=-1$ we have from Eq. (8.44)

$$\begin{aligned}(e_{12},e_{12}) = (e_{13},e_{13}) = (e_{23},e_{23}) = 1\\ (e_{14},e_{14}) = (e_{24},e_{24}) = (e_{34},e_{34}) = -1\end{aligned}$$

and all other $(e_{ij},e_{kl})=0$. Hence the index of this inner product space is $3-3=0$.

(b) Assume

$$\eta_1=\eta_2=\dots=\eta_r=1,\qquad \eta_{r+1}=\eta_{r+2}=\dots=\eta_{r+s}=-1$$

where $r+s=n$, $r-s=t$. The $\binom{n}{2}$-dimensional space $\Lambda^2(V)$ is spanned by e_{ij} where $1 \le i < j \le n$. The only non-vanishing products among the basis elements are

$$(e_{ab}, e_{ab}) = 1 \qquad \text{if } 1 \le a < b \le r \text{ or } r+1 \le a < b \le r+s = n$$
$$(e_{ab}, e_{ab}) = -1 \qquad \text{if } 1 \le a \le r \text{ and } r+1 \le s \le r+s.$$

Hence there are

$$\frac{r(r-1)}{2} + \frac{(s(s-1)}{2} = \frac{r^2+s^2-r-s}{2}$$

$+$ signs, and rs $-$ signs in this inner product space. The index is therefore

$$\frac{r^2+s^2-r-s-2rs}{2} = \frac{(r-s)^2-(r+s)}{2} = \frac{t^2-n}{2}.$$

Problem 8.12 Show that in an arbitrary basis the component representation of the dual of a p-form α is

$$(*\alpha)_{j_1 j_2 \dots j_{n-p}} = \frac{\sqrt{|g|}}{(n-p)!}\epsilon_{i_1 \dots i_p j_1 j_2 \dots j_{n-p}}\alpha^{i_1 \dots i_p}.$$

Solution: This problem is essentially identical to Problem 8.8. The components of Eq. (8.57),

$$(\alpha \wedge \beta)_{i_1 \dots i_p j_1 \dots j_{n-p}} = (-1)^s (*\alpha, \beta)\Omega_{i_1 \dots i_p j_1 \dots j_{n-p}}$$

are

$$\alpha_{[i_1 \dots i_p}\beta_{j_1 \dots j_{n-p}]} = (*\alpha, \beta)\frac{(-1)^s\sqrt{|g|}}{n!}\epsilon_{i_1 \dots i_p j_1 \dots j_{n-p}}.$$

Contract this equation with

$$\Omega^{i_1 \dots i_p j_1 \dots j_{n-p}} = \frac{(-1)^s}{n!\sqrt{|g|}}\epsilon^{i_1 \dots i_p j_1 \dots j_{n-p}}$$

gives

$$\alpha_{i_1 \dots i_p}\frac{(-1)^s}{n!\sqrt{|g|}}\epsilon^{i_1 \dots i_p j_1 \dots j_{n-p}}\beta_{j_1 \dots j_{n-p}} = \frac{((n-p)!}{n!}(*\alpha)_{j_1 \dots j_{n-p}}\beta^{j_1 \dots j_{n-p}}.$$

The LHS can be written

$$\alpha^{i_1 \dots i_p}\frac{(-1)^s g}{n!\sqrt{|g|}}\epsilon_{i_1 \dots i_p j_1 \dots j_{n-p}}\beta^{j_1 \dots j_{n-p}}.$$

As $(-1)^s g = |g|$ and the coefficients $B_{j_1 \dots j_{n-p}}$ are arbitrary, we arrive at

$$(*\alpha)_{j_1 j_2 \dots j_{n-p}} = \frac{\sqrt{|g|}}{(n-p)!} \epsilon_{i_1 \dots i_p j_1 j_2 \dots j_{n-p}} \alpha^{i_1 \dots i_p}.$$

Problem 8.13 **If u is any vector, and α any p-form show that**

$$\mathrm{i}_u * \alpha = *(\alpha \wedge \bar{g}(u)).$$

Solution: We prove this by making use of the component expression for $*\alpha$ derived in Problem 8.12.

$$(*\alpha)_{j_1 j_2 \dots j_{n-p}} = \frac{\sqrt{|g|}}{(n-p)!} \epsilon_{i_1 \dots i_p j_1 j_2 \dots j_{n-p}} \alpha^{i_1 \dots i_p}.$$

Then

$$\begin{aligned}
(\mathrm{i}_u * \alpha)_{j_2 \dots j_{n-p}} &= (n-p) \frac{\sqrt{|g|}}{(n-p)!} u^{j_1} \epsilon_{i_1 \dots i_p j_1 j_2 \dots j_{n-p}} \alpha^{i_1 \dots i_p} \\
&= \frac{\sqrt{|g|}}{(n-p-1)!} \epsilon_{i_1 \dots i_p j_1 j_2 \dots j_{n-p}} \alpha^{i_1 \dots i_p} u^{j_1} \\
&= \frac{\sqrt{|g|}}{(n-p-1)!} \epsilon_{i_1 \dots i_p j_1 j_2 \dots j_{n-p}} \alpha^{[i_1 \dots i_p} u^{j_1]} \\
&= *(\alpha \wedge \bar{g}(u))_{j_2 \dots j_{n-p}} \, ,
\end{aligned}$$

since

$$(\alpha \wedge \bar{g}(u))_{i_1 \dots i_p j_1} = \alpha_{[i_1 \dots i_p} u_{j_1]}.$$

The desired result follows since it holds componentwise in any basis.

Chapter 9.

Problem 9.1 **Show that**

$$[L^{\mu'}_{\nu}] = \begin{pmatrix} 1 & 0 & -\alpha & \alpha \\ 0 & 1 & -\beta & \beta \\ \alpha & \beta & 1-\gamma & \gamma \\ \alpha & \beta & -\gamma & 1+\gamma \end{pmatrix} \qquad \textbf{where} \qquad \gamma = \tfrac{1}{2}(\alpha^2+\beta^2)$$

is a Lorentz transformation for all values of α and β. Find those 4-vectors V^μ whose components are unchanged by all Lorentz transformations of this form.

Solution: From Eq. (9.5) $\mathsf{L} = [L^{\mu'}_{\nu'}]$ is a Lorentz transformation if

$$\mathsf{G} = \mathsf{L}^T \mathsf{G} \mathsf{L},$$

i.e.

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & \alpha & \alpha \\ 0 & 1 & \beta & \beta \\ -\alpha & -\beta & 1-\gamma & -\gamma \\ \alpha & \beta & \gamma & 1+\gamma \end{pmatrix} \begin{pmatrix} 1 & 0 & -\alpha & \alpha \\ 0 & 1 & -\beta & \beta \\ \alpha & \beta & 1-\gamma & \gamma \\ -\alpha & -\beta & \gamma & -1-\gamma \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \alpha^2+\beta^2+1-2\gamma & -\alpha^2-\beta^2+2\gamma \\ 0 & 0 & -\alpha^2-\beta^2+2\gamma & \alpha^2+\beta^2-1-2\gamma \end{pmatrix}$$

which holds iff $\gamma = \frac{1}{2}(\alpha^2+\beta^2)$.

The components V^μ of a 4-vector remain unchanged under such a Lorentz transformation iff

$$V'^{\mu'} = L^{\mu'}_{\nu} V^\nu = V^{\mu'}$$

i.e.

$$\begin{aligned} V'^1 &= V^1 - \alpha V^3 + \alpha V^4 = V^1 \\ V'^2 &= V^2 - \beta V^3 + \beta V^4 = V^2 \\ V'^3 &= \alpha V^1 + \beta V^2 + (1-\gamma)V^3 + \gamma V^4 = V^3 \\ V'^4 &= \alpha V^1 + \beta V^2 - \gamma V^3 + (1+\gamma)V^4 = V^4 \end{aligned}$$

Since α and β are arbitrary the first two equations imply $V^3 = V^4$. The V^3 equation (and/or the V^4 equation) imply that

$$\alpha V^1 + \beta V^2 = 0 \qquad \text{for all } \alpha \text{ and } \beta.$$

This is only possible if $V^1 = V^2 = 0$. Hence the most general vector whose components are invariant under this class of Lorentz transformations is of the form

$$V^\mu = (0,\, 0,\, v,\, v),$$

Problem 9.2 **Show that for *any* Lorentz transformation $L^{\mu'}_{\nu}$ one must have either**

$$L^4_{\ 4} \geq 1 \qquad \textbf{or} \qquad L^4_{\ 4} \leq -1.$$

(a) Show that those transformations having $L^4_{\ 4} \geq 1$ have the property that they preserve the concept of 'before' and 'after' for timelike separated events by demonstrating that they preserve the sign of Δx^4.

(b) What is the effect of a Lorentz transformation having $L^4_{\ 4} \leq -1$?

(c) Is there any meaning, independent of the inertial frame, to the concepts of "before" and "after" for space-like separated events?

Solution: In $g_{\rho\sigma} = g_{\mu'\nu'} L^{\mu'}_{\ \rho} L^{\nu'}_{\ \sigma}$ set $\rho = \sigma = 4$,

$$\sum_{i=1}^{3} (L^i_{\ 4})^2 - (L^4_{\ 4})^2 = -1.$$

Hence

$$(L^4_{\ 4})^2 = 1 + \sum_{i=1}^{3} (L^i_{\ 4})^2 \geq 1.$$

Therefore $L^4_{\ 4} \geq 1$ or $L^4_{\ 4} \leq -1$.

(a) For a pair of timelike separated events,

$$\Delta x^\mu \Delta x^\nu g_{\mu\nu} = \sum_{i=1}^{3} (\Delta x^i)^2 - (\Delta x^4)^2 < 0.$$

Assuming $\Delta x^4 > 0$ this means

$$\Delta x^4 > \sqrt{\sum_{i=1}^{3} (\Delta x^i)^2} = \sqrt{(\Delta bf x^i)^2}$$

where $\Delta \mathbf{x}^i = (\Delta x^1, \Delta x^2, \Delta x^3)$. Under a Lorentz transformation with $L^4_{\ 4} \geq 1$ set $\mathbf{l} = (L^4_{\ 1}, L^4_{\ 2}, L^4_{\ 3})$, and the time difference between events transforms as

$$\begin{aligned} \Delta x'^4 &= L^4_{\ i} \Delta x^i + L^4_{\ 4} \Delta x^4 \\ &= \mathbf{l} \cdot \Delta \mathbf{x} + L^4_{\ 4} \Delta x^4 \\ &> \mathbf{l} \cdot \Delta \mathbf{x} + |\mathbf{l}|\, |\Delta \mathbf{x}| > 0 \end{aligned}$$

since

$$|\mathbf{l} \cdot \Delta \mathbf{x}| = |\mathbf{l}|\, |\Delta \mathbf{x}|\, |\cos\theta| \leq |\mathbf{l}|\, |\Delta \mathbf{x}|.$$

(b) Lorentz transformations have the effect of reversing the direction of time: if $\Delta x^4 > 0$ then $\Delta x'^4 < 0$.

(c) There is no invariant meaning the concept of "before" and "after" for space-like separated events. For example, let $\Delta x^\mu = (1, 0, 0, \frac{1}{2})$. The events are space-like separated, for $\Delta x^\mu \Delta x_\mu = 1 - (\frac{1}{2})^2 = \frac{3}{4} > 0$, and $\Delta x^4 = \frac{1}{2} > 0$. Under a boost with velocity $v > \frac{1}{2}c$,

$$\mathsf{L} = [L^{\mu'}_{\ \nu}] = \begin{pmatrix} \gamma & 0 & 0 & -\gamma\frac{v}{c} \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\gamma\frac{v}{c} & 0 & 0 & \gamma \end{pmatrix}, \qquad \text{where } \gamma = \frac{1}{\sqrt{1-(v/c)^2}}$$

we have

$$\Delta x'^4 = \gamma\big(\frac{1}{2} - \frac{v}{c}\big) < 0.$$

Problem 9.3 **Show that (i) if T^α is a timelike 4-vector it is always possible to find a Lorentz transformation such that $T'^{\alpha'}$ will have components $(0, 0, 0, a)$, and (ii) if N^α is a null vector then it is always possible to find a Lorentz transformation such that $N'^{\alpha'}$ has components $(0, 0, 1, 1)$.**

Let U^α and V^α be 4-vectors. Show the following:

(a) If $U^\alpha V_\alpha = 0$ and U^α is timelike, then V^α is spacelike.

(b) If $U^\alpha V_\alpha = 0$ and U^α and V^α are both null vectors, then they are proportional to each other.

(c) If U^α and V^α are both timelike future-pointing then $U_\alpha V^\alpha < 0$, and $U^\alpha + V^\alpha$ is timelike.

(d) Find other statements similar to the previous assertions when U^α and V^α are taken to be various combinations of null, future-pointing null, timelike future-pointing, spacelike etc.

Solution: This problem is similar to Problem 5.5.
(i) By Gram-Schmidt orthonormalization it is always possible to find an o.n. basis such that $e_4 \propto T$. In this basis, which must be related by a Lorexntz transformation to the original basis, $T = (0, 0, 0, a)$.
(ii) As in Problem 5.5 the null vector requirement is

$$N^\mu N_\mu = 0 = \sum_{i=1}^{3} (N^i)^2 \ - \ (N^4)^2.$$

A rotation of the first 3 vectors can always be performed such that $(N'^1, N'^2, N'^3) \propto (0, 0, 1)$. Then there exists a real number $\alpha > 0$ such that $N = (0, 0, \alpha, \alpha)$. A further Lorentz transformation of the form

$$\mathsf{L} = [L^{\mu'}_{\ \nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \cosh\psi & -\sinh\psi \\ 0 & 0 & -\sinh\psi & \cosh\psi \end{pmatrix}$$

has the result $N'^{\mu} = (0, 0, 1, 1)$ if we select ψ such that

$$(\cosh\psi - \sinh\psi)\alpha = \mathrm{e}^{-\psi}\alpha = 1,$$

i.e. choose $\psi = \ln\alpha$.

(a) If U^α is a time-like vector, pick a basis such that $U^\alpha = (0, 0, 0, a)$. Then the equation $U^\alpha V_\alpha = 0$ reads $V_4 = 0$, whence

$$V^\alpha V_\alpha = \sum_{i=1}^{3} (V^i)^2 > 0.$$

Hence V^α is a space-like 4-vector.

(b) If U^α is a null vector, pick a basis such that $U^\alpha = (0, 0, 1, 1)$. In this basis, $U^\alpha V_\alpha = V_1 + V_4 = 0$ implies that $V^1 - V^4 = 0$, so that we can write

$$V^\alpha = (v,\ a,\ b,\ v).$$

The 4-vector V^α can only be a null vector if

$$0 = V^\alpha V_\alpha = v^2 + a^2 + b^2 - v^2 = a^2 + b^2,$$

i.e. iff $a = b = 0$. In other words $V^\alpha \propto U^\alpha$.

(c) Choosing a frame such that $U^\alpha = (0, 0, 0, U^4)$ where $U^4 > 0$, then

$$U^\alpha V_\alpha = U^4 V_4 = -U^4 V^4 < 0$$

if V^α is future-pointing ($V^4 > 0$). Similarly

$$(U^\alpha + V^\alpha)(U_\alpha + V_\alpha) = U^\alpha U_\alpha + V^\alpha V_\alpha + 2U^\alpha V_\alpha < 0$$

since each of the three terms in the RHS sum is negative.

(d) Similar arguments can be used to show, for example, (i) if U^α and V^α are null future-pointing 4-vectors ($U^4 > 0$, $V^4 > 0$) then $U^\alpha V_\alpha \leq 0$, and $U^\alpha V_\alpha = 0$ only if $V^\alpha \propto U^\alpha$.
(ii) If U^α is a timelike future-pointing 4-vector, and V^α is null future-pointing then $U^\alpha V_\alpha < 0$.

Problem 9.4 If the 4-component of a 4-vector equation $A^4 = B^4$ is shown to hold in all inertial frames, show that all components are equal in all frames, $A^\mu = B^\mu$.

Solution: Assume $A^4 - B^4 = 0$ in all frames. Then under an arbitrary Lorentz transformation

$$A'^4 - B'^4 = L^4_{\ \mu}(A^\mu - B^\mu) = 0.$$

Set

$$\mathsf{L} = [L^{\mu'}_{\ \nu}] = \begin{pmatrix} \cosh\psi & 0 & 0 & -\sinh\psi \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\sinh\psi & 0 & 0 & \cosh\psi \end{pmatrix}$$

which is easily verified to be a Lorentz transformation for any ψ, and we have

$$0 = A'^4 - B'^4 = -\sinh\psi(A^1 - B^1) + \cosh\psi(A^4 - B^4) = -\sinh\psi(A^1 - B^1).$$

Hence $A^1 = B^1$. Similarly $A^2 = B^2$ and $A^3 = B^3$.

Problem 9.5 From the law of transformation of velocities, Eq. (9.19), show that the velocity of light in an arbitrary direction is invariant under boosts.

Solution: The transformation of a velocity $\mathbf{u}$ under a boost in the x-direction is given by Eq. (9.19),

$$u'_x = \frac{u_x - v}{1 - u_x v/c^2}, \qquad u'_y = \frac{u_y}{\gamma(1 - u_x v/c^2)}, \qquad u'_z = \frac{u_z}{\gamma(1 - u_x v/c^2)}.$$

Therefore, if $u_x^2 + u_y^2 + u_z^2 = c^2$ then

$$\begin{aligned} u'^{\,2}_x + u'^{\,2}_y + u'^{\,2}_z &= \frac{1}{(1 - u_x v/c^2)^2}\big[u_x^2 + v^2 - 2u_x v \\ &\qquad + \big(u_y^2 + u_z^2\big)\big(1 - \frac{v^2}{c^2}\big)\big] \\ &= \frac{c^2 + v^2 u_x^2/c^2 - 2u_x v}{(1 - u_x v/c^2)^2} \\ &= \frac{(c - u_x v/c)^2}{(1 - u_x v/c^2)^2} \\ &= c^2. \end{aligned}$$

Problem 9.6 If two intersecting light beams appear to be making a non-zero angle ϕ in one frame K, show that there always exists a frame K' whose motion relative to K is in the plane of the beams such that the beams appear to be directed in opposite directions.

Solution: From the previous problem we have for any beam of light making angle θ with the x-axis,

$$u'_x = c\cos\theta' = \frac{u_x - v}{1 - u_x v/c^2} = \frac{c\cos\theta - v}{1-(v/c)\cos\theta}.$$

For simplicity choose the x-axis to be along the bisector of the angle between the two light beams. Setting K' to be a frame having vecloicty v relative to K in this chosen x-direction,

$$c\cos\theta' = \frac{c\cos(\phi/2) - v}{1-(v/c)\cos(\phi/2)} = 0$$

if $v/c = \cos(\phi/2)$. This achieves that $\theta' = \pi/2$. Similarly for the lower beam $\theta' = -\pi/2$. To check the sign, use Eq. (9.21)

$$u'_y = c\sin\theta' = \frac{-c\sin(\phi/2)\sqrt{1-\cos^2(\phi/2)}}{1-\cos^2(\phi/2)} = -c.$$

In the frame K' the beams are directed in opposite directions.

Problem 9.7 A source of light emits photons uniformly in all directions in its own rest frame.

(a) If the source moves with velocity v with respect to an inertial frame K, show the 'headlight effect': half the photons seem to be emitted in a forward cone whose semi-angle is given by $\cos\theta = v/c$.

(b) In films of the *Star Wars* genre, star fields are usually seen to be swept backwards around a rocket as it accelerates towards the speed of light. What would such a rocketeer really see as his velocity $v \to c$?

Solution: (a) Using the aberration formula in the form given in Problem 9.6,

$$\cos\theta' = \frac{\cos\theta - (v/c)}{1-(v/c)\cos\theta}.$$

The forward half of photons emitted in the frame K' of the emitter has $\cos\theta' > 0$. Since the denominator in the above equation is positive, $1-(v/c)\cos\theta > 0$, these photons are restricted to a forward cone with semi-angle $\cos\theta > v/c$, i.e.

$$\theta < \arccos\frac{v}{c}.$$

As $v \to c$ this cone becomes increasingly narrow, $\theta \to 0$.

(b) For a rocketeer with velocity v relative to a frame K, the situation is essentially reversed. The forward field of view consists of photons having

$$\frac{\pi}{2} < \theta' < \frac{3\pi}{2},$$

i.e. $\cos\theta' < 0$. From the above transformation these photons are a larger section of the sky having $\cos\theta < v/c$ in the rest frame. This includes some areas of the sky which appear to lie in the "forward direction" in the rest frame. The fraction of the sky in front of the rocket frame K' is therefore

$$\frac{1}{4\pi}\int_{\arccos(v/c)}^{\pi} \sin\theta d\theta \int_0^{2\pi} \mathrm{d}\phi = \frac{1}{2}\int_{-1}^{v/c} \mathrm{d}(\cos\theta) = \frac{1}{2}\big(1+\frac{v}{c}\big).$$

As $v \to c$ the fraction of the sky visible to the rocketeer in a forward direction approximates the entire star field.

Problem 9.8 If two separate events occur at the same time in some inertial frame S, prove that there is no limit on the time separations assigned to these events in other frames, but that their space separation varies from infinity to a minimum that is measured in S. With what speed must an observer travel in order that two simultaneous events at opposite ends of a 10-metre room appear to differ in time by 100 years?

Solution: From Eq. (9.14) we have for the difference between a pair of events,

$$\Delta t' = \gamma(\Delta t - \frac{v}{c^2}\Delta x), \quad \Delta x' = \gamma(\Delta x - v\Delta t).$$

If the two events are simultaneous in S then $\Delta t = 0$, so that

$$\Delta t' = -\gamma(v)\frac{v}{c^2}\Delta x).$$

As the velocity varies over its full range $-c < v < c$, the time difference $\Delta t'$ in the boosted frame varies over the complete range $-\infty < \Delta t' < \infty$.
The spatial separation in the boosted frame is

$$\Delta x' = \gamma\Delta x = \frac{\Delta x}{\sqrt{1-v^2/c^2}} \geq \Delta x,$$

which varies over the entire range, $\Delta x \leq \Delta x' < \infty$.
If $\Delta x = 10\,\mathrm{m}$ the velocity needed to achieve a time difference $\Delta t' = 3\times 10^9$ sec (≈ 100 years) is given by

$$\frac{v}{\sqrt{1-v^2/c^2}} = \frac{\Delta t'}{\Delta x}c^2.$$

Set $v = c(1-\epsilon)$ where $\epsilon \ll 1$, then

$$\sqrt{1-v^2/c^2} = \sqrt{(1-v/c)(1+v/c)} = \sqrt{\epsilon(2-\epsilon)} \approx (2\epsilon)^{1/2}.$$

Hence

$$\frac{v/c}{\sqrt{1-v^2/c^2}} \approx (2\epsilon)^{-1/2} \approx \frac{c\Delta t'}{\Delta x} = 9\times 10^{16}.$$

Hence

$$\frac{v}{c} = 1 - \epsilon \approx 1 - \frac{1}{2}(9 \times 10^{16})^{-2} \approx 1 - 6 \times 10^{-35}.$$

Problem 9.9 A supernova is seen to explode on Andromeda galaxy, while it is on the western horizon. Observers A and B are walking past each other, A at 5 km/hr towards the east, B at 5 km/hr towards the west. Given that Andromeda is about a million light years away, calculate the difference in time attributed to the supernova event by A and B. Who says it happened earlier?

Solution: The relative speed of A and B is $v = 10\text{km/hr} \approx 3\text{m/sec}$ Hence the time assigned by B to the supernova (primed frame) is

$$\Delta t' = \gamma\big(\Delta t - \frac{v}{c^2}\Delta x\big)$$

where $\Delta x = 10^6$ light years $\approx 10^{22}$ metres, and $\Delta t = 0$. Hence, since $v/c \ll 1$

$$\Delta t' \approx \Delta t - \frac{v}{c^2}\Delta x \approx -\frac{3 \times 10^{22}}{9 \times 10^{16}} \approx -3 \times 10^5 \text{ sec.}$$

B says that the supernova occurred about 80 hours earlier than A.

Problem 9.10 Twin A on the Earth and twin B who is in a rocketship moving away from him at a speed of $\frac{1}{2}c$ separate from each other at midday on their common birthday. They decide to each blow out candles exactly four years from B's departure.

(a) What moment in B's time corresponds to the event P that consists of A blowing his candle out? And what moment in A's time corresponds to the event Q that consists of B blowing her candle out?

(b) According to A which happened earlier, P or Q? And according to B?

(c) How long will A have to wait before he *sees* his twin blowing her candle out?

Solution: (a) The time since departure assigned to the event P by B is, since it occurs on A's world line $x_P = 0$,

$$t'_P = \frac{t_P}{1 - v^2/c^2} = \frac{4}{\sqrt{3/4}} = \frac{8}{\sqrt{3}} = 4.62 \text{ years.}$$

Similarly, or by symmetry, A assigns a time 4.62 years to the event Q where B blows her candle out.

(b) This means that the moment A regards as simultaneous with Q is 4.62 years. Hence, according to A, B blows her candle out after A. Similarly, according to B, A blows his candle out after she does.

(c) The time at which A sees B blowing her candle out is the time t_Q assigned to event Q by A, plus the time it takes light to reach A from B at this moment. At event Q, the distance from A is vt_Q, where $v = \frac{1}{2}c$. Hence if t_S is the time at which A sees Q is

$$t_S = t_Q + \frac{vt_Q}{c} = \frac{3}{2}t_Q = \frac{3}{2} \cdot \frac{8}{\sqrt{3}} = 4\sqrt{3} = 6.93 \text{ years.}$$

Problem 9.11 **Using the fact that the 4-velocity $V^\mu = \gamma(u)(u_x, u_y, u_z, c)$ transforms as a 4-vector, show from the transformation equation for V'^4 that the transformation of u under boosts is**

$$\frac{\gamma(u')}{\gamma(u)} = \gamma(v)\Big(1 - \frac{vu_x}{c^2}\Big).$$

From the remaining transformation equations for $V'^{i'}$ derive the law of transformation of velocities (9.19).

Solution: The 4-vector transformation law is

$$V'^{\mu'} = \gamma(u')(u'_x,\, u'_y,\, u'_z,\, c) = L^{\mu'}_{nu} V^\nu$$

where

$$\mathsf{L} = [L^{\mu'}_{nu}] = \begin{pmatrix} \gamma(v) & 0 & 0 & -\gamma(v)v/c \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\gamma(v)v/c & 0 & 0 & \gamma(v) \end{pmatrix}.$$

Hence

$$V'^4 = \gamma(u')c = \gamma(u)\gamma(v)\big(c - \frac{v}{c}u_x\big),$$

which immediately gives

$$\frac{\gamma(u')}{\gamma(u)} = \gamma(v)(1 - \frac{vu_x}{c^2}).$$

Setting $\mu' = 1$, we have

$$V'^1 = \gamma(u')u'_x = L^1_\nu V^\nu = \gamma(v)\gamma(u)\big(u_x - \frac{v}{c}c\big),$$

so that

$$u'_x = \frac{\gamma(v)\gamma(u)}{\gamma(u')}(u_x - v) = \frac{u_x - v}{1 - vu_x/c^2}.$$

Setting $\mu' = 2$,

$$V'^2 = \gamma(u')u'_y = V^2 = \gamma(u)u_y$$

whence

$$u'_y = \frac{\gamma(u)}{\gamma(u')}u_y = \frac{u_y}{\gamma(v)(1 - vu_x/c^2)}$$

Similarly for u'_z. These formulae agree with the formulae for relativistic transformation of velocities, Eq. (9.19).

Problem 9.12 **Let K' be a frame with velocity v relative to K in the x-direction.**

(a) Show that for a particle having velocity u', acceleration a' in the x'-direction relative to K', its acceleration in K is

$$a = \frac{a'}{[\gamma(1 + vu'/c^2)]^3}.$$

(b) A rocketeer leaves Earth at $t = 0$ with constant acceleration g at every moment relative to his instantaneous rest frame. Show that his motion relative to the Earth is given by

$$x = \frac{c^2}{g}\left(\sqrt{1 + \frac{g^2}{c^2}t^2} - 1\right).$$

(c) In terms of his own proper time τ show that

$$x = \frac{c^2}{g}\left(\cosh\frac{g}{c}\tau - 1\right).$$

(d) If he proceeds for 10 years of his life, decelerates with $g = 9.80$ m s^2 for another 10 years to come to rest, and returns in the same way, taking 40 years in all, how much will people on Earth have aged on his return? How far, in light years, will he have gone from Earth?

Solution: (a) Differentiate the first of the relativistic formulae for addition of velocities, Eq. (9.20), with respect to times t,

$$u = \frac{u' + v}{1 + vu'/c^2}, \qquad t = \gamma\big(t' + \frac{vx'}{c^2}\big),$$

giving a formula for acceleration a,

$$\begin{aligned} a = \frac{\mathrm{d}u}{\mathrm{d}t} &= \frac{1}{\gamma(\mathrm{d}t' + v dx'/c^2}\mathrm{d}u'\big(\frac{1}{1 + vu'/c^2} - \frac{(u' + v)v/c^2}{(1 + vu'/c^2)^2}\big) \\ &= \frac{\mathrm{d}u'}{\mathrm{d}t'}\frac{1 - v^2/c^2}{\gamma(1 + vu'/c^2)(1 + vu'/c^2)^2} \\ &= \frac{a'}{\gamma^3(1 + vu'/c^2)^3} \end{aligned}$$

(b) In the i.r.f. $u' = 0$, $a' = g$. Let v be the velocity of the rocket in the earth's frame, then

$$a = \frac{\mathrm{d}v}{\mathrm{d}t} = g\big(1 - \frac{v^2}{c^2}\big)^{3/2}$$

and we can integrate to obtain

$$gt = \int_0^v \frac{w\mathrm{d}w}{(1 - w^2/c^2)^{3/2}} = \frac{v}{\sqrt{1 - v^2/c^2}}.$$

Solving this equation for v,

$$v = \frac{\mathrm{d}x}{\mathrm{d}t} = \frac{gt}{\sqrt{1 + g^2t^2/c^2}},$$

and integrating once more gives

$$x = \int_0^t \frac{gt\mathrm{d}t}{\sqrt{1 + g^2t^2/c^2}} = \frac{c^2}{g}\left(\sqrt{1 + \frac{g^2}{c^2}t^2} - 1\right).$$

(c) The proper time is given by

$$\tau = \int_0^t \frac{\mathrm{d}t}{\gamma} = \int_0^t \sqrt{1 - \frac{v^2}{c^2}}\mathrm{d}t = \int_0^t \frac{\mathrm{d}t}{\sqrt{1 + g^2t^2/c^2}} = \frac{c}{g}\sinh^{-1}\frac{gt}{c}.$$

Substituting in the formula derived in (b),

$$x = \frac{c^2}{g}\left(\cosh\frac{g}{c}\tau - 1\right).$$

(d) If $\tau = 10$ yrs, $g = 9.8\mathrm{m\ sec}^{-2}$, we have

$$\frac{c}{g} = 3.1 \times 10^7 \text{ sec} \approx 1 \text{ yr}.$$

Hence

$$t \approx \sinh 10.25 \text{ yrs} \approx 14,400 \text{ yrs}.$$

For the total journey, 10 years accelerating, 10 years decelerating and then a similar return, people on Earth will have aged $4 \times 14,400 = 58,000$ years. The total distance travelled from Earth at the furthest point will be

$$2x = \frac{2c^2}{g}\big(\cosh g\frac{\tau}{c} - 1\big) \approx 29,000 \text{ light years}.$$

Problem 9.13 A particle is in hyperbolic motion along a world-line whose equation is given by

$$x^2 - c^2t^2 = a^2, \qquad y = z = 0.$$

Show that

$$\gamma = \frac{\sqrt{a^2 + c^2t^2}}{a}$$

and that the proper time starting from $t = 0$ along the path is given by

$$\tau = \frac{a}{c} \cosh^{-1} \frac{ct}{a}.$$

Evaluate the particle's 4-velocity V^μ and 4-acceleration A^μ. Show that A^μ has constant magnitude.

Solution:

$$\begin{aligned} x^2 - c^2t^2 = A^2 \quad &\Longrightarrow \quad 2x\frac{\mathrm{d}x}{\mathrm{d}t} - 2c^2t = 0 \\ &\Longrightarrow \quad v = \frac{\mathrm{d}x}{\mathrm{d}t} = \frac{c^2t}{x} = \frac{c^2t}{\sqrt{A^2 + c^2t^2}} \\ &\Longrightarrow \quad \gamma = \frac{1}{\sqrt{1 - v^2/c^2}} = \frac{1}{A}\sqrt{A^2 + c^2t^2}. \end{aligned}$$

The proper time is given by

$$\tau = \int_0^t \frac{\mathrm{d}t}{\gamma} = \int_0^t \frac{A\mathrm{d}t}{\sqrt{A^2 + c^2t^2}} = \frac{A}{c}\sinh^{-1}\frac{ct}{A}.$$

4-velocity and 4-accelerations are

$$V^\mu = \frac{\mathrm{d}x^mu}{\mathrm{d}\tau} = \gamma\frac{\mathrm{d}x^mu}{\mathrm{d}t} = \gamma(v,\, 0,\, 0,\, c) = \big(\frac{c^2t}{A},\, 0,\, 0,\, \frac{c}{A}\sqrt{A^2 + c^2t^2}\big),$$

$$A^\mu = \frac{\mathrm{d}V^mu}{\mathrm{d}\tau} = \gamma\frac{\mathrm{d}V^mu}{\mathrm{d}t} = \frac{c^2}{A^2}\big(\sqrt{A^2 + c^2t^2},\, 0,\, 0,\, ct\big),$$

and the magnitude of the 4-acceleration is

$$A^\mu A_\mu = \frac{c^4}{A^4}(A^2 + c^2t^2 - c^2t^2) = \frac{c^4}{A^4} = \text{const.}$$

Problem 9.14 **For a system of particles it is generally assumed that the conservation of total 4-momentum holds in any localized interaction,**

$$\sum_a P^\mu_{(a)} = \sum_b Q^\mu_{(b)}.$$

Use Problem 9.4 to show that the law of conservation of 4-momentum holds for a given system provided the law of energy conservation holds in all inertial frames. Also show that the law of conservation of momentum in all frames is sufficient to guarantee conservation of 4-momentum.

Solution: Since $P^4_{(a)} = E_{(a)}/c$ and $Q^4_{(b)} = E_{(b)}/c$, conservation of energy of energy

$$\sum_a E_{(a)} = \sum_b E_{(b)} \quad \Longrightarrow \quad \sum_a P^4_{(a)} - \sum_b Q^4_{(b)} = 0.$$

If this holds in all frames, then by Problem 9.4 the 4-vector equation holds in all frames,

$$\sum_a P^\mu_{(a)} - \sum_b Q^\mu_{(b)} = 0.$$

Similarly, if the first component of a 4-vector equation holds in all frames $A^1 - B^1 = 0$ in all frames, then under an arbitrary Lorentz transformation of the form

$$\mathsf{L} = [L^{\mu'}_{\ \nu}] = \begin{pmatrix} \cosh\psi & 0 & 0 & -\sinh\psi \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\sinh\psi & 0 & 0 & \cosh\psi \end{pmatrix}$$

we have

$$0 = A'^1 - B'^1 = \cosh\psi(A^1 - B^1) - \sinh\psi(A^4 - B^4) = -\sinh\psi(A^4 - B^4).$$

Hence $A^4 = B^4$. Hence $A^\mu = B^\mu$, and it holds in all frames since it is a 4-vector equation. The remainder of the proof is essentially identical to that above, since $P^i_{(a)} \equiv p_{i(a)}$, $Q^i_{(b)} \equiv p_{i(b)}$.

Problem 9.15 A particle has momentum p, energy E in a frame K.

(a) If K' is an inertial frame having velocity v relative to K, use the transformation law of the momentum 4-vector $P^\mu = \left(\mathbf{p}, \dfrac{E}{c}\right)$ to show that

$$E' = \gamma(E - \mathbf{v}\cdot\mathbf{p}), \qquad \mathbf{p}'_\perp = \mathbf{p}_\perp \qquad \textbf{and} \qquad \mathbf{p}'_{||} = \gamma\Big(\mathbf{p}_{||} - \frac{E}{c^2}\mathbf{v}\Big),$$

where $\mathbf{p}_\perp$ and $\mathbf{p}_{||}$ are the components of p respectively perpendicular and parallel to v.

(b) If the particle is a photon, use these transformations to derive the aberration formula

$$\cos\theta' = \frac{\cos\theta - v/c}{1 - \cos\theta\,(v/c)}$$

where θ is the angle between p and v.

Solution: (a) From $P'^{\mu'} = L^{\mu'}_{\ \nu}P^\nu$, where $\mathsf{L} = [L^{\mu'}_{\ \nu}]$ is a boost,

$$[L^{\mu'}_{\ nu}] = \begin{pmatrix} \gamma & 0 & 0 & -\gamma v/c \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\gamma v/c & 0 & 0 & \gamma \end{pmatrix}, \qquad \gamma = \frac{1}{\sqrt{1 - v^2/c^2}}$$

we have

$$P'^4 = \frac{E'}{c} = L^4_{\ 1}P^1 + L^4_{\ 4}P^4 = -\gamma\frac{v}{c}p_1 + \gamma\frac{E}{c}.$$

Hence, setting $\mathbf{v} = (v, 0, 0)$,

$$E' = \gamma(E - \mathbf{v}\cdot\mathbf{p}).$$

The component of momentum parallel to $\mathbf{v}$ is $\mathbf{p}_{\|} = (p_1,\, 0,\, 0)$, and its transformation is found from

$$P'^1 = p'_1 = L^1_{\ 1}P^1 + L^1_{\ 4}P^4 = \gamma p_1 - \gamma\frac{v}{c}\frac{E}{c},$$

so that

$$\mathbf{p}'_{\|} = (p'_1,\, 0,\, 0) = \gamma\big(\mathbf{p}_{\|} - \frac{E}{c^2}\mathbf{v}\big).$$

Similarly the transformation of $\mathbf{p}_{\perp} = (0,\, p_2,\, p_3)$ is found from

$$P'^2 = p'_2 = P^2 = p_2, \qquad p'_3 = p_3,$$

i.e. $\mathbf{p}'_{\perp} = \mathbf{p}_{\perp}$.

(b) For a photon, $E = pc$, $E' = p'c$, $p_1 = p\cos\theta$, and $p'_1 = p'\cos\theta'$, from which the energy transformation equation reads

$$p' = \frac{E'}{c} = \frac{\gamma}{c}(E - \mathbf{v}\cdot\mathbf{p}) = \gamma p(1 - \frac{v}{c}\cos\theta),$$

and the parallel momentum equation is

$$p'\cos\theta' = \gamma p\big(\cos\theta - \frac{v}{c}\big).$$

Dividing these two equations gives the result

$$\cos\theta' = \frac{\cos\theta - v/c}{1 - \cos\theta\,(v/c)}.$$

Problem 9.16 **Use $F^{\mu}V_{\mu} = 0$ to show that**

$$\mathbf{f}\cdot\mathbf{v} = \frac{\mathrm{d}E}{\mathrm{d}t}.$$

Also show this directly from the definitions (9.30) and (9.31) of $\mathbf{p}$, E.

Solution: By Eq. (9.28), $A^{\mu}V_{\mu} = 0$, hence

$$F^{\mu}V_{\mu} = mA^{\mu}V_{\mu} = 0.$$

Since

$$F^{\mu} = \gamma\big(\mathbf{v}, \frac{1}{c}\frac{\mathrm{d}E}{\mathrm{d}t}\big), \qquad V^{\mu}\gamma(\mathbf{v},\, c)$$

this equation reads

$$\gamma^2\big(\mathbf{f}\cdot\mathbf{v} - \frac{1}{c}\frac{\mathrm{d}E}{\mathrm{d}t}c\big) = \gamma^2\big(\mathbf{f}\cdot\mathbf{v} - \frac{\mathrm{d}E}{\mathrm{d}t}\big) = 0,$$

whence

$$\mathbf{f}\cdot\mathbf{v} = \frac{\mathrm{d}E}{\mathrm{d}t}.$$

Eqs. (9.30), (9.31) are

$$\mathbf{p} = m\gamma\mathbf{v}, \qquad E = m\gamma c^2 = \frac{mc^2}{\sqrt{1 - v^2/c^2}}.$$

Hence, since $\mathbf{f} = \mathrm{d}\mathbf{p}/\mathrm{d}t$ it is required to show

$$\frac{\mathrm{d}\mathbf{p}}{\mathrm{d}t}\cdot\mathbf{v} = \frac{\mathrm{d}E}{\mathrm{d}t}.$$

Now

$$\begin{aligned}\frac{\mathrm{d}\mathbf{p}}{\mathrm{d}t}\cdot\mathbf{v} - \frac{\mathrm{d}E}{\mathrm{d}t} &= m\mathbf{v}\cdot\frac{\mathrm{d}\gamma\mathbf{v}}{\mathrm{d}t} - mc^2\frac{\mathrm{d}\gamma}{\mathrm{d}t} \\ &= m\big(v^2\frac{\mathrm{d}\gamma}{\mathrm{d}t} + \gamma\mathbf{v}\cdot\frac{\mathrm{d}\mathbf{v}}{\mathrm{d}t} - c^2\frac{\mathrm{d}\gamma}{\mathrm{d}t}\big)\end{aligned}$$

Using

$$\frac{\mathrm{d}\gamma}{\mathrm{d}t} = \frac{\mathrm{d}}{\mathrm{d}t}\Big(\frac{1}{\sqrt{1 - v^2/c^2}}\Big) = \frac{v\mathrm{d}v/\mathrm{d}t}{c^2(1 - v^2/c^2)^3/2}$$

and

$$v^2 = \mathbf{v}\cdot\mathbf{v} \Longrightarrow v\mathrm{d}v/\mathrm{d}t = \mathbf{v}\cdot\frac{\mathrm{d}\mathbf{v}}{\mathrm{d}t}$$

we have

$$\frac{\mathrm{d}\mathbf{p}}{\mathrm{d}t}\cdot\mathbf{v} - \frac{\mathrm{d}E}{\mathrm{d}t} = m\Big(\frac{v^3\mathrm{d}v/\mathrm{d}t}{c^2(1 - v^2/c^2)^{3/2}} + \frac{v\mathrm{d}v/\mathrm{d}t}{(1 - v^2/c^2)^{1/2}} - \frac{v\mathrm{d}v/\mathrm{d}t}{(1 - v^2/c^2)^{3/2}}\Big) = 0.$$

Problem 9.17 Show that with respect to a rotation (9.8) the electric and magnetic fields E and B transform as 3-vectors,

$$E'_i = a_{ij}E_j, \qquad B'_i = a_{ij}B_j.$$

Solution: Under a Lorentz transformation

$$F'_{\mu'\nu'} = L'^{\mu}{}_{\mu'}L'^{\nu}{}_{\nu'}F_{\mu\nu}$$

where

$$\mathsf{L}' = [L'^{\mu}{}_{\mu'}] = \mathsf{L}^{-1} = \begin{pmatrix} a_{11} & a_{21} & a_{31} & 0 \\ a_{12} & a_{22} & a_{32} & 0 \\ a_{13} & a_{23} & a_{33} & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

where $\mathsf{A} = [a_{ij}]$ is a 3×3 orthogonal matrix. That is, adopting cartesian tensor summation conventions

$$L'^{i}{}_{j} = a_{ji} \qquad \text{where} \qquad a_{ki}a_{kj} = \delta_{ij}.$$

From $E_i = F_{i4}$ we have

$$E'_i = F'_{i4} = L'^{\mu}_{i} L'^{\nu}_{4} F_{\mu\nu} = a_{ij} F_{i4} = a_{ij} E_j$$

which is the standard transformation of a 3-vector under a rotation.
From the tensor transformation law

$$F'_{ij} = L'^{\mu}_{i} L'^{\nu}_{j} F_{\mu\nu} = a_{ik} a_{jl} F_{kl}$$

and the components of magnetic field

$$F_{ij} = \epsilon_{ijk} B_k \quad \Longleftrightarrow \quad = B_k = \tfrac{1}{2}\epsilon_{ijk} F_{ij},$$

we have

$$\begin{aligned} B'_k &= \tfrac{1}{2}\epsilon_{ijk} F'_{ij} \\ &= \tfrac{1}{2}\epsilon_{ijk} F'_{ij} \\ &= \tfrac{1}{2}\epsilon_{ijk} a_{im} a_{jl} F_{ml} \\ &= \tfrac{1}{2}\epsilon_{ijk} \epsilon_{mlp} a_{im} a_{jl} B_p \\ &= a_{kp} B_p. \end{aligned}$$

The last step follows from the fact that if $b_{kp} = \frac{1}{2}\epsilon_{ijk}\epsilon_{mlp} a_{im} a_{jl}$ then

$$\begin{aligned} b_{kp} a_{kq} &= \tfrac{1}{2}\epsilon_{ijk}\epsilon_{mlp} a_{im} a_{jl} a_{kq} \\ &= \tfrac{1}{2}\epsilon_{mlp}\epsilon_{mlq} \det \mathsf{A} \\ &= \delta_{pq} \det \mathsf{A} \\ &= \delta_{pq} \end{aligned}$$

since a rotation is an orthogonal matrix A having determinant $+1$. Hence $b_{pq} = a_{pq}$.

Problem 9.18 Under a boost (9.13) show that the 4-tensor transformation law for $F_{\mu\nu}$ or $F^{\mu\nu}$, gives rise to

$$E'_1 = F'_{14} = E_1, \qquad E'_2 = \gamma\Big(E_2 - \frac{v}{c}B_3\Big), \qquad E'_3 = \gamma\Big(E_2 + \frac{v}{c}B_2\Big),$$

$$B'_1 = F'_{23} = B_1, \qquad B'_2 = \gamma\Big(B_2 + \frac{v}{c}E_3\Big), \qquad B'_3 = \gamma\Big(B_2 - \frac{v}{c}E_2\Big).$$

Decomposing E and B into components parallel and perpendicular to $\mathbf{v} = (v, 0, 0)$, show that these transformations can be expressed in vector form:

$$\mathbf{E}'_{\|} = \mathbf{E}_{\|}, \qquad \mathbf{E}'_{\perp} = \gamma\Big(\mathbf{E}_{\perp} + \frac{1}{c}\mathbf{v}\times\mathbf{B}\Big),$$
$$\mathbf{B}'_{\|} = \mathbf{B}_{\|}, \qquad \mathbf{B}'_{\perp} = \gamma\Big(\mathbf{B}_{\perp} - \frac{1}{c}\mathbf{v}\times\mathbf{E}\Big).$$

Solution: Under a boost with velocity $\mathbf{v} = (v, 0, 0)$ the 4-tensor transformation equation of $F_{\mu\nu}$ is

$$F_{\mu'\nu'} = L'^{\mu}{}_{\mu'} L'^{\nu}{}_{\nu'} F_{\mu\nu},$$

where

$$\mathsf{L}' = [L'^{\mu}{}_{\mu'}] = \begin{pmatrix} \gamma & 0 & 0 & \gamma v/c \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \gamma v/c & 0 & 0 & \gamma \end{pmatrix}.$$

Then

$$E'_1 = F'_{14} = L'^{\mu}{}_{1} L'^{\nu}{}_{4} F_{\mu\nu} = \gamma^2\big(1 - \frac{v^2}{c^2}\big)F_{14} = E_1,$$
$$E'_2 = F'_{24} = L'^{\mu}{}_{2} L'^{\nu}{}_{4} F_{\mu\nu} = \gamma\frac{v}{c}F_{21} + \gamma F_{24} = \gamma\big(E_2 - \frac{v}{c}B_3\big)$$

and similarly

$$E'_3 = \gamma\big(E_3 + \frac{v}{c}B_2\big),$$
$$B'_1 = F'_{23} = L'^{\mu}{}_{2} L'^{\nu}{}_{3} F_{\mu\nu} = F_{23} = B_1,$$
$$B'_2 = F'_{31} = L'^{\mu}{}_{3} L'^{\nu}{}_{1} F_{\mu\nu} = \gamma F_{31} + \gamma\frac{v}{c}F_{34} = \gamma\big(B_2 + \frac{v}{c}E_3\big)$$
$$B'_3 = F'_{12} = \gamma\big(B_3 - \frac{v}{c}E_2\big).$$

For any vector $\mathbf{u}$, we set $\mathbf{u}_{\perp} = (u_1, 0, 0)$ and $\mathbf{u}_{\perp} = (0, u_2, u_3)$ for the components parallel and perpendicular to the velocity $\mathbf{v}$. Then these transformation equations are precisely the components of the 3-vector equations

$$\mathbf{E}'_{\|} = \mathbf{E}_{\|}, \qquad \mathbf{E}'_{\perp} = \gamma\big(\mathbf{E}_{\perp} + \frac{1}{c}\mathbf{v}\times\mathbf{B}\big),$$
$$\mathbf{B}'_{\|} = \mathbf{B}_{\|}, \qquad \mathbf{B}'_{\perp} = \gamma\big(\mathbf{B}_{\perp} - \frac{1}{c}\mathbf{v}\times\mathbf{E}\big).$$

for $\mathbf{v}\times\mathbf{B} = (v_2B_3 - v_3B_2,\ v_3B_1 - v_1B_3,\ v_1B_2 - v_2B_1) = (0,\ -vB_3,\ vB_2)$ etc.

Problem 9.19 It is possible to use transformation of E and B under boosts to find the field of a uniformly moving charge. Consider a charge q

travelling with velocity v, which without loss of generality may be taken to be in the x-direction. Let $\mathbf{R} = (x - vt,\, y,\, z)$ be the vector connecting charge to field point $\mathbf{r} = (x,\, y,\, z)$. In the rest frame of the charge, denoted by primes, suppose the field is the coulomb field

$$\mathbf{E}' = \frac{q\mathbf{r}'}{r'^3}, \qquad \mathbf{B}' = \mathbf{0}$$

where

$$\mathbf{r}' = (x',\, y',\, z') = \left(\frac{x - vt}{\sqrt{1 - v^2/c^2}},\, y,\, z\right).$$

Apply the transformation law for E and B derived in Problem 9.18 to show that

$$\mathbf{E} = \frac{q\mathbf{R}(1 - v^2/c^2)}{R^3(1 - (v^2/c^2)\sin^2\theta)^{3/2}} \qquad \textbf{and} \qquad \mathbf{B} = \frac{1}{c}\mathbf{v} \times \mathbf{E},$$

where θ is the angle between R and v. At a given distance R where is most of the electromagnetic field concentrated for highly relativistic velocities $v \approx c$?

Solution: Firstly

$$r'^2 = x'^2 + y'^2 + z'^2 = \frac{(x - vt)^2 + (1 - v^2/c^2)(x^2 + y^2)}{1 - v^2/c^2}.$$

Since $x^2 + y^2 = R^2 \sin^2\theta$ where θ is the angle between $\mathbf{R}$ and $\mathbf{v}$,

$$r'^2 = \frac{1}{1 - v^2/c^2} R^2 \left(1 - \frac{v^2}{c^2}\sin^2\theta\right).$$

Applying the transformation laws of $\mathbf{E}$ and $\mathbf{B}$ derived in Problem 9.18

$$E_1 = E_1' = \frac{qx'}{r'^3} = \frac{q(x - vt)(1 - v^2/c^2)}{R^3(1 - (v^2/c^2)\sin^2\theta)^{3/2}}$$

and

$$E_2 = \gamma E_2' = \gamma\frac{qy'}{r'^3} = \frac{qy(1 - v^2/c^2)}{R^3(1 - (v^2/c^2)\sin^2\theta)^{3/2}},$$

$$E_3 = \gamma E_3' = \gamma\frac{qz'}{r'^3} = \frac{qz(1 - v^2/c^2)}{R^3(1 - (v^2/c^2)\sin^2\theta)^{3/2}}.$$

Combining these equations as a 3-vector equation gives

$$\mathbf{E} = \frac{q\mathbf{R}(1 - v^2/c^2)}{R^3(1 - (v^2/c^2)\sin^2\theta)^{3/2}}.$$

Since $\mathbf{B}' = \mathbf{0}$ the magnetic field is given by

$$\mathbf{B}_{\|} = \mathbf{0}, \qquad \mathbf{B}_{\perp} = \frac{1}{c}\mathbf{v} \times \mathbf{E},$$

and as $\mathbf{v} \times \mathbf{E}$ is necessarily perpendicular to $\mathbf{v}$, we have

$$\mathbf{B} = \mathbf{B}_{\|} + \mathbf{B}_{\perp} = \frac{1}{c}\mathbf{v} \times \mathbf{E}.$$

The maximum value of $|\mathbf{E}|$ is at $\sin\theta = 1$, i.e. $\theta = \pi/2$, where

$$|\mathbf{E}| = \frac{q}{R^2\sqrt{1 - v^2/c^2}} \to \infty \qquad \text{as} \qquad v \to c.$$

In the direction of motion $\theta = 0$, on the other hand $|\mathbf{E}| = q(1 - v^2/c^2)/R^2 \to 0$ as $v \to c$. The closer v is to c the more the electric (and magnetic) field are concentrated in a thin wedge of angles about $\theta = \pi/2$, and almost all the e.m. field is perpendicular to direction of motion as $v \to c$.

Problem 9.20 **A particle of rest mass m, charge q is in motion in a uniform constant magnetic field $\mathbf{B} = (0, 0, B)$. Show from the Lorentz force equation that the energy $\mathcal{E}$ of the particle is constant, and its motion is a helix about a line parallel to B, with angular frequency**

$$\omega = \frac{qcB}{\mathcal{E}}.$$

Solution: Setting $\mathbf{E} = \mathbf{0}$ in the Lorentz force equation (9.46), we have

$$\frac{\mathrm{d}\mathcal{E}}{\mathrm{d}t} = q\mathbf{E} \cdot \mathbf{v} = 0,$$

i.e.

$$\mathcal{E} = \frac{mc^2}{\sqrt{1 - v^2/c^2}} = \text{const.}$$

The rate of change of momentum is given by

$$\frac{\mathrm{d}\mathbf{p}}{\mathrm{d}t} = \frac{q}{c}\mathbf{v} \times \mathbf{B}, \qquad \text{where} \qquad \mathbf{p} = \frac{\mathcal{E}\mathbf{v}}{c^2},$$

which may be written

$$\begin{aligned} \frac{\mathrm{d}v_1}{\mathrm{d}t} &= \frac{qcB}{\mathcal{E}}v_2 \\ \frac{\mathrm{d}v_2}{\mathrm{d}t} &= -\frac{qcB}{\mathcal{E}}v_1 \\ \frac{\mathrm{d}v_3}{\mathrm{d}t} &= 0. \end{aligned}$$

hence $v_3 = \text{const.}$, and we have $z = at + z_0$ for some constants a and z_0. Differentiating the v_1 equation with respect to t and using the v_2 equation gives a simple harmonic oscillator equation,

$$\frac{\mathrm{d}^2 v_1}{\mathrm{d}t^2} = -\omega^2 v_1 \qquad \text{where} \qquad \omega = \frac{qcB}{\mathcal{E}}.$$

Then

$$v_1 = A \sin \omega t$$
$$v_2 = \frac{1}{\omega}\frac{\mathrm{d}v_1}{\mathrm{d}t} = A \cos \omega t,$$

whence

$$x = x_0 - r_0 \cos \omega t, \qquad y = y_0 + r_0 \sin \omega t,$$

where x_0, y_0 and $r_0 = A/\omega$. Combined with the uniform motion in the z-direction the result is a helical spiral winding anti-clockwise about the direction of the magnetic field. The angular frequency (time of return in the $x - y$ plane is $\omega = qcB/\mathcal{E}$.

Problem 9.21 **Let E and B be perpendicular constant electric and magnetic fields, $\mathbf{E} \cdot \mathbf{B} = 0$.**

(a) If $B^2 > E^2$ show that a transformation to a frame K' having velocity $\mathbf{v} = k\mathbf{E} \times \mathbf{B}$ can be found such that $\mathbf{E}'$ vanishes.

(b) What is the magnitude of $\mathbf{B}'$ after this transformation?

(c) If $E^2 > B^2$ find a transformation that makes $\mathbf{B}'$ vanish.

(d) What happens if $E^2 = B^2$?

(e) A particle of charge q is in motion in a crossed constant electric and magnetic field $\mathbf{E} \cdot \mathbf{B} = 0$, $B^2 > E^2$. From the solution of Problem 9.20 for a particle in a constant magnetic field, describe its motion.

Solution: (a) $\mathbf{v} = k\mathbf{E} \times \mathbf{B}$ is perpendicular to both $\mathbf{E}$ and $\mathbf{B}$. Therefore, by the transformations derived in Problem 9.18

$$\mathbf{E}'_{||} = \mathbf{E}_{||} = \mathbf{0}, \qquad \mathbf{B}'_{||} = \mathbf{B}_{||} = \mathbf{0}.$$

The perpendicular components are

$$\begin{aligned}
\mathbf{E}'_{\perp} &= \gamma\big(\mathbf{E}_{\perp} + \frac{1}{c}\mathbf{v} \times \mathbf{B}\big) \\
&= \gamma\big(\mathbf{E} + \frac{k}{c}(\mathbf{E} \times \mathbf{B}) \times \mathbf{B}\big) \\
&= \gamma\big(\mathbf{E} + \frac{k}{c}((\mathbf{E} \cdot \mathbf{B})\mathbf{B}) - \mathbf{B}^2\mathbf{E})\big) \\
&= \gamma\big(1 - \frac{B^2 k}{c}\big)\mathbf{E} \qquad \text{since } \mathbf{E} \cdot \mathbf{B} = 0.
\end{aligned}$$

Similarly,

$$\begin{aligned}\mathbf{B}'_\perp &= \gamma\big(\mathbf{B}_\perp - \frac{1}{c}\mathbf{v}\times\mathbf{E}\big)\\ &= \gamma\big(\mathbf{B} - \frac{k}{c}(\mathbf{E}\times\mathbf{B})\times\mathbf{E}\big)\\ &= \gamma\big(1 - \frac{E^2k}{c}\big)\mathbf{B}.\end{aligned}$$

If we can set $k = c/B^2$ then we would achieve $\mathbf{E}' = \mathbf{0}$. This is possible, since the velocity of the transformation is

$$|\mathbf{v}| = k|\mathbf{E}||\mathbf{B}| = \frac{c|\mathbf{E}|}{|\mathbf{B}|} < c \qquad \text{since } E^2 < B^2.$$

(b) With this choice of k,

$$\mathbf{B}' = \mathbf{B}"_\perp = \gamma\big(1 - \frac{E^2}{B^2}\big)\mathbf{B} = \mathbf{B}\sqrt{1 - \frac{E^2}{B^2}},$$

so that

$$|B'| = |B|\sqrt{1 - \frac{E^2}{B^2}} = \sqrt{B^2 - E^2}.$$

(c) If $E^2 > B^2$ then take $k = c/E^2$. We then have

$$|\mathbf{v}| = c\frac{|\mathbf{B}|}{|\mathbf{E}|} < c$$

and $\mathbf{B}' = \mathbf{0}$.

(d) If $E^2 = B^2$ then this holds in all frames, since by Eq. (9.44) $B^2 - E^2$ is an invariant. It is therefore impossible to transform away either vector $\mathbf{E}$ or $\mathbf{B}$.

(e) By performing a Lorentz transformation (boost) with velocity

$$\mathbf{v} = \frac{c}{B^2}\mathbf{E}\times\mathbf{B}$$

we have $\mathbf{E}'$, and in the primed frame the problem can be solved as in Problem 9.20. The resulting motion is therefore a helical spiral about the $\mathbf{B}'$ field, with a velocity $\mathbf{v}$ superimposed.

Problem 9.22 **An electromagnetic field $F_{\mu\nu}$ is said to be of 'electric type' at an event p if there exists a unit timelike 4-vector U_μ at p, $U_\alpha U^\alpha = -1$, and a spacelike 4-vector field E_μ orthogonal to U^μ such that**

$$F_{\mu\nu} = U_\mu E_\nu - U_\nu E_\mu, \qquad E_\alpha U^\alpha = 0.$$

(a) Show that any purely electric field, i.e. one having $\mathbf{B} = \mathbf{0}$, is of electric type.

(b) If $F_{\mu\nu}$ is of electric type at p, show that there is a velocity v such that

$$\mathbf{B} = \frac{\mathbf{v}}{c} \times \mathbf{E} \qquad (|\mathbf{v}| < c).$$

Using Problem 9.18 show that there is a Lorentz transformation that transforms the electromagnetic field to one that is purely electric at p.

(c) If $F_{\mu\nu}$ is of electric type everywhere with U^μ a constant vector field, and satisfies the Maxwell equations *in vacuo*, $J^\mu = 0$, show that the vector field E^μ is divergence-free, $E^\nu{}_{,\nu} = 0$.

Solution: (a) If $\mathbf{B} = \mathbf{0}$ then

$$F_{\mu\nu} = \begin{pmatrix} 0 & 0 & 0 & E_1 \\ 0 & 0 & 0 & E_2 \\ 0 & 0 & 0 & E_3 \\ -E_1 & -E_2 & -E_3 & 0 \end{pmatrix},$$

Whence

$$F_{\mu\nu} = U_\mu E_\nu - U_\nu E_\mu$$

where

$$U_\mu = (0, 0, 0, -1), \qquad E_\mu = (E_1, E_2, E_3, 0)$$

evidently satisfying $U^\mu E_\mu = 0$.

(b) Suppose $F_{\mu\nu}$ is electric type. Then there exist 4-vector fields U_μ and G_μ such that

$$F_{\mu\nu} = U_\mu G_\nu - U_\nu G_\mu, \qquad U_\alpha U^\alpha = -1, \qquad G_\alpha U^\alpha = 0.$$

(We choose a different notation for G^μ so as not to confuse its spatial components G_i with those of the electric field $E_i = F_{i4}$). Then

$$E_i = F_{i4} = U_i G_4 - U_4 G_i$$

and, using $U_4 \neq 0$ (since U^μ is time-like),

$$\begin{aligned} B_i = \tfrac{1}{2} F_{jk} &= \tfrac{1}{2}\epsilon_{ijk}(U_j G_k - G_j U_k) \\ &= \epsilon_{ijk} U_j G_k \\ &= \epsilon_{ijk} U_j \Big(-\frac{E_i}{U_4} + \frac{U_i G_4}{U_4}\Big) \\ &= -\epsilon_{ijk} U_j \frac{E_i}{U_4} \qquad \text{since } \epsilon_{ijk} U_j U_i = 0 \\ &= \frac{1}{c}(\mathbf{v} \times \mathbf{E})_i \end{aligned}$$

if we set $\mathbf{v}$ to be the veclocity corresponding to 4-velocity U^μ,

$$U^\mu = \gamma\big(\frac{v_1}{c}, \frac{v_2}{c}, \frac{v_3}{c}, 1\big).$$

If we perform a Lorentz transformation with boost velocity $\mathbf{v}$, then by Problem 9.18

$$\mathbf{B}'_{\|} = \mathbf{B}_{\|} = \mathbf{0}, \qquad \mathbf{B}'_{\perp} = \gamma\big(\mathbf{B}_{\perp} - \frac{1}{c}\mathbf{v}\times\mathbf{E}\big) = \mathbf{0}$$

i.e. $\mathbf{B}' = \mathbf{0}$.

(c) $U^\mu = \text{const.}$ implies that $U^\mu_{,\nu} = 0$. Then

$$F^{\mu\nu}{}_{,\nu} = U^\mu E^\nu{}_{,\nu} - U^\nu E^\mu{}_{,\nu} = 0.$$

Multiply this equation through with U_μ gives

$$-E^\nu{}_{,\nu} - U^\nu U_\mu E^\mu{}_{,\nu} = 0.$$

Now

$$U_\mu E^\mu{}_{,\nu} = (U_\mu E^\mu)_{,\nu} - U_{\mu,\nu} E^\mu = 0 - 0 = 0.$$

Hence

$$E^\nu{}_{,\nu} = 0.$$

Problem 9.23 **Use the gauge freedom $\Box\psi = 0$ in the Lorentz gauge to show that it is possible to set $\phi = 0$ and $\nabla\cdot\mathbf{A} = 0$. This is called a *radiation gauge*.**

(a) What gauge freedoms are still available to maintain the radiation gauge?

(b) Suppose $\mathbf{A}$ is independent of coordinates x and y in the radiation gauge. Show that the Maxwell equations have solutions of the form

$$\mathbf{E} = (E_1(u),\, E_2(u),\, 0), \qquad \mathbf{B} = (-E_2(u),\, E_1(u),\, 0)$$

where $u = ct - z$ and $E_i(u)$ are arbitrary differentiable functions.

(c) Show that these solutions may be interpreted as right-travelling electromagnetic waves.

Solution: Let $(\mathbf{A}, \phi)$ be a Lorentz gauge:

$$\nabla\cdot\mathbf{A} + \frac{\partial\phi}{\partial x^4} = 0$$

where, by the vacuum Maxwell equations

$$\Box\phi = \nabla^2\phi - \frac{\partial^2\phi}{\partial(x^4)^2} = 0$$

Under a gauge transformation, $\phi' = \phi - \psi_{,4}$. Set

$$\psi = \int_0^{x^4} \phi \mathrm{d}x^4 + f(\mathbf{x})$$

where $f(\mathbf{x})$ is any solution of the Poisson equation

$$\nabla^2 f = \left.\frac{\partial \phi}{\partial x^4}\right|_{x^4=0}.$$

Then

$$\phi' = \phi - \phi = 0$$

and

$$\begin{aligned}
\Box\psi &= \nabla^2\psi - \frac{\partial^2\psi}{\partial(x^4)^2} \\
&= \int_0^{x^4} \nabla^2\phi \mathrm{d}x^4 + \nabla^2 f(\mathbf{x}) - \frac{\partial\phi}{\partial x^4} \\
&= \int_0^{x^4} \frac{\partial^2\phi}{\partial(x^4)^2}\mathrm{d}x^4 + \left.\frac{\partial\phi}{\partial x^4}\right|_{x^4=0} \\
&= 0
\end{aligned}$$

The Lorentz gauge condition with $\phi = 0$ implies $\nabla \cdot \mathbf{A} = 0$.

(a) The remaining freedom has $\psi_{,4} = 0$, i.e. $\psi = \psi(\mathbf{x})$ and

$$\nabla^2\psi = 0.$$

(c) Suppose that $\mathbf{A} = \mathbf{A}(z,t)$, $\phi = 0$. Maxwell's vacuum equation implies

$$\Box\mathbf{A} = \frac{\partial^2\mathbf{A}}{\partial z^2} - \frac{1}{c^2}\frac{\partial^2\mathbf{A}}{\partial t^2} = 0,$$

which has the general solution

$$\mathbf{A} = \mathbf{A}(ct - z) + \mathbf{A}(ct + z).$$

Setting $u = ct - z$ and ignoring the solution depending on $ct + z$, we have $\mathbf{A} = \mathbf{A}(u)$. Then, writing $f'(u) \equiv \mathrm{d}f/\mathrm{d}u$,

$$0 = \nabla\cdot\mathbf{A} \quad\Longrightarrow\quad \frac{\partial A_3}{\partial z} = 0 \quad\Longrightarrow\quad -A_3'(u) = 0 \quad\Longrightarrow\quad A_3 = \text{const}.$$

Hence $\mathbf{A} = (A_1(u),\, A_2(u),\, A_3)$ and

$$\begin{aligned}
\mathbf{E} &= -\frac{1}{c}\frac{\partial\mathbf{A}}{\partial t} = (-A_1',\, -A_2',\, 0) = (E_1(u),\, E_2(u),\, 0), \\
\mathbf{B} &= \nabla\times\mathbf{A} = (A_2',\, -A_1',\, 0) = (-E_2(u),\, E_1(u),\, 0),
\end{aligned}$$

(c) If $f = f(u) = f(ct - z)$ then any phase $f =$ const. has a path $z = ct + c$. That is, the entire "wave profile" $f(-z)$ at $t = 0$ moves steadily to the right as t increases, when plotted against coordinate z. For this electromagnetic wave, it is right-travelling and both electric and magnetic fields, $\mathbf{E}$ and $\mathbf{B}$ lie in the plane perpendicular to the direction of travel of the wave, $\mathbf{E} \cdot \mathbf{n} = \mathbf{B} \cdot \mathbf{n} = 0$ Furthermore they are orthogonal to each other, $\mathbf{B} \cdot \mathbf{E} = 0$, and equal in magnitude $|\mathbf{B}| = |\mathbf{E}|$. Such waves are known as *plane polarized.*

Problem 9.24 **Show that as a consequence of the Maxwell equations,**

$$T^{\beta}{}_{\alpha,\beta} = -\frac{1}{c} F_{\alpha\gamma} J^{\gamma}$$

where $T^{\beta}{}_{\alpha}$ is the electromagnetic energy-stress tensor (9.59), and when no charges and currents are present it satisfies Eq. (9.56). Show that the $\alpha = 4$ component of this equation has the form

$$\frac{\partial \epsilon}{\partial t} + \nabla \cdot \mathbf{S} = -\mathbf{j} \cdot \mathbf{E}$$

where $\epsilon =$ energy density and $\mathbf{S} =$ Poynting vector. Interpret this equation physically.

Solution: Lowering the second index in Eq. (9.59) gives

$$T^{\beta}{}_{\alpha} = \frac{1}{4\pi}\big(F^{\beta\rho} F_{\alpha\rho} - \frac{1}{4}\delta^{\beta}_{\alpha} F^{\rho\sigma} F_{\rho\sigma}\big).$$

Hence

$$\begin{aligned}
T^{\beta}{}_{\alpha,\beta} &= \frac{1}{4\pi}\big(F^{\beta\rho}{}_{,\beta} F_{\alpha\rho} + F^{\beta\rho} F_{\alpha\rho,\beta} \\
&\qquad - \frac{1}{4}\delta^{\beta}_{\alpha}(F^{\rho\sigma}{}_{,\beta} F_{\rho\sigma} + F^{\rho\sigma} F_{\rho\sigma,\beta}\big) \\
&= \frac{1}{4\pi}\big(-\frac{4\pi}{c} J^{\rho} F_{\alpha\rho} + F^{\beta\rho} F_{\alpha\rho,\beta} - \frac{1}{4}\delta^{\beta}_{\alpha}(F^{\rho\sigma}{}_{,\beta} F_{\rho\sigma} + F^{\rho\sigma} F_{\rho\sigma,\beta}\big) \\
&= -\frac{1}{c} J^{\rho} F_{\alpha\rho} + \frac{1}{4\pi}\big(F^{\beta\rho} F_{\alpha\rho,\beta} - \frac{1}{2} F^{\rho\sigma} F_{\rho\sigma,\alpha}\big) \\
&= \frac{1}{c} F_{\alpha\gamma} J^{\gamma} + \frac{1}{8\pi} F^{\sigma\rho}\big(2F_{\alpha\rho,\sigma} + F_{\rho\sigma,\alpha}\big) \\
&= -\frac{1}{c} F_{\alpha\gamma} J^{\gamma} + \frac{1}{8\pi} F^{\sigma\rho}\big(F_{\alpha\rho,\sigma} - F_{\alpha\sigma,\rho} + F_{\rho\sigma,\alpha}\big) \\
&= -\frac{1}{c} F_{\alpha\gamma} J^{\gamma} + \frac{1}{8\pi} F^{\sigma\rho}\big(F_{\alpha\rho,\sigma} + F_{\sigma\alpha,\rho} + F_{\rho\sigma,\alpha}\big) \\
&= -\frac{1}{c} F_{\alpha\gamma} J^{\gamma}
\end{aligned}$$

on using the source-free Maxwell equation, $F_{\alpha\rho,\sigma} + F_{\sigma\alpha,\rho} + F_{\rho\sigma,\alpha} = 0$.
For $\alpha = 4$ the equation reads

$$T^{4}{}_{4,4} + T^{i}{}_{4,i} = -\frac{1}{c} F_{4i} J^{i},$$

or, on raising the second index 4,

$$-T^{44}{}_{,4} - T^{4i}{}_{,i} = \frac{1}{c} E_i J^i.$$

In terms of the energy density $\epsilon = T^{44}$ and Poynting vector $\mathbf{S}$ whose components are $S_i = cT^{4i}$,

$$-\frac{1}{c}\frac{\partial \epsilon}{\partial t} - \frac{1}{c}\nabla \cdot \mathbf{S} = \frac{1}{c}\mathbf{E} \cdot \mathbf{j}$$

i.e.

$$\frac{\partial \epsilon}{\partial t} + \nabla \cdot \mathbf{S} = -\mathbf{E} \cdot \mathbf{j}.$$

The interpretation of this equation is that the rate of energy increase per unit volume in the electromagnetic field is equal to the work done by the charges per unit volume (i.e. minus the work done on the charges by the electric field).

Problem 9.25 For a plane wave, Problem 9.23, show that

$$T_{\alpha\beta} = \epsilon\, n_\alpha\, n_\beta$$

where $\epsilon = E^2/4\pi$ and $n^\alpha = (\mathbf{n}, 1)$ is the null vector pointing in the direction of propagation of the wave. What pressure does the wave exert on a wall placed perpendicular to the path of the wave?

Solution: For a plane wave, $\mathbf{E} = (E_1,\, E_2,\, 0)$, $\mathbf{B} = (-E_2,\, E_1,\, 0)$, the electromagnetic tensor is

$$[F_{\mu\nu}] = \begin{pmatrix} 0 & 0 & -E_1 & E_1 \\ 0 & 0 & -E_2 & E_2 \\ E_1 & E_2 & 0 & 0 \\ -E_1 & -E_2 & 0 & 0 \end{pmatrix}$$

so that

$$F_{\rho\sigma}F^{\rho\sigma} = 2(B^2 - E^2) = 0.$$

Hence the energy-stress tensor is

$$T_{\alpha\beta} = \frac{1}{4\pi} F_{\alpha\rho} F_\beta{}^\rho$$

and component by component computation results in

$$[T_{\alpha\beta}] = \frac{1}{4\pi}\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & E^2 & -E^2 \\ 0 & 0 & -E^2 & E^2 \end{pmatrix}$$

(e.g. $T_{11} \propto F_{12}F_{12} + F_{13}F_{13} - F_{14}F_{14} = 0 + (E_1)^2 - (E_1)^2$ etc.) Hence, setting $n^\alpha = (\mathbf{n}, 1) = (0,\, 0,\, 1,\, 1)$ we have $n_\alpha = (0,\, 0,\, 1,\, -1)$ and

$$T_{\alpha\beta} = \epsilon\, n_\alpha\, n_\beta$$

where $\epsilon = E^2/4\pi$.
The pressure on a plane perpendicular to the z-direction is $T_{33} = E^2/4\pi = \epsilon$.

Chapter 10

Problem 10.1 **Give an example in $\mathbb{R}^2$ of each of the following:**

(a) A family of open sets whose intersection is a closed set that is not open.

(b) A family of closed sets whose union is an open set that is not closed.

(c) A set that is neither open nor closed.

(d) A countable dense set.

(e) A sequence of continuous functions $f_n : \mathbb{R}^2 \to \mathbb{R}$ whose limit is a discontinuous function.

Solution: (a) The open balls $\{B_{1+1/n}(\mathbf{0})$ where $n = 1, 2, \ldots$ has intersection the closed ball

$$\overline{B_1(\mathbf{0})} = \{\mathbf{r} \in \mathbb{R}^2 \,|\, |\mathbf{r}| \leq 1\}.$$

This set is closed, being the closure of an open set, and not open since every ball $B_\epsilon(\mathbf{r})$ centred on a point on the circumference, $|\mathbf{r}| = 1$ intersects both $\overline{B_1(\mathbf{0})}$ and its complement.

(b) The union of closed balls $\overline{B_{1-1/n}(\mathbf{0})}$, where $n = 1, 2, \ldots$ is the open ball $B_1(\mathbf{0})$, which is not closed since its closure is the closed ball and includes points on the circumference, having $|\mathbf{r}| = 1$.

(c) There are, of course, many such examples. Consider for example the union of two disjoint sets, one open, the other closed such as

$$B_1((0,0)) \cup \overline{B_1((0,2))}.$$

(d) The set of all rational points (x, y) such that $x = p/q$, $y = r/s$ where p, q, r, s are integers. This set is dense since there is a rational point in any open ball neighbourhood of any point of $\mathbb{R}$, and it is countable by Theorem 1.2 and Corollary 1.3.

(e) Let $h_n : \mathbb{R}^+ \to \mathbb{R}$ be the continuous function on the positive real numbers, $\mathbb{R}^+ = \{x \,|\, x \geq 0\}$, defined by

$$h_n(x) = \begin{cases} 0 & \text{if } x \geq 1/n \\ 1 - nx & \text{if } 0 \leq x \leq 1/n \end{cases}$$

The sequence of continuous functions $f_n : \mathbb{R}^2 \to \mathbb{R}$ defined by $f_n(\mathbf{r}) = h_n(|\mathbf{r}|)$ then has the limit

$$h_n(\mathbf{r}) \longrightarrow f(\mathbf{r}) = \begin{cases} 1 & \text{if } \mathbf{r} = \mathbf{0} \\ 0 & \text{else} \end{cases} \quad .$$

The function f is not continuous at $\mathbf{r} = \mathbf{0}$, since the inverse image of the neighbourhood of 1, $f^{-1}((\frac{1}{2}, \frac{3}{2}))$ consists of the isolated point $\mathbf{0}$, which is clearly not an open set.

Problem 10.2 If $\mathcal{U}$ generates the topology on X show that $\{A \cap U \mid U \in \mathcal{U}\}$ generates the relative topology on A.

Solution: The relative topology on A includes all sets of the the form $A \cap O$ such that O is open in X. If $\mathcal{O}$ is the topology on X generated by the family of sets $\mathcal{U}$, then every set $U \in \mathcal{U}$ is open in this topology. Hence every $A \cap U$ where $U \in \mathcal{U}$ is open in the topology on A induced by $\mathcal{O}$. It remains to show that it is the weakest such topology. If $B \subset A$ is in the topology generated by all sets of the form $\{A \cap U \mid U \in \mathcal{U}\}$ then it is a union of finite intersections of sets of the form $A \cap U_i$ where $U_i \in \mathcal{U}$. Since, by Problem 1.2, for any family (finite or infinite) of such sets U_i we have

$$\bigcap_i (A \cap U_i) = A \cap \big(\bigcap U_i\big), \qquad \bigcup (A \cap U_i) = A \cap \big(\bigcup U_i\big),$$

it follows that every B in the topology generated by $\{A \cap U \mid U \in \mathcal{U}\}$ is of the form $A \cap V$ where V is open in the topology on X generated by $\mathcal{U}$. Hence it lies in the topology induced on A by this topology, which is therefore the weakest topology on A including all sets of the form $\{A \cap U \mid U \in \mathcal{U}\}$.

Problem 10.3 Let X be a topological space and $A \subset B \subset X$. If B is given the relative topology, show that the relative topology induced on A by B is identical to the relative topology induced on it by X.

Solution: If $U \subset A$ is open in the topology induced by B then $U = A \cap V$ where V is open in B. Since B is given the relative topology induced by the topology on X, we have $V = B \cap W$ where W is open in X. Hence

$$U = A \cap (B \cap W) = (A \cap B) \cap W$$

and U is open in the relative topology induced on A by X. Conversely if $U \subset A$ is open in the relative topology induced by X then $U = A \cap W$ where W is open in X. Hence $U = (A \cap B) \cap W = A \cap (B \cap W)$ and U is open in the topology induced by B, since $B \cap W$ is open in the relative topology on B induced by X.

Problem 10.4 Show that for any subsets U, V of a topological space $\overline{U \cup V} = \overline{U} \cup \overline{V}$. Is it true that $\overline{U \cap V} = \overline{U} \cap \overline{V}$? What corresponding statements hold for the interior and boundaries of unions and intersections of sets?

Solution: Let x be an accumulation point of $U \cup V$. Then any open neighbourhood S of x contains a point $y \in U \cup V$ such that $y \neq x$. The point belongs to either U or V and therefore x is an accumulation point of either U or of V. Hence $\overline{U \cup V} \subseteq \overline{U} \cup \overline{V}$. Conversely, every accumulation point x of U is evidently an accumulation point of $U \cup V$, since every open neighbourhood of x contains a point $y \neq x$ which lies in U, and therefore $y \in U \cup V$. Similarly, for accumulation points of V, so that $\overline{U} \cup \overline{V} \subseteq \overline{U \cup V}$, and we have shown the equality of these two sets.
It is not true that $\overline{U \cap V} = \overline{U} \cap \overline{V}$, for let

$$U = \{(x, y) \,|\, x < 0\} \subset \mathbb{R}^2, \qquad V = \{(x, y) \,|\, x > 0\} \subset \mathbb{R}^2.$$

Then $U \cap V = \emptyset$ and therefore $\overline{U \cap V} = \emptyset$. However $\overline{U} = \{(x, y) \,|\, x \leq 0\}$ and $\overline{V} = \{(x, y) \,|\, x \geq 0\}$, so that

$$\overline{U} \cap \overline{V} = \{(0, y) \,|\, y \in \mathbb{R}\} \neq \emptyset.$$

Any open subset of $U \cap V$ is an open subset both of U and V, and conversely any open set A which lies entirely in U and entirely in V is an open subset of $U \cap V$. Hence $(U \cap V)^o = U^o \cap V^o$. However, $(U \cup V)^o \neq U^o \cup V^o$, for let

$$U = \{(x, y) \,|\, x \leq 0\} \subset \mathbb{R}^2, \qquad V = \{(x, y) \,|\, x \geq 0\} \subset \mathbb{R}^2.$$

Then $U \cup V = \mathbb{R}^2$ and $(U \cup)^o = \mathbb{R}^2$, but $U^o \cup V^o$ does not include the line $x = 0$.

For boundaries, $b(U \cup V) = b(U) \cup b(U)$ and $b(U \cap V) \neq b(U) \cap b(U)$; the demonstration is essentially identical to that for the closure of sets given above.

Problem 10.5 **If A is a dense set in a topological space X and $U \subseteq X$ is open, show that $U \subseteq \overline{A \cap U}$.**

Solution: Let $x \in U$. If $x \in A \cap U$ then $x \in \overline{A \cap U}$. If $x \notin A$, then let V be any open neighbourhood of x. Then $W = U \cap V$ is an open set containing x and $W \subseteq U$. Since A is a dense set, there must be an element $y \in A \cap W$, and $y \neq x$ since $x \notin A$. Hence the arbitrary open neighbourhood V of x contains a point $y \neq x$ which belongs to $A \cap W$ and therefore to $A \cap U$. Hence x is a point of accumulation of $A \cap U$, and we have shown $U \subseteq \overline{A \cap U}$.

Problem 10.6 **Show that a map $f : X \to Y$ between two topological spaces X and Y is continuous if and only if $f(\overline{U}) \subseteq \overline{f(U)}$ for all sets $U \subseteq X$. Show that f is a homeomorphism only if $f(\overline{U}) = \overline{f(U)}$ for all sets $U \subseteq X$.**

Solution: If f is continuous then for any open set $A \in Y$ the inverse image $f^{-1}(A)$ is an open subset of X. Taking complements, this implies that the inverse image under f of every closed set is closed. Let U be any subset of X. The closure of its image in Y, $\overline{f(U)}$ is a closed set whose inverse image under the continuous map f

is a closed set such that

$$f^{-1}\big(\overline{f(U)}\big) \supseteq f^{-1}\big(f(U)\big) \supseteq U.$$

Since $\overline{U}$ is the small closed set containing U, we have, as required,

$$f^{-1}\big(\overline{f(U)}\big) \supseteq \overline{U}.$$

Conversely, suppose $f(\overline{U}) \subseteq \overline{f(U)}$ for all sets $U \subseteq X$. Let V be any open subset of Y. To prove that f is continuous we must show that $W = f^{-1}(V)$ is an open subset of X, or equivalently that $X - W$ is closed, $\overline{X - W} = X - W$. Let $x \in \overline{X - W}$; our aim is to show that $x \in X - W$. Suppose, to the contrary, that $x \in W$. Then $f(x) \in V$. However, by supposition,

$$f(\overline{X - W}) \subseteq \overline{f(X - W)} = \overline{f(X) - V}.$$

Hence $f(x)$ is in the closure of $f(X) - V$; i.e. it either lies in $f(x) - V$ or is a point of accumulation of $f(X) - V$. In either case, every open neighbourhood containing $f(x)$ intersects $f(X) - V$. However this is not true of the open neighbourhood V. Hence $X - W$ is closed and $W = f^{-1}(V)$ is open.

If f is a homeomorphism then it is continuous and therefore $f(\overline{U}) \subseteq \overline{f(U)}$ for all sets $U \subseteq X$. Since the map $f^{-1} : Y \to X$ is also continuous $f^{-1}(\overline{V}) \subseteq \overline{f^{-1}(V)}$ for all sets $V \subseteq Y$. Setting $V = f(U)$, we have $f^{-1}(U)$ since f is one-to-one. Hence

$$f^{-1}\big(\overline{f(U)}\big) \subseteq \overline{U},$$

or equivalently $\overline{f(U)} \subseteq f(\overline{U})$. Combining with the earlier, we have $\overline{f(U)} = f(\overline{U})$.

Problem 10.7 Show the following:

(a) In the trivial topology, every sequence x_n converges to every point of the space $x \in X$.

(b) In R^2 the family of open sets consisting of all open balls centred on the origin $B_r(0)$ is a topology. Any sequence $\mathbf{x}_n \to \mathbf{x}$ converges to all points on the circle of radius $|\mathbf{x}|$ centred on the origin.

(c) If C is a closed set of a topological space X it contains all limit points of sequences $x_n \in C$.

(d) Let $f : X \to Y$ be a continuous function between topological spaces X and Y. If $x_n \to x$ is any convergent sequence in X then $f(x_n) \to f(x)$ in Y.

Solution: (a) In the trivial topology the only open neighbourhood of any point x is the entire space X. Hence every sequence x_n converges to x since it lies entirely in all open neighbourhoods of x.

(b) Any finite intersection and all unions of such balls centered on the origin is an open ball center on $\mathbf{0}$, or it is the entire space $\mathbb{R}^2$. Thus these sets form a topology on $\mathbb{R}^2$. A sequence $\mathbf{x}_n \to \mathbf{x}$ in this topology iff for every open ball with radius $r > |\mathbf{x}|$ there is an element $\mathbf{x}_n$ such that $|\mathbf{x}_n| < r$. This is true of all other points $\mathbf{x}'$ having the same distance from the origin, $|\mathbf{x}'| = |\mathbf{x}|$, i.e. points lying on the circle of radius $|\mathbf{x}|$ centre $\mathbf{0}$.

(c) Let $x_n \to x$ be a sequence of points such that $x_n \in C$. Suppose $x \notin C$, i.e. $x \in X - C$. Then $U = X - C$ is an open neighbourhood of x which contains no points of x_n, contradicting the assumption that x is a limit point of the sequence. Hence $x \in C$.

(d) Let U be any open neighbourhood of $f(x)$ in Y. Then $f^{-1}(U)$ is an open neighbourhood of x, and since $x_n \to x$ there exists integer N such that $x_n \in f^{-1}(U)$ for all $n \geq N$; i.e. $f(x_n) \in U$ for all $n \geq N$. Hence $f(x_n) \to f(x)$.

Problem 10.8 **If W, X and Y are topological spaces and the functions $f : W \to X$, $g : X \to Y$ are both continuous, show that the function $h = g \circ f : W \to Y$ is continuous.**

Solution: Let U be any open subset of Y. Then by continuity of g, $g^{-1}(U)$ is an open subset of X, and since f is continuous $f^{-1}(g^{-1}(U))$ is an open subset of W. Now

$$(g \circ f)^{-1}(U) = f^{-1}(g^{-1}(U))$$

for

$$\begin{aligned} x \in (g \circ f)^{-1}(U) &\iff g(f(x)) \in U \\ &\iff f(x) \in g^{-1}(U) \\ &\iff x \in f^{-1}(g^{-1}(U)) \end{aligned}$$

Hence $(g \circ f)^{-1}(U)$ is an open subset of W, showing that $g \circ f$ is continuous.

Problem 10.9 **Show that every metric space is first countable. Hence show that every subset of a metric space can be written as the intersection of a countable collection of open sets.**

Solution: For each $x \in X$ let $\mathcal{U}_x = \{U_n(x) = B_{1/n}(x) \mid n = 1, 2, 3, \dots\}$. This is a countable family of open neighbourhoods of x. If U is an open neighbourhood of x if there is an open ball $B_a(x) \subset U$. Set n such that $1/n < a$, and we have $U_n(x) \subset U$. If A is any subset of X, let A_n be the open set

$$A_n = \bigcup_{x \in A} U_n(x).$$

Using Problem 1.2

$$\begin{aligned}\bigcap_{n=1}^{\infty} A_n &= A_1 \cap A_2 \cap A_3 \cap \dots \\ &= \bigcup_{x\in A} \bigcap_{n=1}^{\infty} U_n(x) \\ &= \bigcup_{x\in A} x \\ &= A\end{aligned}$$

since $\bigcap_{n=1}^{\infty} U_n(x) = U_1(x) \cap U_2(x) \cap U_3(x) \cap \dots = x$.

Problem 10.10 **If $\mathcal{U}_1$ and $\mathcal{U}_2$ are two families of subsets of a set X, show that the topologies generated by these families are homeomorphic if every member of $\mathcal{U}_2$ is a union of sets from $\mathcal{U}_1$ and vice versa. Use this property to show that the metric topologies on $\mathbb{R}^n$ defined by the metrics d, d_1 and d_2 are all homeomorphic.**

Solution: Every member of $\mathcal{U}_2$ is an open set in the topology $\mathcal{O}_1$ generated by $\mathcal{U}_1$ since it is a union of sets from $\mathcal{U}_1$. The intersection $U \cap V$ of any two members of $\mathcal{U}_2$ is also open in the topology $\mathcal{O}_1$, for if

$$U = \bigcup_{i\in I} U_{(1)i} \qquad \text{and} \qquad V = \bigcup_{j\in J} V_{(1)j}$$

then

$$U \cap V = \bigcup_{i\in I} \bigcup_{j\in J} U_{(1)i} \cap V_{(1)j}$$

This argument extends to all finite intersections of members of $\mathcal{U}_2$. The general member of the topology $\mathcal{O}_2$ generated by $\mathcal{U}_2$ is a union of such finite intersections, which must therefore be open in the topology $\mathcal{O}_1$. Hence every open set in the topology $\mathcal{O}_2$ is open in the topology $\mathcal{O}_1$, and by an identical argument all open sets in $\mathcal{O}_1$ are open in $\mathcal{O}_2$. Hence the the map $\mathrm{id}_X : X \to X$, regarded as a map from the topological space $(X, \mathcal{O}_1)$ to the topological space $(X, \mathcal{O}_2)$, is continuous as

$$\mathrm{id}_X^{-1}(U) = U \qquad \text{for all sets } U$$

and if U is open in $\mathcal{O}_2$ then its inverse image under id_X is continuous in $\mathcal{O}_1$. As id_X is one-to-one and the same properties apply to id_X^{-1}, the map id_X is a homeomorphism.

The open ball $B_a(\mathbf{r})$ defined by metric d is an open region bounded by a sphere of radius a. For the metric d_1 the open region

$$B'_a(\mathbf{0}) = \{\mathbf{x} \mid d_1(\mathbf{x}) < a\} = \{\mathbf{x} \mid |x_1| + |x_2| + \dots + |x_n| < a\}$$

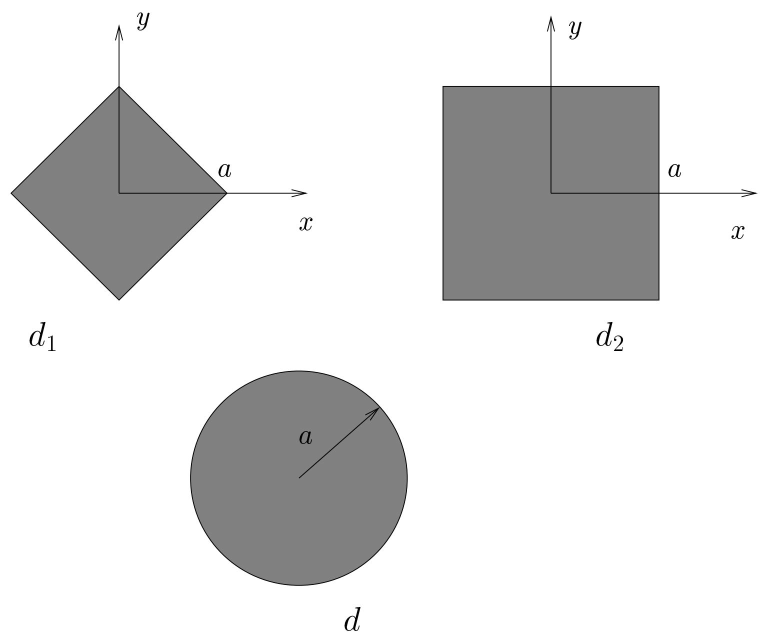
is the intersection of open half-spaces determined by the inequalities

$$\pm x_1 \pm x_2 \cdots \pm x_n < a.$$

This is an open diamond shaped region (see figure below). For $B'_a(\mathbf{y}) = \{\mathbf{x} \,|\, d_1(\mathbf{x} - \mathbf{y}) < a\}$ is centred on the point $\mathbf{y}$. It is an open set, and therefore every point $\mathbf{x} \in B'_a(\mathbf{y})$ lies in an open ball $B_a(\mathbf{x}) \subset B'_a(\mathbf{y})$. Conversely it can be seen by simple inspection that any point x in an open spherical ball can be placed at the centre of a diamond shaped region within the ball. For d_2 the open region

$$B''_a(\mathbf{0}) = \{\mathbf{x} \,|\, d_2(\mathbf{x}) < a\} = \{\mathbf{x} \,|\, \max(|x_1|, |x_2|, \ldots, |x_n|) < a\}$$

is the intersection of open half-spaces determined by the inequalities $x_i < a$, $x_i > -a$ $(i = 1, 2, \ldots, n)$. This is an open cubical region. Every point in such a region lies within an open spherical ball, and conversely every point in a spherical ball can be placed at the center of a cubical region centred on this point. Hence the topologies generated by metrics d and d_1 are homeomorphic, as are the topologies generated by d and d_2. It follows that the topologies generated by d_1 and d_2 are also homeomorphic, since homeomorphism is an equivalence relation between topologies.



Problem 10.11 **A topological space X is called normal if for every pair of disjoint closed subsets A and B there exist disjoint open sets U and V such that $A \subset U$ and $B \subset V$. Show that every metric space is normal.**

Solution: For any set $A \subset X$ and point $x \in X$ we may define the distance between x and A to be the least upper bound of distances between x and points of A,

$$d(x, A) = \inf_{y \in A} d(x, y).$$

If A is closed then $x \in A$ iff $d(x, A) = 0$, for (i) $x \in A \Longrightarrow d(x, A)$ since $d(x, x) = 0$, and (ii) since $X - A$ is an open set, if $x \notin A$ there exists $\epsilon > 0$ such that $d(x, y) > \epsilon$ for all $y \in A$.
If A and B are two disjoint closed sets, $A \cap B = \emptyset$, let U and V be the sets

$$U = \{x \in X \,|\, d(x, A) < d(x, B)\}, \qquad V = \{x \in X \,|\, d(x, A) > d(x, B)\}.$$

U and V are clearly disjoint, and $A \subset U$ for

$$x \in A \Rightarrow x \notin B \Rightarrow (d(x, A) = 0 \text{ and } d(x, B) > 0) \Longrightarrow x \in U.$$

Similarly $B \subset V$.
The set U is open, for if $x, y \in U$ then by (Met4)

$$d(y, z) \le d(x, y) + d(x, z) \qquad \text{for all } z \in A.$$

Hence

$$d(y, A) \le d(x, y) + d(x, A).$$

Similarly,

$$d(x, B) \le d(x, y) + d(y, B).$$

Setting $r = d(x, B) - d(x, A) > 0$ we can combine these two equations to give

$$d(y, A) \le 2d(x, y) + d(y, B) - r < d(y, B)$$

if $d(y, x) < \frac{1}{2}r$. Thus for any $x \in U$ the entire open ball $B_{r/2}(x) \subset U$, which shows that U is open. Similarly V is open.

Problem 10.12 **If $f : X \to Y$ is a continuous map between topological spaces, we define its graph to be the set $G = \{(x, f(x)) \,|\, x \in X\} \subseteq X \times Y$. Show that if G is given the relative topology induced by the topological product $X \times Y$ then it is homeomorphic to the topological space X.**

Solution: The induced topology on G is such that any subset $A \subset G$ is open if it is the union of sets of the type $(U \times V) \cap G$, where $U \subset X$ and $V \subset Y$ are open subsets of X and Y resp. Let $\varphi : X \to G$ be the map defined by $\varphi(x) = (x, f(x))$. This map is clearly one-to-one and onto, therefore invertible. Our aim is to show that both φ and φ^{-1} are continuous.
(i) Let A be any open subset of G. The inverse image of any set $(U \times V) \cap G$, where U is an open subset of X and V an open subset of Y, is

$$\varphi^{-1}((U \times V) \cap G) = U \cap f^{-1}(V).$$

This is an open subset of X since f is continuous (so that $f^{-1}(V)$ is open for all open $V \subset Y$). Hence $\varphi^{-1}(A)$, being a union of sets of the type $\varphi^{-1}((U \times V) \cap G)$ is an open subset of X. Hence φ is a continuous map.
(ii) The inverse map φ^{-1} is also continuous, for let U be any open set in X. Then

$$(\varphi^{-1})^{-1}(U) = \varphi(U) = \mathrm{pr}_1^{-1}(U) \cap G$$

is an open subset of G in the induced topology since it is the intersection of G with the open subset $\mathrm{pr}_1^{-1}(U)$ of $X \times Y$.
Hence φ is a homeomorphism bewteen X and G.

Problem 10.13 **Let X and Y be topological spaces and $f : X \times Y \to X$ a continuous map. For each fixed $a \in X$ show that the map $f_a : Y \to X$ defined by $f_a(y) = f(a, y)$ is continuous.**

Solution: We may write $f_a(y) = f(\iota_a(y)$ where $\iota_a : Y \to X \times Y$ is the injection map defined in Theorem 10.3. Hence f_a is continuous since it is the composition of two continuous maps (see Problem 10.8)

$$f_a = f \circ \iota_a.$$

Problem 10.14 **If Y is a Hausdorff topological space show that every continuous map $f : X \to Y$ from a topological space X with indiscrete topology into Y is a *constant map*; that is, a map of the form $f(x) = y_0$ where y_0 is a fixed element of Y.**

Solution: Let Y be Hausdorff and x, x' any two points in a space X having indiscrete topology. Set $y = f(x) \in Y$, $y' = f(x') \in Y$. For any open set $U \subset Y$ such that $y \in U$, continuity implies $f^{-1}(U)$ is an open set in X. Hence $f^{-1}(U) = \emptyset$ or X. The first possibility is out since $f(x) = y$, and the second implies $y' = f(x') \in U$. Hence every open set containing y also must contain y', contradicting the Hausdorff property (actually this only needs a slightly weaker separation property than Haurdorff called the T_0 separation property).

Problem 10.15 **Show that if $f : X \to Y$ and $g : X \to Y$ are continuous maps from a topological space X into a Hausdorff space Y then the set of points A on which these maps agree, $A = \{x \in X \mid f(x) = g(x)\}$, is closed. If A is a dense subset of X show that $f = g$.**

Solution: To show A is closed we need to show that its complement, the set of points on which f and g disagree,

$$X - A = \{x \in X \mid f(x) \neq g(x)\}$$

is open. If $x \in X - A$ then $g(x) \neq f(x)$, and since Y is Hausdorff, it is possible to find open disjoint sets U and V in Y such

$$g(x) \in U, \quad f(x) \in V \qquad U \cap V = \emptyset.$$

Since g and f are continuous maps the set $B = g^{-1}(U) \cap f^{-1}(V)$ is an open subset of X. For all $x \in B$ we must have $g(x) \in U$ and $f(x) \in V$; hence $g(x) \neq f(x)$, since $U \cap V = \emptyset$. Hence every $x \in X - A$ has an open neighbourhood B such that $x \in B \subset X - A$, which proves that $X - A$ is open, since it is the union of such open neighbourhoods at every one of its points. Therefore the complementary set A is closed.

If A is a dense set every point of X either belongs to A or is an accumulation point of A. Since A is closed it must contain all its accumulation points, hence $A = X$. Hence $f(x) = g(x)$ for all $x \in X$.

Problem 10.16 Show that every compact Hausdorff space is normal (see Problem 10.11).

Solution: Let A and B be any pair of disjoint closed sets, $A \cap B = \emptyset$, in a compact Hausdorff space X. Let x be any point of of the space X. By the Hausdorff property, for each $y \in B$ there exists an open neighbourhood U_x of x and V_y of y such that $U_x \cap V_y = \emptyset$. The sets V_y for all $y \in B$ form an open cover of B. By compactness, we may pick a finite subcover V_{y_i} $(i = 1, \dots n)$, such that $B \subset W_x = \bigcup_{i=1}^{n} V_{y_i}$, where W_x an open set. Furthermore the set $Y_x = \bigcap_{i=1}^{n} U_n$, being a finite intersection of open sets, is an open set covering x. That is, for every point $x \in X$

$$x \in Y_x, \qquad B \subset W_x, \qquad Y_x \cap W_x = \emptyset.$$

For each point $x \in A$ let us define such a pair of open sets (Y_x, W_x). Since the sets Y_x form an open cover of A it is possible, by compactness, to find a finite subcover Y_i $(i = 1, \dots m)$. Setting $Y = \bigcup_{i=1}^{m} Y_i$ and $W = \bigcap_{i=1}^{m} W_i$ we have that both Y and W are open sets such that

$$A \subset Y, \qquad B \subset W, \qquad Y \cap W = \emptyset,$$

showing that X is a normal space.

Problem 10.17 Show that every one-to-one continuous map $f : X \to Y$ from a compact space X onto a Hausdorff space Y is a homeomorphism

Solution: Let U be any open subset of X. To show that $f^{-1} : Y \to X$ is continuous it is necessary show that $(f^{-1})^{-1}(U) = f(U)$ is an open subset of Y. The complement $X - U$ is a closed set and therefore, by Theorem 10.9, it is a compact subspace of X. By Theorem 10.10 its image $f(X - U)$ is a compact subspace of Y, and by Theorem (10.13) it must be a closed subset. Hence

$$f(U) = f(X - (X - U)) = Y - f(X - U)$$

is an open subset of X. The map f^{-1} is therefore continuous, and since f is a continuous invertible map it is a homeomorphism.

Problem 10.18 Show that a topological space X is connected if and only if every continuous map $f : X \to Y$ of X into a discrete topological space Y consisting of at least two points is a constant map (see Problem 10.14).

Solution: If X is connected let $y = f(x) \in X$. Since the singleton $\{y\}$ is an open set in the discrete topology of Y, we have $U = f^{-1}(\{y\})$ is an open subset of X and $x \in U$. Since the complement $Y - \{y\}$ is also an open set in the discrete topology of Y, its inverse image $U' = f^{-1}(Y - \{y\}) = X - U$ is an open set such that $X = U \cup U'$. Since X is connected and $U \neq \emptyset$ we must have $U' = \emptyset$. Hence $U = X$ and every point of X is mapped to the same point y; i.e. f is a constant map.
Conversely, if suppose X not disconnect, then $X = U \cup U'$ where U and U' are disjoint open sets, $U \cap U' = \emptyset$. Let X be the space of two points $\{0, 1\}$ with discrete topology generated by singletons $\{0\}$ and $\{1\}$, and f any map such that $f(0) = y \in U$, $f(1) = y' \in U'$. Such a map is clearly not constant; however it is continuous, for the inverse of any open set in Y is either $\emptyset$, $\{0\}$, $\{1\}$ or $\{0, 1\}$, i.e. an open set in the discrete topology of $(0, 1)$. Hence there exists a continuous non-constant map $f : X \to Y$ where Y is a discrete topological space consisting of at least two points.

Problem 10.19 From Theorem 10.16 show that the unit circle S^1 is connected, and that the punctured n-space $\dot{\mathbb{R}}^n = \mathbb{R}^n - \{\mathbf{0}\}$ is connected for all $n > 1$. Why is this not true for $n = 1$?

Solution: The demonstration that the unit circle $S^1 = \{(x, y) \,|\, x^2 + y^2 = 1\}$ is connected proceeds essentially as for S^2 in Example 10.19. Consider the "punctured" circles $S' = S^2 - \{(0, 1)\}$ and $S'' = S^2 - \{(0, -1)\}$. Each of these is homeomorphic to the real line $\mathbb{R}$; e.g. S' is homeomorphic to $\mathbb{R}$ under the map

$$(x, y) \mapsto x' = \frac{x}{1 - y}, \qquad \text{where} \qquad y = \sqrt{1 - x^2}$$

and similarly for S''. Clearly $S' \cap S'' \neq \emptyset$ and $S^1 = S' \cup S''$ whence, by Theorem 10.16, S^1 is a connected subset of $\mathbb{R}^2$.

The punctured n-space $\dot{\mathbb{R}}^n = \mathbb{R}^n - \{\mathbf{0}\}$ is the union of all half lines

$$A_{\mathbf{k}} = \{t\mathbf{k} \,|\, t > 0\}$$

where $\mathbf{k}$ is a unit vector,

$$(k_1)^2 + (k_2)^2 + \cdots + (k_n)^2 = 1.$$

Each of these half lines is clearly connected, since $t \mapsto (t-1)/t$ provides a homeomorphism with the entire real line. Since each half line intersects the unit sphere S^{n-1} in precisely one point (where $t = 1$) and S^{n-1} is connected, we have by Theorem 10.16 that the puncture n-space $\dot{\mathbb{R}}^n = S^{n-1} \cup \bigcup_{\mathbf{n}} A_{\mathbf{n}}$ is a connected subset of $\mathbb{R}^n$.

The punctured line $\dot{\mathbb{R}}$ is not connected since it is the union of two disjoint open sets

$$\mathbb{R} = \{x \,|\, x < 0\} \cup \{x \,|\, x > 0\}.$$

The above argument fails because the 0-sphere, $S^0 = \{1\} \cup \{-1\}$ is clearly a disconnected space.

Problem 10.20 Show that the real projective space defined in Example 10.15 is connected, Hausdorff and compact.

Solution: The map $\phi : S^n \to P^n$ which assigns to each point of the sphere the line passing through it is continuous (but not one-to-one, opposite points of the sphere being mapped to the same point of P^n). Continuity follows, because a set $U \subset P^n$ is open iff the union of lines making up U is an open subset U' of $\mathbb{R}^{n+1}$. Hence the inverse image $\phi^(U) = U' \cap S^n$ which by definition of induced topology is an open subset of S^n. By Theorems 10.10 and 10.17 this implies that P^n is compact and connected, since these properties pass to images under a continuous map and both hold for S^n.

To show that P^n is Hausdorff, we first remark that S^n is Hausdorff by Corallary 10.6, since it is a subspace of $\mathbb{R}^{n+1}$. Let $[\mathbf{u}]$ and $[\mathbf{v}]$ be two distinct points, corresponding to points $\mathbf{u}$ and $\mathbf{v}$ of S^n which are not diametrically opposite ($\mathbf{u} \neq \pm\mathbf{v}$). It is possible to find open neighbourhoods U of $\mathbf{u}$ and V of $\mathbf{v}$ in S^n such that $U \cap V = \emptyset$. Similarly there exist open neighbourhoods U' of $\mathbf{u}$ and V' of $-\mathbf{v}$ such that $U' \cap V' = \emptyset$. Then $U \cap U'$ is an open neighbourhood of $\mathbf{u}$ and $V \cap (-V')$ an open neighbourhood of $\mathbf{v}$, where $-V' = \{-\mathbf{x} \,|\, \mathbf{x} \in V'\}$. These two neighbourhoods on S^n do not intersect,

$$(U \cap U') \cap (V \cap (-V')) = U \cap V \cap U' \cap (-V') = \emptyset$$

and neither do any of their opposites, e.g.

$$-(U \cap U') \cap (V \cap (-V')) = (-U) \cap (-U') \cap V \cap (-V') = -(U \cap V') \cap (-U') \cap V = \emptyset$$

etc. Hence the sets $\phi(U \cap U')$ and $\phi(V \cap (-V'))$ are non-intersecting open neighbourhoods in P^n of $[\mathbf{u}]$ and $[\mathbf{v}]$ respectively.

Problem 10.21 Show that the rational numbers $\mathbb{Q}$ are a disconnected subset of the real numbers. Are the irrational points a disconnected subset of $\mathbb{R}$? Show that the connected components of the rational numbers $\mathbb{Q}$ consist of singleton sets $\{x\}$.

Solution: Let x_0 be any irrational number, say $x_0 = \sqrt{2}$. The rational numbers $\mathbb{Q}$ are then the union of two disjoint open sets in the induced topology from $\mathbb{R}$,

$$\mathbb{Q} = \big(\mathbb{Q} \cap \{x \in \mathbb{Q} \,|\, x < \sqrt{2}\}\big) \cup \big(\mathbb{Q} \cap \{x \in \mathbb{Q} \,|\, x > \sqrt{2}\}\big).$$

It is therefore a disconnected subset of $\mathbb{R}$.
Similarly the irrational numbers are disconnected by dividing them into those less than 0 and those greater than 0, both of which are induced open sets in the relative topology.
The connected component of any $x \in \mathbb{Q}$ is the singleton set $\{x\}$, for
(i) it is clearly connected since it cannot be divided into two disjoint non-empty sets of any kind and
(ii) if C is any subset of $\mathbb{Q}$ containing two rational numbers $x < y$, then it is always possible to find an irrational number s between them, $x < s < y$ and

$$C_1 = C \cap \{u \in \mathbb{Q} \,|\, u < s\} \qquad \text{and} \qquad C_2 = C \cap \{u \in \mathbb{Q} \,|\, u > s\}$$

are disjoint open subsets of C in the induced topology such that $C = C_1 \cap C_2$. Hence C is not a connected subset of $\mathbb{Q}$.

Problem 10.22 If G_0 is the component of the identity of a locally connected topological group G, the factor group G/G_0 is called the group of components of G. Show that the group of components is a discrete topological group with respect to the topology induced by the natural projection map $\pi : g \mapsto gG_0$.

Solution: By Theorem G_0 is a closed normal subgroup. Hence the factor space G/G_0, consisting of all cosets gG_0 forms a group. Its topology is the finest such that the natural projection map $\pi : g \mapsto gG_0$ defined by $\pi(g) = gG_0$ is continuous, and is such that π is an open continuous map (see paragraph before Theorem 10.22). Let g be any point of G, and V a connected open neighbourhood of g (which exists because G is assumed to be locally connected). Then $g^{-1}V$ is a connected open nbhd of e, since it is the image under V of the homeomorphism $L_{g^{-1}} : G \to G$ (see Theorem 10.17). Hence $g^{-1}V \subseteq G_0$ since G_0 is the connected component of the identity e. Applying the map L_g we have $V \subseteq gG_0$. The canonical projection π therefore sends V to the single coset gG_0. Since π is an open map it follows that each coset gG_0 is an open set, and the topology of the space of these cosets, G/G_0 has the discrete topology.

Problem 10.23 Prove the properties (10.3)–(10.7)

Solution: (10.3): $v_n \to v$ and $v_n \to v' \quad \Longrightarrow \quad v = v'$.
If $v \neq v'$ then, because V is a Hausdorff space, there exist open nbhds $U \ni v$, $U' \ni v'$ such that $U \cap U' = \emptyset$. If $v_n \to v$ then there exist N such that $v_n \in U$ for all $n \geq N$.

Hence $v_n \notin U'$ for all $n \geq N$ and it is not possible that $v_n \to v'$. Hence we must have $v = v'$.

(10.4): $v_n = v$ for all $n \implies v_n \to v$. Since for every open neighbourhood U of v, every $v_n = v \in U$ the result is trivial.

(10.5): If $\{v'_n\}$ is a subsequence of $v_n \to v$ then $v'_n \to v$.
A subsequence is determined by an infinite sequence of integers $n_1 < n_2 < n_3 < \dots$ such that $v'_1 = v_{n_1}$, $v'_2 = v_{n_2}$, etc. If for all $n \geq N$ we have $v \in U$ then it is clearly true of the subsequence $v'_i = v_{n_i} \in U$ for all $n_i \geq N$. Hence $v'_i \to v$.

(10.6) $u_n \to u, \quad v_n \to v \implies u_n + \lambda v_n \to u + \lambda v$, where $\lambda \in \mathbb{K}$ is any scalar. Let U be any open nbhd of $u + \lambda v$. The map $\phi : V \times V \to V$ defined by $\phi(u, v) = u + \lambda v$ is a continuous map since it can be written as the composition

$$\phi = \psi \circ (\mathrm{id} \times \tau_\lambda)$$

where $\tau_\lambda : V \to V$ is defined by $\tau_\lambda(v) = \tau(\lambda, v) = \lambda v$. Hence there exist open neighbourhoods A and B of u and v in V such that for for all $u' \in A$ and $v' \in B$ we have $u' + \lambda v' \in U$. Define $N > 0$ such that for all $n \geq N$ we have $u_n \in A$, and $M > 0$ such that for all $m \geq M$ we have $v_m \in B$. Then for all $k \geq K = \max(N, M)$ we have $u_k \in A$ and $v_k \in B$. Hence $u_k + \lambda v_k \in U$, showing that $u_k + \lambda v_k \to u + \lambda v$.

(10.7): If λ_n is a convergent sequence of scalars in $\mathbb{K}$ then $\lambda_n \to \lambda \implies \lambda_n u \to \lambda u$.
The proof is similar to the previous example: the continuity of the map τ means that for any neighbourhood U of λv there exist neighbourhoods A of λ and B of v such that $\lambda' v' \in U$ for all $\lambda' \in A$, $v' \in B$. Let N be such that $\lambda_n \in A$ for all $n \geq N$, and it follows at once that $\lambda_n v \in U$ for all $n \geq N$.

Problem 10.24 Show that a linear map $T : V \to W$ between topological vector spaces is continuous everywhere on V if and only if it is continuous at the origin $0 \in V$.

Solution: If T is continuous everywhere it is evidently continuous at 0.
In any topological vector space V, the map $\psi_u : V \to V$ defined by $\psi_u(v) = \psi(u, v) = u + v$ is a homeomomorphism since it is continuous, as is its inverse ψ_{-u}. Assume that the linear map $T : V \to W$ is continuous at the origin 0_V of V (and note by linearity that $T0_V = 0_W$). If v is any vector in V, our aim is to show that T is continuous at v. If A is any open neighbourhood of $Tv \in W$, then the set $A - Tv = \{a - Tv \,|\, a \in A\}$ is an open neighbourhood of the origin 0_W of W. Since T is continuous at 0_V there exists an open neighbourhood B of 0_V such that $T(B) \subset A - Tv$. Then $\psi_v(B) = B + v$ is an open neighbourhood of v and

$$T(B + v) = T(B) + Tv \subset A - Tv + Tv = A,$$

which shows that T is continuous at v. Hence T is continuous everywhere on V.

Problem 10.25 Give an example of a linear map $T : V \to W$ between topological vector spaces V and W that is not continuous.

Solution: By Theorem 10.23 the key to such an example is that it be an unbounded linear map. Such an example can only occur in an infinite dimensional vector space. Consider, for example, the vector space $\hat{\mathbb{R}}^\infty$ defined in Example 3.10 consisting of finite sequences of real numbers

$$\hat{\mathbb{R}}^\infty = \{\mathbf{a} = (a_1, a_2, \dots, a_n, 0, 0, \dots)\}.$$

This is a topological vector space with respect a norm such as

$$\|\mathbf{a}\| = |a_1| + |a_2| + \dots + |a_n|.$$

Consider the linear map $\varphi : \hat{\mathbb{R}}^\infty \to \mathbb{R}$ (where $\mathbb{R}$ is treated as a one-dimensional topological vector space) defined by

$$\varphi((a_1, a_2, \dots, a_n, 0, 0, \dots)) = a_1 + 2a_2 + \dots + na_n.$$

This is not continuous at the origin, for if U is the open nbhd of $0 \in \mathbb{R}$ defined by $U = \{x \mid |x| < 1\}$, then in every neigbhourhood of origin in $\hat{\mathbb{R}}^\infty$ there exists a vector $\mathbf{a}$ such that $|\varphi(\mathbf{a})| > 1$. For example, given $\epsilon > 0$, set $N > 2/\epsilon$ and let all $a_n = 0$ except for $a_N = 2/N$. Then $\|\mathbf{a}\| = 2/N < \epsilon$ but $|\varphi(\mathbf{a})| = Na_N = 2 > 1$. Thus the linear map φ is not continuous at $\mathbf{0} = (0, 0, 0, \dots)$.

Problem 10.26 Complete the proof that a normed vector space is a topological vector space with respect to the metric topology induced by the norm.

Solution: The continuity of the vector addition function ψ is shown in the text. To show continuity of the scalar multiplication function $\tau : \mathbb{C} \times V \to V$ defined by $\tau(\lambda, v)\lambda v$, we proceed as in Example 10.23. Let $u \in V$ and $\epsilon > 0$ any positive real number. Let $M = \tau^{-1}(B_\epsilon(\lambda u)) \subset \mathbb{C} \times V$ where $B_\epsilon(x) = \{v \in V \mid \|x - v\| < \epsilon\}$. Given positive real numbers $\delta, \delta' > 0$, for any $v \in B_\delta(u)$ and $\mu \in \mathbb{C}$ such that $|\mu - \lambda| < \delta'$, we have

$$\begin{aligned} \|\mu v - \lambda u\| &= \|\mu v - \mu u + \mu u \lambda u\| \\ &\leq |\mu| \|v - u\| + |\mu - \lambda| \|u\| \qquad \text{by (Norm2) and (Norm3)} \\ &< (|\lambda| + \delta')\delta + \delta' \|u\| \\ &< \epsilon \end{aligned}$$

if we choose

$$\delta' = \frac{\epsilon}{2\|u\|} \quad \text{and} \quad \delta = \frac{\epsilon}{2|\lambda| \|u\| + \epsilon}.$$

This shows continuity at any vector $u \neq 0$. It is easily modified to show continuity at $u = 0$ ($\|u\| = 0$) by selecting

$$\delta' = \epsilon, \qquad \delta = \frac{\epsilon}{|\lambda| + \epsilon}.$$

Problem 10.27 **Show that a real vector space V of dimension ≥ 1 is not a topological vector space with respect to either the discrete or indiscrete topology.**

Solution: If V has the discrete topology then the product topology on $V \times V$ is also discrete (every point (u, v) is the intersection of a horizontal and vertical line, which are open sets in the product topology if isolated points of V are open sets). Hence vector addition is continuous, for every set of the form

$$\psi^{-1}(\{v\}) = \{(x, v - x) \mid x \in V\} \subset V \times V$$

is necessarily open.
However scalar product is not a continuous operation. The product topology on $\mathbb{K} \times V$ consists of unions of sets of the form $U_x \times \{x\}$ where U_x is any open set of $\mathbb{K}$ and $x \in V$; i.e. the general open set of $\mathbb{K} \times V$ is of the form

$$\bigcup_{x \in A} U_x \times \{x\}$$

where A is an arbitrary subset of V. Now if $u \in V$ then $\{u\}$ is open in the discrete topology but

$$\tau^{-1}(\{u\}) = \big\{(\lambda, \frac{u}{\lambda}) \;\big|\; \lambda \neq 0 \in \mathbb{K}\big\}$$

is not an open subset of $\mathbb{K} \times V$ since it is not of the above form. For example, the set of points in $\tau^{-1}(u)$ of the form (λ, u) is the singleton $\{(1, u)\}$, which is clearly not of the form $U \times \{u\}$ since $\mathbb{K} = \mathbb{C}$ of $\mathbb{R}$ does not have the discrete topology (U must contain an entire open neighbourhood of 1 in the standard topology).

In the indiscrete topology on V, both $V \times V$ and $\mathbb{K} \times V$ have the indiscrete topology, and both vector addition and scalar multiplication are continuous. However the space V is obviously not Hausdorff, as it is impossible to separate points u and v with disjoint open sets since the only open sets available are $\emptyset$ and V.

Problem 10.28 **Show that the following are all norms in the vector space $\mathbb{R}^2$:**

$$\|\mathbf{u}\|_1 = \sqrt{(u_1)^2 + (u_2)^2},$$
$$\|\mathbf{u}\|_2 = \max\{|u_1|, |u_2|\},$$
$$\|\mathbf{u}\|_3 = |u_1| + |u_2|.$$

What are the shapes of the open balls $B_a(\mathbf{u})$? Show that the topologies generated by these norms are the same.

Solution: For $\|\cdot\|_1$ the conditions (Norm1) and (Norm(2) are trivial to verify:

$$\|\mathbf{u}\|_1 = \sqrt{(u_1)^2 + (u_2)^2} \geq 0 \qquad \text{and} \qquad \|\mathbf{u}\|_1 = 0 \quad \text{iff } u_1 = u_2 = 0$$
$$\|\lambda\mathbf{u}\|_1 = \sqrt{\lambda^2(u_1)^2 + \lambda^2(u_2)^2} = |\lambda|\sqrt{(u_1)^2 + (u_2)^2} = |\lambda|\|\mathbf{u}\|_1.$$

To show (Norm3) we must show

$$\sqrt{(u_1+v_1)^2 + (u_2+v_2)^2} \leq \sqrt{(u_1)^2 + (u_2)^2} + \sqrt{(v_1)^2 + (v_2)^2}.$$

Squaring both sides of this equation, leads to

$$2u_1v_1 + 2u_2v_2 \leq 2\sqrt{(u_1)^2 + (u_2)^2}\sqrt{(v_1)^2 + (v_2)^2}.$$

This inequality either follows by noting $\mathbf{u}\cdot\mathbf{v} = |\mathbf{u}||\mathbf{v}|\cos\theta$ or more algebraically from

$$\begin{aligned}\big((u_1)^2 + (u_2)^2\big)\big((v_1)^2 + (v_2)^2\big) &= u_1^2v_1^2 + u_2^2v_2^2 + u_1^2v_2^2 + u_2^2v_1^2 \\ &= (u_1v_1 + u_2v_2)^2 - 2u_1v_1u_2v_2 + u_1^2v_2^2 + u_2^2v_1^2 \\ &= (u_1v_1 + u_2v_2)^2 + (u_1v_2 - u_2v_1)^2 \\ &\geq (u_1v_1 + u_2v_2)^2.\end{aligned}$$

For $\|\cdot\|_2$ and $\|\cdot\|_3$ conditions (Norm1) and (Norm(2) are again trivial while (Norm3) follows immediately from the (real or complex number) inequality

$$|a+b| \leq |a| + |b|.$$

The open balls have been described in Problem 10.10, and the argument that the topologies generated by these norms are homeomorphic is also follows the same lines as that problem.

Problem 10.29 **Show that if $x_n \to x$ in a normed vector space then**

$$\frac{x_1 + x_2 + \cdots + x_n}{n} \to x.$$

Solution: Firstly, note the inequality

$$\begin{aligned}\Big\|x - \frac{x_1 + x_2 + \cdots + x_n}{n}\Big\| &= \Big\|\frac{x - x_1}{n} + \frac{x - x_2}{n} + \cdots + \frac{x - x_n}{n}\Big\| \\ &\leq \frac{1}{n}\sum_{i=1}^{n}\|x - x_i\|\end{aligned}$$

For every $\epsilon > 0$ there exists $N > 0$ such that $\|x - x_n\| < \frac{1}{2}\epsilon$ for all $n \geq N$. Then

$$\begin{aligned}\Big\|x - \frac{x_1 + x_2 + \cdots + x_n}{n}\Big\| &\leq \frac{1}{n}\sum_{i=1}^{n}\|x - x_i\| + \frac{1}{n}\big(\|x - x_{N+1}\| + \cdots + \|x - x_n\|\big) \\ &< \frac{1}{n}\sum_{i=1}^{N}\|x - x_i\| + \frac{n-N}{n}\frac{\epsilon}{2}\end{aligned}$$

Since the first term on the RHS $\to 0$ as $n \to \infty$ (for N fixed) it is possible to select $M \geq N$ such that for all $n \geq M$

$$\frac{1}{n}\sum_{i=1}^{N}\|x - x_i\| < \frac{\epsilon}{2}$$

and for all $n \geq M$ we have

$$\left\|x - \frac{x_1 + x_2 + \cdots + x_n}{n}\right\| < \frac{2n - N}{2n}\epsilon < \epsilon.$$

This proves the result.

Problem 10.30 **Show that if x_n is a sequence in a normed vector space V such that every subsequence has a subsequence convergent to x, then $x_n \to x$.**

Solution: The wording of this question may seem a little strange. However to have a subsequence converging to x is clearly not enough; e.g. the sequence $1, \frac{1}{2}, 1, \frac{1}{3}, \dots, 1, \frac{1}{n}, 1, \dots$ has a subsequence approaching 0 and another converging to 1, but the sequence itself however approaches neither of these.
Let x_n be a sequence such that every subsequence has a subsequence convergent to x. Suppose, to the contrary, that $x_n \not\to x$. Then there exists $\epsilon > 0$ such that for every $N > 0$, there exists $n_N \geq N$ such that $\|x - x_{n_N}\| > \epsilon$. The subsequence $x'_N = x_{n_N}$ $(N = 1, 2, 3, \dots)$ does not converge to x. As all elements of this sequence lie outside the ball of radius ϵ this is evidently true of every subsequence of x'_N, contradicting the assumption. Hence we must have $x_n \to x$, if every subsequence has a subsequence converging to x.

Problem 10.31 **Let V be a Banach space and W be a vector subspace of V. Define its *closure* $\overline{W}$ to be the union of W and all limits of Cauchy sequences of elements of W. Show that $\overline{W}$ is a closed vector subspace of V in the sense that the limit points of all Cauchy sequences in $\overline{W}$ lie in $\overline{W}$ (note that the Cauchy sequences may include the newly added limit points of W).**

Solution: W is a vector subspace, for if $u = \lim u_n$ and $v = \lim v_n$ then

$$\alpha u + \beta v = \lim_{n\to\infty}(\alpha u_n + \beta v_n) \in \overline{W}.$$

To show that it is closed, let $\{u_n\}$ be any Cauchy sequence in $\overline{W}$. Since V is a Banach space, the sequence $u_n \to u$ for some $u \in V$. For each u_n let u_{ni} be a sequence in W such that $u_{ni} \to u_n$ as $i \to \infty$ (if $u_n \in W$ we may pick $u_{ni} = u_n$ for all i).

For each $N > 0$ there exists $M > 0$ and i_M such that

$$\|u - u_M\| \leq \frac{1}{2N} \qquad \text{and} \qquad \|u_M - u_{Mi_M}\| < \frac{1}{2N}.$$

Then

$$\|u - u_{Mi_M}\| < \|u - u_M\| + \|u_M - u_{Mi_M}\| < \frac{1}{N}.$$

It follows that $u_{Mi_M} \to u$ as $M \to \infty$, whence $u \in \overline{W}$ since it is a limit of a sequence of elements of W. Hence the vector space $\overline{W}$ is closed.

Problem 10.32 **Show that every space $\mathcal{F}(S)$ is complete with respect to the supremum norm of Example 10.26. Hence show that the vector space ℓ_∞ of bounded infinite complex sequences is a Banach space with respect to the norm $\|\mathbf{x}\| = \sup(x_i)$.**

Solution: Let f_i be a Cauchy sequence of functions in $\mathcal{F}(S)$,

$$\|f_i - f_j\| = \sup_{x\in S} |f_i(x) - f_j(x)| \to 0 \qquad \text{as } i, j \to \infty.$$

Then for each $x \in S$ we must have $|f_i(x) - f_j(x)| \to 0$, i.e. $f_i(x)$ is a Cauchy sequence of real numbers. Hence for each $x \in S$ there exists $f(x)$ such that $f_i(x) \to f(x)$. It remains to show that the function $f(x)$ is bounded. Suppose not, then there exists a sequence of points x_n such that $|f(x_n)| > n$ for $n = 1, 2, 3, \dots$. For each n choose an N such that

$$|f(x_n) - f_N(x_n)| < \frac{1}{2}.$$

Then

$$\|f_N\| \geq |f_N(x_n)| > n - \tfrac{1}{2}.$$

Setting $n = 1, 2, 3, \dots$ we arrive at a sequence $N_1, N_2, N_3, \dots$ such that the subseqence of functions $f_i = f_{N_i}$ has unbounded norms, $\|f_i\| \to \infty$ as $i \to \infty$. However

$$\|f_i - f_j\| \geq \big|\|f_i\| - \|f_j\|\big|,$$

hence $\|f_i\|$ is a Cauchy sequence of real numbers $\|f_i\| - \|f_j\| \to 0$ as $i, j \to \infty$. Thus $\|f_i\| \to A < \infty$ in contradiction to the above. Hence the limit $f(x)$ is a bounded function on S.

The vector space ℓ_∞ of bounded infinite complex sequences $\|\mathbf{x}\|$ can be regarded as the space of bounded functions from the positive integers $S = \{1, 2, 3, \dots\}$ to $\mathbb{C}$, defined by $\mathbf{x}(i) = x_i$. The discussion, with minor modification to the complex numbers as arguments, shows that this is complete, and therefore a Banach space.

Problem 10.33 **Show that the set V' consisting of bounded linear functionals on a Banach space V is a normed vector space with respect**

to the norm

$$\|\varphi\| = \sup\{M \mid |\varphi(x)| \leq M\|x\| \textbf{ for all } x \in V\}.$$

Show that this norm is complete on V'.

Solution: The set of all linear functionals on V is a vector space by the usual discussion as in Chapter 3,

$$(\varphi + \lambda\psi)(u + \mu v) = (\varphi + \lambda\psi)(u) + \mu(\varphi + \lambda\psi)(v).$$

The norm $\|\varphi\|$ defined in the question is a norm on V' for it satisfies (Norm1):

$$\|\varphi\| = \sup_{x \neq 0 \in V} \frac{|\varphi(x)|}{\|x\|} > 0$$

and $\|\varphi\| = 0$ iff $|\varphi(x)| = 0$ for all $x \in V$, i.e. iff $\varphi = 0$.
(Norm2) follows trivially from $\|\lambda\varphi(x)\| = |\lambda||\varphi(x)| \leq |\lambda|M\|x\|$.
To show (Norm3) let φ and ψ be two bounded linear functionals,

$$\|\varphi(x)\| \leq \|\varphi\|\|x\|, \qquad \|\psi(x)\| \leq \|\psi\|\|x\|.$$

Then

$$\begin{aligned}\|(\varphi + \psi)(x)\| &= \|(\varphi(x) + \psi(x)\| \\ &\leq \|\varphi(x)\| + \|\psi(x)\| \\ &\leq (\|\varphi\| + \|\psi\|)\|x\|\end{aligned}$$

Hence $\varphi + \psi$ is a bounded linear functional with

$$\|\varphi + \psi\| \leq \|\varphi\| + \|\psi\|$$

since the maximum M such that $\|(\varphi + \psi)(x)\| < M\|x\|$ is $\|\varphi\| + \|\psi\|$.
To prove completeness we must show that if φ_n is a Cauchy sequence of bounded linear functionals, $\|\varphi_i - \varphi_j\| \to 0$ then there exists a bounded linear functional $\varphi : V \to \mathbb{C}$ such that $\varphi_n \to \varphi$. For any $\epsilon > 0$ there exists $N > 0$ such that for all $i, j \geq N$

$$|\varphi_i(x) - \varphi_j(x)| < \epsilon\|x\| \qquad \text{for all } x \in V.$$

Hence the sequence of complex numbers $\varphi_n(x)$ is a Cauchy sequence and has a limit $\varphi(x)$. The map $x \mapsto \varphi(x)$ is a linear functional since

$$\varphi(u + \lambda v) = \lim \varphi_n(u + \lambda v) = \lim(\varphi_n(u) + \lambda\varphi_n(v)) = \varphi(u) + \lambda\varphi(v).$$

To show boundedness let $i \to \infty$ in the above equation, gives

$$|\varphi(x) - \varphi_j(x)| < \epsilon\|x\|$$

whence

$$\begin{aligned}|\varphi(x)| &\le \|\varphi(x) - \varphi_j(x) + \varphi_j(x)\| \\ &\le \|\varphi(x) - \varphi_j(x)\| + \|\varphi_j(x)\| \\ &\le (\epsilon + M_j)\|x\|.\end{aligned}$$

This proves that V' is a complete normed vector space.

Problem 10.34 **We say two norms $\|u\|_1$ and $\|u\|_2$ on a vector space V are *equivalent* if there exist constants A and B such that**

$$\|u\|_1 \le A\|u\|_2 \qquad \text{and} \qquad \|u\|_2 \le B\|u\|_1$$

for all $u \in V$. If two norms are equivalent then show the following:

(a) If $u_n \to u$ with respect to one norm then this is also true for the other norm.

(b) Every linear functional that is continuous with respect to one norm is continuous with respect to the other norm.

(c)Let $V = \mathcal{C}[0,1]$ be the vector space of continuous complex functions on the interval $[0,1]$. By considering the sequence of function

$$f_n(x) = \frac{n}{\sqrt{\pi}}\mathrm{e}^{-n^2x^2}$$

show that the norms

$$\|f\|_1 = \sqrt{\int_0^1 |f|^2\,\mathrm{d}x} \qquad \text{and} \qquad \|f\|_2 = \max\{|f(x)|\,|\,0 \le x \le 1\}$$

are not equivalent.

(d)Show that the linear functional defined by $F(f) = f(1)$ is continuous with respect to $\|\cdot\|_2$ but not with respect to $\|\cdot\|_1$.

Solution: (a) If $\|u_n - u\|_1 \to 0$ then $\|u_n - u\|_1 \le B\|u_n - u\|_1 \to 0$. Hence convergence with respect to $\|\cdot\|_1$ implies convergence with respect to $\|\cdot\|_2$ and vice versa.

(b) If φ is continuous with respect to $\|\cdot\|_1$, then for every sequence u_n such that $\|u_n - u\|_1 \to 0$ we have $|\varphi(u_n) - \varphi(u)| \to 0$. But by part (a), every sequence such that $\|u_n - u\|_2 \to 0$ also converges with respect to the first norm, whence $|\varphi(u_n) - \varphi(u)| \to 0$. Hence φ is continuous with respect to the norm $\|\cdot\|_2$.

(c) To show inequivalence of these two norms, we seek a sequence of continuous functions f_n such that $\|f\|_1$ is bounded but $\|f\|_2$ is unbounded. The chopped-off error functions,

$$f_n = \frac{n}{\sqrt{\pi}}\mathrm{e}^{-n^2x^2} \qquad (0 <\le x \le 1)$$

all have $\|f_n\|_1 \leq \frac{1}{2}$, but $\|f_n\|_2 \to \infty$ since the value at $x = 0$ becomes arbitrarily large.

The linear functional F is continuous with respect to $\|\cdot\|_2$ for if $|f_n - f| \to 0$ then by uniform convergence

$$f_n(x) \to f(x) \qquad \text{for all } x \in [0,1].$$

In particular $f_n(1) \to f(1)$, i.e. $F(f_n) \to F(f)$.
Consider now the functions $f_n(x) = x^n$, having limit $f_n \to 0$ with respect to $\|\cdot\|_1$, for

$$\|f_n\|_1 = \frac{1}{2n+1} \to 0 \qquad \text{as} \qquad n \to \infty.$$

However, $F(f_n) = 1$ for all $f_n = x^n$, so $F(f_n) \not\to F(0) = 0$. Hence this linear functional F is not continuous with respect to to $\|\cdot\|_1$.

Chapter 11

Problem 11.1 **If $(X, \mathcal{M})$ and $(Y, \mathcal{N})$ are measurable spaces, show that the projection maps $\mathbf{pr}_1 : X \times Y \to X$ and $\mathbf{pr}_2 : X \times Y \to Y$ defined by $\mathbf{pr}_1(x, y) = x$ and $\mathbf{pr}_2(x, y) = y$ are measurable functions.**

Solution: If $A \subseteq X$ is a measurable set in $(X, \mathcal{M})$, then $\mathrm{pr}_1^{-1}(A) = A \times Y$ is measurable in the product measurable space $X \times Y, \mathcal{M} \times \mathcal{N}$, since Y is a measurable set in $(Y, \mathcal{N})$. Hence pr_1 is a measurable function.
Similarly, pr_2 is a measurable function, since $\mathrm{pr}_1^{-1}(B) = X \times B$ is measurable in $X \times Y$ for any measurable subset $B \subseteq Y$.

Problem 11.2 **Find a step function $s(x)$ that approximates $f(x) = x^2$ uniformly to within $\varepsilon > 0$ on $[0, 1]$, in the sense that $|f(x) - s(x)| < \varepsilon$ everywhere in $[0, 1]$.**

Solution: Divide $[0, 1]$ into N intervals of equal length,

$$I_n = [\frac{n-1}{N}, \frac{n}{N}], \quad (n = 1, 2, \ldots, N-1), \qquad \text{and } I_N = [\frac{N-1}{N}, 1].$$

On I_n set $s(x) = ((n-1)/N)^2$, and we have

$$\begin{aligned} |f(x) - s(x)| &\le \left(\frac{n}{N}\right)^2 - \left(\frac{n-1}{N}\right)^2 \\ &= \frac{2n-1}{N^2} \\ &\le \frac{2N-1}{N^2}. \end{aligned}$$

If we choose $N > 2/\epsilon$, then

$$|f(x) - s(x)| \le \frac{2}{N} - \frac{1}{N^2} < \frac{2}{N} < \epsilon.$$

Problem 11.3 **Let $f : X \to \mathbb{R}$ and $g : X \to \mathbb{R}$ be measurable functions and $E \subset X$ a measurable set. Show that**

$$h(x) = \begin{cases} f(x) & \textbf{if } x \in E \\ g(x) & \textbf{if } x \notin E \end{cases}$$

is a measurable function on X.

Solution: If $A \subseteq \mathbb{R}$ is any Borel set, then

$$h^{-1}(A) = \big(f^{-1}(A) \cap E\big) \cup \big(g^{-1}(A) \cap E^c\big)$$

which is measurable since $f^{-1}(A)$ and $g^{-1}(A)$, E and E^c are all measurable sets in X.

Problem 11.4 **If $f, g : \mathbb{R} \to \mathbb{R}$ are Borel measurable real functions show that $h(x, y) = f(x)g(y)$ is a measurable function $h : \mathbb{R}^2 \to \mathbb{R}$ with respect to the product measure on $\mathbb{R}^2$.**

Solution: From exercises after Theorem 11.1 and Problem 11.1, the product function

$$h = (f \circ \mathrm{pr}_1)(g \circ \mathrm{pr}_2)$$

is measurable.

More directly, let $A = (a, \infty)$ be any semi-infinite open interval on $\mathbb{R}$. From the comments after Theorem 11.1 h is measurable if $h^{-1}(A) \subseteq \mathbb{R}^2$ is a measurable set for all $a \in \mathbb{R}$. Since the product function $\pi : (u, v) \mapsto uv$ is clearly continuous, the inverse image

$$\pi^{-1}(A) = \{(u, v) \mid uv \in A\}$$

is an open subset of $\mathbb{R}^2$. For any $(u, v) \in \pi^{-1}(A)$ there exist open intervals $I, J \subset \mathbb{R}$ such that $u \in I$, $v \in J$ and $xy \in A$ for all $x \in I$, $y \in J$ (i.e. $I \times J \subset \pi^{-1}(A)$). Since f and g are measurable functions, the sets $f^{-1}(I)$ and $g^{-1}(J)$ are measurable (Borel) subsets of $\mathbb{R}$. Hence $f^{-1}(I) \times g^{-1}(J)$ is a measurable neighbourhood of (u, v), and $h(f^{-1}(I) \times g^{-1}(J)) \subset A$, i.e.

$$f^{-1}(I) \times g^{-1}(J) \subset h^{-1}(A).$$

Since A can be written as a countable union of interval products $I \times J$, the set $h^{-1}(A)$ is a countable union of products of measurable sets in $\mathbb{R}$, and is therefore a measurable set.

Problem 11.5 **Show that every countable subset of $\mathbb{R}$ is measurable and has Lebesgue measure zero.**

Solution: Every singleton $\{x\}$ where $x \in \mathbb{R}$ is a Lebsegue measurable set, for if $I = (a, b)$ is any open interval then

$$\mu(I) = b - a = \mu^*(I \cap \{x\}) + \mu^*(I \cap \{x\}^c).$$

For example, if $x \notin (a, b)$ then $\mu^*(I \cap \{x\}) = \mu^*(\emptyset) = 0$ and $\mu^*(I \cap \{x\}^c) = \mu^*(I) = b - a$. If, on the other hand, $x \in I$ then $\mu^*(I \cap \{x\}) = \mu^*(\{x\}) = 0$ for the inf of the lengths of open intervals covering x is 0, and

$$\mu^*(I \cap \{x\}^c) = b - x + x - a = b - a.$$

Since a countable set E is the countable union of disjoint singletons consisting of its elements, it is measurable by Theorem 11.6 and (Meas5), while Corollary 11.5

ensures that

$$\mu(E) = \mu^*(E) = \bigcup_{i=1}^{\infty} 0 = 0.$$

Problem 11.6 Show that the union of a sequence of sets of measure zero is a set of Lebesgue measure zero.

Solution: Let E_1, E_2, $E_3, \ldots$ be a sequence of sequence of sets of measure zero, $\mu^*(E_n) = 0$ for all $n = 1, 2, \ldots$. As in the concluding stages of the proof of Theorem 11.6, set

$$\begin{aligned}
&F_1 = E_1 \\
&F_2 = E_2 - E_1 \\
&\ldots \\
&F_n = E_n - (E_1 \cup E_2 \cup \cdots \cup E_{n-1}) \\
&\ldots
\end{aligned}$$

By Eq. (11.3) we have $\mu^* F_n \leq \mu^*(E_n)$, since $F_n \subseteq E_n$ for each $n = 1, 2, \ldots$. The sets F_n are mutually disjoint, and

$$\bigcup_{i=1}^{\infty} E_i = \bigcup_{i=1}^{\infty} F_i.$$

Hence, by Problem 11.8 and Corollary 11.5,

$$\mu^*\big(\bigcup_{i=1}^{\infty} E_i\big) = \sum_{i=1}^{\infty} \mu^*(F_i) = \sum_{i=1}^{\infty} 0 = 0.$$

Problem 11.7 If $\mu^*(N) = 0$ show that for any set E, $\mu^*(E \cup N) = \mu^*(E - N) = \mu^*(E)$. Hence show that $E \cup N$ and $E - N$ are Lebesgue measurable if and only if E is measurable.

Solution: Since

$$E \cup N = E \cup (N \cap E^c).$$

where E and $N \cap E^c$ are disjoint sets, we have by subadditivity, Eq. (11.2), and the inequality (11.3)

$$\mu^*(E \cup N) \leq \mu^*(E) + \mu^*(N \cap E^c) \leq \mu^*(E) + \mu^*(N) = \mu^*(E).$$

Also, by (11.3),

$$\mu^*(E) \leq \mu^*(E \cup N).$$

Hence $\mu^*(E) = \mu^*(E \cup N)$.
Similarly, $E = (E - N) \cup (E \cap N)$ gives

$$\mu^*(E) \leq \mu^*(E - N) + \mu^*(E \cap N) \leq \mu^*(E - N) + \mu^*(N) = \mu^*(E - N)$$

and

$$\mu^*(E - N) \leq \mu^*(E)$$

since $E - N \subseteq E$. Hence $\mu^*(E - N) = \mu^*(E)$.

Let I be any open interval.

$$\mu^*(I \cap (E \cup N)) = \mu^*((I \cap E) \cup (I \cap N)) = \mu^*(I \cap E)$$

since $\mu^*(I \cap N) \leq \mu^*(N) = 0$. Similarly

$$\mu^*(I \cap (E \cup N)^c) = \mu^*(I \cap E^c \cap N^c) = \mu^*((I \cap E^c) - N) = \mu^*(I \cap E^c).$$

Hence

$$\mu^*(I \cap (E \cup N)) + \mu^*(I \cap (E \cup N)^c) = \mu^*(I \cap E) + \mu^*(I \cap E^c) = \mu(I)$$

if and only if E is measurable.
The argument that $E - N$ is measurable iff E is measurable follows on similar lines:

$$\mu^*(I \cap (E - N)) = \mu^*(I \cap E \cap N^c) = \mu^*((I \cap E) - N) = \mu^*(I \cap E),$$

and

$$\begin{aligned} \mu^*(I \cap (E - N)^c) &= \mu^*(I \cap (E^c \cup N)) \\ &= \mu^*((I \cap E^c) \cup (I \cap N)) = \mu^*(I \cap E^c), \end{aligned}$$

since $\mu^*(I \cap N) \leq \mu^*(N) = 0$.

Problem 11.8 A measure is said to be complete if every subset of a set of measure zero is measurable. Show that if $A \subset \mathbb{R}$ is a set of outer measure zero, $\mu^*(A) = 0$, then A is Lebesgue measurable and has measure zero. Hence show that Lebesgue measure is complete.

Solution: If $\mu^*(A) = 0$ then for all intervals I, $\mu^*(I \cap A) = 0$ since $I \cap A \subseteq A$. Hence $\mu^*(I) = \mu^*(I \cap A^c)$, for

$$\begin{aligned} \mu^*(I \cap A^c) &\leq \mu^*(I) && \text{by (11.3)} \\ &\leq \mu^*(I \cap A^c) + \mu^*(I \cap A) && \text{by (11.2)} \\ &= \mu^*(I \cap A^c). \end{aligned}$$

Hence A is measurable, for

$$\mu(I) = \mu^*(I) = \mu^*(I \cap A) + \mu^*(I \cap A^c).$$

Its Lebesgue measure is $\mu(A) = \mu^*(A) = 0$.

Every subset of $B \subseteq A$ must have outer measure zero, $\mu^* B = 0$ by (11.3). Hence B is measurable by the above, and Lebesgue measure is complete.

Problem 11.9 Show that a subset E of $\mathbb{R}$ is measurable if for all $\epsilon > 0$ there exists an open set $U \supset E$ such that $\mu^*(U - E) < \epsilon$.

Solution: Let I be any open interval in $\mathbb{R}$. Then $I \cap E \subseteq I \cap U$, and by the inequality (11.3)

$$\mu^*(I \cap E) \leq \mu^*(I \cap U).$$

Since $U^c \subseteq E^c$ we have $E^c = U^c \cup (U - E)$ and it follows that

$$I \cap E^c = (I \cap U^c) \cup (I \cap (U - E)).$$

The two sets on the RHS are disjoint, and by subadditivity (11.2), and (11.3)

$$\mu^*(I \cap E^c) \leq \mu^*(I \cap U^c) + \mu^*(U - E) < \mu^*(I \cap U^c) + \epsilon.$$

Hence, since $I = (I \cap E) \cup (I \cap E^c)$ where the sets on the RHS are disjoint, we have by (11.2)

$$\begin{aligned}\mu(I) = \mu^*(I) &\leq \mu^*(I \cap E) + \mu^*(I \cap E^c)\\ &\leq \mu^*(I \cap U) + \mu^*(I \cap U^c) + \epsilon\\ &= \mu(I) + \epsilon\end{aligned}$$

since U is measurable, as it is an open set (and therefore a disjoint union of open intervals). As $\epsilon > 0$ is arbitrary, we must have

$$\mu(I) = \mu^*(I \cap U) + \mu^*(I \cap U^c),$$

i.e. E is a measurable set.

Problem 11.10 If E is bounded and there exists an interval $I \supset E$ such that

$$\mu^*(I) = \mu^*(I \cap E) + \mu^*(I - E)$$

then this holds for all intervals, possibly even those overlapping E.

Solution: If J is any interval, then as it is a measurable set we have, using Theorem 11.4

$$\mu^*(E) = \mu^*(E \cap J) + \mu * (E - J)$$

and

$$\mu^*(I - E) = \mu^*((I - E) \cap J) + \mu^*((I - E) - J).$$

Hence

$$\begin{aligned}\mu^*(I) &= \mu^*(E) + \mu^*(I - E)\\ &= \mu^*(E \cap J) + \mu * (E - J) + \mu^*((I - E) \cap J) + \mu^*((I - E) - J)\\ &= \mu^*(J \cap E) + \mu * (E - J) + \mu^*((I \cap J) - E) + \mu^*((I - J) - E)\end{aligned}$$

Now, by subadditivity, Eq. (11.2)

$$\mu^*(I \cap J) \le \mu^*(I \cap J \cap E) + \mu^*((I \cap J) - E) = \mu^*(J \cap E) + \mu^*((I \cap J) - E)$$
$$\mu^*(I - J) \le \mu^*((I - J) \cap E) + \mu^*((I - J) - E) = \mu^*(E - J) + \mu^*((I - J) - E).$$

Now since $I \cap J$ and $I - J$ are disjoint measurable sets whose union is I, we have

$$\mu^*(I) = \mu^*(I \cap J) + \mu^*(I - J)$$

so that the previous two inequalities must in fact be equalities; in particular

$$\mu^*(I \cap J) = \mu^*(I \cap J \cap E) + \mu^*((I \cap J) - E)$$

Since $J \cap I$ and $J - I$ are disjoint measurable sets whose union is J,

$$\begin{aligned}\mu^*(J) &= \mu^*(J \cap I) + \mu^*(J - I)\\ &= \mu^*(J \cap I \cap E) + \mu^*((J \cap I) - E) + \mu^*(J - I)\\ &= \mu^*(J \cap E) + \mu^*((J - E) \cap I) + \mu^*((J - E) - I)\end{aligned}$$

since $I \cup E = E$. Using Theorem 11.4 on the measurable set I, with $A = J - E$, we have

$$\mu^*(J - E) = \mu^*((J - E) \cap I) + \mu^*((J - E) - I),$$

whence

$$\mu^*(J) = \mu^*(J \cap E) + \mu^*(J - E).$$

Problem 11.11 **The *inner measure* $\mu_*(E)$ of a set E is defined as the least upper bound of the measures of all measurable subsets of E. Show that $\mu_*(E) \le \mu^*(E)$.**

For any open set $U \supset E$, show that

$$\mu(U) = \mu_*(U \cap E) + \mu^*(U - E)$$

and that E is measurable with finite measure if and only if $\mu_*(E) = \mu^*(E) < \infty$.

Solution: If A is measurable and U is an open set such that $A \subseteq E \subseteq U$. Hence $\mu(A) = \mu^*(A) \le \mu^*(U)$ for all open sets U such that $E \subset U$. In particular $\mu(A) \le \inf\{\mu^*(U) \,|\, E \subset U,\, U \text{ open}\} = \mu^*(E)$, whence

$$\mu_*(E) = \sup\{\mu(A) \,|\, A \subset E,\, A \text{ measurable}\} \le \mu^*(E).$$

Given $\epsilon > 0$, let W be an open set such that $U - E \subset W$ and $\mu^*(U - E) > \mu(W) - \epsilon$. Then $U - E \subset W \cap U$ and

$$\mu^*(U - E) > \mu(W) - \epsilon > \mu(U \cap W) - \epsilon.$$

The complementary set $U - W$ is a closed (and therefore measurable) subset of E, and therefore

$$\mu_*(E) \geq \mu(U - W).$$

Combining these two equations and using $\mu(U \cap W) + \mu(U - W) = \mu(U)$, since $U \cap W$ and $U - W$ are disjoint measurable sets whose union is U, we have

$$\begin{aligned}\mu_*(U \cap E) + \mu^*(U - E) &= \mu_*(E) + \mu^*(U - E)\\ &> \mu(U - W) + \mu(U \cap W) - \epsilon = \mu(U) - \epsilon\end{aligned}$$

for all $\epsilon > 0$. Hence

$$\mu_*(U \cap E) + \mu^*(U - E) \geq \mu(U).$$

On the other hand, since $(U \cap E) \cup (U - E) = U$ we must have

$$\mu_*(U \cap E) + \mu^*(U - E) \leq \mu^*(U \cap E) + \mu^*(U - E) \leq \mu^*(U) = \mu(U).$$

This implies equality,

$$\mu_*(U \cap E) + \mu^*(U - E) = \mu(U).$$

If $\mu_*(E) = \mu^*(E) < \infty$ then for any bounded open interval $I \supset E$

$$\mu(I) = \mu^*(E) + \mu^*(I - E)$$

whence E is measurable and $\mu(E) = \mu^*(E) = \mu_*(E)$. Conversely if E is measurable and bounded then $\mu(E) = \mu^*(E) < \infty$ and for any bounded interval I

$$\mu(I) = \mu_*(E) + \mu(I - E) = \mu(E) + \mu(I - E)$$

since E and $I - E$ are disjoint measurable sets such that $I = E \cup (I - E)$. Since $\mu(I)$ and $\mu(I - E)$ are finite, this implies $\mu_*(E) = \mu(E) = \mu^*(E) < \infty$.

Problem 11.12 **Show that if f and g are Lebesgue integrable on $E \subset \mathbb{R}$ and $f \geq g$ a.e., then**

$$\int_E f d\mu \geq \int_E g \mathrm{d}\mu.$$

Solution: Let $h = f - g$. The $h(x) \geq 0$ a.e. Dividing into positive and negative parts, $h = h^+ - h^-$, we must have $h^- = 0$ a.e. Then, by Theorem 11.8, $\int_E h^- \mathrm{d}\mu = 0$, and

$$\int_E h \mathrm{d}\mu = \int_E h^+ \mathrm{d}\mu - \int_E h^- \mathrm{d}\mu = \int_E h^+ \mathrm{d}\mu \geq 0.$$

Hence

$$\int_E f \mathrm{d}\mu - \int_E g \mathrm{d}\mu = \int_E f - g \mathrm{d}\mu = \int_E h \mathrm{d}\mu \geq 0,$$

whence $\int_E f \mathrm{d}\mu \geq \int_E g \mathrm{d}\mu$, as required.

Problem 11.13 Prove Theorem 11.9.

Solution: If f and g are non-negative Lebesgue integrable functions and $a \geq 0$, $b \geq 0$ non-negative real numbers then

$$\int_E (af + bg) \mathrm{d}\mu = \sup \int_E h \, \mathrm{d}\mu$$

where h is any simple function such that $h \leq af + bg$. Since, by Eq. (11.8)

$$\sup \int_E af \mathrm{d}\mu = a \sup \int_E k_1 \mathrm{d}\mu, \qquad \sup \int_E bg \mathrm{d}\mu = b \sup \int_E k_2 \mathrm{d}\mu$$

where k_1 and k_2 are simple functions such that $k_1 \leq f$, $k_2 \leq g$ we have using (11.9)

$$\int_E (af + bg) \mathrm{d}\mu = a \sup \int_E k_1 \mathrm{d}\mu + b \sup \int_E k_2 \mathrm{d}\mu = a \int_E f \mathrm{d}\mu + b \int_E g \mathrm{d}\mu.$$

If a or b is negative and/or f or g have non-vanishing negative parts, we must break the integrals into positive parts. For example, if $a > 0$ and $b < 0$ then

$$af + bg = a(f^+ - f^-) + b(g^+ - g^-) = (af^+ - bg^-) - (af^- + bg^+)$$

and we can apply the above conclusion to the positive and negative parts $(af^+ - bg^-)$ and $(af^- + bg^+)$ separately to arrive at the general conclusion.

Problem 11.14 If f is a Lebesgue integrable function on $E \subset \mathbb{R}$ then show that the function ψ defined by

$$\psi(a) = \mu\big(\{x \in E \,|\, |f(x)| > a\}\big) = O(a^{-1}) \quad \textbf{as } a \to \infty.$$

Solution: Let $E_a = \{x \in E \,|\, |f(x)| > a\}$, so that $\psi(a) = \mu(E_a)$. Suppose $\psi(a) \neq O(a^{-1})$ as $a \to \infty$. Then for all $K > 0$, there exists $a > 0$ such that $\mu(E_a) > Ka^{-1}$. Hence

$$\int_E |f(x)| \mathrm{d}\mu > a \frac{K}{a} = K,$$

which contradicts f being Lebesgue integrable (for that would implies $|f(x)|$ is also integrable).

Chapter 12

Problem 12.1 **Construct a test function such that $\phi(x) = 1$ for $|x| \le 1$ and $\phi(x) = 0$ for $|x| \ge 2$.**

Solution: The function $\phi(x) = 1$ for $|x| \le 1$ and $\phi(x) = 0$ for $|x| \ge 2$, and

$$\phi(x) = \exp(1)\Big(1 + \exp\big(\frac{1}{1-x^2}\big)\Big)\exp\big(\frac{3}{x^2-4}\big) \qquad \text{for } 1 < x < 2$$

is C^∞ at $x = 1$ and $x = 2$, and is a test function with the required properties. Alternatively, see Chapter 16, Lemma 16.1 for a similar construction.

Problem 12.2 **For every compact set $K \subset \mathbb{R}^n$ let $\mathcal{D}(K)$ be the space of C^∞ functions of compact support within K. Show that if all integer vectors $\underline{k}$ are set out in a sequence where $N(\underline{k})$ denotes the position of $\underline{k}$ in the sequence, then**

$$\|f\|_K = \sup_{\mathbf{x}\in K}\sum_{|\underline{k}|}\frac{1}{2^{N(\underline{k})}}\frac{|D_{\underline{k}}f(\mathbf{x})|}{1+|D_{\underline{k}}f(\mathbf{x})|}$$

is a norm on $\mathcal{D}(K)$. Let a set U be defined as open in $\mathcal{D}(\mathbb{R}^n)$ if it is a union of open balls $\{g \in K \,|\, \|g - f\|_K < a\}$. Show that sequence convergence with respect to this topology is identical with convergence of sequences of functions of compact support to all orders.

Solution: The three criteria for a norm are
(Norm1): $\|f\|_K \ge 0$, which is evident since all terms in the sums are positive, and

$$\|f\|_K = 0 \quad \Longrightarrow \quad |D_{\underline{k}}f(\mathbf{x})| = 0 \qquad \text{for all } \underline{k}, \text{and all } \mathbf{x} \in K.$$

That is, $\|f\|_K = 0$ iff $f = 0$.
(Norm2):

$$\|\lambda f\|_K = \sup_{\mathbf{x}\in K}\sum_{|\underline{k}|}\frac{1}{2^{N(\underline{k})}}\frac{|\lambda||D_{\underline{k}}f(\mathbf{x})|}{1+|\lambda||D_{\underline{k}}f(\mathbf{x})|} \le |\lambda|\,\|f\|_K.$$

Since $f(\mathbf{x}) \to 0$ as $\mathbf{x}$ approaches the boundary of K, the denominator becomes of no significance in that limit and, taking the sup over K, we find that $\|\lambda f\|_K = |\lambda|\,\|f\|_K$.
(Norm3): $\|f + g\|_K \le \|f\|_K + \|g\|_K$ follows from the inequality

$$\frac{|A+B|}{1+|A+B|} \le \frac{|A|}{1+|A|} + \frac{|B|}{1+|B|}$$

which is computed straightforwardly:

$$\frac{|A+B|}{1+|A+B|} - \frac{|A|}{1+|A|} - \frac{|B|}{1+|B|} = -\frac{|A+B||AB|}{(1+|A+B|)(1+|A|)(1+|B|)} \le 0.$$

A sequence $f_n \to f$ with respect to the norm topology iff $\|f - f_n\|_K \to 0$, i.e.

$$\sup_{\mathbf{x}\in K} \frac{|D_{\underline{k}}(f(\mathbf{x}) - f_n(\mathbf{x}))|}{1 + |D_{\underline{k}}(f(\mathbf{x}) - f_n(\mathbf{x}))|} \to 0$$

for derivatives of all orders $\underline{k}$. This condition is equivalent to convergence to all orders,

$$\sup_{\mathbf{x}\in K} |D_{\underline{k}}(f(\mathbf{x}) - f_n(\mathbf{x}))| \to 0.$$

Problem 12.3 Which of the following is a distribution?

(a) $T(\phi) = \sum_{n=1}^{m} \lambda_n \phi^{(n)}(0) \quad (\lambda_n \in \mathbb{R})$.

(b) $T(\phi) = \sum_{n=1}^{m} \lambda_n \phi(x_n) \quad (\lambda_n, x_n \in \mathbb{R})$.

(c) $T(\phi) = \big(\phi(0)\big)^2$.

(d) $T(\phi) = \sup \phi$.

(e) $T(\phi) = \int_{-\infty}^{\infty} |\phi(x)| \mathrm{d}x$.

Solution: (a) This is a distribution since it is linear,

$$\begin{aligned} T(a\phi + b\psi) &= \sum_{n=1}^{m} \lambda_n (a\phi + b\psi)^{(n)}(0) \\ &= a\sum_{n=1}^{m} \lambda_n (\phi)^{(n)}(0) + b\sum_{n=1}^{m} \lambda_n (\psi)^{(n)}(0) \\ &= aT(\phi) + bT(\psi) \end{aligned}$$

and continuous with respect to convergence of test functions on a compact domain,

$$\phi_i \to \phi \quad \Longrightarrow \quad \phi_i^{(n)}(0) \to \phi^{(n)}(0).$$

(b) is again a distribution; in fact it is

$$T = \sum_{n=1}^{m} \lambda_n \delta_{x_n} .$$

(c) $T(\phi) = \big(\phi(0)\big)^2$ is not a distribution, for

$$T(a\phi) = a^2\big(\phi(0)\big)^2 \neq aT(\phi).$$

(d) $T(\phi) = \sup \phi$ is not a distribution for in general

$$\sup(\phi + \psi) \neq \sup(\phi) + \sup(\psi).$$

For example, if $\phi(x) = x$ and $\psi(x) = -x$ on $-1 \le x \le 1$, then on this interval

$$\sup(\phi) = \sup(\psi) = 1, \quad \sup(\phi + \psi) = 0 \neq \sup(\phi) + \sup(\psi).$$

(e) $T(\phi) = \displaystyle\int_{-\infty}^{\infty} |\phi(x)| \mathrm{d}x$ is not a distribution, for $T(\phi + \psi) \neq T(\phi) + T(\psi)$ is general, as

$$|\phi(x) + \psi(x)| \neq |\phi(x)| + |\psi(x)|.$$

Problem 12.4 **We say a sequence of distributions T_n converges to a distribution T, written $T_n \to T$, if $T_n(\phi) \to T(\phi)$ for all test functions $\phi \in \mathcal{D}$ (this is sometimes called *weak convergence*). If a sequence of continuous functions f_n converges uniformly to a function $f(x)$ on every compact subset of $\mathbb{R}$, show that the associated regular distributions $T_{f_n} \to T_f$.**

In the distributional sense, show that we have the following convergences:

$$f_n(x) = \frac{n}{\pi(1 + n^2 x^2)} \to \delta(x),$$
$$g_n(x) = \frac{n}{\sqrt{\pi}} \mathrm{e}^{-n^2 x^2} \to \delta(x).$$

Solution: By uniform convergence on any compact set K, or every $\epsilon > 0$ there exists N such that for all $n > N$ $|f(x) - f(x_n)| < \epsilon$ for all $n \ge N$. Hence for any test function on K,

$$\begin{aligned}
\big|T(\phi) - T_{f_n}(\phi)\big| &= \Big|\int_{\mathbb{R}} \phi(x)(f(x) - f_n(x)) \mathrm{d}x\Big| \\
&\le \int_{\mathbb{R}} |\phi(x)|\, |f(x) - f_n(x)| \mathrm{d}x \\
&\le \epsilon \int_{\mathbb{R}} |\phi(x)| \mathrm{d}x
\end{aligned}$$

for all $n \ge N$. Thus $T_{f_n} \to T_f$.

We expect that

$$f_n(x) = \frac{n}{\pi(1 + n^2 x^2)} \to \delta(x)$$

since for all $x \neq 0$, $f_n(x) \to 0$ and for all n,

$$\int_{-\infty}^{\infty} f_n(x) \mathrm{d}x = \Big[\frac{1}{\pi} \arctan nx\Big]_{-\infty}^{\infty} = 1.$$

Similarly for the sequence of error functions $g_n(x)$.
To show $f_n \to \delta$ in the distributional sense, let ϕ be any test function (compact support, and C^∞). Then

$$T_{f_n}(\phi) = I_n^{(-)} + I_n^{(0)} + I_n^{(+)}$$

where

$$\begin{aligned} I_n^{(-)} &= \int_{-\infty}^{-1/\sqrt{n}} f_n(x)\phi \mathrm{d}x \\ I_n^{(0)} &= \int_{-1/\sqrt{n}}^{1/\sqrt{n}} f_n(x)\phi \mathrm{d}x \\ I_n^{(+)} &= \int_{1/\sqrt{n}}^{\infty} f_n(x)\phi \mathrm{d}x. \end{aligned}$$

As $n \to \infty$ we have

$$|I_n^{(-)}| \le \left[\frac{|\sup(\phi)|}{\pi} \arctan nx\right]_{-\infty}^{1/\sqrt{n}} = \frac{|\sup(\phi)|}{\pi}(\pi - 2\arctan\sqrt{n}) \to 0.$$

Similarly $|I_n^{(-)}| \to 0$ as $n \to \infty$. On the other hand, $I_n^{(0)} \to \phi(0)$, for any $\epsilon > 0$ there exists N such that for all $n \ge N$ we have $|\phi(0) - \phi(x)| < \epsilon$ for all $-1/\sqrt(n) \le x \le 1/\sqrt{n}$, and

$$\int_{-1/\sqrt{n}}^{1/\sqrt{n}} f_n(x)\mathrm{d}x = \frac{2}{\pi}\arctan\sqrt{n} = 1 - \epsilon.$$

Hence

$$(\phi(0) - \epsilon)(1 - \epsilon) < I_n^{(0)} < (\phi(0) + \epsilon)(1 + \epsilon)$$

for all $n \ge N$, from which it follows that $I_n^{(0)} \to \phi(0)$ as $n \to \infty$, and in the distributional sense$T_{f_n}(\phi) \to \delta(\phi)$ for all $\phi \in \mathcal{D}$.
The proof that $T_{g_n} \to \delta$ is similar, dividing the interval $(-\infty, \infty)$ into three sections, and noting that, on making the substitution $y = nx$,

$$\frac{n}{\sqrt{\pi}} \int_{-1/\sqrt{n}}^{1/\sqrt{n}} \mathrm{e}^{-n^2x^2}\mathrm{d}x = \frac{1}{\sqrt{\pi}} \int_{-\sqrt{n}}^{\sqrt{n}} \mathrm{e}^{-y^2}\mathrm{d}y \to 1$$

as $n \to \infty$.

Problem 12.5 In the sense of convergence defined in Problem 12.4 show that if $T_n \to T$ then $T_n' \to T'$.

In the distributional sense, show that we have the following convergences:

$$f_n(x) = -\frac{2n^3x}{\sqrt{\pi}}\mathrm{e}^{-n^2x^2} \to \delta'(x).$$

Solution: The limit $T_n \to T$ implies that $T_n(\phi) \to T(\phi)$ for all test functions $\phi \in \mathcal{D}$. Hence

$$T_n'(\phi) = -T_n(\phi') \to -T(\phi') = T'(\phi)$$

for all test functions ϕ. Hence $T_n' \to T'$.

We have

$$-\frac{2n^3x}{\sqrt{\pi}}\mathrm{e}^{-n^2x^2} \to \delta'(x) = \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{n}{\sqrt{\pi}}\mathrm{e}^{-n^2x^2}\right) = g_n'(x)$$

where $g_n(x)$ is the sequence of error functions in Problem 12.4. Hence

$$T_n \equiv T_{g_n'} = T_{g_n}' \to \delta'$$

since $T_{g_n} \to \delta$. In terms of effect on any test function ϕ

$$T_{g_n'}(\phi) \to -\phi'(0).$$

Problem 12.6 **Evaluate**

(a) $\displaystyle\int_{-\infty}^{\infty} \mathrm{e}^{at}\,\sin bt\,\delta^{(n)}(t)\,\mathrm{d}t$ **for** $n = 0, 1, 2.$

(b) $\displaystyle\int_{-\infty}^{\infty} (\cos t + \sin t)\delta^{(n)}(t^3 + t^2 + t)\,\mathrm{d}t$ **for** $n = 0, 1.$

Solution: (a) For $n = 0$

$$\int_{-\infty}^{\infty} \mathrm{e}^{at}\,\sin bt\,\delta(t)\,\mathrm{d}t = \mathrm{e}^{at}\sin bt\big|_{t=0} = 0.$$

For $n = 1$,

$$\begin{aligned}\int_{-\infty}^{\infty} \mathrm{e}^{at}\,\sin bt\,\delta'(t)\,\mathrm{d}t &= -(\mathrm{e}^{at}\sin bt)'\big|_{t=0}\\ &= -(a\mathrm{e}^{at}\sin bt + b\mathrm{e}^{at}\cos bt)\big|_{t=0}\\ &= -b.\end{aligned}$$

For $n = 2$,

$$\begin{aligned}\int_{-\infty}^{\infty} \mathrm{e}^{at}\,\sin bt\,\delta''(t)\,\mathrm{d}t &= (\mathrm{e}^{at}\sin bt)''\big|_{t=0}\\ &= (a^2\mathrm{e}^{at}\sin bt + 2ab\mathrm{e}^{at}\cos bt - b^2\mathrm{e}^{at}\sin bt)\big|_{t=0}\\ &= 2ab.\end{aligned}$$

(b)

$$\int_{-\infty}^{\infty} (\cos t + \sin t)\delta(t^3 + t^2 + t)\,\mathrm{d}t = \frac{\cos t + \sin t}{3t^2 + 2t + 1}\Big|_{t=0} = 1.$$

$$\begin{aligned}\int_{-\infty}^{\infty} (\cos t + \sin t)\delta'(t^3 + t^2 + t)\,\mathrm{d}t &= \frac{-1}{3t^2 + 2t + 1}\frac{\mathrm{d}}{\mathrm{d}t}\Big(\frac{\cos t + \sin t}{t^3 + t^2 + t}\Big)\Big|_{t=0} \\ &= -\Big(-\frac{(\cos t + \sin t)(6t + 2)}{(t^3 + t^2 + t)^2} + \frac{-\sin t + \cos t}{t^3 + t^2 + t}\Big)\Big|_{t=0} \\ &= -(-2 + 1) = 1.\end{aligned}$$

Problem 12.7 **Show the following identities:**

(a) $\delta((x-a)(x-b)) = \dfrac{1}{b-a}(\delta(x-a) + \delta(x-b)).$

(b) $\dfrac{\mathrm{d}}{\mathrm{d}x}\theta(x^2 - 1) = \delta(x-1) - \delta(x+1) = 2x\delta(x^2 - 1).$

(c) $\dfrac{\mathrm{d}}{\mathrm{d}x}\delta(x^2 - 1) = \frac{1}{2}(\delta'(x-1) + \delta'(x+1)).$

(d) $\delta'(x^2 - 1) = \frac{1}{4}(\delta'(x-1) - \delta'(x+1) + \delta(x-1) + \delta(x+1)).$

Solution: (a) The argument closely follows that in Example 12.7 (note however we must assume that $a < b$). We give here a direct argument. Let $\varphi(x)$ be any test function, then by the argument leading to Eq. (12.9)

$$\begin{aligned}\int_{-\infty}^{\infty} \delta((x-a)(x-b))\varphi(x)\mathrm{d}x &= \frac{\varphi(x)}{|2x - a - b|}\Big|_{x=a} + \frac{\varphi(x)}{|2x - a - b|}\Big|_{x=b} \\ &= \frac{\varphi(a)}{|a-b|} + \frac{\varphi(b)}{|b-a|} \\ &= \frac{1}{|b-a|}\int_{-\infty}^{\infty} (\delta(x-a) + \delta)x - b))\varphi(x)\,\mathrm{d}x\end{aligned}$$

Since $\varphi(x)$ is an arbitrary test function we have for $b > a$ the required identity

$$\delta((x-a)(x-b)) = \frac{1}{b-a}(\delta(x-a) + \delta(x-b)).$$

(b) Again, let $\varphi(x)$ be any test function and we have

$$\begin{aligned}\int_{-\infty}^{\infty} \frac{\mathrm{d}}{\mathrm{d}x}\big(\theta(x^2 - 1)\big)\varphi(x)\,\mathrm{d}x &= -\int_{-\infty}^{\infty} \theta(x^2 - 1)\varphi'(x)\,\mathrm{d}x \\ &= -\int_{1}^{\infty} \varphi'(x)\,dx - \int_{-\infty}^{-1} \varphi'(x)\,\mathrm{d}x\end{aligned}$$

since $\theta(x^2 - 1) = 0$ for $-1 < x < 1$

$$= \varphi(1) - \varphi(-1).$$

On setting $a=-1$, $b=1$ in part (a) (or using Example 12.7) we have

$$\begin{aligned}\int_{-\infty}^{\infty} 2x\delta(x^2-1)\varphi(x)\mathrm{d}x &= \int_{-\infty}^{\infty} 2x\varphi(x)\frac{1}{2}\big(\delta(x-1)+\delta(x+1)\big)\mathrm{d}x \\ &= \varphi(1)-\varphi(-1).\end{aligned}$$

Hence

$$\frac{\mathrm{d}}{\mathrm{d}x}\theta(x^2-1) = 2x\delta(x^2-1) = \delta(x-1)-\delta(x+1).$$

(c) For any test function $\varphi(x)$

$$\int_{-\infty}^{\infty} \frac{\mathrm{d}}{\mathrm{d}x}\big\{\delta(x^2-1)\big\}\varphi(x)\,\mathrm{d}x = -\int_{-\infty}^{\infty} \delta(x^2-1)\varphi'(x)\,\mathrm{d}x.$$

Using $\delta(x^2-1) = \frac{1}{2}\{(\delta(x-1)+\delta(x+1)\}$ we have

$$\begin{aligned}\int_{-\infty}^{\infty} \frac{\mathrm{d}}{\mathrm{d}x}\big\{\delta(x^2-1)\big\}\varphi(x)\,\mathrm{d}x &= -\frac{1}{2}\varphi'(1)-\frac{1}{2}\varphi'(-1) \\ &= \frac{1}{2}\int_{-\infty}^{\infty}\big\{\delta'(x-1)+\delta'(x+1)\big\}\varphi(x)\,\mathrm{d}x\end{aligned}$$

and since φ is an arbitrary test function,

$$\frac{\mathrm{d}}{\mathrm{d}x}\delta(x^2-1) = \frac{1}{2}(\delta'(x-1)+\delta'(x+1)).$$

(d) In the integral

$$\int_{-\infty}^{\infty} \delta'\{(x^2-1)\}\varphi(x)\,\mathrm{d}x$$

divide the range of integration into two small intervals $[-1-\epsilon,-1+\epsilon]$ and $[1-\epsilon,1+\epsilon]$, all other parts being treated as contributing zero to the integral since we may set $\delta'\{(x^2-1)\}=0$ at all points $x\neq\pm1$. Make the change of variable to $y=x^2-1$ we have

$$\begin{aligned}\int_{-\infty}^{\infty} \delta'\{(x^2-1)\}\varphi(x)\,\mathrm{d}x &= \int_{-\infty}^{\infty} \frac{\mathrm{d}\delta(y)}{\mathrm{d}y}\varphi(x)\frac{\mathrm{d}y}{|2x|} \\ &= -\frac{\mathrm{d}}{\mathrm{d}y}\Big(\frac{\varphi(x)}{|2x|}\Big)\Big|_{x=-1} - \frac{\mathrm{d}}{\mathrm{d}y}\Big(\frac{\varphi(x)}{|2x|}\Big)\Big|_{x=1} \\ &= -\frac{1}{2x}\frac{\mathrm{d}}{\mathrm{d}x}\Big(\frac{\varphi(x)}{-2x}\Big)\Big|_{x=-1} - \frac{1}{2x}\frac{\mathrm{d}}{\mathrm{d}x}\Big(\frac{\varphi(x)}{2x}\Big)\Big|_{x=1} \\ &= \frac{1}{4}\big(\varphi'(-1)+\varphi(-1)-\varphi'(1)+\varphi(1)\big) \\ &= \frac{1}{4}\int_{-\infty}^{\infty}\{-\delta'(x+1)+\delta(x+1)+\delta'(x-1)+\delta(x-1)\}\varphi(x)\,\mathrm{d}x\end{aligned}$$

which gives the desired result.

Problem 12.8 **Show that for a monotone function $f(x)$ such that $f(\pm\infty) = \pm\infty$ with $f(a) = 0$**

$$\int_{-\infty}^{\infty} \varphi(x)\delta'\bigl(f(x)\bigr)\,\mathrm{d}x = -\frac{1}{f'(x)}\frac{\mathrm{d}}{\mathrm{d}x}\Bigl(\frac{\varphi(x)}{|f'(x)|}\Bigr)\Big|_{x=a}.$$

For a general function $f(x)$ that is monotone on a neighbourhood of all its zeros, find a general formula for the distribution $\delta' \circ f$.

Solution: Making a change of variable $y = f(x)$ in the LHS we have

$$\begin{aligned}\int_{-\infty}^{\infty} \varphi(x)\delta'\bigl(f(x)\bigr)\,\mathrm{d}x &= \int_{-\infty}^{\infty} \frac{\varphi(x)}{|f'(x)|}\delta'(y)\mathrm{d}y \\ &= -\frac{\mathrm{d}}{\mathrm{d}y}\Bigl(\frac{\varphi(x)}{|f'(x)|}\Bigr)\Big|_{y=0} \\ &= -\frac{1}{f'(x)}\frac{\mathrm{d}}{\mathrm{d}x}\Bigl(\frac{\varphi(x)}{|f'(x)|}\Bigr)\Big|_{x=a}.\end{aligned}$$

For a general function $f(x)$ having zeros at $x = a_1, a_2, \dots$, at all of which it is locally monotone, an identical argument to that leading to Eq. (12.9) gives

$$\int_{-\infty}^{\infty} \varphi(x)\delta'\bigl(f(x)\bigr)\,\mathrm{d}x = -\sum_i \frac{1}{f'(a_i)}\frac{\mathrm{d}}{\mathrm{d}x}\Bigl(\frac{\varphi(x)}{|f'(x)|}\Bigr)\Big|_{x=a_i}.$$

Using

$$\frac{\mathrm{d}}{\mathrm{d}x}\Bigl(\frac{\varphi(x)}{f'(x)}\Bigr)\Big|_{x=a_i} = \frac{\varphi'(a_i)}{f'(a_i)} - \frac{\varphi(a_i)f''(a_i)}{(f'(a_i))^2}$$

and the fact that at points $x = a_i$ where f is monotone increasing $|f'(a_i)| = f'(a_i)$, while at those where it is decreasing $|f'(a_i)| = -f'(a_i)$, we have

$$\delta'\bigl(f(x)\bigr) = \sum_i \frac{1}{f'(a_i)|f'(a_i)|}\delta'(x - a_i) + \frac{f''(a_i)}{(f'(a_i))^3}\delta(x - a_i).$$

Problem 12.9 **Show the identities**

$$\frac{\mathrm{d}}{\mathrm{d}x}\bigl(\delta(f(x))\bigr) = f'(x)\delta'\bigl(f(x)\bigr)$$

and

$$\delta\bigl(f(x)\bigr) + f(x)\delta'\bigl(f(x)\bigr) = 0.$$

Hence show that $\phi(x, y) = \delta(x^2 - y^2)$ is a solution of the partial differential equation

$$x\frac{\partial \phi}{\partial x} + y\frac{\partial \phi}{\partial y} + 2\phi(x, y) = 0.$$

Solution: For any test function $\varphi(x)$, and monotone function $f(x)$ with $f(a) = 0$, we have by Problem 12.8

$$\begin{aligned}\int_{-\infty}^{\infty} f'(x)\delta'\big(f(x)\big)\varphi(x)\,\mathrm{d}x &= \int_{-\infty}^{\infty} \delta'(y) f'(x)\varphi(x)\frac{\mathrm{d}y}{|f'(x)|} \\ &= -\frac{1}{f'(x)}\frac{\mathrm{d}}{\mathrm{d}x}\Big(\frac{\varphi(x)f'(x)}{|f'(x)|}\Big)\Big|_{x=a} \\ &= -\frac{1}{|f'(x)|}\frac{\mathrm{d}}{\mathrm{d}x}\big(\varphi(x)\big)\Big|_{x=a}\end{aligned}$$

and, on integration by parts

$$\begin{aligned}\int_{-\infty}^{\infty} \frac{\mathrm{d}}{\mathrm{d}x}\big(\delta(f(x))\big)\varphi(x)\,\mathrm{d}x &= -\int_{-\infty}^{\infty} \delta(f(x))\varphi'(x)\,\mathrm{d}x \\ &= -\int_{-\infty}^{\infty} \delta(y)\frac{\varphi'(x)}{|f'(x)|}\,\mathrm{d}y \\ &= -\frac{\varphi'(x)}{|f'(x)|}\Big|_{x=a}\end{aligned}$$

Hence

$$\frac{\mathrm{d}}{\mathrm{d}x}\big(\delta(f(x))\big) = f'(x)\delta'\big(f(x)\big).$$

The result extends to non-monotone functions as in Problem 12.8, or Eq. (12.9).

The second identity follows from $y\delta(y) = 0$, on setting $y = f(x)$ and using the Leibnitz rule and the above identity,

$$\begin{aligned}f(x)\delta(f(x)) = 0 \quad &\Longrightarrow \quad \frac{\mathrm{d}}{\mathrm{d}x}(f(x)\delta(f(x)) = 0 \\ &\Longrightarrow \quad f'(x)\delta(f(x)) + f(x)\frac{\mathrm{d}}{\mathrm{d}x}\delta(f(x)) = 0 \\ &\Longrightarrow \quad f'(x)\delta(f(x)) + f(x)f'(x)\delta'(f(x)) = 0 \\ &\Longrightarrow \quad \delta(f(x)) + f(x)\delta'(f(x)) = 0.\end{aligned}$$

Setting $f = x^2 - y^2$ we have

$$\begin{aligned}\frac{\partial}{\partial x}\delta(x^2-y^2) &= \frac{\partial x^2 - y^2}{\partial x}\delta'(x^2-y^2) = 2x\delta'(x^2-y^2) \\ \frac{\partial}{\partial y}\delta(x^2-y^2) &= \frac{\partial x^2 - y^2}{\partial y}\delta'(x^2-y^2) = -2y\delta'(x^2-y^2)\end{aligned}$$

Hence, using the second identity above,

$$2x\frac{\partial}{\partial x}\delta(x^2-y^2) + 2y\frac{\partial}{\partial y}\delta(x^2-y^2) = 2(x^2-y^2)\delta'(x^2-y^2) = -2\delta(x^2-y^2),$$

and $\phi = \delta(x^2 - y^2)$ is a distributional solution of the partial differential equation

$$x\frac{\partial\phi}{\partial x} + y\frac{\partial\phi}{\partial y} + 2\phi(x, y) = 0.$$

Problem 12.10 **Find the Fourier transforms of the functions**

$$f(x) = \begin{cases} 1 & \textbf{if } -a \le x \le a \\ 0 & \textbf{otherwise} \end{cases}$$

and

$$g(x) = \begin{cases} 1 - \frac{|x|}{2} & \textbf{if } -a \le x \le a \\ 0 & \textbf{otherwise}. \end{cases}$$

Solution:

$$\begin{aligned} \mathcal{F}f(y) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \mathrm{e}^{-ixy} f(x)\,\mathrm{d}x \\ &= \frac{1}{\sqrt{2\pi}} \int_{-a}^{a} \mathrm{e}^{-ixy}\,\mathrm{d}x \\ &= \frac{1}{\sqrt{2\pi}} \left[\frac{\mathrm{e}^{-ixy}}{-iy}\right]_{-a}^{a} \\ &= \frac{1}{\sqrt{2\pi}} \frac{\mathrm{e}^{iay} - \mathrm{e}^{-iay}}{iy} \\ &= \sqrt{\frac{2}{\pi}} \frac{\sin ay}{y}. \end{aligned}$$

$$\begin{aligned} \mathcal{F}g(y) &= \frac{1}{\sqrt{2\pi}} \int_{-a}^{a} \mathrm{e}^{-ixy}\big(1 - \frac{|x|}{2}\big)\,\mathrm{d}x \\ &= \frac{1}{\sqrt{2\pi}} \Big(\int_{0}^{a} \mathrm{e}^{-ixy}\big(1 - \frac{x}{2}\big)\,\mathrm{d}x + \int_{-a}^{0} \mathrm{e}^{-ixy}\big(1 + \frac{x}{2}\big)\,\mathrm{d}x \\ &= \sqrt{\frac{2}{\pi}} \frac{\sin ay}{y} - \frac{1}{2\sqrt{2\pi}} \Big[\frac{\mathrm{e}^{-ixy}x}{-iy} + \frac{\mathrm{e}^{-ixy}x}{y^2}\Big]_{0}^{a} \\ &\quad + \frac{1}{2\sqrt{2\pi}} \Big[\frac{\mathrm{e}^{-ixy}x}{-iy} + \frac{\mathrm{e}^{-ixy}x}{y^2}\Big]_{-a}^{0} \\ &= \sqrt{\frac{2}{\pi}} \frac{\sin ay}{y} + \frac{ia\cos ay}{\sqrt{2\pi}y} + \frac{1 - \cos ay}{\sqrt{2\pi}y^2}. \end{aligned}$$

Problem 12.11 **Show that**

$$\mathcal{F}\big(\mathrm{e}^{-a^2x^2/2}\big) = \frac{1}{|a|}\mathrm{e}^{-k^2/2a^2}.$$

Solution: The meaning of this Fourier integral is that if $f(x) = \mathrm{e}^{-a^2x^2/2}$ then

$$\mathcal{F}f(k) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} \mathrm{e}^{-ikx}\mathrm{e}^{-a^2x^2/2}\,\mathrm{d}x.$$

We may write

$$-\frac{a^2}{2}x^2 - ikx = -\frac{a^2}{2}\Big[\Big(x + \frac{ik}{a^2}\Big)^2 + \frac{k^2}{a^4}\Big]$$

and if we make the change of variable $y = x + ik/a^2$ then

$$\mathcal{F}f(k) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} \mathrm{e}^{-a^2y^2/2-k^2/2a^2}\,\mathrm{d}y$$

(of course the path of integration is on a line in the complex plane shifted upward and parallel to the real axis by a constant amount k/a^2, but this has no effect on the integral as it is a regular function which vanishes strongly at infinity). Using the standard error function integral

$$\int_{-\infty}^{\infty} \mathrm{e}^{-a^2y^2/2}\mathrm{d}y = \int_{-\infty}^{\infty} \mathrm{e}^{-z^2/2}\,\frac{\mathrm{d}z}{|a|} = \frac{\sqrt{2\pi}}{|a|}$$

we have

$$\mathcal{F}f(k) = \frac{1}{\sqrt{2\pi}}\frac{\sqrt{2\pi}}{|a|}\mathrm{e}^{-k^2/2a^2} = \frac{1}{|a|}\mathrm{e}^{-k^2/2a^2}.$$

Problem 12.12 **Evaluate Fourier transforms of the following distributional functions.**

(a) $\delta(x-a)$.

(b) $\delta'(x-a)$.

(c) $\delta^{(n)}(x-a)$.

(d) $\delta(x^2-a^2)$.

(e) $\delta'(x^2-a^2)$.

Solution: (a) If $f(x) = \delta(x-a)$ then

$$\mathcal{F}f(y) = \sqrt{1}\sqrt{2\pi}\int_{-\infty}^{\infty} \mathrm{e}^{-ixy}\delta(x-a)\,\mathrm{d}x = \frac{\mathrm{e}^{-iay}}{\sqrt{2\pi}}.$$

(b) If $f(x) = \delta'(x-a)$ then

$$\begin{aligned}\mathcal{F}f(y) &= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} \mathrm{e}^{-ixy}\delta'(x-a)\,\mathrm{d}x \\ &= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} -\frac{\mathrm{d}}{\mathrm{d}x}\big(\mathrm{e}^{-ixy}\big)\delta(x-a)\,\mathrm{d}x \\ &= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} iy\mathrm{e}^{-ixy}\delta(x-a)\,\mathrm{d}x \\ &= \frac{iy\mathrm{e}^{-iay}}{\sqrt{2\pi}}\,.\end{aligned}$$

(c) Similarly if $f(x) = \delta^{(n)}(x-a)$ we have, by induction,

$$\begin{aligned}\mathcal{F}f(y) &= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} \mathrm{e}^{-ixy}\delta^{(n)}(x-a)\,\mathrm{d}x \\ &= \frac{(-1)^n}{\sqrt{2\pi}}\frac{\mathrm{d}^n}{\mathrm{d}x^n}\big(\mathrm{e}^{-ixy}\big)\Big|_{x=a} \\ &= \frac{(-1)^n}{\sqrt{2\pi}}(-iy)^n\mathrm{e}^{-iay} \\ &= \frac{i^n y^n \mathrm{e}^{-iay}}{\sqrt{2\pi}}\,.\end{aligned}$$

(d) On performing a change of variable $z = x^2 - a^2$ as in Example 12.7, the Fourier transform of $f(x) = \delta(x^2 - a^2)$ is

$$\begin{aligned}\mathcal{F}f(y) &= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} \mathrm{e}^{-ixy}\delta(x^2-a^2)\,\mathrm{d}x \\ &= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} \mathrm{e}^{-ixy}\delta(z)\,\frac{\mathrm{d}z}{|2x|} \\ &= \frac{1}{\sqrt{2\pi}}\bigg(\frac{\mathrm{e}^{-ixy}}{|2x|}\bigg|_{x=-a} + \frac{\mathrm{e}^{-ixy}}{|2x|}\bigg|_{x=a}\bigg) \\ &= \frac{1}{\sqrt{2\pi}}\bigg(\frac{\mathrm{e}^{iay}}{2a} + \frac{\mathrm{e}^{-iay}}{2a}\bigg) \\ &= \frac{\sqrt{\cos ay}}{a\sqrt{2\pi}}\,.\end{aligned}$$

(e) For $f(x)\delta'(x^2 - a^2)$ the change of variable $z = x^2 - a^2$ results in (see Problem

12.8)

$$\begin{aligned}
\mathcal{F}f(y) &= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} \mathrm{e}^{-ixy}\delta'(x^2-a^2)\,\mathrm{d}x \\
&= \frac{1}{\sqrt{2\pi}}\Big\{\frac{\mathrm{d}}{\mathrm{d}z}\Big(\frac{\mathrm{e}^{-ixy}}{-2x}\Big)\Big|_{x=-a} + \frac{\mathrm{d}}{\mathrm{d}z}\Big(\frac{\mathrm{e}^{-ixy}}{2x}\Big)\Big|_{x=a}\Big\} \\
&= \frac{1}{\sqrt{2\pi}}\Big\{\frac{1}{2x}\frac{\mathrm{d}}{\mathrm{d}x}\Big(\frac{\mathrm{e}^{-ixy}}{-2x}\Big)\Big|_{x=-a} + \frac{1}{2x}\frac{\mathrm{d}}{\mathrm{d}x}\Big(\frac{\mathrm{e}^{-ixy}}{2x}\Big)\Big|_{x=a}\Big\} \\
&= \frac{1}{4\sqrt{2\pi}}\Big\{\frac{(1+ixy)\mathrm{e}^{-ixy}}{x^3}\Big|_{x=-a} - \frac{(1+ixy)\mathrm{e}^{-ixy}}{x^3}\Big|_{x=a}\Big\} \\
&= \frac{1}{4a^2\sqrt{2\pi}}\{iy\mathrm{e}^{iay} - \frac{1}{a}\mathrm{e}^{iay} - iy\mathrm{e}^{-iay} - \frac{1}{a}\mathrm{e}^{-iay}\} \\
&= \frac{-1}{2a^2\sqrt{2\pi}}\big(y\sin ay + \frac{1}{a}\cos ay\big)\,.
\end{aligned}$$

Problem 12.13 **Prove that**

$$x^m\delta^{(n)}(x) = (-1)^m\frac{n!}{(n-m)!}\delta^{(n-m)}(x) \qquad \textbf{for } n \geq m.$$

Hence show that the Fourier transform of the distribution

$$\sqrt{2\pi}\frac{k!}{(m+k)!}x^m\delta^{(m+k)}(-x) \quad (m,k\geq 0)$$

is $(-iy)^k$**.**

Solution: For any test function $\varphi(x)$,

$$\int_{-\infty}^{\infty} x^m\delta^{(n)}(x)\varphi(x)\,\mathrm{d}x = (-1)^n\frac{\mathrm{d}^n}{\mathrm{d}x^n}(x^m\varphi(x))\Big|_{x=0}.$$

From the binomial expansion

$$(f(x)g(x))^{(n)} = f^{(n)}g + nf^{(n-1)}g' + \cdots + fg^{(n)} = \sum_{k=0}^{n}\binom{n}{k}f^{(n-k)}g^{(k)}.$$

Setting $f = x^m$ and $g(x) = \varphi(x)$ the highest non-vanishing derivative of f in this expansion corresponds to $k = m \leq n$,

$$\frac{\mathrm{d}^n}{\mathrm{d}x^n}(x^m\varphi(x)) = \binom{n}{n-m}m!\varphi^{(n-m)} + \binom{n}{n-m+1}\frac{m!}{1!}x\varphi^{(n-m+1)} + \cdots + x^m\varphi^{(n)}.$$

Hence

$$\int_{-\infty}^{\infty} x^m\delta^{(n)}(x)\varphi(x)\,\mathrm{d}x = (-1)^n\binom{n}{n-m}m!\varphi^{(n-m)}(0) = (-1)^n\frac{n!}{(n-m)!}\varphi^{(n-m)}(0).$$

Since

$$\int_{-\infty}^{\infty} \delta^{(n-m)}\varphi(x)\mathrm{d}x = (-1)^{n-m}\varphi^{(n-m)}(0),$$

we have the identity

$$x^m\delta^{(n)}(x) = (-1)^m \frac{n!}{(n-m)!}\delta^{(n-m)}(x).$$

If

$$f(x) = \sqrt{2\pi}\frac{k!}{(m+k)!}x^m\delta^{(m+k)}(-x)$$

then

$$\begin{aligned}\mathcal{F}f(y) &= \int_{-\infty}^{\infty} \mathrm{e}^{-tiyx}\frac{k!}{(m+k)!}x^m\delta^{(m+k)}(-x)\,\mathrm{d}x \\ &= \int_{-\infty}^{\infty} \mathrm{e}^{tiyx}\frac{k!}{(m+k)!}(-1)^m x^m\delta^{(m+k)}(x)\,\mathrm{d}x\end{aligned}$$

on making the substitution $x \to -x$

$$\begin{aligned}&= \int_{-\infty}^{\infty} \mathrm{e}^{tiyx}\delta^{(k)}(x)\,\mathrm{d}x \qquad \text{by above identity} \\ &= (-1)^k \frac{\mathrm{d}^k}{\mathrm{d}x^k}\mathrm{e}^{iyx}\Big|_{x=0} \\ &= (-iy)^k\,.\end{aligned}$$

Problem 12.14 **(a) Show that the Fourier transform of the distribution**

$$\delta_0 + \delta_a + \delta_{2a} + \cdots + \delta_{(2n-1)a}$$

is a distribution with density

$$\frac{1}{\sqrt{2\pi}}\frac{\sin(nay)}{\sin(\frac{1}{2}ay)}\mathrm{e}^{-(n-\frac{1}{2})iay}.$$

(b) Show that

$$\mathcal{F}^{-1}(f(y)\mathrm{e}^{iby}) = (\mathcal{F}^{-1}f)(x+b).$$

Hence find the inverse Fourier transform of

$$g(y) = \frac{\sin nay}{\sin(\frac{1}{2}ay)}.$$

Solution: (a)

$$\mathcal{F}\delta_a(y) = \sqrt{1}\sqrt{2\pi}\int_{-\infty}^{\infty} \mathrm{e}^{-ixy}\delta_a(x)\,\mathrm{d}x = \frac{\mathrm{e}^{-iay}}{\sqrt{2\pi}}.$$

Hence

$$\begin{aligned}
&\mathcal{F}\big(\delta_0 + \delta_a + \delta_{2a} + \cdots + \delta_{(2n-1)a}\big)\\
&= \frac{1}{\sqrt{2\pi}}\big(1 + \mathrm{e}^{-iay} + \mathrm{e}^{-2iay} + \cdots + \mathrm{e}^{-(2n-1)iay}\big)\\
&= \frac{1}{\sqrt{2\pi}}\frac{1-\mathrm{e}^{-2niay}}{1-\mathrm{e}^{-iay}}\\
&= \frac{1}{\sqrt{2\pi}}\mathrm{e}^{-(n-\frac{1}{2})iay}\frac{\sin(nay)}{\sin(\frac{1}{2}ay)}
\end{aligned}$$

(b)
$$\mathcal{F}^{-1}\big(f(y)\mathrm{e}^{iby}\big) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} \mathrm{e}^{iyx}f(y)\mathrm{e}^{iby}\,\mathrm{d}y = \big(\mathcal{F}^{-1}f\big)(x+b).$$

Take $b = (n - \frac{1}{2}a$ gives

$$\begin{aligned}
\mathcal{F}^{-1}\Big(\frac{\sin(nay)}{\sin(\frac{1}{2}ay)}\Big) &= \sqrt{2\pi}\mathcal{F}^{-1}\big(\mathrm{e}^{iby}\frac{\mathrm{e}^{-(n-\frac{1}{2})iay}}{\sqrt{2\pi}}\frac{\sin(nay)}{\sin(\frac{1}{2}ay)}\big)\\
&= \sqrt{2\pi}\big(\delta(x+b) + \delta(x-a+b) + \delta(x-2a+b) + \dots\big)\\
&= \sqrt{2\pi}\big(\delta(x+(n-\tfrac{1}{2})a) + \delta(x+(n-\tfrac{3}{2})a) + \cdots + \delta(x-(n-\tfrac{1}{2})a)\big)\,.
\end{aligned}$$

Problem 12.15 Show that the Green's function for the time-independent Klein-Gordon equation

$$(\nabla^2 - m^2)\phi = \rho(\mathbf{r})$$

can be expressed as the Fourier integral

$$G(\mathbf{x}-\mathbf{x}') = -\frac{1}{(2\pi)^3}\iiint \mathrm{d}^3k\,\frac{\mathrm{e}^{i\mathbf{k}.(\mathbf{x}-\mathbf{x}')}}{\mathbf{k}^2+m^2}.$$

Evaluate this integral and show that it results in

$$G(\mathbf{R}) = -\frac{\mathrm{e}^{-mR}}{4\pi R} \qquad \textbf{where} \qquad \mathbf{R} = \mathbf{x}-\mathbf{x}', \quad R = |\mathbf{R}|.$$

Find the solution ϕ corresponding to a point source

$$\rho(\mathbf{r}) = q\delta^3(\mathbf{r}).$$

Solution: We seek a solution to the distributional equation

$$(\nabla^2 - m^2)G(\mathbf{x}-\mathbf{x}') = \delta^3(\mathbf{x}-\mathbf{x}').$$

so that a solution of the time-independent K-G equation is

$$\phi(\mathbf{x}) = \iiint \rho(\mathbf{x}')G(\mathbf{x}-\mathbf{x}')\mathrm{d}^3x'.$$

Set $g(\mathbf{k}) = \mathcal{F}G$ to be the Fourier transform of G, so that

$$G(\mathbf{y}) = \frac{1}{(2\pi)^{3/2}} \iiint_{-\infty}^{\infty} \mathrm{e}^{i\mathbf{k}\cdot\mathbf{y}} g(\mathbf{k})\mathrm{d}^3k,$$

and

$$(\nabla^2 - m^2)G(\mathbf{x}-\mathbf{x}') = \frac{1}{(2\pi)^{3/2}} \iiint_{-\infty}^{\infty} -(\mathbf{k}^2+m^2)\mathrm{e}^{i\mathbf{k}\cdot(\mathbf{x}-\mathbf{x}')} g(\mathbf{k})\mathrm{d}^3k.$$

The Green's equation

$$(\nabla^2 - m^2)G(\mathbf{x}-\mathbf{x}') = \delta^3(\mathbf{x}-\mathbf{x}') = \frac{1}{(2\pi)^3} \iiint_{-\infty}^{\infty} \mathrm{e}^{i\mathbf{k}\cdot(\mathbf{x}-\mathbf{x}')}\mathrm{d}^3k$$

implies that

$$g(\mathbf{k}) = -\frac{1}{(2\pi)^{3/2}} \frac{1}{\mathbf{k}^2+m^2},$$

and

$$G(\mathbf{x}-\mathbf{x}') = -\frac{1}{(2\pi)^3} \iiint_{-\infty}^{\infty} \frac{\mathrm{e}^{i\mathbf{k}\cdot(\mathbf{x}-\mathbf{x}')}}{\mathbf{k}^2+m^2}\mathrm{d}^3k.$$

In $\mathbf{k}$-space use polar coordinates (k,θ,ϕ) with the k_3-axis pointing along the direction $\mathbf{R} = \mathbf{x}-\mathbf{x}'$, so that

$$\mathbf{k}\cdot(\mathbf{x}-\mathbf{x}') = kR\cos\theta \qquad (k = \sqrt{\mathbf{k}\cdot\mathbf{k}})$$

and

$$\mathrm{d}^3k = k^2\sin\theta\mathrm{d}k\mathrm{d}\theta\mathrm{d}\phi.$$

The integral then evaluates to

$$\begin{aligned}
G(\mathbf{R}) &= -\frac{1}{(2\pi)^3}\int_0^\infty \mathrm{d}k \int_0^\pi \mathrm{d}\theta \int_0^{2\pi} \mathrm{d}\phi \frac{\mathrm{e}^{ikR\cos\theta}}{k^2+m^2}k^2\sin\theta \\
&= -\frac{2\pi}{(2\pi)^3}\int_0^\infty dk \int_0^\pi \mathrm{d}\theta \frac{k^2}{k^2+m^2}\frac{-1}{ikR}\frac{\mathrm{d}}{\mathrm{d}\theta}\Big(\frac{-\mathrm{e}^{ikR\cos\theta}}{ikR}\Big) \\
&= \frac{1}{2(2\pi)^2}\int_{-\infty}^\infty \mathrm{d}k \frac{k}{iR(k^2+m^2)}(\mathrm{e}^{-ikR}-\mathrm{e}^{ikR})
\end{aligned}$$

on making a change of variable $k \to -k$ and averaging the integrals before and after this change of variable. The integrand has poles at $k = \pm im$ in the complex k-plane,

$$\frac{k}{k^2+m^2} = \frac{1}{2}\Big(\frac{1}{k-im} + \frac{1}{k+im}\Big).$$

For the e^{-ikR} integral, complete the the integral along the real line with a circuit in the lower half plane encompassing the pole at $k = -im$, while for the e^{ikR} integral complete it with a circuit in the upper half plane (see figure below). The residue theorem then gives

$$\int_{-\infty}^{\infty} \mathrm{d}k \frac{k}{k^2 + m^2} \left(\mathrm{e}^{-ikR} - \mathrm{e}^{ikR}\right) = 2\pi i \left(-\frac{\mathrm{e}^{-mR}}{2} - \frac{\mathrm{e}^{-mR}}{2}\right) = -2\pi i \mathrm{e}^{-mR}.$$

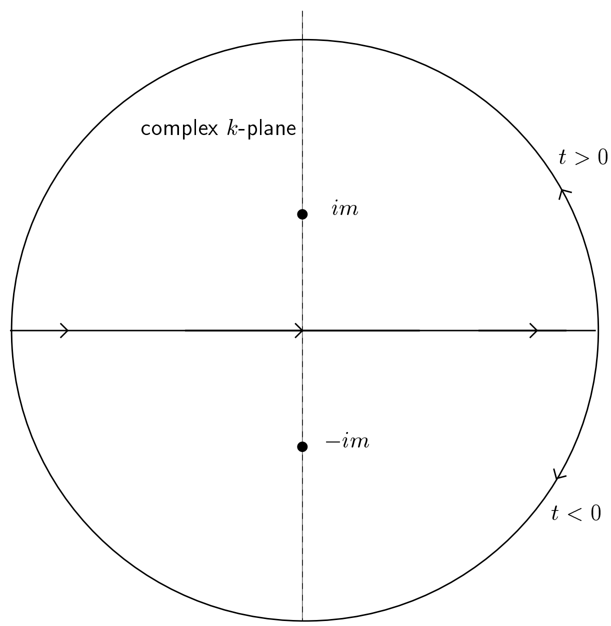
Hence

$$G(\mathbf{R}) = \frac{1}{2(2\pi)^2 R} \frac{-2\pi i}{iR} \mathrm{e}^{-mR} = -\frac{\mathrm{e}^{-mR}}{4\pi R}.$$

For a point source $\rho = q\delta^3(\mathbf{r})$ we have

$$\phi(\mathbf{x}) = -\iiint \frac{\rho(\mathbf{x}')\mathrm{e}^{-m|\mathbf{x}-\mathbf{x}'|}}{4\pi|\mathbf{x}-\mathbf{x}'|} \mathrm{d}^3 x' = -\frac{q\mathrm{e}^{-mr}}{4\pi r}$$

where $r = |\mathbf{x}|$.



Problem 12.16 Show that the Green's function for the one-dimensional diffusion equation,

$$\frac{\partial^2 G(x,t)}{\partial x^2} - \frac{1}{\kappa} \frac{\partial G(x,t)}{\partial t} = \delta(x - x')\delta(t - t')$$

is given by

$$G(x - x', t - t') = -\theta(t - t') \sqrt{\frac{\kappa}{4\pi(t - t')}} \mathrm{e}^{-(x-x')^2/4\kappa(t-t')},$$

and write out the corresponding solution of the inhomogeneous equation

$$\frac{\partial^2 \psi(x,t)}{\partial x^2} - \frac{1}{\kappa}\frac{\partial \psi(x,t)}{\partial t} = F(x,t).$$

Do the same for the two- and three-dimensional diffusion equations

$$\nabla^2 G(x,t) - \frac{1}{\kappa}\frac{\partial G(x,t)}{\partial t} = \delta^n(\mathbf{x}-\mathbf{x}')\delta(t-t') \qquad (n = 2,3).$$

Solution: We look for a solution $G(x,t)$ of

$$\frac{\partial^2 G(x,t)}{\partial x^2} - \frac{1}{\kappa}\frac{\partial G(x,t)}{\partial t} = \delta(x)\delta(t).$$

Then the Green's function is clearly $G(x-x', t-t')$. Decomposing as a Fourier integral,

$$G(x,t) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(k,\omega)\mathrm{e}^{ikx}\mathrm{e}^{i\omega t}\mathrm{d}k\mathrm{d}\omega$$

and

$$\delta(x)\delta(t) = \frac{1}{(2\pi)^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathrm{e}^{ikx}\mathrm{e}^{i\omega t}\mathrm{d}k\mathrm{d}\omega,$$

we find that

$$g(k,\omega) = \frac{-1}{2\pi(k^2 + i\omega/\kappa)}.$$

Hence

$$G(x,t) = \frac{-1}{(2\pi)^2}\int_{-\infty}^{\infty}\mathrm{d}k\mathrm{e}^{ikx}\int_{-\infty}^{\infty}\mathrm{d}\omega\frac{\mathrm{e}^{i\omega t}}{(k^2 + i\omega/\kappa)}.$$

In the complex ω-plane there is a pole at $\omega = ik^2\kappa$. The integral is completed with a circle at infinity around the pole in upper half plane for $t > 0$, and in the lower half plane for $t < 0$. Hence, by the residue theorem

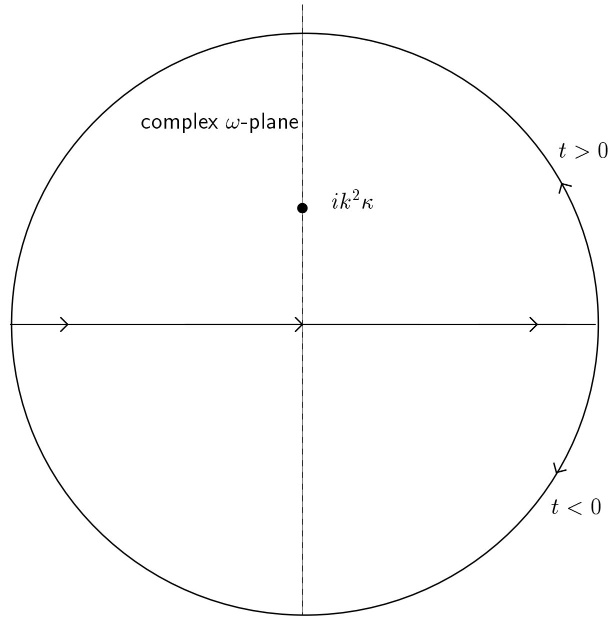
$$\int_{-\infty}^{\infty}\mathrm{d}\omega\frac{\mathrm{e}^{i\omega t}}{(k^2 + i\omega/\kappa)} = \theta(t)2\pi\kappa\mathrm{e}^{-k^2\kappa t}$$

and

$$G(x,t) = \frac{-1}{(2\pi)^2}2\pi\kappa\theta(t)\int_{-\infty}^{\infty}\mathrm{d}k\mathrm{e}^{ikx - k^2\kappa t}.$$

Using Problem 12.11 with $a^2 = 2\kappa t$ this integral evaluates to

$$\begin{aligned} G(x,t) &= \frac{-\kappa}{2\pi}\theta(t)\frac{\sqrt{2\pi}}{\sqrt{2\kappa t}}\mathrm{e}^{-x^2/4\kappa t} \\ &= -\sqrt{\frac{\kappa}{4\pi t}}\theta(t)\mathrm{e}^{-x^2/4\kappa t} \end{aligned}$$



and the Green's function of the one-dimensional diffusion equation is

$$G(x-x',t-t') = -\theta(t-t')\sqrt{\frac{\kappa}{4\pi(t-t')}}\mathrm{e}^{-(x-x')^2/4\kappa(t-t')}.$$

The corresponding solution of the inhomogeneous diffusion equation is

$$\begin{aligned}\psi(x,t) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} G(x-x',t-t')F(x',t')\mathrm{d}x'\mathrm{d}t' \\ &= -\sqrt{\frac{\kappa}{4\pi}}\int_{-\infty}^{\infty}\mathrm{d}x'\int_{-\infty}^{t}\mathrm{d}t' F(x',t')\frac{\mathrm{e}^{-(x-x')^2/4\kappa(t-t')}}{\sqrt{t-t'}}.\end{aligned}$$

In the two-dimensional diffusion equation set

$$G(x,y,t) = \frac{1}{(2\pi)^{3/2}}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(k_1,k_2,\omega)\mathrm{e}^{ik_1x}\mathrm{e}^{ik_2x}\mathrm{e}^{i\omega t}\mathrm{d}k_1\mathrm{d}k_2\mathrm{d}\omega$$

and

$$\delta(x)\delta)y)\delta(t) = \frac{1}{(2\pi)^3}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\mathrm{e}^{ik_1x}\mathrm{e}^{ik_2x}\mathrm{e}^{i\omega t}\mathrm{d}k_1\mathrm{d}k_2\mathrm{d}\omega.$$

Then

$$g(k_1,k_2,\omega) = \frac{-1}{(2\pi)^{3/2}(\mathbf{k}^2+i\omega/\kappa)},$$

and using the same pole integration argument as in the one-dimensional case

$$\begin{aligned}
G &= \frac{-\kappa}{(2\pi)^2}\theta(t)\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathrm{d}k_1\mathrm{d}k_2 \mathrm{e}^{ik_1x}\mathrm{e}^{ik_2x}\mathrm{e}^{-k^2\kappa t} \\
&= \frac{-\kappa}{(2\pi)^2}\theta(t)\int_{-\infty}^{\infty} \mathrm{d}k_1 \mathrm{e}^{ik_1x}\mathrm{e}^{-(k_1)^2\kappa t}\int_{-\infty}^{\infty} \mathrm{d}k_2 \mathrm{e}^{ik_2x}\mathrm{e}^{-(k_2)^2\kappa t} \\
&= \frac{-\kappa}{(2\pi)^2}\theta(t)\frac{2\pi}{2\kappa t}\mathrm{e}^{-x^2/4\kappa t}\mathrm{e}^{-y^2/4\kappa t} \\
&= -\theta(t)\frac{\mathrm{e}^{-r^2/4\kappa t}}{4\pi t}
\end{aligned}$$

where $r^2 = x^2 + y^2$. Hence

$$G(x-x', y-y', t-t') = -\theta(t-t')\frac{\mathrm{e}^{-((x-x')^2+(y-y')^2)/4\kappa t}}{4\pi(t-t')}.$$

For the three-dimensional diffusion equation, set

$$G(\mathbf{r}, t) = \frac{1}{(2\pi)^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(\mathbf{k}, \omega)\mathrm{e}^{i\mathbf{k}\cdot\mathbf{x}}\mathrm{e}^{i\omega t}\mathrm{d}^3k\mathrm{d}\omega$$

and we find

$$g(\mathbf{k}, \omega) = \frac{-1}{(2\pi)^2(\mathbf{k}^2 + i\omega/\kappa)},$$

so that

$$\begin{aligned}
G(\mathbf{r}, t) &= \frac{-\kappa}{(2\pi)^3}\theta(t)\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathrm{d}k_1\mathrm{d}k_2\mathrm{d}k_3 \mathrm{e}^{ik_1x-(k_1)^2t}\mathrm{e}^{ik_2y-(k_2)^2t}\mathrm{e}^{ik_3z-(k_3)^2t} \\
&= \frac{-\kappa}{(2\pi)^3}\theta(t)\frac{(2\pi)^{3/2}}{(2\kappa t)^{3/2}}\mathrm{e}^{-\mathbf{r}^2/4\kappa t}.
\end{aligned}$$

Hence

$$G(\mathbf{r}-\mathbf{r}', t-t') = -\frac{\theta(t-t')\mathrm{e}^{-(\mathbf{r}-\mathbf{r}')^2/4\kappa t}}{(4\pi)^{3/2}\kappa^{1/2}t^{3/2}}.$$

Chapter 13

Problem 13.1 **Let E be a Banach space in which the norm satisfies the parallelogram law (13.2). Show that it is a Hilbert space with inner product given by**

$$\langle x \,|\, y\rangle = \tfrac{1}{4}\big(\|x+y\|^2 - \|x-y\|^2 + i\|x-iy\|^2 - i\|x+iy\|^2\big).$$

Solution: This has equation has been shown to be an identity in Hilbert space in Problem 5.7. To show (IP1) holds for this definition of $\langle x \,|\, y\rangle$:

$$\begin{aligned}\langle u \,|\, v\rangle - \overline{\langle v \,|\, u\rangle} &= \tfrac{1}{4}\big(\|u+v\|^2 - \|u-v\|^2 + i\|u-iv\|^2 - i\|u+iv\|^2 \\ &\quad - \|v+u\|^2 + \|v-u\|^2 + i\|v-iu\|^2 - i\|v+iu\|^2\big) \\ &= \tfrac{1}{4}i\big(\|u-iv\|^2 - \|v+iu\|^2 - \|u+iv\|^2 + \|v-iu\|^2\big) \\ &= 0\end{aligned}$$

since

$$\|v+iu\|^2 = \|i(u-iv)\|^2 = |i|^2\|u-iv\|^2 = \|u-iv\|^2$$

and similarly $\|v-iu\|^2 = \|u+iv\|^2$.
(IP3) follows on setting $y = x$:

$$\langle x \,|\, x\rangle = \tfrac{1}{4}\big(4\|x\|^2 + i|1-i|^2\|x\|^2 - i|1+i|^2\|x\|^2\big) = \|x\|^2.$$

The argument for (IP2) is considerably more complicated. It is best to divide the proof of linearity in the second argument into two parts.
(i) We show

$$\langle u \,|\, v+w\rangle = \langle u \,|\, v\rangle + \langle u \,|\, w\rangle$$

by expanding

$$\begin{aligned}&\|u+v+w\|^2 - \|u-v-w\|^2 + i\|u-iv-iw\|^2 - i\|u+iv+iw\|^2 \\ &\quad - \|u+v\|^2 + \|u-v\|^2 - i\|u-iv\|^2 + i\|u+iv\|^2 \\ &\quad - \|u+w\|^2 + \|u-w\|^2 - i\|u-iw\|^2 + i\|u+iw\|^2 = 0.\end{aligned}$$

As the imaginary part is identical with the real part with v and w replaced by $-iv$ and $-iw$ resp., it is only necessary to show the real part of this equation:

$$\|u+v+w\|^2 + \|u-v\|^2 + \|u-w\|^2 = \|u-v-w\|^2 + \|u+v\|^2 + \|u+w\|^2.$$

Using the parallelogram law (13.2) with $x = u+v$ and $y = u+w$ gives

$$\|u+v\|^2 + \|u+w\|^2 = \tfrac{1}{2}\|2u+v+w\|^2 + \tfrac{1}{2}\|v-w\|^2$$

and similarly

$$\|u-v\|^2 + \|u-w\|^2 = \tfrac{1}{2}\|2u-v-w\|^2 + \tfrac{1}{2}\|w-v\|^2.$$

Thus it is required to show

$$\|u+x\|^2 + \tfrac{1}{2}\|2u - x\|^2 = \|u - x\|^2 + \tfrac{1}{2}\|2u + x\|^2$$

where $x = v + w$, i.e.

$$\|u+x\|^2 + 2\|u - \frac{x}{2}\|^2 = \|u - x\|^2 + 2\|u + \frac{x}{2}\|^2.$$

From the parallelogram law,

$$\|u+x\|^2 = 2\|u + \frac{x}{2}\|^2 + 2\|\frac{x}{2}\|^2 - \|u\|^2$$

and

$$\|u-x\|^2 = 2\|u - \frac{x}{2}\|^2 + 2\|\frac{x}{2}\|^2 - \|u\|^2$$

whence the result follows immediately.
(ii) It remains to show that $\langle u \,|\, av\rangle = a\langle u \,|\, v\rangle$ for all complex numbers a, i.e.

$$\begin{aligned}\|u+av\|^2 - \|u-av\|^2 + i\|u - iav\|^2 - i\|u + iav\|^2 \\ - a\|u+v\|^2 + a\|u-v\|^2 - ia\|u-iv\|^2 + ia\|u+iv\|^2 = 0.\end{aligned}$$

It is trivial to verify that this equation holds for $a = i$ and $a = -1$, so it only remains to show this equation for real positive numbers $a > 0$. From (i), setting $w = v$, we have

$$\langle u \,|\, 2v\rangle = \langle u \,|\, v+v\rangle = \langle u \,|\, v\rangle + \langle u \,|\, v\rangle = 2\langle u \,|\, v\rangle.$$

Continuing by induction it follows simply that for any positive integer m

$$\langle u \,|\, mv\rangle = m\langle u \,|\, v\rangle.$$

Setting $w = nv$, we have then

$$\langle u \,|\, w\rangle = n\langle u \,|\, \tfrac{1}{n}w\rangle$$

whence

$$\langle u \,|\, \tfrac{1}{n}w\rangle = \tfrac{1}{n}\langle u \,|\, w\rangle$$

and for any positive rational number $a = m/n$

$$\langle u \,|\, \tfrac{m}{n}w\rangle = \tfrac{m}{n}\langle u \,|\, w\rangle.$$

As every real number a is a limit of a sequence of rational numbers, $r_1, r_2, \dots, r_i \to a$, and $r_i v \to av$ for any $v \in E$ it follows that

$$\langle u \,|\, av\rangle = \lim_{i\to\infty}\langle u \,|\, r_i v\rangle = \lim_{i\to\infty} r_i\langle u \,|\, v\rangle = a\langle u \,|\, v\rangle.$$

The result now follows for any complex number $a = a_1 + ia_2$ where a_1 and a_2 are real.

Problem 13.2 **On the vector space $\mathcal{F}^1[a,b]$ of complex continuous differentiable functions on the interval $[a,b]$, set**

$$\langle f \,|\, g\rangle = \int_a^b \overline{f'(x)}g'(x)\mathrm{d}x \qquad \textbf{where} \qquad f' = \frac{\mathrm{d}f}{\mathrm{d}x}, g' = \frac{\mathrm{d}g}{\mathrm{d}x}.$$

Show that this is not an inner product, but becomes one if restricted to the space of functions $f \in \mathcal{F}^1[a,b]$ having $f(c) = 0$ for some fixed $a \le c \le b$. Is it a Hilbert space?

Give a similar analysis for the case $a = -\infty$, $b = \infty$, and restricting functions to those of compact support.

Solution: While (IP1) and (IP2) are satisfied, (IP3) is not for if

$$\|f\|^2 = \langle f \,|\, f\rangle = \int_a^b |f'(x)|^2 \mathrm{d}x = 0$$

then $f'(x) = 0$ a.e. Hence

$$f(x) = f(a) + \int_a^x f'(x)\mathrm{d}x = f(a) = \text{const.}$$

and in general $f(x) \neq 0$ a.e. If we restrict $\mathcal{F}^1[a,b]$ to functions such that $f(c) = 0$ for some fixed $c \in [a,b]$, then it is still a vector space (this is why c must be a fixed point), and (IP3) follows. It is not, however, a Hilbert space since it is not complete as the limit of a Cauchy sequence of differentiable functions need not be differentiable; e.g.

$$f_n(x) = x(1-x)\big(1 - \sqrt{\frac{1}{n^2} + (x - \frac{1}{2})^2}\big) \;\to\; f(x) = x(1-x)(1 - |x - \tfrac{1}{2}|)$$

is a sequence of differentiable functions all having $f_n(0) = 0$ whose limit is not differentiable at $x = \frac{1}{2}$.
For functions of compact support in $\mathbb{R} = (-\infty, \infty)$ there is no need for the extra restriction $f(c) = 0$ since all such functionsd vanish at some point $c \in \mathbb{R}$. Hence it is an inner product vector space, but again it is not a Hilbert space for the same reason as above.

Problem 13.3 **In the space $L^2([0,1])$ which of the following sequences of functions (i) is a Cauchy sequence, (ii) converges to 0, (iii) converges everywhere to 0, (iv) converges almost everywhere to 0, and (v) converges almost nowhere to 0?**

(a) $f_n(x) = \sin^n(x)$, $n = 1, 2, \ldots$

(b) $f_n(x) = \begin{cases} 0 & \textbf{for } x < 1 - \frac{1}{n}, \\ nx + 1 - n & \textbf{for } 1 - \frac{1}{n} \le x \le 1. \end{cases}$

(c) $f_n(x) = \sin^n(nx)$.

(d) $f_n(x) = \chi_{U_n}(x)$**, the characteristic function of the set**

$$U_n = \left[\frac{k}{2^m}, \frac{k+1}{2^m}\right] \qquad \textbf{where} \qquad n = 2^m + k,\ m = 0, 1, \ldots \ \textbf{and}\ k = 0, \ldots, 2^m - 1.$$

Solution: (a) The functions $f_n(x) = \sin^n(x)$ approach 0 everywhere in the interval $[0,1]$ since $1 < \frac{\pi}{2}$. Hence $\int_0^1 |f_n(x)|^2 \mathrm{d}x \to 0$ and the functions converge to the zero function in the norm sense, and therefore form a Cauchy sequence.
(b) These functions approach zero everywhere except at $x = 1$, hence they also converge to the zero function in the sense that $\int_0^1 |f_n(x)|^2 \mathrm{d}x \to 0$.
(c) $f_n(x) = \sin^n(nx) \to 0$ at all points $x \neq \pi/2k$ where k is an integer. Hence

$$\int_0^1 |sin^n nx| \mathrm{d}x = \frac{1}{n} \int_0^n |\sin^n y| \mathrm{d}y = \frac{\mathrm{o}(n)}{n} \to 0$$

and the sequence converge to the zero function. (d) The function $f(n)(x) = 0$ except on a set of measure 2^{-m}. The intervals on which $f(n)(x) \neq 0$ move back and forth over the range $[0,1]$ so that at no point x does $f_n(x) \to 0$. However the integrals $\int_0^1 |f_n(x)|^2 \mathrm{d}x < 2^{-m} \to 0$ as $n = 2^m + k \to \infty$. The sequence f_n therefore converges to the zero function and is a Cauchy sequence. In summary, the sequences which
(i) are Cauchy sequences: (a), (b) , (c), (d),
(ii) converge to zero: (a), (b) , (c), (d),
(iii) converge everywhere to 0: (a)
(iv) converge to 0 a.e.: (a), (b), (c)
(v) converge almost nowhere to 0: (d).

Problem 13.4 **Show that a vector subspace V is a closed subset of $\mathcal{H}$ with respect to the norm topology iff the limit of every sequence of vectors in V belongs to V.**

Solution: If V is closed with respect to the norm topology then $V^c = \mathcal{H} - V$ is open. Hence a sequence $u_n \in V$ cannot converge to a vector $u \in V^c$, for there exists an open ball

$$B_\epsilon(u) = \{v \in \mathcal{H} \mid \|v - u\| < \epsilon\} \subset V^c$$

and $\|u_n - u\| \geq \epsilon$ for all n, which is a contradiction to $\|u - u_n\| \to 0$.
Conversely, if V is not closed then, since V^c is not open, there exists $v \in V^c$ such that every open ball $B_\epsilon(v) \cap V \neq \emptyset$. For each $n = 1, 2, 3, \ldots$ there exists $v_n \in V$ such that $\|v - v_n\| < \frac{1}{n}$. In other words, if V is not closed then there exists a sequence $v_n \in V$ which converges to a vectors $v \notin V$. This proves the *if* part of the problem.

Problem 13.5 **Let ℓ_0 be the subset of ℓ^2 consisting of sequences with only finitely many terms different from zero. Show that ℓ_0 is a vector subspace of ℓ^2, but that it is not closed. What is its closure $\overline{\ell_0}$?**

Solution: If

$$u = (u_1, u_2, \dots, u_n, 0, 0, \dots),\ v = (v_1, v_2, \dots, v_m, 0, 0, \dots) \in \ell_0$$

then for any complex number a, $u+av \in \ell_0$ since $(u+av)_i = 0$ for all $i > \max(n, m)$. Hence ℓ_0 is a vector subspace of ℓ^2.
It is not closed, and therefor not a subspace of ℓ^2 for the sequence of vectors

$$u_1 = (1, 0, 0, \dots),\ u_2 = (1, \tfrac{1}{2}, 0, 0 \dots),\ u_3 = (1, \tfrac{1}{2}, \tfrac{1}{3} 0, 0 \dots) \quad \dots$$

has limit

$$u = (1, \tfrac{1}{2}, \tfrac{1}{3}, \tfrac{1}{4}, \tfrac{1}{5}, \dots).$$

Furthermore, every vector $v = (v_1, v_2, v_3, \dots) \in \ell^2$ is the limit of a sequence of vectors in ℓ_0,

$$v_{(1)} = (v_1, 0, 0, \dots),\ v_{(2)} = (v_1, v_2, 0, 0, \dots),\ v_{(3)} = (v_1, v_2, v_3, 0, \dots),\ \dots.$$

Hence the closure of ℓ_0 is the entire space ℓ^2.

Problem 13.6 **We say a sequence $\{x_n\}$ *converges weakly* to a point x in a Hilbert space $\mathcal{H}$, written $x_n \rightharpoonup x$ if $\langle x_n \mid y \rangle \to \langle x \mid y \rangle$ for all $y \in \mathcal{H}$. Show that every strongly convergent sequence, $\|x_n - x\| \to 0$ is weakly convergent to x. In finite dimensional Hilbert spaces show that every weakly convergent sequence is strongly convergent.**

Give an example where $x_n \rightharpoonup x$ but $\|x_n\| \not\to \|x\|$. Is it true in general that the weak limit of a sequence is unique?

Show that if $x_n \rightharpoonup x$ and $\|x_n\| \not\to \|x\|$ then $x_n \not\to x$.

Solution: Lemma 13.3 is essentially the statement that if $x_n \to x$, i.e. $\|x_n - x\| \to 0$, then $x_n \rightharpoonup x$.
In a finite dimensional Hilbert space of dimension m, let $e_1, \dots e_m$ be any o.n. basis. Then if $x_n \rightharpoonup x$ we have

$$\langle e_i \,|\, x_n \rangle = x_{ni} \to x_1 = \langle e_i \,|\, x \rangle \qquad (i = 1, 2, \dots, m)$$

where $x_n = \sum_{j=1}^m x_{nj} e_j$, $x = \sum_{j=1}^m x_j e_j$. hence

$$\|x_n - x\| = \sqrt{|x_{n1} - x_1|^2 + \dots + |x_{nm} - x_m|^2} \to 0$$

That is, in finite dimensional Hilbert spaces, $x_n \rightharpoonup x$ implies $x_n \to x$.

In ℓ^2, let $x_1 = e_1 = (1, 0, 0, \dots)$, $x_2 = e_2 = (0, 1, 0, 0, \dots)$ etc. Then $x_n \rightharpoonup 0$ since, for any $v = (v_1, v_2, v_3, \dots) \in \ell^2$, $\langle x_n \,|\, v \rangle \to \langle 0 \,|\, v \rangle$:

$$\langle x_n \,|\, v \rangle = \langle e_n \,|\, v \rangle = v_n \to 0.$$

However, $\|x_n\| = 1 \not\to \|0\| = 0$.

The weak limit of a sequence, if it exists, is unique, for if $x_n \rightharpoonup x$ and $x_n \rightharpoonup x'$ then for all $y \in \mathcal{H}$

$$\langle x_n \,|\, y\rangle \to \langle x \,|\, y\rangle = \langle x' \,|\, y\rangle.$$

Hence $\langle x - x' \,|\, y\rangle = 0$ for all $y \in \mathcal{H}$, and setting $y = x - x'$ gives $x = x'$ by (IP3).

On setting $u_n = v_n = x_n$ in Lemma 13.5, we have $x_n \to x \Longrightarrow \|x_n\| \to \|x\|$. Hence $\|x_n\| \not\to \|x\| \Longrightarrow x_n \not\to x$.

Problem 13.7 **In the Hilbert space $L^2([-1,1])$ let $\{f_n(x)\}$ be the sequence of functions $1, x, x^2, \dots, f_n(x) = x^n, \dots$**

(a) Apply Schmidt orthonormalization to this sequence, writing down the first three polynomials so obtained.

(b) The nth *Legendre polynomial* $P_n(x)$ is defined as

$$P_n(x) = \frac{1}{2^n n!}\frac{\mathrm{d}^n}{\mathrm{d}x^n}\left(x^2 - 1\right)^n.$$

Prove that

$$\int_{-1}^{1} P_m(x)P_n(x)\mathrm{d}x = \frac{2}{2n+1}\delta_{mn}.$$

(c) Show that the nth member of the o.n. sequence obtained in (a) is $\sqrt{n+\frac{1}{2}}P_n(x)$.

Solution: (a) $\|f_0\|^2 = \int_{-1}^{1} 1^2 \mathrm{d}x = 2$. Therefore set

$$e_0 = \frac{1}{\sqrt{2}} f_0 = \frac{1}{\sqrt{2}}, \qquad \|e_0\|^2 = 1.$$

Now

$$\langle f_1 \,|\, f_0\rangle = \int_{-1}^{1} x \mathrm{d}x = 0, \qquad \|f_1\|^2 = \int_{-1}^{1} x^2 \mathrm{d}x = \frac{2}{3},$$

hence we must set

$$e_1(x) = \sqrt{\frac{3}{2}}x.$$

From

$$\langle f_2 \,|\, e_0\rangle = \int_{-1}^{1} x^2 \frac{1}{\sqrt{2}}\mathrm{d}x = \frac{\sqrt{2}}{3}$$
$$\langle f_2 \,|\, e_1\rangle = \int_{-1}^{1} x^2 \frac{\sqrt{3}}{\sqrt{2}} x\mathrm{d}x = 0$$

we set

$$\begin{aligned} g_2 &= f_2 - \langle e_0 \,|\, f_2\rangle e_0 - \langle e_1 \,|\, f_2\rangle e_1 \\ &= x^2 - \frac{1}{3}. \end{aligned}$$

Since

$$\|g_2\|^2 = \int_{-1}^{1} (x^2 - \frac{1}{3})^2 \mathrm{d}x = \frac{8}{45}$$

the normalized function e_2 is

$$e_2(x) = \frac{3\sqrt{5}}{2\sqrt{2}}(x^2 - \frac{1}{3}) = \frac{\sqrt{5}}{2\sqrt{2}}(3x^2 - 1).$$

(b) Since $(x^2-1)^m$ is a polynomial of degree $2m$, we have

$$\frac{\mathrm{d}^{m+p}}{\mathrm{d}x^{m+p}}(x^2-1)^m = \begin{cases} 0 & \text{for } p > m \\ (2n)! & \text{for } p = m \end{cases}$$

and

$$\frac{\mathrm{d}^{m-p}}{\mathrm{d}x^{m-p}}(x^2-1)^n \Big|_{x=\pm 1} = 0 \qquad \text{for } p = 1, 2, \ldots m-1.$$

Hence m repeated integrations by parts (assuming $m \geq n$) results in

$$\begin{aligned}\int_{-1}^{1} P_m(x)P_n(x)\mathrm{d}x &= \frac{1}{2^m m!}\frac{1}{2^n n!}\int_{-1}^{1}(x^2-1)^m \frac{\mathrm{d}x^{n+m}}{\mathrm{d}x^{n+m}}(t^2-1)^n \mathrm{d}x \\ &= \frac{1}{2^{n+m}m!n!}(2n)!\int_{-1}^{1}(x^2-1)^n \mathrm{d}x \delta_{mn}\end{aligned}$$

The change of variables $x = \sin\theta$ gives

$$\int_{-1}^{1}(x^2-1)^n \mathrm{d}x = (-1)^n \int_{-\pi/2}^{\pi/2} \cos^{2n+1}\theta \mathrm{d}\theta = (-1)^n 2\int_{0}^{\pi/2} \cos^{2n+1}\theta \mathrm{d}\theta.$$

On integration by parts,

$$\begin{aligned}\int_{0}^{\pi/2} \cos^{2n+1}\theta \mathrm{d}\theta &= \int_{0}^{\pi/2} 2n\cos^{2n-1}\theta \sin^2\theta \mathrm{d}\theta \\ &= 2n\int_{0}^{\pi/2} \cos^{2n-1}\theta - \cos^{2n+1}\theta \mathrm{d}\theta.\end{aligned}$$

Hence

$$\int_{0}^{\pi/2} \cos^{2n+1}\theta \mathrm{d}\theta = \frac{2n}{2n+1}\int_{0}^{\pi/2} \cos^{2n-1}\theta \mathrm{d}\theta$$

and, continuing by induction,

$$\int_{0}^{\pi/2} \cos^{2n+1}\theta \mathrm{d}\theta = \frac{2^n n!}{(2n+1)(2n-1)\ldots 2}\int_{0}^{\pi/2} \cos\theta \mathrm{d}\theta = \frac{(2^n n!)^2}{(2n+1)!}.$$

Thus we have

$$\langle P_m \,|\, P_n \rangle = \int_{-1}^{1} P_m(x)P_n(x)\mathrm{d}x = \frac{2}{2n+1}\delta_{mn}.$$

For $m < n$ the conclusion holds by reversing roles of m and n.

(c) The functions $\phi_n(x) = \sqrt{n+\frac{1}{2}}P_n(x)$ are an orthonormal set, $\langle \phi_m \mid \phi_n \rangle = \delta_{mn}$. Since $\phi_n(x)$ is a polynomial of degree n it can be written in the form

$$\phi_n(x) = a_0 f_0(x) + a_1 f_1(x) + \cdots + a_n f_n(x)$$

and since $a_n \neq 0$ these may be reversed,

$$f_n(x) = b_0\phi_0(x) + b_1\phi_1(x) + \cdots + b_n\phi_n(x).$$

Clearly $e_0(x) = \phi_0(x) = 1/\sqrt{2}$, and

$$e_1 = A_0 f_0 + A_1 f_1 = B_0\phi_0 + B_1\phi_1.$$

Since $\langle \phi_0 \mid e_1 \rangle = 0$ we have $B_0 = 0$, and $\langle \phi_1 \mid \phi_1 \rangle = 1$ implies we may set $B_1 = 1$. Continuing in this way

$$\begin{aligned} e_n(x) &= A_0 f_0(x) + A_1 f_1(x) + \cdots + A_n f_n(x) \\ &= B_0\phi_0(x) + B_1\phi_1(x) + \cdots + B_n\phi_n(x). \end{aligned}$$

Since $\langle \phi_i \,|\, e_n \rangle = 0$ for $i = 0, 1, \dots, n-1$, all $B_i = 0$ except for B_n, and $\langle e_n \,|\, e_n \rangle = 1$ implies that $|B_n|^2 = 1$. Apart from a redundant phase factor we may set $B_n = 1$ and the Schmidt orthonormalization procedure results precisely in the normalized Legendre polynomials, $e_n(x) = \phi_n(x) = \sqrt{n+\frac{1}{2}}P_n(x)$.

Problem 13.8 Show that Schmidt orthonormalization in $L^2(\mathbb{R})$, applied to the sequence of functions

$$f_n(x) = x^n \mathrm{e}^{-x^2/2},$$

leads to the normalized hermite functions (13.6) of Example 13.7.

Solution: From Example 13.7, we define the n-th Hermite polynomial to be

$$H_n(x) = (-1)^n \mathrm{e}^{x^2} \frac{\mathrm{d}^n \mathrm{e}^{-x^2}}{\mathrm{d}x^n},$$

a polynomial of degree n with leading term $(-2x)^n$. The related normalized functions on $(-\infty, \infty)$ are shown in Eq. (13.6) to be

$$\phi_n(x) = \frac{\mathrm{e}^{-(1/2)x^2}}{\sqrt{2^n n! \sqrt{\pi}}} H_n(x),$$

and form a complete orthonormal set. Since $f_0(x) = \mathrm{e}^{-x^2/2}$, the first function arrived at by Schmidt orthonormalization is

$$e_0(x) = \phi_0(x) = \frac{\mathrm{e}^{-(1/2)x^2}}{\pi^{1/4}}.$$

The functions $\phi_n(x)$ can be expressed as linear combinations

$$\phi_n(x) = \sum_{i=0}^{n} a_i f_i(x) \implies f_n(x) = \sum_{i=0}^{n} b_i \phi_i(x)$$

from which as in Problem 13.7 (c),

$$e_n(x) = \sum_{i=0}^{n} A_i f_i(x) = \sum_{i=0}^{n} B_i \phi_i(x)$$

and $\langle \phi_i \mid e_n \rangle = 0$ implies $B_i = 0$ for $i = 0, 1, \dots, n-1$, while it follows from $\langle e_n \mid e_n \rangle = 1$ that, apart from a phase factor, we must set $B_n = n$. Hence the functions $\phi_n(x)$ are the o.n. set arrived by Schmidt orthonormalization from the set of functions $f_n(x)$ on the full real line $\mathbb{R} = (-\infty, \infty)$.

Problem 13.9 **Show that applying Schmidt orthonormalization in $L^2([0,\infty])$ to the sequence of functions**

$$f_n(x) = x^n \mathrm{e}^{-x/2}$$

leads to a normalized sequence of functions involving the *Laguerre polynomials*

$$L_n(x) = \mathrm{e}^x \frac{\mathrm{d}^n}{\mathrm{d}x^n}\left(x^n \mathrm{e}^{-x}\right).$$

Solution: The function $L_n(x)$ is clearly a polynomial of degree n with leading coefficient $(-x)^n$. Set $\psi_n = \mathrm{e}^{-x/2} L_n(x)$, and by a similar argument to that in Example 13.7,

$$\begin{aligned}
\langle \psi_m \,|\, \psi_n \rangle &= \int_0^\infty \mathrm{e}^x \frac{\mathrm{d}^m (x^m \mathrm{e}^{-x})}{\mathrm{d}x^m} \frac{\mathrm{d}^n (x^n \mathrm{e}^{-x})}{\mathrm{d}x^n} \mathrm{d}x \\
&= \left\{ \left[\mathrm{e}^x \frac{\mathrm{d}^{m-1} (x^m \mathrm{e}^{-x})}{\mathrm{d}x^{m-1}} \frac{\mathrm{d}^n (x^n \mathrm{e}^{-x})}{\mathrm{d}x^n} \right]_0^\infty \right. \\
&\qquad \left. - \int_0^\infty \frac{\mathrm{d}^{m-1} (x^m \mathrm{e}^{-x})}{\mathrm{d}x^{m-1}} \frac{\mathrm{d}}{\mathrm{d}x} \left(\mathrm{e}^x \frac{\mathrm{d}^n (x^n \mathrm{e}^{-x})}{\mathrm{d}x^n} \right) \mathrm{d}x \right\} \\
&= \dots \\
&= (-1)^m \int_0^\infty x^m \mathrm{e}^{-x} \frac{\mathrm{d}^m}{\mathrm{d}x^m} \left(\mathrm{e}^x \frac{\mathrm{d}^n (x^n \mathrm{e}^{-x})}{\mathrm{d}x^n} \right) \mathrm{d}x
\end{aligned}$$

after m successive integration by parts. Since the term to be differentiated m times in the final integrand is the n-th Laguerre polynomial and has degree n, this term vanishes if $m > n$. Hence, as the result is similar for $m < n$ on reversing roles of m and n,

$$\langle \psi_m \,|\, \psi_n \rangle = 0 \qquad \text{if } m \neq n.$$

For $m = n$, we have

$$\begin{aligned}
\langle \psi_n \,|\, \psi_n \rangle &= (-1)^n \int_0^\infty x^n \mathrm{e}^{-x} \frac{\mathrm{d}^n}{\mathrm{d}x^n}\Big(\mathrm{e}^x \frac{\mathrm{d}^n (x^n \mathrm{e}^{-x})}{\mathrm{d}x^n}\Big)\mathrm{d}x \\
&= (-1)^n \int_0^\infty x^n \mathrm{e}^{-x} (-1)^n n! \mathrm{d}x \\
&= n! \Big\{ \Big[-x^n \mathrm{e}^{-x}\Big]_0^\infty + \int_0^\infty n x^{n-1} \mathrm{e}^{-x} \mathrm{d}x \Big\} \\
&= \dots \\
&= (n!)^2 \int_0^\infty \mathrm{e}^{-x} \mathrm{d}x \\
&= (n!)^2
\end{aligned}$$

after n integrations by parts. Hence

$$\langle \psi_m \,|\, \psi_n \rangle = (n!)^2 \delta_{mn}$$

and the functions

$$\phi_n(x) = \frac{\mathrm{e}^{-x/2}}{n!} L_n(x) = \frac{\mathrm{e}^{x/2}}{n!} \frac{\mathrm{d}^n}{\mathrm{d}x^n}(x^n \mathrm{e}^{-x})$$

form a complete orthonormal set on the real half-line $[0, \infty)$.
The argument that these arise by Schmidt orthonormaliation on this interval from the functions $f_n(x) = x^n \mathrm{e}^{-x/2}$ is essentially identical to that in the Problems 13.7 and 13.8.

Problem 13.10 **If S is any subset of $\mathcal{H}$, and V the closed subspace generated by S, $V = \overline{L(S)}$, show that $S^\perp = \{u \in \mathcal{H} \,|\, \langle u \mid x \rangle = 0$ for all $x \in S\} = V^\perp$.**

Solution: If $u \in V^\perp$ then $\langle u \,|\, v \rangle = 0$ for all $v \in V$, and therefore $\langle u \,|\, x \rangle = 0$ for all $x \in S \subset V$. Hence $V^\perp \subseteq S^\perp$.
Conversely, if $u \in S^\perp$, then $\langle u \,|\, x \rangle = 0$ for all $x \in S$. Hence u is orthogonal to all (finite) linear combinations of elements in S,

$$\langle u \,|\, a_1 x_1 + \dots a_n x_n \rangle = \sum_i^n a_i \langle u \,|\, x_i \rangle = 0 \qquad (x_i \in S),$$

i.e. $u \in L(S)$. Since V is the topological closure of $L(S)$ (see Problem 13.4), every element v of V either belongs to $L(S)$ or is a limit of a convergent sequence of vectors $v_1, v_2, \dots \to v$, where $v_i \in L(S)$. By Lemma 13.3

$$\langle u \,|\, v_i \rangle = 0 \implies \langle u \,|\, v \rangle = 0.$$

Hence $u \in V^\perp$, so that $S^\perp \subseteq V^\perp$. Combined with the earlier result, we have that $S^\perp = V^\perp$.

Problem 13.11 Which of the following is a vector subspace of ℓ^2, and which are closed? In each case find the space of vectors orthogonal to the set.

(a) $V_N = \{(x_1, x_2, \dots) \in \ell^2 \mid x_i = 0$ **for** $i > N\}$.

(b)$V = \bigcup_{N=1}^{\infty} V_N = \{(x_1, x_2, \dots) \in \ell^2 \mid x_i = 0$ **for** $i >$ **some** $N\}$.

(c) $U = \{(x_1, x_2, \dots) \in \ell^2 \mid x_i = 0$ **for** $i = 2n\}$.

(d) $W = \{(x_1, x_2, \dots) \in \ell^2 \mid x_i = 0$ **for some** $i\}$.

Solution: (a) V_N is a the finite dimensional Hilbert subspace. It is therefore a closed vector subspace. Its orthogonal complement consists of vectors whose first$N$ components vanish

$$V_N^\perp = \{x = (0, 0, \dots, 0, x_{n+1}, x_{n+2}, \dots)\}.$$

(b) V_N is the space ℓ_0 discussed in Problem 13.5. It is a vector subspace but is not closed. Its orthogonal complement consists only of the zero vector, $V^\perp = \{0\}$, for if $x \in V^\perp$ then

$$\langle x \mid e_1 \rangle = \langle x \mid e_2 \rangle = \dots = \langle x \mid e_n \rangle = \dots = 0$$

where e_i is the vector all of whose components vanish except for $(e_i)_i = 1$ (i.e. $(x_i)_j = \delta_{ij}$). Since, $x = \sum_{i=1}^{\infty} x_i e_i$, we have $\|x\|^2 = \langle x \mid x \rangle = 0$, whence $x = 0$ by (IP3).

(c) U is clearly a closed vector subspace (since limits of sequences of vectors having even components vanish, must again have these components vanish.
The orthogonal complement consists vectors having odd components vanish,

$$U^\perp = \{(0, x_2, 0, x_4, 0 \dots) \in \ell^2\}.$$

(d) W is not a vector space, e.g. if $u = (0, \frac{1}{2}, \frac{1}{3}, \dots) \in W$ and $v = (1, 0, \frac{1}{3}, \frac{1}{4}, \dots) \in W$ then $u + v \notin W$ since it has no vanishing component. It is also not a closed set, for similar obvious reasons. Its orthogonal complement is $W^\perp = \{0\}$ by a similar argument to that given in part (b).

Problem 13.12 Show that the real Banach space $\mathbb{R}^2$ with the norm $\|(x, y)\| = \max\{|x|, |y|\}$ does not have the closest point property of Theorem 13.8. Namely for a given point x and one-dimensional subspace L, there does not in general exist a unique point in L that is closest to x.

Solution: Consider the one-dimensional subspace L consisting of the y-axis:

$$L = \{(0, y) \mid -\infty < y < \infty\}$$

and the point $\mathbf{x} = (1, 0)$. With respect to the maimum modulus norm, the distance between a point $\mathbf{y} = (0, y) \in L$ and $\mathbf{x}$ is

$$d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\| = \max(1, |y|).$$

Hence all points on the line L having $-1 \leq y \leq 1$ have the same distance 1 from $\mathbf{x}$ with respect to this norm, while points with $|y| > 1$ have distance greater than 1. Thus there are infinite number of points on L which are "closest" to $\mathbf{x}$ with respect to the maximum modulus norm.

Problem 13.13 **If $A : \mathcal{H} \to \mathcal{H}$ is an operator such that $Au \perp u$ for all $u \in \mathcal{H}$, show that $A = 0$.**

Solution: Let u and v be any two vectors in $\mathcal{H}$. Since $\langle Ax \,|\, x\rangle = 0$ for all $x \in \mathcal{H}$ we have

$$\begin{aligned} 0 &= \langle A(u+v) \,|\, u+v\rangle \\ &= \langle Au + Av \,|\, u+v\rangle \\ &= \langle Au \,|\, u\rangle + \langle Av \,|\, v\rangle + \langle Au \,|\, v\rangle + \langle Av \,|\, u\rangle \\ &= \langle Au \,|\, v\rangle + \langle Av \,|\, u\rangle \end{aligned}$$

Similarly setting $x = u + iv$ we have

$$\begin{aligned} 0 &= \langle Au + iAv \,|\, u + iv\rangle \\ &= \langle Au \,|\, u\rangle + \langle Av \,|\, v\rangle + i\langle Au \,|\, v\rangle - i\langle Av \,|\, u\rangle \\ &= i(\langle Au \,|\, v\rangle - \langle Av \,|\, u\rangle) \end{aligned}$$

Combining these two equations we have

$$\langle Au \,|\, v\rangle = 0$$

for all $u, v \in \mathcal{H}$. In particular, setting $v = Au$,

$$\|Au\|^2 = \langle Au \,|\, Au\rangle = 0$$

whence $Au = 0$ for all $u \in \mathcal{H}$, by (IP3). Thus the operator A is the zero operator, $A = 0$.

Problem 13.14 **The norm $\|\phi\|$ of a bounded linear operator $\phi : \mathcal{H} \to \mathbb{C}$ is defined as the greatest lower bound of all M such that $|\phi(u)| \leq M\|u\|$ for all $u \in \mathcal{H}$. If $\phi(u) = \langle v \,|\, u\rangle$ show that $\|\phi\| = \|v\|$. Hence show that the bounded linear functional norm satisfies the parallelogram law**

$$\|\phi + \psi\|^2 + \|\phi - \psi\|^2 = 2\|\phi\|^2 + 2\|\psi\|^2.$$

Solution: From the Cauchy-Schwarz inequality (5.13),

$$\langle v \,|\, u\rangle \leq \|v\|\|u\|.$$

Hence $\|\phi\| \leq \|v\|$. Furthermore $M = \|v\|$ is the least possible value such that $|\phi(u)| \leq M\|u\|$ for all $u \in \mathcal{H}$ since

$$\langle v \,|\, v\rangle = \|v\|^2 = \|v\|\|v\|.$$

Hence $\|v\|$ is the greatest lower bound of such M, so that $\|\phi\| = \|v\|$.

By the Riesz representation theorem, Theorem 13.10, we may set $\phi(u) = \phi_v(u) = \langle v \,|\, u\rangle$, $\psi = \phi_w$ for some pair of vectors v, w. Hence, as clearly $\phi + \psi = \phi_{v+w}$ and $\phi - \psi = \phi_{v-w}$, we have

$$\begin{aligned}\|\phi+\psi\|^2 + \|\phi-\psi\|^2 &= \|\phi_{v+w}\|^2 + \|\phi_{v-w}\|^2 \\ &= \|v+w\|^2 + \|v-w\|^2 \\ &= 2\|v\|^2 + 2\|w\|^2 \qquad \text{by the parallelogram law (13.2)} \\ &= 2\|\phi\|^2 + 2\|\psi\|^2.\end{aligned}$$

Problem 13.15 **If $\{e_n\}$ is a complete o.n. set in a Hilbert space $\mathcal{H}$, and α_n a bounded sequence of scalars, show that there exists a unique bounded operator A such that $Ae_n = \alpha_n e_n$. Find the norm of A.**

Solution: Let $M = \sup\{|\alpha_n| \,|\, n = 1, 2, 3, \dots\}$. For any $u = \sum_n u_n e_n$, we have by Parseval's identity (13.7),

$$\|u\|^2 = \sum_{n=1}^{\infty} |u_n|^2 < \infty.$$

The element

$$Au = \sum_{n=1}^{\infty} (\alpha_n u_n) e_n$$

belongs to $\mathcal{H}$ for $\|Au\|^2 \leq M^2\|u\|^2 < \infty$. Furthermore this inequality shows that the linear operator A so defined is bounded and therfore continuous. Hence, it is unique, for if $Ae_n = A'e_n = \alpha_n e_n$ for all n then$(A - A')e_n = 0$ for all n and

$$(A - A')\sum_{n=1}^{N} u_m e_m = 0$$

for arbitrary $u_n \in \mathbb{C}$ and integers $N > 0$. As every $u \in \mathcal{H}$ is a limit of such finite sums, and $A - A'$ is a continuous operator it follows that $(A - A')u = 0$ for all $u \in \mathcal{H}$. Hence $A = A'$ and the operator A defined above is unique.

The norm of A is $\|A\| = M$ for $\|Au\|^2 \leq M^2\|u\|^2$ by the above argument, so that $\|A\| \leq M$. Furthermore M is the least upper bound of all positive numbers with this property, for if there exists $\epsilon > 0$ such that $\|A\| < M - \epsilon$, then $\|Au\| < (M-\epsilon)\|u\|$ for all $u \in \mathcal{H}$. In particular, setting $u = e_n$ we have $|\alpha_n| < M - \epsilon$ for all $n = 1, 2, 3, \dots$, in contradiction to M being the least upper bound of $\{|\alpha_n| \,|\, n = 1, 2, \dots\}$.

Problem 13.16 **For bounded linear operators A, B on a normed vector space V show that**

$$\|\lambda A\| = |\lambda|\,\|A\|, \qquad \|A+B\| \le \|A\| + \|B\|, \qquad \|AB\| \le \|A\|\,\|B\|.$$

Hence show that $\|A\|$ is a genuine norm on the set of bounded linear operators on V.

Solution:

$$\begin{aligned}\|\lambda A\| &= \sup\bigl\{\|\lambda Au\| \bigm| \|u\| \le 1\bigr\}\\ &= \sup\bigl\{|\lambda|\|Au\| \bigm| \|u\| \le 1\bigr\} = |\lambda|\|A\|.\end{aligned}$$

For all $u\ in \mathcal{H}$ such that $\|u\| \le 1$

$$\|(A+B)u\| \le \|Au\| + \|Bu\| \le \|A\| + \|B\|,$$

hence

$$\|A+B\| = \sup\bigl\{\|(A+B)u\| \bigm| \|u\| \le 1\bigr\} \le \|A\| + \|B\|.$$

Similarly $\|AB\| \le \|A\|\|B\|$ since for all $\|u\| \le 1$

$$\|A(Bu)\| \le \|A\|\|Bu\| \le \|A\|\|B\|.$$

[Note that, since $\|Au\| \le \|A\|$ for all $\|u\| \le 1$, it follows for all $v \in \mathcal{H}$ that $\|Av\| = \|v\|\|A(v/\|v\|)\| \le \|v\|\|A\|$ since $\|(v/\|v\|)\| = 1$.]

Problem 13.17 **Prove properties (iii)–(v) of Theorem 13.13. Show that $\|A^*\| = \|A\|$.**

Solution: (iii) For arbitrary $u, v \in \mathcal{H}$

$$\langle (AB)^* u \,|\, v\rangle = \langle u \,|\, ABv\rangle = \langle A^* \,|\, Bv\rangle = \langle B^*A^*u \,|\, v\rangle$$

and since u and v are arbitrary, we have $(AB)^* = B^*A^*$.

(iv) For arbitrary $u, v \in \mathcal{H}$

$$\langle A^{**}u \,|\, v\rangle = \langle u \,|\, A^*v\rangle = \overline{\langle A^*v \,|\, u\rangle} = \overline{\langle v \,|\, Au\rangle} = \langle Au \,|\, v\rangle$$

hence $A^{**} = A$.

(v) If A is invertible, there exists a linear operator A^{-1} such that $AA^{-1} = I$, where $I = \mathrm{id}_{\mathcal{H}}$. Hence, using (iii) and $I^* = I$ (since $\langle u \,|\, Iv\rangle = \langle Iu \,|\, v\rangle = \langle u \,|\, v\rangle$),

$$I = I^* = (AA^{-1})^* = (A^{-1})^*A^*$$

so that $(A^*)^{-1} = (A^{-1})^*$.

From the proof that the definition of the adjoint operator in Eq. (13.9) defines a bounded operator, we have that $\|A^*u\| \le \|A\|\|u\|$. Hence, for all $\|u\| \le 1$ we see that $\|A^*u\| \le \|A\|$, giving $\|A^*\| \le \|A\|$. Using (iv) we then have $\|A\| = \|A^{**}\| \le \|A^*\|$, whence $\|A^*\| = \|A\|$.

Problem 13.18 **Let A be a bounded operator on a Hilbert space $\mathcal{H}$ with a one-dimensional range.**

(a) Show that there exist vectors u, v such that $Ax = \langle v \,|\, x\rangle u$ for all $x \in \mathcal{H}$.

(b) Show that $A^2 = \lambda A$ for some scalar λ, and that $\|A\| = \|u\|\|v\|$.

(c) Prove that A is hermitian, $A^* = A$, if and only if there exists a real number a such that $v = au$.

Solution: (a) If $Ax = \alpha u$ then

$$\langle u \,|\, Ax\rangle = \langle u \,|\, \alpha u\rangle = \alpha \|u\|^2$$

so that

$$\alpha = \frac{\langle u \,|\, Ax\rangle}{\|u\|^2} = \frac{\langle A^*u \,|\, x\rangle}{\|u\|^2} = \langle v \,|\, x\rangle$$

where

$$v = \frac{A^*u}{\|u\|^2}.$$

(b) We now have for arbitrary $x \in \mathcal{H}$

$$A^2x = \langle v \,|\, Ax\rangle u = \langle v \,|\, \langle v \,|\, x\rangle u\rangle u = \langle v \,|\, x\rangle\langle v \,|\, u\rangle u = \langle v \,|\, u\rangle Ax$$

so that $A^2 = \lambda A$ where $\lambda = \langle v \,|\, u\rangle$. Using the Cauchy-Schwarz inequality

$$\|Ax\|^2 = \langle Ax \,|\, Ax\rangle = \|u\|^2 |\langle v \,|\, x\rangle|^2 \le \|u\|^2\|v\|^2\|x\|^2.$$

Hence, if $\|x\|^2 \le 1$ then

$$\|Ax\|^2 \le \|u\|^2\|v\|^2$$

while equality is achieved if we set $x = v/\|v\|$. Thus $\|A\|^2 = \|u\|^2\|v\|^2$.

(c) If $A = A^*$ then

$$v = \frac{A^*u}{\|u\|^2} = \frac{Au}{\|u\|^2} = \frac{\langle v \,|\, u\rangle u}{\|u\|^2} = au$$

where

$$a = \frac{\langle v \,|\, u\rangle u}{\|u\|^2} = \frac{\langle A^*u \,|\, u\rangle u}{\|u\|^4} = \frac{\langle Au \,|\, u\rangle u}{\|u\|^4}$$

whence $a = \overline{a}$ since $\langle Au \,|\, u\rangle = \overline{\langle u \,|\, Au\rangle} = \overline{\langle A^*u \,|\, u\rangle}$.
Conversely if $v = au$ then for all $x, y \in \mathcal{H}$

$$\langle Ax \,|\, y\rangle = a\overline{\langle v \,|\, x\rangle}\langle u \,|\, y\rangle = a\langle x \,|\, u\rangle\langle u \,|\, y\rangle,$$

while

$$\langle A^* x \,|\, y \rangle = \langle x \,|\, Ay \rangle = \langle v \,|\, y \rangle \langle x \,|\, u \rangle = a \langle u \,|\, y \rangle \langle x \,|\, u \rangle = \langle Ax \,|\, y \rangle .$$

Hence $A^* x = Ax$ for all x, i.e. A is Hermitian, $A = A^*$.

Problem 13.19 For every bounded operator A on a Hilbert space $\mathcal{H}$ show that the exponential operator

$$\mathrm{e}^A = \sum_{n=0}^{\infty} \frac{A^n}{n!}$$

is well-defined and bounded on $\mathcal{H}$. Show that

(a) $\mathrm{e}^0 = I$.

(b) For all positive integers n, $(\mathrm{e}^A)^n = \mathrm{e}^{nA}$.

(c) e^A is invertible for all bounded operators A (even if A is not invertible) and $\mathrm{e}^{-A} = (\mathrm{e}^A)^{-1}$.

(d) If A and B are commuting operators then $\mathrm{e}^{A+B} = \mathrm{e}^A \mathrm{e}^B$.

(e) If A is hermitian then e^{iA} is unitary.

Solution: The proof that e^A is a bounded linear operator follows along similar lines to the proof of Theorem 13.12. Let x be any vector in $\mathcal{H}$. Since $\|A^k x\| \le \|A\| (A^{k-1} x)$ it follows by induction that A^k is bounded and has norm $\|A^k\| \le (\|A\|)^k$. The vectors

$$u_n = \big(I + A + \frac{A^2}{2!} + \dots + \frac{A^n}{n!}\big) x$$

form a Cauchy sequence, for

$$\begin{aligned} \|u_n - u_m\| &= \Big\| \big(\frac{A^{m+1}}{(m+1)!} + \dots + \frac{A^n}{n!}\big) x \Big\| \\ &\le \Big\| \frac{A^{m+1}}{(m+1)!} + \dots + \frac{A^n}{n!} \Big\| \|x\| \\ &\le \big(\frac{\|A\|^{m+1}}{(m+1)!} + \dots + \frac{\|A\|^n}{n!}\big) \|x\| \\ &\le \Big(\mathrm{e}^{\|A\|} - \big(I + \|A\| + \frac{\|A\|^2}{2!} + \dots + \frac{\|A\|^m}{m!}\big)\Big) \|x\| \\ &\longrightarrow 0 \quad \text{as } m \to \infty. \end{aligned}$$

Since $\mathcal{H}$ is complete, $u_m \to u$ for some $u \in \mathcal{H}$, so there is a linear operator $\mathrm{e}^A : \mathcal{H} \to \mathcal{H}$ such that $u = \mathrm{e}^A x$. We write $\mathrm{e}^A = \sum_{n=0}^{\infty} A^n / n!$, in the sense that

$$\lim_{N \to \infty} \Big(\mathrm{e}^A - \sum_{n=1}^{N} A^n\Big) x = 0$$

for all $x \in \mathcal{H}$.

(a) If $A = 0$ then $A^n = 0$ for all integers $n \geq 1$, hence

$$u_n = \big(I + 0 + \frac{0^2}{2!} + \cdots + \frac{0^n}{n!}\big)x = x = Ix$$

so that the limit $e^0 x = \lim_{n\to\infty} u_n = x = Ix$ for all $x \in \mathcal{H}$.

(b) and (c) are most easily proved by first showing part (d) (alternatively we can show them specifically in a completely analogous manner).

(d) Since all operators are bounded and therefore continous, there is no ambiguity in reorganizing the infinite series,

$$e^A e^B = \sum_{m=0}^{\infty} \frac{A^m}{m!} \sum_{k=0}^{\infty} \frac{B^k}{k!} = \sum_{p=0}^{\infty} \sum_{k=0}^{p} \frac{A^{p-k}}{(p-k)!} \frac{B^k}{k!}.$$

The p-th term in this sum is

$$C_p = \frac{A^p}{p!}\frac{B^0}{0!} + \frac{A^{p-1}}{(p-1)!}\frac{B^1}{1!} + \cdots + \frac{A^0}{0!}\frac{B^p}{p!}.$$

Since A and B commute, $AB = BA$, we have the binomial expansion

$$(A+B)^p = \sum_{k=0}^{p} \binom{p}{k} A^{p-k} B^k = p! C_p$$

whence

$$e^A e^B = \sum_{p=0}^{\infty} C_p = \sum_{p=0}^{\infty} \frac{(A+B)^p}{p!} = e^{A+B}.$$

Part (b) now follows by induction. For, the case $n = 0$ is essentially part (a), while if we assume $(e^A)^n = e^{nA}$ then set $B = A$, which evidently commutes with nA, we have by part (d)

$$\big(e^A\big)^{n+1} = \big(e^A\big)^n e^A = e^{nA} e^A = e^{nA+A} = e^{(n+1)A}.$$

To show (c), set $B = -A$, and we have

$$e^A e^{-A} = e^{A-A} = e^0 = I.$$

Hence $e^{-A} = (e^A)^{-1}$.

(e) From Theorem 13.13 (iii) it follows at once by induction that $(A^n)^* = (A^*)^n$. Hence

$$\big(e^A\big)^* = e^{A^*}.$$

Hence if A is Hermitian, $A = A^*$, then e^{iA} is unitary, for

$$e^{iA}\big(e^{iA}\big)^* = e^{iA} e^{(iA)^*} = e^{iA} e^{-iA^*} = e^{iA} e^{-iA} = I.$$

Problem 13.20 **Show that the sum of two projection operators $P_M + P_N$ is a projection operator iff $P_M P_N = 0$. Show that this condition is equivalent to $M \perp N$.**

Solution: $P_M + P_N$ is a projection operator iff

$$\begin{aligned}(P_M + P_N)^2 &= (P_M)^2 + (P_N)^2 + P_M P_N + P_N P_M \\ &= P_M + P_N + P_M P_N + P_N P_M \\ &= P_M + P_N\end{aligned}$$

i.e. iff

$$P_M P_N = -P_N P_M. \tag{$*$}$$

Multiplying this equation on the right and left by P_N gives (using $(P_N)^2 = P_N$)

$$P_N P_M P_N = -P_N P_M P_N$$

whence $P_N P_M P_M = 0$. Multiplying Eq. $(*)$ on the left by $(I - P_N)$ and using $(I - P_N)P_N = P_N - (P_N)^2 = 0$ we have

$$0 = -(I - P_N)P_M P_N = -P_M P_N + P_N P_M P_N = -P_M P_N.$$

Hence a necessary and sufficient condititon for $P_M + P_N$ to be a projection operator is that $P_M P_N = 0$. Similarly we must have $P_N P_M = 0$.

If M and N are orthogonal subspaces then for all $u \in \mathcal{H}$,

$$P_M(P_N u) = 0$$

for $P_N u \in N \subset M^\perp$. Hence $P_M P_N = 0$.
Conversely, if $P_M P_N u = 0$ for all $u \in \mathcal{H}$, then for all $v \in M$, $w \in N$ we have (using the fact that P_M is Hermitian)

$$\langle v \,|\, w\rangle = \langle P_M v \,|\, P_N w\rangle = \langle v \,|\, P_M^* P_N w\rangle = \langle v \,|\, P_M P_N w\rangle = \langle v \,|\, 0\rangle = 0.$$

Hence $M \perp N$.

Problem 13.21 **Verify that the operator on three-dimensional Hilbert space, having matrix representation in an o.n. basis**

$$\begin{pmatrix} \frac{1}{2} & 0 & \frac{i}{2} \\ 0 & 1 & 0 \\ -\frac{i}{2} & 0 & \frac{1}{2} \end{pmatrix}$$

is a projection operator, and find a basis of the subspace it projects onto.

Solution: The operator is clearly Hermitian, since its matrix satisfies $P^\dagger = P$. Furthermore it is idempotent

$$P^2 = \begin{pmatrix} \frac{1}{2} & 0 & \frac{i}{2} \\ 0 & 1 & 0 \\ -\frac{i}{2} & 0 & \frac{1}{2} \end{pmatrix}^2 = \begin{pmatrix} \frac{1}{2} & 0 & \frac{i}{2} \\ 0 & 1 & 0 \\ -\frac{i}{2} & 0 & \frac{1}{2} \end{pmatrix} = P.$$

We seek eigenvectors $u = xe_1 + ye_2 + ze_3$, such that $Pu = u$:

$$\begin{pmatrix} \frac{1}{2} & 0 & \frac{i}{2} \\ 0 & 1 & 0 \\ -\frac{i}{2} & 0 & \frac{1}{2} \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \frac{1}{2}x + \frac{i}{2}z \\ y \\ \frac{-i}{2}x + \frac{1}{2}z \end{pmatrix} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$

The general solution is $z = -ix$, so that the space M projected into by P is spanned by the orthonormal vectors

$$\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ -i \end{pmatrix} = \frac{1}{\sqrt{2}}(e_1 - ie_3), \qquad \text{and} \qquad \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = e_2.$$

Problem 13.22 **Let $\omega = \mathrm{e}^{2\pi i/3}$. Show that $1 + \omega + \omega^2 = 0$.**

(a) In Hilbert space of 3-dimensions let V be the subspace spanned by the vectors $(1, \omega, \omega^2)$ and $(1, \omega^2, \omega)$. Find the vector u_0 in this subspace that is closest to the vector $u = (1, -1, 1)$.

(b) Verify that $u - u_0$ is orthogonal to V.

(c) Find the matrix representing the projection operator P_V into the subspace V.

Solution: $\omega = \mathrm{e}^{4\pi i/3} = \mathrm{e}^{-2\pi i/3} = \overline{\omega}$. Hence

$$1 + \omega + \omega^2 = 1 + 2\cos\frac{2\pi}{3} = 1 - 2\cos\frac{\pi}{3} = 1 - 2.\frac{1}{2} = 0.$$

(a) Let $u_0 = \alpha(1, \omega, \omega^2) + \beta(1, \omega^2, \omega)$, then

$$\begin{aligned} \|u_0 - u\|^2 &= |\alpha + \beta - 1|^2 + |\alpha\omega + \beta\omega^2 + 1|^2 + |\alpha\omega^2 + \beta\omega - 1|^2 \\ &= |\alpha|^2 + |\beta|^2 + 1 + \alpha\overline{\beta} + \overline{\alpha}\beta - \alpha - \overline{\alpha} - \beta - \overline{\beta} \\ &\quad + |\alpha|^2 + |\beta|^2 + 1 + \alpha\overline{\beta}\omega\overline{\omega}^2 + \overline{\alpha}\beta\overline{\omega}\omega^2 + \alpha\omega + \overline{\alpha\omega} + \beta\omega^2 + \overline{\beta}\overline{\omega}^2 \\ &\quad + |\alpha|^2 + |\beta|^2 + 1 + \alpha\overline{\beta}\overline{\omega}\omega^2 + \overline{\alpha}\beta\omega\overline{\omega}^2 - \alpha\omega^2 - \overline{\alpha\omega}^2 - \beta\omega - \overline{\beta}\overline{\omega}. \end{aligned}$$

Using $\overline{\omega} = \omega^2$, $\overline{\omega}^2 = \omega$ and, from the above identity,

$$-1 + \omega - \omega^2 = 2\omega \qquad \text{etc.}$$

we have

$$\|u_0 - u\|^2 = 3 + 3|\alpha|^2 + 3|\beta|^2 + 2\alpha\omega + 2\overline{\alpha\omega} + 2\beta\overline{\omega} + 2\overline{\beta}\omega.$$

This has a minimum at

$$\frac{\partial}{\partial\alpha} = \frac{\partial}{\partial\overline{\alpha}} = \frac{\partial}{\partial\beta} = \frac{\partial}{\partial\overline{\beta}} = 0$$

(where α and $\overline{\alpha}$ etc. are taken as independent variables), i.e.

$$\begin{aligned} 3\overline{\alpha} + 2\omega &= 0 \\ 3\alpha + 2\overline{\omega} &= 0 \\ 3\overline{\beta} + 2\overline{\omega} &= 0 \\ 3\overline{\beta} + 2\omega &= 0 \end{aligned}$$

whence

$$\alpha = -\frac{2}{3}\overline{\omega} = -\frac{2}{3}\omega^2, \qquad \beta = -\frac{2}{3}\omega.$$

The vector closest to $u = (1, -1, 1)$ is thus

$$\begin{aligned} u_0 &= (-\tfrac{2}{3}\omega^2,\, -\tfrac{2}{3},\, -\tfrac{2}{3}\omega) + (-\tfrac{2}{3}\omega,\, -\tfrac{2}{3},\, -\tfrac{2}{3}\omega^2) \\ &= (\tfrac{2}{3},\, -\tfrac{4}{3},\, \tfrac{2}{3}). \end{aligned}$$

(b) Let

$$e_1 = \frac{1}{\sqrt{3}}\big(1, \omega, \omega^2\big), \qquad e_2 = \frac{1}{\sqrt{3}}\big(1, \omega^2, \omega\big)$$

so that $\langle e_i \mid e_j\rangle\delta_{ij}$ $(i, j = 1, 2)$. This is an orthonormal basis spanning V. Then $u - u_0 = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$ and we have

$$\begin{aligned} \langle u - u_0 \,|\, e_1\rangle &= \frac{1}{3\sqrt{3}}(1 + \omega + \omega^2) = 0, \\ \langle u - u_0 \,|\, e_2\rangle &= \frac{1}{3\sqrt{3}}(1 + \omega^2 + \omega) = 0, \end{aligned}$$

so that $u - u_0$ is orthogonal to V.

(c) The projection operator onto V is (in bra-ket notation)

$$\begin{aligned} P_V &= |e_1\rangle\langle e_1| + |e_2\rangle\langle e_2| \\ &= \frac{1}{3}\begin{pmatrix} 2 & \omega + \omega^2 & \omega + \omega^2 \\ \omega + \omega^2 & 2 & \omega^2 + \omega \\ \omega^2 + \omega & \omega^2 + \omega & 2 \end{pmatrix} \\ &= \begin{pmatrix} \frac{2}{3} & -\frac{1}{3} & -\frac{1}{3} \\ -\frac{1}{3} & \frac{2}{3} & -\frac{1}{3} \\ -\frac{1}{3} & -\frac{1}{3} & \frac{2}{3} \end{pmatrix}. \end{aligned}$$

Problem 13.23 **An operator A is called normal if it is bounded and commutes with its adjoint, $A^*A = AA^*$. Show that the operator**

$$A\psi(x) = c\psi(x) + i\int_a^b K(x,y)\psi(y)\mathrm{d}y$$

on $L^2([a,b])$, where c is a real number and $K(x,y) = \overline{K(y,x)}$, is normal.

(a) Show that an operator A is normal if and only if $\|Au\| = \|A^*u\|$ for all vectors $u \in \mathcal{H}$.

(b) Hence show that if A and B are commuting normal operators, AB and $A + \lambda B$ are normal for all $\lambda \in \mathbb{C}$.

Solution: The operator A is bounded since,

$$\begin{aligned}
\|A\psi\|^2 &= \int_a^b \overline{A\psi(x)}A\psi(x)\mathrm{d}x \\
&= \int_a^b \mathrm{d}x\Big(c^2\overline{\psi(x)}\psi(x) + ic\overline{\psi(x)}\int_a^b K(x,y)\psi(y)\mathrm{d}y - ic\psi(x)\int_a^b \overline{K(x,y)\psi(y)}\mathrm{d}y\Big) \\
&\quad + \int_a^b \mathrm{d}x\int_a^b dy\int_a^b \mathrm{d}z\overline{K(x,y)\psi(y)}K(x,z)\psi(z) \\
&= \int_a^b \mathrm{d}x\Big(c^2\overline{\psi(x)}\psi(x) + \int_a^b \mathrm{d}y\int_a^b \mathrm{d}z\overline{K(x,y)\psi(y)}K(x,z)\psi(z)\Big) \\
&\le \big(c^2 + \big|\int_a^b \mathrm{d}x\int_a^b \mathrm{d}y\int_a^b \mathrm{d}zK(y,x)K(x,z)\big|\big)\|\psi\|^2
\end{aligned}$$

Its adjoint is found form $\langle A^*\psi \,|\, \phi\rangle = \langle \psi \,|\, A\phi\rangle$, i.e.

$$\begin{aligned}
\int_a^b \overline{A^*\psi(x)}\phi(x)\mathrm{d}x &= \int_a^b \overline{\psi(x)}\big(c\phi(x) + i\int_a^b K(x,y)\phi(y)\mathrm{d}y\big)\mathrm{d}x \\
&= \int_a^b c\overline{\psi(x)}\phi(x) + i\int_a^b \mathrm{d}x\int_a^b \mathrm{d}yK(y,x)\overline{\psi(x)}\phi(x)
\end{aligned}$$

on interchanging the variables of integration x and y in the double integral. Hence

$$\overline{A^*\psi(x)} = c\overline{\psi(x)} + i\int_a^b K(y,x)\overline{\psi(y)},\mathrm{d}y$$

i.e.

$$\begin{aligned}
A^*\psi(x) &= c\psi(x) - i\int_a^b \overline{K(y,x)}\psi(y)\mathrm{d}y \\
&= c\psi(x) + i\int_a^b K(x,y)\psi(y)\mathrm{d}y = A\psi(x)
\end{aligned}$$

Hence A is Hermitian, $A = A^*$ and therefore normal since $AA^* = A^*A = A^2$.

(a) If A is normal then, using $A = A^{**}$,

$$\|Au\|^2 = \langle Au \,|\, Au\rangle = \langle u \,|\, A^*Au\rangle = \langle u \,|\, A^*Au\rangle = \langle A^*u \,|\, A^*u\rangle = \|A^*u\|^2.$$

Conversely, if $\|Au\|^2 = \|A^*u\|^2$ then we have for all $u \in \mathcal{H}$

$$\langle u \,|\, A^*Au\rangle = \langle u \,|\, AA^*u\rangle.$$

Replacing u by $v + w$ and $v + iw$ and expanding we find

$$\begin{aligned}\langle v \,|\, AA^*w\rangle + \langle w \,|\, AA^*v\rangle &= \langle v \,|\, A^*Aw\rangle + \langle w \,|\, A^*Av\rangle \\ i\langle v \,|\, AA^*w\rangle - i\langle w \,|\, AA^*v\rangle &= i\langle v \,|\, A^*Aw\rangle - i\langle w \,|\, A^*Av\rangle.\end{aligned}$$

and adding and subtracting these equations as in Problems 13.2 and 13.13, we have $\langle v \mid AA^*w\rangle = \langle v \mid A^*Aw\rangle$ for arbitrary vectors v and w. Hence A is normal, $AA^* = A^*A$.

(b) This result is much easier if we assume T commutes with $S*$,

$$TS^* = S^*T$$

which on taking the adjoint gives

$$ST^* = (TS^*)^* = (S^*T)^* = T^*S.$$

Then, since S and T are normal, $SS^* = S^*S$ and $TT^* = T^*T$,

$$\begin{aligned}(TS)(TS)^* = TSS^*T^* &= TS^*ST^* \\ &= S^*TT^*S = S^*T^*TS = (TS)^*TS\end{aligned}$$

and

$$\begin{aligned}(T + \lambda S)(T + \lambda S)^* &= (T + \lambda S)(T^* + \overline{\lambda}S^*) \\ &= TT^* + \lambda ST^* + \overline{\lambda}TS^* + \lambda\overline{\lambda}SS^* \\ &= T^*T + \lambda T^*S + \overline{\lambda}S^*T + \lambda\overline{\lambda}S^*S \\ &= (T^* + \overline{\lambda}S^*)(T + \lambda S) \\ &= (T + \lambda S)^*(T + \lambda S)\end{aligned}$$

The tricky part is to show that $TS^* = S^*T$ implies $TS = ST$. Any operator may be written $T = A + iB$ where A and B are hermitian ($A = A^*$, $B = B^*$), on setting $A = \frac{1}{2}(T + T^*)$ and $B = -\frac{1}{2}i(T - T^*)$. The operator T is normal iff A and B commute, $[A, B] = 0$, for

$$\begin{aligned}TT^* = T^*T \iff& (A + iB)(A - iB) = (A - iB)(A + iB) \\ \iff& 2i(BA - AB) = 0.\end{aligned}$$

Using the fact that any commuting pair of hermitian operators have a complete set of eigenvectors $\{e_i\}$ (see Section 13.5 for definition of complete set of eigenvectors and Chapter 14, Theorem 14.2) we have

$$Ae_i = a_ie_i, \quad Be_i = b_ie_i$$

where a_i and b_i are real numbers. Hence

$$Te_i = \alpha_i e_i, \qquad \alpha_i = a_i + ib_i,$$

i.e. we may assume every normal operator is diagonalizable. Set

$$Se_i = \sum_j S_{ij} e_j, \qquad S^* e_i = \sum_j \overline{S_{ij}} e_j$$

and we have

$$(TS^* - S^*T)e_i = \sum_j (\overline{S_{ji}}\alpha_j - \alpha_i \overline{S_{ji}})e_j = 0.$$

Hence

$$\overline{S_{ji}}(\alpha_j - \alpha_i) = 0$$

so that for each pair $i \neq j$

$$\overline{S_{ji}} = 0 \qquad \text{or} \qquad \alpha_j = \alpha_i.$$

This implies that

$$S_{ji} = 0 \qquad \text{or} \qquad \alpha_j = \alpha_i$$

which is equivalent to $TS = ST$. This proves the result.

Problem 13.24 **Show that a non-zero vector u is an eigenvector of an operator A if and only if $|\langle u \,|\, Au\rangle| = \|Au\|\|u\|$.**

Solution: If u is an eigenvector of, $Au = \lambda u$, then

$$\|Au\| = \sqrt{\langle Au \,|\, Au\rangle} = \sqrt{\lambda\overline{\lambda}\langle u \,|\, u\rangle} = |\lambda|\|u\|.$$

Hence

$$|\langle u \,|\, Au\rangle| = |\langle u \,|\, \lambda u\rangle| = |\lambda|\|u\|^2 = \|Au\|\|u\|.$$

By the Cauchy-Schwarz inequality (Theorem 5.4), we have

$$|\langle u \,|\, Au\rangle| \leq \|u\|\|Au\|$$

and Corollary 5.13 asserts that equality can only arise if Au is proportional to u, i.e. there exists $\lambda \in \mathbb{C}$ such that $Au = \lambda u$. Hence $|\langle u \,|\, Au\rangle| = \|Au\|\|u\| \Longrightarrow Au = \lambda u$.

Problem 13.25 **For any projection operator P_M show that every value $\lambda \neq 0, 1$ is a regular value, by showing that $(P_M - \lambda I)$ has a bounded inverse.**

Solution: To show that $P_M - \lambda I$ has an inverse it is required that for every vector $u \in \mathcal{H}$ there exists a vector v such that $u = (P_M - \lambda I)v$. Then

$$u' = P_M u = (1-\lambda)P_M v$$
$$u'' = (I - P_M)u = -\lambda(I - P_M)v.$$

Hence, if $\lambda \neq 0$ or 1,

$$v' = P_M v = \frac{1}{1-\lambda}u'$$
$$v'' = (I - P_M)v = -\frac{1}{\lambda}u''$$

Hence,

$$v = \frac{1}{1-\lambda}P_M u - \frac{1}{\lambda}(I - P_M)u$$

and it is simple to verify that $(P_M - \lambda I)v = u$. Thus

$$(P_M - \lambda I)^{-1} = \frac{1}{1-\lambda}P_M - \frac{1}{\lambda}(I - P_M).$$

To show $(P_M - \lambda I)^{-1}$ is bounded,

$$\begin{aligned}\|(P_M - \lambda I)^{-1}u\|^2 &= \langle (P_M - \lambda I)^{-1}u \,|\, (P_M - \lambda I)^{-1}u\rangle \\ &= \Big\langle \frac{u'}{1-\lambda} - \frac{u''}{\lambda} \Big| \frac{u'}{1-\lambda} - \frac{u''}{\lambda} \Big\rangle \\ &= \frac{1}{|1-\lambda|^2}\|u'\|^2 + \frac{1}{|\lambda|^2}\|u''\|^2.\end{aligned}$$

Since $u = u' + u''$ and $\langle u' \,|\, u''\rangle = 0$ we have

$$\|u\|^2 = \|u'\|^2 + \|u''\|^2,$$

whence $\|u'\|^2 \leq \|u\|^2$ and $\|u''\|^2 \leq \|u\|^2$. We conclude that

$$\|(P_M - \lambda I)^{-1}u\|^2 \leq \big(\frac{1}{|1-\lambda|^2} + \frac{1}{|\lambda|^2}\big)\|u\|^2,$$

so that $(P_M - \lambda I)^{-1}$ is a bounded linear operator on $\mathcal{H}$.

Problem 13.26 Show that every complex number λ in the spectrum of a unitary operator has $|\lambda| = 1$.

Solution: If U is a unitary operator, $U^*U = I$, all eigenvalues have absolute value 1, for

$$Uu = \lambda u \implies \|Uu\|^2 = \|u\|^2 = |\lambda|^2\|u\|^2.$$

Hence if $u \neq 0$ then $\|u\| \neq 0$ and $|\lambda|^2 = 1$, i.e. $\lambda = e^{i\theta}$ for a real angle θ.
To show that the spectrum lies entirely on the unit circle we must show that all $|\lambda| \neq 1$ are regular values, i.e. $(U - \lambda I)^{-1}$ exists and is bounded. The proof follows on similar lines to that of Theorem 13.18. Firstly, the operator $U - \lambda I$ is one-to-one for $|\lambda| \neq 1$, since there are no non-trivial solutions of the equation $(U - \lambda I)u = 0$. If $Uu - \lambda u = v$ and $|\lambda| < 1$ then, by the triangle inequality,

$$Uu = \lambda u + v \implies \|Uu\| = \|u\| \leq |\lambda|\|u\| + \|v\| \implies \|u\| \leq \frac{\|v\|}{1 - |\lambda|},$$

while if $|\lambda| > 1$

$$\lambda u = Uu - v \implies |\lambda|\|u\| \leq \|u\| + \|v\| \implies \|u\| \leq \frac{\|v\|}{|\lambda| - 1}.$$

Hence for all $|\lambda| \neq 1$,

$$\|u\| \leq \frac{\|v\|}{\big||\lambda| - 1\big|}. \qquad (*)$$

Let $V = \{U - \lambda I)u \,|\, u \in \mathcal{H}\}$. This vector subspace is closed for if $v_n = (U - \lambda I)u_n \in V$ is a convergent subsequence, $v_n \to v$ then is is a Cauchy sequence, and by the inequality $(*)$ it follows u_n (uniquely defined because $U - \lambda I$ is one-to-one) is also a Cauchy sequence and therefore converges, $u_n \to u$. Since $U - \lambda I$ is a continuous linear operator (since it is bounded), we have $(U - \lambda)u = v$ and therefore $v \in V$. Thus V is a closed subspace. Furthermore $V = \mathcal{H}$, for if not there exists a vector $w \perp V$, i.e.

$$0 = \langle (U - \lambda I)u \,|\, w\rangle = \langle u \,|\, (U^* - \overline{\lambda})w\rangle \qquad \text{for all } u \in \mathcal{H}.$$

Setting $u = (U^* - \overline{\lambda})w$ we have $(U^* - \overline{\lambda})w = 0$, which is impossible since $U^* = U^{-1}$ is also unitary and $|\overline{\lambda}| = |\lambda| \neq 1$. Thus $w = 0$, and we conclude $V = \mathcal{H}$, the operator $(U - \lambda I)$ is onto and one-to-one, i.e. invertible.

Problem 13.27 Prove that every hermitian operator A on a finite dimensional Hilbert space can be written as

$$A = \sum_{i=1}^{k} \lambda_i P_i \qquad \textbf{where} \qquad \sum_{i=1}^{k} P_i = I, \qquad P_i P_j = P_j P_i = \delta_{ij} P_i.$$

Solution: From Chapter 4 we have that every operator A on a finite dimension vector space $\mathcal{H}$ has an eigenvalue, being a solution of the polynomial equation $\det(A - \lambda I) = 0$. By Theorem 13.15 all eigenvalues of a Hermitian operator are real, displayed in ascending order

$$\lambda_1 < \lambda_2 < \cdots < \lambda_k,$$

and eigenvectors corresponding to different eigenvalues are orthogonal,

$$Au = \lambda_i u, \quad Av = \lambda_j v, \quad (i \neq j) \implies \langle u \,|\, v \rangle = 0.$$

Let M_i be the i-th eigenspace

$$M_i = \{u \in \mathcal{H} \,|\, Au = \lambda_i u\}$$

so that $M_i \perp M_j$ for $i \neq j$.
A Hermitian operator A on a finite dimensional Hilbert space is always complete, $\mathcal{H} = M_1 + M_2 + \cdots + M_k$, for if not let $W = (M_1 + M_2 + \cdots + M_k)^{\perp}$. The subspace W is invariant under A, for if $w \in W$ and $u \in M_1 + M_2 + \cdots + M_k$ then

$$\langle Aw \,|\, u \rangle = \langle w \,|\, Au \rangle = 0 \implies Aw \in W.$$

Hence the restricted operator A_W has an eigenvalue α with eigenvector $w \in W$, so that

$$Aw = A_W w = \alpha w.$$

But this would imply $\alpha = \lambda_i$ for some i, and gives the contradiction $w \in M_i$.
Let $P_k = P_{M_k}$ be the projection operator into the k-th subspace, then $P_i P_j = P_j P_i = 0$ for $i \neq j$ since the subspaces M_i and M_j are orthogonal to each other (see Problem 13.20). Since $(P_i)^2 = P_i$ we have the equation

$$P_i P_j = P_j P_i = \delta_{ij} P_i.$$

By Theorem 13.8 every $u \in \mathcal{H}$ has a unique decomposition $u = P_1 u + u'$ where $u' \in M_1^{\perp}$, and

$$u' = P_2 u' + u'' = P_2 u + u'' \qquad \text{where} \qquad u'' \in (M_1 + M_2)^{\perp} \qquad \text{etc.}$$

Continuing in this way we find every vector u has the decomposition

$$u = Iu = P_1 u + P_2 + P_3 u + \cdots + P_k u$$

so that

$$I = P_1 + P_2 + \cdots + P_k.$$

Since $P_i u \in M_i$ we have $AP_i u = \lambda_i P_i u$, whence

$$Au = A(P_1 u + P_2 u + \cdots + P_k u) = \lambda_1 P_1 u + \lambda_2 P_2 u + \cdots + \lambda_k P_k u,$$

giving

$$Au = \lambda_1 P_1 + \lambda_2 P_2 + \cdots + \lambda_k P_k.$$

Problem 13.28 **For any pair of hermitian operators A and B on a Hilbert space $\mathcal{H}$, define $A \leq B$ iff $\langle u | Au \rangle \leq \langle u | Bu \rangle$ for all $u \in \mathcal{H}$. Show that**

this is a partial order on the set of hermitian operators–pay particular attention to the symmetry property, $A \le B$ and $B \le A$ implies $A = B$.

(a) For multiplication operators on $L^2(X)$ show that $A_\alpha \le A_\beta$ iff $\alpha(x) \le \beta(x)$ a.e. on X.

(b) For projection operators show that the definition given here reduces to that given in the text, $P_M \le P_N$ iff $M \subseteq N$.

Solution: Since A and B are Hermitian operators, $\langle u \mid Au\rangle$ and $\langle u \mid Bu\rangle$ are real numbers, e.g.

$$\langle u \mid Au\rangle = \langle A^* u \mid u\rangle = \langle Au \mid u\rangle = \overline{\langle u \mid Au\rangle},$$

and can therefore be compared by $\le$. Referring to the definition of a partial order in Section 1.3, we must show that the defined relation is symmetric transitive and symmetric. Reflexitivity, $A \le A$ and transitivity $A \le B$, $B \le C \Longrightarrow A \le C$ are both trivial, but for symmetry we must argue as follows: If $A \le B$ and $B \le A$ then $\langle u \mid Au\rangle = \langle u \mid Bu\rangle$ for all $u \in \mathcal{H}$, i.e. setting $C = A - B$, $\langle u \mid Cu\rangle = 0$ for all $u \in \mathcal{H}$. replace u by $u + v$ and $u + iv$ respectively gives

$$\begin{aligned} \langle v \mid Cu\rangle + \langle u \mid Cv\rangle &= 0 \\ -\langle v \mid Cu\rangle + \langle u \mid Cv\rangle &= 0. \end{aligned}$$

Adding these two equations gives $\langle u \mid Cv\rangle$ for all $u, v \in \mathcal{H}$, which implies $C = 0$, or equivalently $A = B$.

(a) For $A_\alpha : f \mapsto \alpha g$ where α is a bounded measurable function on X, we have $A_\alpha \le A_\beta$ iff

$$\int_X \alpha |f|^2 \mathrm{d}\mu \le \int_X \beta |f|^2 \mathrm{d}\mu \qquad \text{for all } f \in L^2(X).$$

Suppose that the set $A = \{x \in X \mid \beta(x) < \alpha(x)\}$ has positive measure, $\mu(A) > 0$, then set $f = \chi_A$ and we have, using Problem 11.12 and Theorem 11.8,

$$\int (\alpha - \beta)|f|^2 \mathrm{d}\mu > 0,$$

in contradiction with the assumption. Hence $\alpha(x) \le \beta(x)$ a.e. on X.

(b) If P is a projection operator then any vector u can be decomposed $u = Pu + (I - P)u$, and we have

$$\langle u \mid Pu\rangle = \langle Pu + (I - P)u \mid Pu\rangle = \langle Pu \mid Pu\rangle = \|Pu\|^2,$$

and

$$\|u\|^2 = \langle Pu + (I - P)u \mid Pu + (I - P)u\rangle = \|Pu\|^2 + \|(I - P)u\|^2.$$

Thus $\|u\| \ge \|Pu\|$ with equality iff $u = Pu$.
Assume $P_M \le P_N$, i.e. $\langle u \mid P_M u\rangle \le \langle u \mid P_N u\rangle$ for all vectors u. If there exists a vector in M which does not belong to N then $P_M u = u$ and $P_N u \ne u$, hence

$$\langle u \mid P_M u\rangle = \|u\|^2 > \|P_N u\|^2 = \langle P_N u \mid P_N u\rangle = \langle u \mid (P_N)^2 u\rangle = \langle u \mid P_N u\rangle.$$

Hence we must have $M \subseteq N$.
Conversely if $M \subseteq N$ then for all $u \in \mathcal{H}$,

$$P_M u = P_M(P_N u + (I - P_N)u) = P_M P_N u$$

since $(I - P_N)u \in N^\perp \subset M^\perp$. Hence, for all vectors u,

$$\langle u \,|\, P_M u \rangle = \|P_M u\|^2 = \|P_M P_N u\|^2 \leq \|P_N u\|^2 = \langle u \,|\, P_N u \rangle,$$

i.e. $PM \leq P_N$.

Problem 13.29 For unbounded operators, show that

(a) $(AB)C = A(BC)$**.**

(b) $(A + B)C = AC + BC$**.**

(c) $AB + AC \subseteq A(B + C)$**. Give an example where** $A(B + C) \neq AB + AC$**.**

Solution: If A is an operator with with domain D_A and B has domain D_B, then $A + B$ is defined on domain $D_A \cap D_B$, and AB has domain $B^{-1}(R_B \cap D_A) = D_B \cap B^{-1}(D_A)$. Also note that for any operator A and subsets $U, V\mathcal{H}$, we have $A^{-1}(U \cap V) = A^{-1}(U) \cap A^{-1}(V)$ and $(AB)^{-1}(U) = B^{-1}(A^{-1}(U))$.

(a) On the common domains of these operators it is clear by associativity of functions in general that $(AB)Cx = A(BC)x$. It is therefore only necessary to show that the domains are equal:

$$\begin{aligned} D_{(AB)C} &= D_C \cap C^{-1}(D_{AB}) \\ &= D_C \cap C^{-1}(D_B \cap B^{-1}(D_A)) \\ &= D_C \cap C^{-1}(D_B) \cap C^{-1}(B^{-1}(D_A)). \end{aligned}$$

and

$$\begin{aligned} D_{A(BC)} &= D_{BC} \cap (BC)^{-1}(D_A) \\ &= D_C \cap C^{-1}(D_B) \cap C^{-1}(B^{-1}(D_A)) = D_{(AB)C}. \end{aligned}$$

(b) Since $D_{(A+B)C} = D_C \cap C^{-1}(D_A \cap D_B)$ we have

$$\begin{aligned} D_{AC+BC} &= D_{AC} \cap D_{BC} \\ &= D_C \cap C^{-1}(D_A) \cap D_C \cap C^{-1}(D_B) \\ &= D_C \cap C^{-1}(D_A \cap D_B) = D_{(A+B)C}, \end{aligned}$$

and for all x in this common domain, it is clear that $(A + B)Cx = ACx + BCx$.

(c) Since D_A is a vector space, then for any $x \in D_B \cap D_C$ such that $Bx \in D_A$ and $Cx \in D_A$ we evidently have $Bx + Cx = (B + C)x \in D_A$, i.e. $x \in (B + C)^{-1}(D_A)$.

Hence $B^{-1}(D_A) \cap C^{-1}(D_A) \subset (B+C)^{-1}(D_A)$, and it follows that

$$\begin{aligned} D_{AB+AC} &= D_{AB} \cap D_{AC} \\ &= D_B \cap B^{-1}(D_A) \cap D_C \cap C^{-1}(D_A) \\ &= D_B \cap D_C \cap B^{-1}(D_A) \cap C^{-1}(D_A) \\ &\subseteq D_{B+C} \cap (B+C)^{-1}(D_A) = D_{A(B+C)}. \end{aligned}$$

Hence $AB + AC \subseteq A(B+C)$.

To find an example where this is a strict set inclusion, let A be any unbounded operator on ℓ^2 such as the operator A^{-1} given in Example 13.16. This operator has $Ae_n = ne_n$, and $D_A = \{y \mid \sum_{n=1}^{\infty} n^2 |y_n|^2 < \infty\}$. Let B be the identity operator on ℓ^2, $B = I$ and $C = -I$. Then $B + C = 0$, and since $0 \in D_A$ it is clear that $(B+C)^{-1}(D_A) = \mathcal{H}$. Hence

$$D_{A(B+C)} = D_{B+C} \cap (B+C)^{-1}(D_A) = \mathcal{H} \cap \mathcal{H} = \mathcal{H}$$

while

$$D_{AB+AC} = D_{AB} \cap D_{AC} = D_A \cap D_A = D_A \neq \mathcal{H}.$$

Problem 13.30 Show that a densely defined bounded operator A in $\mathcal{H}$ has a unique extension to an operator $\hat{A}$ defined on all of $\mathcal{H}$. Show that $\|\hat{A}\| = \|A\|$.

Solution: Let D_A be the domain of the bounded operator A. Boundedness means that there exists $K > 0$ such that $\|Ax\| \leq K\|x\|$ for all $x \in D_A$. Let u be any vector in $\mathcal{H}$, and $u_n \in D_A$ a sequence such that $u_n \to u$ (exists since D_A is a dense domain in $\mathcal{H}$). Since u_n is a Cauchy sequence, it follows that so is Au_n,

$$\|Au_n - Au_m\| < K\|U_n - u_m\| \to 0.$$

Hence, by completeness of $\mathcal{H}$, Au_n converges to a vector $v \in \mathcal{H}$. Set $\hat{A}u = v$. This definition is independent of the choice of sequence approaching u, for if $u'_n \in D_A$ is any other sequence , $u'_n \to u$, then

$$\|Au_n - Au'_n\| \leq K\|u_n - u'_n\| \leq K\big(\|u_n - u\| + \|u'_n - u\|\big) \to 0 + 0 = 0.$$

Hence $Au'_n - Au_n \to 0$, and $\lim_{n\to\infty} Au'_n = \lim_{n\to\infty} Au_n$. Furthermore, if $u \in D_A$, take $u_n = u$ for all $n = 1, 2, \dots$, and we clearly have $\hat{A}u = Au$, so that $\hat{A}$ is an extension of A. The operator $\hat{A} : \mathcal{H} \to \mathcal{H}$ is bounded since, on using Lemma 13.5,

$$\|Au_n\| \leq K\|u_n\| \Longrightarrow \|\hat{u}\| = \lim\|Au_n\| \leq K \lim\|u_n\| = K\|u\|.$$

Uniqueness follows from the fact that every bounded operator on H is continuous, so the limit of Au_n is unique.

The norm of A is defined as the least upper bound of all K such that $\|Au\| \le K$ for all $u \in D_A$ with $\|u\| \le 1$. Consider any $u \in \mathcal{H}$ such that $\|u\| \le 1$. We may take a sequence $u_n \to u$ in D_A such that all $\|u_n\| = \|u\|$: e.g. for any sequence $u_n \to u$ set

$$u'_n = \frac{u_n}{\|u_n\|}\|u\| \;\to\; \frac{u}{\|u\|}\|u\| \to u$$

and clearly $\|u'_n\| = \|u\|$ for all n. Thus

$$\|\hat{A}u\| = \lim Au_n \le \|A\|\|u_n\| = \|A\|\|u\| \le \|A\|.$$

Therefore $\|\hat{A}\| \le \|A\|$. On the other hand it is impossible that $\|\hat{A}\| < \|A\|$, else for all $u \in D_A \subseteq \mathcal{H}$ with $\|u\| \le 1$ we have $\|\hat{A}u\| = \|Au\| \le \|\hat{A}\|$. This contradicts $\|A\|$ being the least upper bound of all values of $\|Au\|$ with $\in D_A$ and $\|u\| \le 1$. Hence $\|\hat{A}\| = \|A\|$.

Problem 13.31 If A is self-adjoint and B a bounded operator, show that B^*AB is self-adjoint.

Solution: Since A is self-adjoint we have $D_A = D_{A^*}$ and $\langle u \,|\, Av\rangle = \langle Au \,|\, v\rangle$ for all $u, v \in D_A$. Now since $D_B = \mathcal{H}$,

$$\begin{aligned} D_{B^*AB} &= D_B \cap B^{-1}(D_{B^*A}) \\ &= \mathcal{H} \cap B^{-1}(D_A \cap A^{-1}(D_{B^*})) \\ &= \cap B^{-1}(D_A) \cap B^{-1}(A^{-1}(\mathcal{H})) \\ &= \cap B^{-1}(D_A) \cap \cap B^{-1}(D_A) \\ &= \cap B^{-1}(D_A). \end{aligned}$$

Hence

$$\begin{aligned} u \in D_{B^*AB} &\iff Bu \in D_A = D_{A^*} \\ &\iff \langle Bu \,|\, Aw\rangle = \langle ABu \,|\, w\rangle \qquad \forall w \in D_A \\ &\iff \langle Bu \,|\, ABv\rangle = \langle ABu \,|\, Bv\rangle \qquad \forall v \in B^{-1}(D_A) \\ &\iff \langle u \,|\, B^*ABv\rangle = \langle B^*ABu \,|\, v\rangle \qquad \forall v \in D_{B^*AB} \\ &\iff u \in D_{(B^*AB)^*} \end{aligned}$$

We have then for all $u \in D_{B^*AB} = D_{(B^*AB)^*}$,

$$(B^*AB)^*u = B^*ABu$$

giving the desired result.

Problem 13.32 Show that if (A, D_A) and (B, D_B) are operators on dense domains in $\mathcal{H}$ then $B^*A^* \subseteq (AB)^*$.

Solution: Let $x \in D_{B^*A^*} = D_{A^*} \cap (A^*)^{-1}(D_{B^*})$. Then $x \in D_{A^*}$ and $A^*x \in D_{B^*}$. Hence for all $y \in D_B$

$$\langle B^*A^*x \,|\, y\rangle = \langle A^*x \,|\, By\rangle$$

In particular, for all $y \in D_{AB} = D_B \cap B^{-1}(D_A)$ we have $By \in D_A$, and

$$\langle x \,|\, ABy\rangle = \langle A^*x \,|\, By\rangle = \langle B^*A^*x \,|\, y\rangle$$

whence $x \in D_{(AB)^*}$ and

$$\langle (AB)^*x \,|\, y\rangle = \langle B^*A^*x \,|\, y\rangle.$$

Thus, $D_{B^*A^*} \subseteq D_{(AB)^*}$ and $(AB)^*x = B^*A^*x$ for all $x \in D_{B^*A^*}$; i.e. $B^*A^* \subseteq (AB)^*$.

Problem 13.33 For unbounded operators, show that $A^*+B^* \subseteq (A+B)^*$.

Solution: For all $x \in D_{A^*+B^*} = D_{A^*} \cap D_{B^*}$ and $y \in D_{A+B} = D_A \cap D_B$ we have

$$\langle x \,|\, Ay\rangle = \langle A^*x \,|\, y\rangle, \qquad \langle x \,|\, By\rangle = \langle B^*x \,|\, y\rangle,$$

since $x \in D_{A^*}$, $y \in D_A$ and $x \in D_{B^*}$, $y \in D_B$. Hence

$$\begin{aligned}\langle x \,|\, (A+B)y\rangle &= \langle x \,|\, Ay\rangle + \langle x \,|\, By\rangle \\ &= \langle A^*x \,|\, y\rangle + \langle B^*x \,|\, y\rangle \\ &= \langle A^*x + B^*x \,|\, y\rangle.\end{aligned}$$

Hence $x \in D_{(A+B)^*}$, i.e. $D_{A^*+B^*} \subseteq D_{(A+B)^*}$, and

$$(A^* + B^*)x = A^*x + B^*x = (A+B)^*x$$

for all $x \in D_{A^*+B^*}$. This is the required conclusion $A^* + B^* \subseteq (A+B)^*$.

Problem 13.34 If (A, D_A) is a densely defined operator and D_{A^*} is dense in $\mathcal{H}$, show that $A \subseteq A^{**}$.

Solution: If $v \in D_A$ and $u \in D_{A^*}$ then

$$\langle u \,|\, Av\rangle = \langle A^*u \,|\, v\rangle.$$

Using (IP1),

$$\overline{\langle Av \,|\, u\rangle} = \overline{\langle v \,|\, A^*u\rangle}$$

and taking complex conjugate

$$\langle v \,|\, A^*u\rangle = \langle Av \,|\, u\rangle.$$

Hence $v \in D_{A^{**}}$, i.e. $D_A \subseteq D_{A^{**}}$, and $A^{**}v = Av$ for all $v \in D_A$. In summary, $A \subseteq A^{**}$.

Problem 13.35 **If A is a symmetric operator, show that A^* is symmetric if and only if it is self-adjoint, $A^* = A^{**}$.**

Solution: Let A be a symmetric operator,

$$\langle Au \,|\, v\rangle = \langle u \,|\, Av\rangle \qquad \forall u, v \in D_A.$$

From Theorem 13.23 we have $A \subset A^*$.
If A^* is self-adjoint, $A^* = A^{**}$, then it is symmetric by Theorem 13.23, since $A^* \subseteq A^{**}$.
Conversely, if A^* is symmetric,

$$\langle A^*u \,|\, v\rangle = \langle u \,|\, A^*v\rangle \qquad \forall u, v \in D_{A^*}$$

then by Theorem 13.23 that $A^* \subset A^{**}$. Setting $B = A^*$ in Theorem 13.22, we have $A \subset A^* \Longrightarrow A^{**} \subset A^*$. Combining these two, $A^* = A^{**}$, A is self-adjoint.

Problem 13.36 **If A_1, A_2, $\dots$, A_n are operators on a dense domain such that**

$$\sum_{i=1}^{n} A_i^* A_i = 0,$$

show that $A_1 = A_2 = \cdots = A_n = 0$.

Solution: Let $B = \sum_{i=1}^{n} A_i^* A_i = 0$ on a dense domain D_B. For any operator A on a dense domain D_A and $u \in D_{A^*A} = D_A \cap (A^*)^{-1}(D_A)$ we have $u \in D_A \subset D_{A^{**}}$ and $A^{**}u = Au$ (see Problem 13.34). Hence

$$\langle u \,|\, A^*Au\rangle = \langle A^{**} \,|\, Au\rangle = \langle Au \,|\, Au\rangle = \|Au\|^2.$$

Hence, for all $u \in D_B$,

$$\langle u \,|\, Bu\rangle = \sum_{i=1}^{n} \|A_i u\|^2 = 0$$

whence, since all summands are non-negative, $\|A_i u\|^2 = 0$ for all $i = 1, 2, \dots, n$. Thus $A_i u = 0$ for all $u \in D_B$, and since D_B is a dense domain, we have $A_i = 0$ on $\mathcal{H}$.

Problem 13.37 **If A is a self-adjoint operator show that**

$$\|(A + iI)u\|^2 = \|Au\|^2 + \|u\|^2$$

and that the operator $A + iI$ is invertible. show that the operator $U = (A - iI)(A + iI)^{-1}$ is unitary (called the *Cayley transform* of A).

Solution:

$$\begin{aligned}\|(A+iI)u\|^2 &= \langle A+iI)u \,|\, A+iI)u\rangle \\ &= \langle Au \,|\, Au\rangle - i\langle u \,|\, Au\rangle + i\langle Au \,|\, u\rangle + (-i)i\langle u \,|\, u\rangle \\ &= \|Au\|^2 + \|u\|^2\end{aligned}$$

The proof that $A + iI$ is invertible follows along similar lines to that of Theorem 13.18. Let $V = \{(A+iI)u \,|\, u \in D_A\}$, a vector subspace of $\mathcal{H}$. It is in fact a closed subspace, for if $v_n = (A+iI)u_n$ is any convergent sequence of vectors in V, $v_n \to v$, then it is a Cauchy sequence, $\|v_m - v_n\| \to 0$ as $m, n \to \infty$. Hence, by the above identity,

$$\|Au_m - Au_n\|^2 + \|u_m - u_n\|^2 = \|v_m - v_n\|^2 \to 0$$

and it follows that both u_n and Au_n are Cauchy sequences, $\|u_m - u_n\|^2 \to 0$, $\|Au_m - Au_n\|^2 \to 0$, so that there are limits $u_n \to u$, $Au_n \to u' = v - iu$. Since A is self-adjoint it is a closed operator (use $A = A^*$ and Theorem 13.21), hence $u' = Au$, and $v = (A+iI)u$. Hence V is a closed subspace of $\mathcal{H}$.
In fact, $V = \mathcal{H}$, for if $w \perp V$, then

$$0 = \langle (A+iI)u \,|\, w\rangle = iu(A - iI)w \qquad \forall u \in D_A.$$

As D_A is a dense domain, $(A - iI)w = 0$, and by the argument for the identity at the beginning of this problem,

$$\|Aw\|^2 + \|w\|^2 = 0$$

so that $w = 0$. Hence $V = \mathcal{H}$, and the operator $A + iI$ is invertible—for any $v \in \mathcal{H}$ there exists a unique $u \in D_A$ such that $(A+iI)u = v$ (uniquencess follows from the fact that $\ker(A+iI) = \{0\}$.

For any pair of vectors $u, v \in \mathcal{H}$,

$$\begin{aligned}\langle Uu \,|\, Uv\rangle &= \langle (A-iI)(A+iI)^{-1}u \,|\, (A-iI)(A+iI)^{-1}v\rangle \\ &= \langle (A+iI)^{-1}u \,|\, (A+iI)(A-iI)(A+iI)^{-1}v\rangle \\ &= \langle (A+iI)^{-1}u \,|\, (A-iI)(A+iI)(A+iI)^{-1}v\rangle \\ &= \langle (A+iI)^{-1}u \,|\, (A-iI)v\rangle \\ &= \langle (A+iI)(A+iI)^{-1}u \,|\, v\rangle \\ &= \langle u \,|\, v\rangle.\end{aligned}$$

Thus U is a unitary transformation of $\mathcal{H}$.

Chapter 14

Problem 14.1 Verify for each direction

$$\mathbf{n} = \sin\theta\cos\phi\mathbf{e}_x + \sin\theta\sin\phi\mathbf{e}_y + \cos\theta\mathbf{e}_z$$

the spin operator

$$\sigma_{\mathbf{n}} = \begin{pmatrix} \cos\theta & \sin\theta e^{-i\phi} \\ \sin\theta e^{i\phi} & -\cos\theta \end{pmatrix}$$

has eigenvalues ± 1. Show that up to phase, the eigenvectors can be expressed as

$$|+\mathbf{n}\rangle = \begin{pmatrix} \cos\frac{1}{2}\theta e^{-i\phi} \\ \sin\frac{1}{2}\theta \end{pmatrix}, \qquad |-\mathbf{n}\rangle = \begin{pmatrix} -\sin\frac{1}{2}\theta e^{-i\phi} \\ \cos\frac{1}{2}\theta \end{pmatrix}$$

and compute the expectation values for spin in the direction of the various axes

$$\langle\sigma_i\rangle_{\pm\mathbf{n}} = \langle\pm\mathbf{n}|\sigma_i|\pm\mathbf{n}\rangle.$$

For a beam of particles in a pure state $|+\mathbf{n}\rangle$ show that after a measurement of spin in the $+x$ direction the probability that the spin is in this direction is $\frac{1}{2}(1+\sin\theta\cos\phi)$.

Solution: Eigenvalues λ are solutions of the equation

$$\det(\sigma_{\mathbf{n}} - \lambda I) = (\cos\theta - \lambda)(-\cos\theta - \lambda) - \sin^2\theta = \lambda^2 - 1 = 0,$$

i.e. $\lambda = \pm 1$.

The $+1$ eigenvector,

$$|+\mathbf{n}\rangle = \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix}, \qquad \sigma_{\mathbf{n}}|+\mathbf{n}\rangle = |+\mathbf{n}\rangle$$

has

$$\psi_1 = \cos\theta\psi_1 + \sin\theta e^{-i\phi}\psi_2$$

i.e.

$$\psi_1 = \frac{\sin\theta e^{-i\phi}}{1-\cos\theta}\psi_2 = \frac{2\sin\frac{1}{2}\theta\cos\frac{1}{2}\theta e^{-i\phi}}{2\sin^2\frac{1}{2}\theta}\psi_2 = \cot\tfrac{1}{2}\theta e^{-i\phi}\psi_2.$$

Normalizing,

$$|\psi_1|^2 + |\psi_2|^2 = |\psi_2|^2(1+\cot^2\tfrac{1}{2}\theta) = \frac{|\psi_2|^2}{\sin^2\frac{1}{2}\theta} = 1$$

gives, up to an arbitrary phase factor,

$$\psi_2 = \sin\tfrac{1}{2}\theta, \qquad \psi_1 = \cos\tfrac{1}{2}\theta e^{-i\phi}.$$

For the -1 eigenvector, $\sigma_{\mathbf{n}}|-\mathbf{n}\rangle = -|-\mathbf{n}\rangle$, we have

$$\psi_1 = \frac{-\sin\theta e^{-i\phi}}{1+\cos\theta}\psi_2 = -\tan\tfrac{1}{2}\theta e^{-i\phi}\psi_2,$$

and normalizing gives (up to a phase)

$$\psi_1 = -\sin\tfrac{1}{2}\theta e^{-i\phi}, \qquad \psi_2 = \cos\tfrac{1}{2}\theta.$$

The expection direction in the x-direction in the $|+\mathbf{n}\rangle$ direction is

$$\begin{aligned}
\langle\sigma_1\rangle_{+\mathbf{n}} &= \langle+\mathbf{n}|\sigma_1|+\mathbf{n}\rangle \\
&= \begin{pmatrix}\cos\tfrac{1}{2}\theta e^{i\phi} & \sin\tfrac{1}{2}\theta\end{pmatrix}\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}\begin{pmatrix}\cos\tfrac{1}{2}\theta e^{-i\phi}\\ \sin\tfrac{1}{2}\theta\end{pmatrix} \\
&= \cos\tfrac{1}{2}\theta\sin\tfrac{1}{2}\theta e^{i\phi} + \sin\tfrac{1}{2}\theta\cos\tfrac{1}{2}\theta e^{-i\phi} \\
&= \sin\theta\cos\phi.
\end{aligned}$$

Similarly

$$\begin{aligned}
\langle\sigma_2\rangle_{+\mathbf{n}} &= \langle+\mathbf{n}|\sigma_2|+\mathbf{n}\rangle \\
&= i(-\cos\tfrac{1}{2}\theta\sin\tfrac{1}{2}\theta e^{i\phi} + \sin\tfrac{1}{2}\theta\cos\tfrac{1}{2}\theta e^{-i\phi}) \\
&= \sin\theta\sin\phi,
\end{aligned}$$

and

$$\langle\sigma_3\rangle_{+\mathbf{n}} = \langle+\mathbf{n}|\sigma_3|+\mathbf{n}\rangle = \cos^2\tfrac{1}{2}\theta - \sin^2\tfrac{1}{2}\theta = \cos\theta.$$

The expection values in the $|-\mathbf{n}\rangle$ direction are

$$\begin{aligned}
\langle\sigma_1\rangle_{-\mathbf{n}} &= \langle-\mathbf{n}|\sigma_1|-\mathbf{n}\rangle = -\sin\theta\cos\phi, \\
\langle\sigma_2\rangle_{-\mathbf{n}} &= \langle-\mathbf{n}|\sigma_2|-\mathbf{n}\rangle = -\sin\theta\sin\phi, \\
\langle\sigma_3\rangle_{-\mathbf{n}} &= \langle-\mathbf{n}|\sigma_3|-\mathbf{n}\rangle = -\cos\theta.
\end{aligned}$$

For a beam of particles in a state $|+\mathbf{n}\rangle$ the probability P that the spin is in the x-direction after a measurement of spin in the $+x$ direction is

$$P = |\langle\mathbf{e}_1|+\mathbf{n}\rangle|^2$$

where (setting $\theta = \pi/2, \phi = 0$ for the $+x$ direction)

$$|\mathbf{e}_1\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix}, \qquad |+\mathbf{n}\rangle = \begin{pmatrix}\cos\tfrac{1}{2}\theta e^{-i\phi}\\ \sin\tfrac{1}{2}\theta\end{pmatrix}.$$

Hence

$$\begin{aligned}
P &= \left|\frac{1}{\sqrt{2}}\cos\tfrac{1}{2}\theta e^{-i\phi} + \frac{1}{\sqrt{2}}\sin\tfrac{1}{2}\theta\right|^2 \\
&= \frac{1}{2}\left(\cos^2\tfrac{1}{2}\theta + \sin^2\tfrac{1}{2}\theta + \sin\tfrac{1}{2}\theta\cos\tfrac{1}{2}\theta(e^{-i\phi} + e^{i\phi}\right) \\
&= \frac{1}{2}(1 + \sin\theta\cos\phi)
\end{aligned}$$

Problem 14.2 **If A and B are vector observables that commute with the Pauli spin matrices, $[\sigma_i, A_j] = [\sigma_i, B_j] = 0$ (but $[A_i, B_j] \neq 0$ in general) show that**

$$(\sigma \cdot \mathbf{A})(\sigma \cdot \mathbf{B}) = \mathbf{A} \cdot \mathbf{B} + i(\mathbf{A} \times \mathbf{B}) \cdot \sigma$$

where $\sigma = (\sigma_1, \sigma_2, \sigma_3)$.

Solution: A straightforward computation shows that

$$\begin{aligned}
\sigma_1\sigma_2 &= -\sigma_2\sigma_1 = i\sigma_3 \\
\sigma_2\sigma_3 &= -\sigma_3\sigma_2 = i\sigma_1 \\
\sigma_3\sigma_1 &= -\sigma_1\sigma_3 = i\sigma_2 \\
(\sigma_1)^2 &= (\sigma_2)^2 = (\sigma_3)^2 = I
\end{aligned}$$

which may be summarized in the single equation

$$\sigma_i\sigma_j = \delta_{ij} I + i\epsilon_{ijk}\sigma_k.$$

Hence, using the Cartesian summation convention,

$$\begin{aligned}
(\sigma \cdot \mathbf{A})(\sigma \cdot \mathbf{B}) &= \sigma_i A_i \sigma_j B_j \\
&= \sigma_i \sigma_j A_i B_j \\
&= (\delta_{ij} I + i\epsilon_{ijk}\sigma_k) A_i B_j \\
&= A_i B_i + i\epsilon_{ijk} A_i \sigma_k B_j \\
&= A_i B_i + i\epsilon_{ijk} A_i B_j \sigma_k \\
&= \mathbf{A} \cdot \mathbf{B} + i(\mathbf{A} \times \mathbf{B}) \cdot \sigma
\end{aligned}$$

Problem 14.3 **Prove the following commutator identities:**

$$[A, [B, C]] + [B, [C, A]] + [C, [A, B]] = 0 \qquad (\textbf{Jacobi identity})$$

$$\begin{aligned}
[AB, C] &= A[B, C] + [A, C]B \\
[A, BC] &= [A, B]C + B[A, C]
\end{aligned}$$

Solution:

$$\begin{aligned}
[A, [B, C]] + [B, [C, A]] + [C, [A, B]] = A(BC - CB) &- (BC - CB)A \\
+ B(CA - AC) &- (CA - AC)B \\
+ C(AB - BA) &- (AB - BA)C \\
= 0 &
\end{aligned}$$

as all terms cancel in pairs.

$$\begin{aligned}
[AB, C] = ABC - CAB &= A(BC - CB) + ACB - CAB \\
&= A[B, C] + [A, C]B \\
[A, BC] = ABC - BCA &= (AB - BA)C + BAC - BCA \\
&= [A, B]C + B[A, C].
\end{aligned}$$

Problem 14.4 Using the identites of Problem 14.3 show the following identities:

$$\begin{aligned}
[Q^n, P] &= ni\hbar Q^{n-1}, \\
[Q, P^m] &= mi\hbar P^{m-1}, \\
[Q^n, P^2] &= 2ni\hbar Q^{n-1}P + n(n-1)\hbar^2 Q^{n-2}, \\
[L_m, Q_k] &= i\hbar\epsilon_{mkj}Q_j, \qquad [L_m, P_k] = i\hbar\epsilon_{mkj}P_j,
\end{aligned}$$

where $L_m = \epsilon_{mij}Q_iP_j$ are the angular momentum operators.

Solution: The first three identities are all assumed to be one-dimensional. We prove the first and second identities by induction. For $n = 1$ and $m = 1$ both reduce to Eq. (14.7), $[Q, P] = i\hbar I$. Assume the first identity holds for $n - 1$, i.e. assume

$$[Q^{n-1}, P] = (n-1)i\hbar Q^{n-2}.$$

Then, using the second pair of idenities in Problem 14.3,

$$\begin{aligned}
[Q^n, P] = [Q^{n-1}Q, P] &= Q^{n-1}[Q, P] + [Q^{n-1}, P]Q \\
&= i\hbar Q^{n-1} + (n-1)i\hbar Q^{n-2}Q \\
&= ni\hbar Q^{n-1}.
\end{aligned}$$

Similarly, assuming $[Q, P^{m-1}] = (m-1)i\hbar P^{m-2}$, we have

$$\begin{aligned}
[Q, P^m] = [Q, P^{m-1}P] &= [Q, P^{m-1}]P + P^{m-1}[Q, P] \\
&= (m-1)i\hbar P^{m-1} + i\hbar P^{m-1} = mi\hbar P^{m-1}.
\end{aligned}$$

$$\begin{aligned}
[Q^n, P^2] = [Q^n, PP] &= [Q^n, P]P + P[Q^n, P] \\
&= ni\hbar Q^{n-1}P + ni\hbar PQ^{n-1} \\
&= ni\hbar[Q^{n-1}, P] + 2ni\hbar PQ^{n-1} \\
&= ni\hbar[Q^{n-1}, P] + 2ni\hbar([P, Q^{n-1}] + Q^{n-1}P) \\
&= -ni\hbar[Q^{n-1}, P] + 2ni\hbar Q^{n-1}P \\
&= n(n-1)\hbar^2 Q^{n-2} + 2ni\hbar Q^{n-1}P.
\end{aligned}$$

In three dimensions we have

$$[Q_i, Q_j] = [P_i, P_j] = 0, \qquad [Q_i, P_j] = i\hbar\delta_{ij}I.$$

Hence

$$\begin{aligned}[L_m, Q_k] &= [\epsilon_{mij} Q_i P_j, Q_k] \\ &= \epsilon_{mij}\big(Q_i[P_j, Q_k] + [Q_i, Q_j]P_j\big) \\ &= \epsilon_{mij}\big(-i\hbar\delta_{jk} Q_i + 0\big) \\ &= -i\hbar\epsilon_{mik} Q_i \\ &= i\hbar\epsilon_{mkj} Q_j\end{aligned}$$

and

$$\begin{aligned}[L_m, P_k] &= [\epsilon_{mij} Q_i P_j, P_k] \\ &= \epsilon_{mij}[Q_i, P_k]P_j \\ &= i\hbar\epsilon_{mij}\delta_{ik} P_j \\ &= i\hbar\epsilon_{mkj} P_j.\end{aligned}$$

Problem 14.5 **Consider a one-dimensional wave packet**

$$\psi(x,t) = \frac{1}{\sqrt{2\pi\hbar}}\int_{-\infty}^{\infty} \mathrm{e}^{i(xp - p^2 t/2m)/\hbar}\,\Psi(p)\mathrm{d}p$$

where

$$\Psi(p) \propto \mathrm{e}^{-(p-p_0)^2/2(\Delta p)^2}.$$

Show that $|\psi(x,t)|^2$ is a Gaussian normal distribution whose peak moves with velocity p/m and whose spread Δx increases with time, always satisfying $\Delta x \Delta p \geq \hbar/\sqrt{2}$.

If an electron ($m = 9 \times 10^{-28}$ g) is initially within an atomic radius $\Delta x_0 = 10^{-8}$ cm, after how long will Δx be (a) 2×10^{-8} cm, (b) the size of the solar system (about 10^{14} cm)?

Solution: A Gaussian normal distribution with peek at $y = y_0$ and root mean square deviation σ is

$$f(y) = \frac{1}{\sigma\sqrt{2\pi}}\mathrm{e}^{-(y-y_0)^2/2\sigma^2},$$

where

$$\begin{aligned}\sigma^2 &= \int_{-\infty}^{\infty}(y-y_0)^2 f(y)\mathrm{d}y \\ &= \frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\infty}(y-y_0)^2\mathrm{e}^{-(y-y_0)^2/2\sigma^2}\mathrm{d}y.\end{aligned}$$

The wave packet

$$\begin{aligned}\psi(x,t) &\propto \frac{1}{\sqrt{2\pi\hbar}}\int_{-\infty}^{\infty}\mathrm{e}^{i(xp-p^2t/2m)/\hbar}\mathrm{e}^{-(p-p_0)^2/2(\Delta p)^2}\mathrm{d}p \\ &\propto \mathrm{e}^{ixp_0/\hbar}\mathrm{e}^{-ip_0^2 t/2m\hbar}\int_{-\infty}^{\infty}\exp\big[-s^2\big(\frac{it}{2m\hbar} + \frac{1}{2(\Delta p)^2}\big) + is\big(\frac{x}{\hbar} - \frac{tp_0}{m\hbar}\big)\big]\mathrm{d}s\end{aligned}$$

where $p = s + p_0$. Now

$$\int_{-\infty}^{\infty} \mathrm{e}^{-as^2+bs} \mathrm{d}s = \int_{-\infty}^{\infty} \mathrm{e}^{-a(s-b/2a)^2+b^2/4a} \mathrm{d}s = \frac{\sqrt{\pi}}{\sqrt{a}} \mathrm{e}^{b^2/4a}.$$

Setting

$$a = \frac{1}{2}\Big(\frac{it}{m\hbar} + \frac{1}{(\Delta p)^2}\Big), \qquad b = i\Big(\frac{x}{\hbar} - \frac{tp_0}{m\hbar}\Big)$$

we have

$$\psi(x,t) = \mathrm{e}^{ixp_0/\hbar} \mathrm{e}^{-ip_0^2 t/2m\hbar} \frac{\sqrt{\pi}}{\big(it/2m\hbar + 1/2(\Delta p)^2\big)^{1/2}} \mathrm{e}^{-(x/h - tp_0/m\hbar)^2/2(it/m\hbar + 1/(\Delta p)^2)}$$

whence

$$|\psi|^2 = \psi\overline{\psi} \propto \mathrm{e}^{-(x-vt)^2/2(\Delta x)^2}$$

where $v = p_0/m$ is the velocity of the packet and

$$\begin{aligned}(\Delta x)^2 &= \hbar^2\Big(\frac{1}{it/m\hbar + 1/(\Delta p)^2} + \frac{1}{-it/m\hbar + 1/(\Delta p)^2}\Big)^{-1} \\ &= \frac{\hbar}{\sqrt{2}}\frac{1}{\Delta p}\sqrt{1 + \frac{t^2}{m^2\hbar^2}(\Delta p)^2},\end{aligned}$$

i.e. the spread Δx increases with time, and

$$\Delta x \Delta p \geq \frac{\hbar}{\sqrt{2}}.$$

At $t = 0$,

$$\Delta x_0 = \frac{\hbar}{\sqrt{2}\Delta p}$$

whence

$$\Delta x = \Delta x_0 \sqrt{1 + \frac{t^2}{m^2\hbar^2}(\Delta p)^2}.$$

It follows that the spread grows with time and $\Delta x = A\Delta x_0$ at time

$$\begin{aligned} t &= \frac{2m(\Delta x_0)^2}{\hbar}\sqrt{A^2 - 1} \\ &= 1.7 \times 10^{-16}\sqrt{A^2 - 1} \text{ sec} \qquad \text{if } \Delta x_0 = 10^{-8} \text{ cm}.\end{aligned}$$

(a) We have $A = 2$ after time $t = 1.7 \times 10^{-16} \times \sqrt{3} = 3.0 \times 10^{-16}$ sec.
(b) $A = 1.0 \times 10^{22}$ at $t \approx 1.7 \times 10^6$ sec ≈ 20 days.

Problem 14.6 In the Heisenberg picture show that the time evolution of the expection value of an operator A is given by

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle A' \rangle_{\psi'} = \frac{1}{i\hbar}\langle [A', H'] \rangle_{\psi'} + \langle \frac{\partial A'}{\partial t} \rangle_{\psi'}.$$

Convert this to an equation in the Schrödinger picture for the time evolution of $\langle A\rangle_\psi$.

Solution:

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\langle A'\rangle_{\psi'} &= \frac{\mathrm{d}}{\mathrm{d}t}\big(\langle\psi'|A'|\psi'\rangle\big)\\ &= \big(\frac{\mathrm{d}}{\mathrm{d}t}\langle\psi'|\big)A'|\psi'\rangle + \langle\psi'|\frac{\mathrm{d}A'}{\mathrm{d}t}|\psi'\rangle + \langle\psi'|A'\big(\frac{\mathrm{d}}{\mathrm{d}t}|\psi'\rangle\big)\\ &= \langle\psi'|\frac{1}{i\hbar}[A',H'] + \frac{\partial A'}{\partial t}|\psi'\rangle\end{aligned}$$

since

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle\psi'| = \frac{\mathrm{d}}{\mathrm{d}t}|\psi'\rangle = 0.$$

Hence

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\langle A'\rangle_{\psi'} &= \frac{1}{i\hbar}\langle\psi'|[A',H']|\psi'\rangle + \langle\psi'|\frac{\partial A'}{\partial t}|\psi'\rangle\\ &= \frac{1}{i\hbar}\langle[A',H']\rangle_{\psi'} + \langle\frac{\partial A'}{\partial t}\rangle_{\psi'}.\end{aligned}$$

In the Schrödinger picture we have, using $[A',B'] = [A,B]'$ and $\langle A'\rangle_{\psi'} = \langle A\rangle_\psi$,

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\langle A\rangle_\psi &= \frac{\mathrm{d}}{\mathrm{d}t}\langle A'\rangle'_\psi\\ &= \frac{1}{i\hbar}\langle[A,H]'\rangle_{\psi'} + \Big\langle\Big(\frac{\mathrm{d}A}{\mathrm{d}t}\Big)'\Big\rangle_{\psi'}\\ &= \frac{1}{i\hbar}\langle[A,H]\rangle_\psi + \Big\langle\Big(\frac{\mathrm{d}A}{\mathrm{d}t}\Big)\Big\rangle_\psi.\end{aligned}$$

Problem 14.7 For a particle of spin half in a magnetic field with Hamiltonian given in Example 14.4, show that in the Heisenberg picture

$$\begin{aligned}\langle\sigma_x(t)\rangle_{\mathbf{n}} &= \sin\theta\cos(\phi-\omega t),\\ \langle\sigma_y(t)\rangle_{\mathbf{n}} &= \sin\theta\sin(\phi-\omega t),\\ \langle\sigma_z(t)\rangle_{\mathbf{n}} &= \cos\theta.\end{aligned}$$

Solution: From Example 14.4 we have

$$\sigma_x(t) = \cos\omega t\sigma_1 + \sin\omega t\sigma_2 = \begin{pmatrix}0 & \mathrm{e}^{-i\omega t}\\ \mathrm{e}^{i\omega t} & 0\end{pmatrix}$$

$$\sigma_y(t) = -\sin\omega t\sigma_1 + \cos\omega t\sigma_2 = \begin{pmatrix}0 & -i\mathrm{e}^{-i\omega t}\\ i\mathrm{e}^{i\omega t} & 0\end{pmatrix}$$

$$\sigma_z(t) = \sigma_3 = \begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}$$

From Problem 14.1 we have

$$|+\mathbf{n}\rangle = \begin{pmatrix} \cos\frac{1}{2}\theta e^{-i\phi} \\ \sin\frac{1}{2}\theta \end{pmatrix}$$

and

$$\begin{aligned}
\langle\sigma_x(t)\rangle_{\mathbf{n}} &= \langle +\mathbf{n}|\sigma_x(t)|+\mathbf{n}\rangle \\
&= \left(\cos\tfrac{1}{2}\theta e^{i\phi} \quad \sin\tfrac{1}{2}\theta\right) \begin{pmatrix} 0 & e^{-i\omega t} \\ e^{i\omega t} & 0 \end{pmatrix} \begin{pmatrix} \cos\frac{1}{2}\theta e^{-i\phi} \\ \sin\frac{1}{2}\theta \end{pmatrix} \\
&= \cos\tfrac{1}{2}\theta e^{i\phi} \sin\tfrac{1}{2}\theta e^{-i\omega t} + \sin\tfrac{1}{2}\theta \cos\tfrac{1}{2}\theta e^{-i(\phi-\omega t)} \\
&= \sin\tfrac{1}{2}\theta \cos\tfrac{1}{2}\theta \sin(e^{i(\phi-\omega t)} + e^{-i(\phi-\omega t)}) \\
&= \sin\theta\cos(\phi-\omega t).
\end{aligned}$$

Similarly

$$\begin{aligned}
\langle\sigma_y(t)\rangle_{\mathbf{n}} &= \left(\cos\tfrac{1}{2}\theta e^{i\phi} \quad \sin\tfrac{1}{2}\theta\right) \begin{pmatrix} 0 & -tie^{-i\omega t} \\ ie^{i\omega t} & 0 \end{pmatrix} \begin{pmatrix} \cos\frac{1}{2}\theta e^{-i\phi} \\ \sin\frac{1}{2}\theta \end{pmatrix} \\
&= \cos\tfrac{1}{2}\theta \sin\tfrac{1}{2}\theta(-ie^{i(\phi-\omega t)} + \sin\tfrac{1}{2}\theta\cos\tfrac{1}{2}\theta e^{-i(\phi-\omega t)} \\
&= \sin\tfrac{1}{2}\theta \cos\tfrac{1}{2}\theta \sin(-ie^{i(\phi-\omega t)} + ie^{-i(\phi-\omega t)}) \\
&= \sin\theta\sin(\phi-\omega t),
\end{aligned}$$

and from Problem 14.1

$$\langle\sigma_z(t)\rangle_{\mathbf{n}} = \langle +\mathbf{n}|\sigma_3|+\mathbf{n}\rangle = \cos\theta.$$

Problem 14.8 A particle of mass m is confined by an infinite potential barrier to remain within a box $0 \le x, y, z \le a$, so that the wave function vanishes on the boundary of the box. Show that the energy levels are

$$E = \frac{1}{2m}\frac{\pi^2\hbar^2}{a^2}\left(n_1^2 + n_2^2 + n_3^2\right),$$

where n_1, n_2, n_3 are positive integers, and calculate the stationary wave functions $\psi_E(\mathbf{x})$. Verify that the lowest energy state is non-degenerate, but the next highest is triply degenerate.

Solution: Let $\psi_E(\mathbf{x})$ be a solution of the time-independent Schrödinger equation with energy E,

$$H\psi_E = \frac{-\hbar^2}{2m}\nabla^2\psi_E = E\psi_E.$$

Look for a solution by separation of variables, $\psi_E(x,y,z) = X(x)Y(y)Z(z)$, (i.e. a common eigenfunction of the commuting hermitian operators $(P_1)^2$, $(P_2)^2$ and $(P_3)^2$. This results in an equation

$$\frac{-\hbar^2}{2m}\frac{d^2X}{dx^2} = E_1 X \qquad (E_1 = \text{const}),$$

whence

$$X = A\sin\sqrt{\frac{2mE_1}{\hbar^2}}x + B\cos\sqrt{\frac{2mE_1}{\hbar^2}}x.$$

For an infinite potential barrier the eigenfunctions ψ_E vanish on the boundary of the box, so that $X(0) = X(a) = 0$, i.e. $B = 0$ and

$$\sqrt{\frac{2mE_1}{\hbar^2}}a = n_1\pi$$

where n_1 is a positive integer (≥ 1). Thus

$$E_1 = \frac{1}{2m}\left(\frac{n_1\pi}{a}\hbar\right)^2.$$

Similarly for $Y(y)$ and $Z(z)$, resulting in

$$\psi_E = A\sin\frac{n_1\pi x}{a}\sin\frac{n_2\pi y}{a}\sin\frac{n_3\pi z}{a}$$

where, on substituting in the original time-independent Schrödinger equation

$$E = E_1 + E_2 + E_3 = \frac{\pi^2}{a^2}\hbar^2\frac{1}{2m}\left(n_1^2 + n_2^2 + n_3^2\right).$$

Normalizing the wave function,

$$\int_0^a\int_0^a\int_0^a |\psi_E|^2 \mathrm{d}x\mathrm{d}y\mathrm{d}z = |A|^2\left(\frac{a}{2}\right)^2 = 1$$

results in the normalized wave function

$$\psi_E = \frac{2\sqrt{2}}{a^{3/2}}\sin\frac{n_1\pi x}{a}\sin\frac{n_2\pi y}{a}\sin\frac{n_3\pi z}{a}.$$

The lowest energy state is

$$E = \frac{3}{2m}\frac{\pi^2\hbar^2}{a^2}$$

which arises from $n_1^2 + n_2^2 + n_3^2 = 1$. This is non-degenerate, since it can only arise from $n_1 = n_2 = n_3 = 1$.
The next energy level is $n_1^2 + n_2^2 + n_3^2 = 1$,

$$E = \frac{3}{m}\frac{\pi^2\hbar^2}{a^2}$$

which arises from the three possibilities, $n_1 = n_2 = 1$, $n_3 = 2$ or $n_1 = n_3 = 1$, $n_2 = 2$ or $n_2 = n_3 = 1$, $n_1 = 2$.

Problem 14.9 For a particle with Hamiltonian

$$H = \frac{\mathbf{p}^2}{2m} + V(\mathbf{r})$$

show from the equation of motion in the Heisenberg picture that

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle \mathbf{r}\cdot\mathbf{p}\rangle = \left\langle \frac{\mathbf{p}^2}{m}\right\rangle - \langle \mathbf{r}\cdot\nabla V\rangle.$$

This is called the *Virial theorem.* For stationary states, show that

$$2\langle T\rangle = \langle \mathbf{r}\cdot\nabla V\rangle$$

where T is the kinetic energy. If $V \propto r^n$ this reduces to the classical result $\langle 2T - nV\rangle = 0$.

Solution: In the Heisenberg picture we have for $H = \mathbf{p}^2/2m + V(\mathbf{r})$, where $\mathbf{r} = (x_1, x_2, x_3)$ are the position and $\mathbf{p} = (p_i, p_2, p_3)$ the momentum operators,

$$\begin{aligned}\frac{\mathrm{d}x_i}{\mathrm{d}t} &= \frac{1}{i\hbar}[x_i, \frac{p_j p_j}{2m} + V(x_1, x_2, x_3)] + \frac{\partial x_i}{\partial t}\\ &= \frac{1}{2i\hbar m}([x_i, p_j]p_j + p_j[x_i, p_j]) \qquad \text{by Problem 14.3}\\ &= \frac{1}{2i\hbar m}(2i\hbar\delta_{ij}p_j)\\ &= \frac{p_i}{m}\end{aligned}$$

and

$$\frac{\mathrm{d}p_i}{\mathrm{d}t} = \frac{1}{i\hbar}[p_i, V(\mathbf{x})] = -\frac{\partial V}{\partial x_i}.$$

The last step is most easily seen in the Schrödinger representation where $p_i = -i\hbar\partial/\partial x_i$, so that for arbitrary wave functions ψ

$$[p_i, V(\mathbf{x})]\psi = -i\hbar\frac{\partial}{\partial x_i}(V\psi) + i\hbar V\frac{\partial\psi}{\partial x_i} = -i\hbar\frac{\partial V}{\partial x_i}\psi.$$

Hence, in the Heisenberg representation,

$$\frac{\mathrm{d}}{\mathrm{d}t}(\mathbf{r}\cdot\mathbf{p}) = \frac{\mathrm{d}}{\mathrm{d}t}(x_i p_i) = \frac{p_i}{m}p_i - x_i\frac{\partial V}{\partial x_i} = \frac{\mathbf{p}^2}{m} - \mathbf{r}\cdot\nabla V.$$

and

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\langle \mathbf{r}\cdot\mathbf{p}\rangle_\psi &= \frac{d}{dt}\langle\psi|\mathbf{r}\cdot\mathbf{p}|\psi\rangle\\ &= \langle\psi|\frac{d}{dt}(\mathbf{r}\cdot\mathbf{p})|\psi\rangle\\ &= \langle\psi|\frac{\mathbf{p}^2}{m} - \mathbf{r}\cdot\nabla V|\psi\rangle\\ &= \langle\frac{\mathbf{p}^2}{m} - \mathbf{r}\cdot\nabla V\rangle_\psi.\end{aligned}$$

For a stationary solution $\psi = \mathrm{e}^{-iEt/\hbar}\phi(\mathbf{r})$ we have

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle\psi|\mathbf{r}\cdot\mathbf{p}|\psi\rangle = \frac{\partial}{\partial t}\langle\phi|\mathbf{r}\cdot\mathbf{p}|\phi\rangle = \langle\phi|\frac{\partial\mathbf{r}\cdot\mathbf{p}}{\partial t}|\phi\rangle = 0.$$

Hence, since the kinetic energy is $T = \mathbf{p}^2/2m$,

$$2\langle T\rangle = \langle\mathbf{r}\cdot\nabla V\rangle.$$

For $V = Ar^n$,

$$\nabla V = \frac{\partial V}{\partial r}\hat{\mathbf{r}} = Anr^{n-1}\hat{\mathbf{r}} = Anr^{n-2}\mathbf{r}$$

and

$$\mathbf{r}\cdot\nabla V = Anr^{n-2}\mathbf{r}\cdot\mathbf{r} = Anr^{n-2}r^2 = Anr^n = nV,$$

which gives the desired result,

$$2\langle T\rangle - n\langle V\rangle = 0.$$

Problem 14.10 Show that the nth normalized eigenstate of the harmonic oscillator is given by

$$|\psi_n\rangle = \frac{1}{(n!)^{1/2}}A^n|\psi_0\rangle.$$

Show from $A^*\psi_0 = 0$ that

$$\psi_0 = c\mathrm{e}^{-\sqrt{km}x^2/2\hbar} \qquad \textbf{where} \qquad c = \left(\frac{m\omega}{\pi\hbar}\right)^{1/4}.$$

and the nth eigenfunction is

$$\psi_n(x) = \frac{i^n}{(2^n n!)^{1/2}}\left(\frac{m\omega}{\pi\hbar}\right)^{1/4}\mathrm{e}^{-m\omega x^2/2\hbar}H_n\left(\sqrt{\frac{m\omega}{\hbar}}x\right),$$

where $H_n(y)$ is the nth hermite polynomial (see Example 13.7).

Solution: In the discussion of the harmonic oscillator in the text, the normalized energy eigenstates $|\psi_n\rangle$ are related by

$$|\psi_{n+1}\rangle = \frac{1}{\sqrt{n+1}}A|\psi_n\rangle$$

where

$$A = \frac{1}{\sqrt{\omega\hbar}}\left(\frac{1}{\sqrt{2m}}P + i\sqrt{\frac{k}{2}}Q\right),$$

so that, inductively,

$$|\psi_n\rangle = \frac{1}{\sqrt{n!}} A^n |\psi_0\rangle$$

where $A^*|\psi_0\rangle = 0$. In the Schrödinger representation the 0 eigenfunction is thus determined by the differential equation

$$A^*\psi_0(x) = \frac{-i\hbar}{\sqrt{2m}} \frac{\mathrm{d}\psi_0(x)}{\mathrm{d}x} - i\sqrt{\frac{k}{2}} x\psi_0(x),$$

which has solution

$$\psi_0(x) = c\mathrm{e}^{-\sqrt{km}x^2/2\hbar} = c\mathrm{e}^{-m\omega x^2/2\hbar}.$$

Normalizing,

$$1 = \int_{-\infty}^{\infty} |\psi_0(x)|^2 \mathrm{d}x = \int_{-\infty}^{\infty} |c|^2 \mathrm{e}^{-m\omega x^2/\hbar} \mathrm{d}x = |c|^2 \sqrt{\frac{\pi\hbar}{m\omega}}$$

gives, up to a phase,

$$c = \left(\frac{m\omega}{\pi\hbar}\right)^{1/4}.$$

From the above discussion,

$$\begin{aligned}
\psi_n(x) &= \frac{1}{\sqrt{n!}} A^n |\psi_0\rangle \\
&= \frac{1}{\sqrt{n!}} \frac{1}{(\omega\hbar)^{n/2}} \left(\frac{-i\hbar}{\sqrt{2m}} \frac{\mathrm{d}}{\mathrm{d}x} + i\sqrt{\frac{k}{2}} x\right)^n \left(\frac{m\omega}{\pi\hbar}\right)^{1/4} \mathrm{e}^{-m\omega x^2/2\hbar}.
\end{aligned}$$

Set $y = \sqrt{m\omega/\hbar}x$ and this can be written

$$\psi_n(x) = \frac{1}{\sqrt{n!}} \left(\frac{m\omega}{\pi\hbar}\right)^{1/4} (-i)^n \frac{1}{2^{n/2}} \left(\frac{\mathrm{d}}{\mathrm{d}y} - y\right)^n \mathrm{e}^{-y^2/2}.$$

Now for any function $f(y)$,

$$\mathrm{e}^{y^2/2} \frac{\mathrm{d}}{\mathrm{d}y} \left(\mathrm{e}^{-y^2/2} f(y)\right) = \left(\frac{\mathrm{d}f}{\mathrm{d}y} - y\right) f(y).$$

Hence, applying the right hand operator again,

$$\left(\frac{\mathrm{d}f}{\mathrm{d}y} - y\right)^2 f(y) = \mathrm{e}^{y^2/2} \frac{\mathrm{d}^2}{\mathrm{d}y^2} \left(\mathrm{e}^{-y^2/2} f(y)\right),$$

and continuing n times we find

$$\left(\frac{\mathrm{d}f}{\mathrm{d}y} - y\right)^n f(y) = \mathrm{e}^{y^2/2} \frac{\mathrm{d}^n}{\mathrm{d}y^n} \left(\mathrm{e}^{-y^2/2} f(y)\right).$$

Applying to the function $f(y) = \mathrm{e}^{-y^2/2}$ results in

$$\Big(\frac{\mathrm{d}f}{\mathrm{d}y} - y\Big)^n \mathrm{e}^{-y^2/2} = \mathrm{e}^{y^2/2}\frac{\mathrm{d}^n}{\mathrm{d}y^n}\mathrm{e}^{-y^2} = (-1)^n \mathrm{e}^{-y^2/2} H_n(y)$$

where $H_n(y)$ are the Hermite polynomials defined in Example 13.7.Expressing in terms of x we have

$$\psi_n(x) = \frac{i^n}{(2^n n!)^{1/2}}\Big(\frac{m\omega}{\pi\hbar}\Big)^{1/4} \mathrm{e}^{-m\omega x^2/2\hbar} H_n\Big(\sqrt{\frac{m\omega}{\hbar}}x\Big).$$

Problem 14.11 **For the two-dimensional harmonic oscillator define operators A_1, A_2 such that**

$$[A_i, A_j] = [A_i^*, A_j^*] = 0, \qquad [A_i, A_j^*] = -\delta_{ij}, \qquad H = \hbar\omega(2N + I)$$

where $i, j = 1, 2$ and N is the number operator

$$N = \tfrac{1}{2}(A_1 A_1^* + A_2 A_2^*).$$

Let J_1, J_2 and J_3 be the operators

$$J_1 = \tfrac{1}{2}(A_2 A_1^* + A_1 A_2^*), \qquad J_2 = \tfrac{1}{2}i(A_2 A_1^* - A_1 A_2^*), \qquad J_3 = \tfrac{1}{2}(A_1 A_1^* - A_2 A_2^*),$$

and show that:

(a) The J_i satisfy the angular momentum commutation relations $[J_1, J_2] = iJ_3$, etc.

(b) $J^2 = J_1^2 + J_2^2 + J_3^2 = N(N+1)$.

(c) $[J^2, N] = 0, \qquad [J_3, N] = 0$.

From the properties of angular momentum deduce the energy levels and their degeneracies for the two-dimensional harmonic oscillator.

Solution: The Hamiltonian of the two-dimensional harmonic oscillator is

$$H = \frac{1}{2m}\big(P_1^2 + P_2^2\big) + \frac{k}{2}\big(Q_1^2 + Q_2^2\big).$$

Set

$$A_i = \frac{1}{\omega\hbar}\Big(\frac{1}{\sqrt{2m}}P_i + i\sqrt{\frac{k}{2}}Q_i\Big) \qquad (i = 1, 2)$$

where $\omega = \sqrt{k/m}$, which, just as in the one-dimensional case, can be shown to satisfy

$$[A_i, A_j^*] = -\delta_{ij}.$$

Set $N_i = A_i A_i^*$ for $i = 1, 2$, and $N = \frac{1}{2}(N_1 + N_2)$, then the Hamiltonian can be written

$$H = \hbar\omega(2N + I).$$

(a)

$$\begin{aligned}
[J_1, J_2] &= \tfrac{1}{4}i[A_2A_1^* + A_1A_2^*,\ A_2A_1^* - A_1A_2^*] \\
&= \tfrac{1}{4}i\big([A_2A_1^*,\ -A_1A_2^*] + [A_1A_2^*,\ A_2A_1^*]\big) \\
&= \tfrac{1}{4}i\big(-A_2[A_1^*, A_1]A_2^* - A_1[A_2, A_1^*]A_1^* + A_1[A_2^*, A_2]A_1^* + A_2[A_1, A_1^*]A_2^*\big) \\
&= \tfrac{1}{4}i\big(-A_2A_2^* + A_1A_1^* + A_1A_1^* - A_2A_2^*\big) \\
&= iJ_3.
\end{aligned}$$

The identities $[J_2,\ J_3] = iJ_1$ and $[J_3,\ J_1] = iJ_2$ follow in a very similar manner.

(b)

$$\begin{aligned}
J^2 = {J_1}^2 + {J_2}^2 + {J_3}^2 &= \tfrac{1}{4}\big(A_2A_1^*A_2A_1^* + A_1A_2^*A_1A_2^* + A_2A_1^*A_1A_2^* + A_1A_2^*A_2A_1^* \\
&\quad - A_2A_1^*A_2A_1^* - A_1A_2^*A_1A_2^* + A_2A_1^*A_1A_2^* + A_1A_2^*A_2A_1^* \\
&\quad + A_1A_1^*A_1A_1^* + A_2A_2^*A_2A_2^* - A_1A_1^*A_2A_2^* - A_2A_2^*A_1A_1^*\big) \\
&= \tfrac{1}{4}\big(2(N_1+1)N_2 + 2N_1(N_2+1) + {N_1}^2 + {N_2}^2 - N_1N_2 - N_2N_1\big) \\
&= \tfrac{1}{4}\big({N_1}^2 + {N_2}^2 + 2N_1 + 2N_2 + N_1N_+2 + N_2N_1\big) \\
&= \tfrac{1}{4}(N_1 + N_2)(N_1 + N_2 + 2I) \\
&= \tfrac{1}{4}2N(2N + 2I) = N(N + I)
\end{aligned}$$

(c) It is immediate that

$$[J^2, N] = [N^2 + N,\ N] = 0$$

and also that

$$[J_3, N] = \tfrac{1}{4}[A_1A_1^* - A_2A_2^*,\ A_1A_1^* + A_2A_2^*] = 0.$$

There is therefore a common set of eigenvectors to J^2, J_3 and N. From the discussion of angular momentum the eigenvalues of J^2 are $j^2 = l^2 + l$ where $l = n/2$ ($n = 0, 1, 2, \dots$), and common eigenvectors to J^2 and J_3 can be labelled $|lm\rangle$ where $m = -l, \dots, l$:

$$J^2|lm\rangle = l(l+1)|lm\rangle, \qquad J_3|lm\rangle = m|lm\rangle.$$

We then must have $N|lm\rangle = \lambda|lm\rangle$ where

$$\lambda(\lambda+1)|lm\rangle = l(l+1)|lm\rangle$$

giving that either $\lambda = l$ of $\lambda = -(l+1)$. The latter possibility is out, since N is positive definite

$$\langle\psi|N_i|\psi\rangle = \langle\psi|A_iA_i^*|\psi\rangle = \|A_i^*|\psi\rangle\|^2 \geq 0.$$

Hence $\lambda = n/2$ and the degeneracy is $n + 1$. The energy level corresponding to this value of λ is $E = \hbar\omega(n + 1)$.

Problem 14.12 **Show that the eigenvalues of the three-dimensional harmonic oscillator have the form $(n+\frac{3}{2})\hbar\omega$ where n is a non-negative integer. Show that the degeneracy of the nth eigenvalue is $\frac{1}{2}(n^2+3n+2)$. Find the corresponding eigenfunctions.**

Solution: The Hamiltonian of the 3-dimensionsal harmonic oscillator is

$$H = \frac{\mathbf{P}^2}{2m} + \frac{k}{2}\mathbf{Q}^2 = \sum_{i=1}^{3} H_i$$

where, each H_i represents a one-dimensional harmonic oscillator,

$$H_i = \omega\hbar(N_i + \tfrac{1}{2}I)$$

where

$$N_i = A_i A_i^* = N_i^*, \qquad A_i = \frac{1}{\sqrt{\omega\hbar}}\Big(\frac{1}{2m}P_i + i\sqrt{\frac{k}{2}}Q_i\Big) \quad \Big(\omega = \sqrt{\frac{k}{m}}\Big).$$

From the discussion of the one-dimensional harmonic oscillator in the text, we have that the eigenvalues of each N_i are $n_i = 0, 1, 2, \dots$. Hence, since the operators N_i (and therefore H_i) commute, the eigenvalues of H are

$$\omega\hbar(n+\tfrac{3}{2}) \qquad \text{where} \qquad n = n_1 + n_2 + n_3.$$

As there is a unique $|\psi_{n_i}\rangle = A^n|\psi_0\rangle$ of H_i corresponding to each eigenvalue $\omega\hbar(n_i + \frac{3}{2})$, the eigenstates corresponding to eigenvalue $\omega\hbar(n+\frac{3}{2})$ of H can be written as a linear combination

$$|\Psi_n\rangle = \sum_{n_1+n_2+n_3=n} c_{n_1 n_2 n_3}|\psi_{n_1}\rangle|\psi_{n_2}\rangle|\psi_{n_3}\rangle.$$

Hence the degeneracy of the eigenvalue $\omega\hbar(n+\frac{3}{2})$ is number of ways an integer $n \geq 0$ can be decomposed as a sum of three non-negative integers, $n = n_1 + n_2 + n_3$. For such a decomposition, n_1 may take any value from 0 to n, while the number of ways in which $n_2 + n_3 = n - n_1$ is $n - n_1 + 1$ (since n_2 may range from 0 to $n - n_1$). Hence the total deneracy of this eigenvalue is

$$(n+1) + n + (n-1) + \cdots + 1 = \frac{n+1)(n+2)}{2} = \frac{n^2+3n+2}{2}.$$

The general eigenvalue $|\Psi_n\rangle$ may be written as the wave function

$$\Psi_n(x,y,z) = \psi_{n_1}(x)\psi_{n_2}(y)\psi_{n_3}(z)$$

where the $\psi_m(x)$ are the eigenfunctions of the one-dimensional harmonic oscillator in terms of Hermite polynomials, described in Problem 14.10. Alternatively, solve the Schrödinger equation

$$\frac{1}{2m}\nabla^2\Psi(\mathbf{x}) + \frac{k}{2}\mathbf{r}^2\Psi(\mathbf{x}) = E\Psi(\mathbf{x})$$

by separation of variables, $\psi(x, y, z) = X(x)Y(y)Z(z)$, which leads to three separate one-dimensional equations

$$\begin{aligned}
\frac{1}{2m}\frac{\mathrm{d}^2X(x)}{\mathrm{d}x^2} + \frac{k}{2}x^2X(x) &= E_1X(x),\\
\frac{1}{2m}\frac{\mathrm{d}^2Y(y)}{\mathrm{d}y^2} + \frac{k}{2}y^2Y(y) &= E_2Y(y),\\
\frac{1}{2m}\frac{\mathrm{d}^2Z(z)}{\mathrm{d}z^2} + \frac{k}{2}z^2Z(z) &= E_3Z(z).
\end{aligned}$$

This will again lead, with the help of problem 14.10, to the same eigenfunctions.

Problem 14.13 If the operator K is complex conjugation with respect to a complete o.n. set,

$$K\big(\sum_i \alpha_i|e_i\rangle\big) = \sum_i \overline{\alpha_i}|e_i\rangle,$$

show that every anti-unitary operator V can be written in the form $V = UK$, where U is a unitary operator.

Solution: The operator K is anti-unitary, since (i) it is antilinear,

$$\begin{aligned}
K(|\psi\rangle + \alpha|\phi\rangle) &= K\big(\sum_i \psi_i|e_i\rangle + \alpha\sum_i \phi_i|e_i\rangle\big)\\
&= K\big(\sum_i(\psi_i + \alpha\phi_i)|e_i\rangle\big)\\
&= \sum_i \overline{(\psi_i + \alpha\phi_i)}|e_i\rangle\\
&= \sum_i \overline{\psi_i}|e_i\rangle + \overline{\alpha}\sum_i \overline{\phi_i}|e_i\rangle\\
&= K|\psi\rangle + \overline{\alpha}K|\phi\rangle,
\end{aligned}$$

and (ii) for any vectors $|\psi\rangle$ and $|\phi\rangle$

$$\begin{aligned}
\langle K\psi\,|\,K\phi\rangle &= \sum_i\sum_j \langle\overline{\psi_i}e_i\,|\,\overline{\phi_j}e_j\rangle\\
&= \sum_i\sum_j \psi_i\overline{\phi_j}\langle e_i\,|\,e_j\rangle\\
&= \sum_i \psi_i\overline{\phi_i} \qquad \text{since } \langle e_i\,|\,e_j\rangle\delta_{ij}\\
&= \langle\phi\,|\,\psi\rangle
\end{aligned}$$

by Parseval's identity, Theorem 13.6 (Eq. (13.7)).

If V is anti-unitary operator,

$$\langle V\psi|V\phi\rangle = \langle\phi|\psi\rangle,$$

then $U = VK$ is unitary for

$$\langle VK\psi|VK\phi\rangle = \langle K\phi|K\psi\rangle = \langle\psi|\phi\rangle.$$

Hence

$$V = UK^{-1} = UK$$

since it is immediate from the definition of K that $K^{-1} = K$.

Problem 14.14 **For any pair of operators A and B show by induction on the coefficients that**

$$\mathrm{e}^{aB}A\mathrm{e}^{-aB} = A + a[B,A] + \frac{a^2}{2!}[B,[B,A]] + \frac{a^3}{3!}[B,[B,[B,A]]] + \dots$$

Hence show the relation (14.32) holds for $T(\mathbf{a}) = \mathrm{e}^{i\mathbf{a}\cdot\mathbf{P}/\hbar}$.

Solution: Expanding as a power series in a,

$$\begin{aligned}\mathrm{e}^{aB}A\mathrm{e}^{-aB} &= \Big(I + aB + \frac{a^2}{2!}B^2 + \frac{a^3}{3!}B^3 + \dots\Big)A.\\ &\quad\Big(I - aB + \frac{a^2}{2!}B^2 - \frac{a^3}{3!}B^3 + \dots\Big)\\ &= A + T_1(A,B)a + T_2(A,B)\frac{a^2}{2!} + T_3(A,B)\frac{a^3}{3!} + \dots\end{aligned}$$

where

$$T_1(A,B) = BA - BA = [B,A]$$

and

$$\begin{aligned}T_n(A,B) = B^nA - nB^{n-1}AB + \binom{n}{2}B^{n-2}AB^2 - \dots\\ - (-1)^r\binom{n}{r}B^{n-r}AB^r + \dots + (-1)^nAB^n.\end{aligned}$$

This implies that

$$[B,T_n(A,B)] = T_{n+1}(A,B)$$

since for each $r = 0,1,\dots,n+1$ the coefficient of $B^{n+1-r}AB^r$ in $[B,T_n(A,B)]$ is

$$\begin{aligned}B(-1)^r\binom{n}{r}B^{n-r}AB^r - (-1)^{r-1}\binom{n}{r-1}B^{n-r+1}AB^{r-1}B\\ = (-1)^rB^{n+1-r}AB^r\Big\{\binom{n}{r} + \binom{n}{r-1}\Big\}\\ = (-1)^rB^{n+1-r}AB^r\binom{n+1}{r}\end{aligned}$$

as

$$\binom{n}{r}+\binom{n}{r-1}=\frac{n!}{r!(n-r)!}+\frac{n!}{r!(n+1-r)!}=\frac{n!(n+1)}{r!(n+1-r)!}=\binom{n+1}{r}.$$

Therefore, by induction we have

$$T_n(A,B)=\underbrace{[B,[B,[B,\dots[B}_{n\text{times}},A]\dots].$$

If $T(\mathbf{a})=e^{i\mathbf{a}\cdot\mathbf{P}/\hbar}$ we have, setting $B=-ia_jP_j/\hbar$ and using $[Q_i,P_j]=i\hbar\delta_{ij}$ in the above relation,

$$\begin{aligned}T^*(\mathbf{a})Q_iT(\mathbf{a})&=Q_i-i\frac{a_j}{\hbar}[P_j,Q_i]+i\frac{a_ja_k}{2!\hbar^2}[P_j[P_k,Q_i]]+\dots\\&=Q_i-a_iI\end{aligned}$$

since $[P_j,\delta_{ki}I]=\delta_{ki}[P_j,I]=0$ etc. Hence Eq. (14.32) follows, $T^*(\mathbf{a})\mathbf{Q}T(\mathbf{a})=\mathbf{Q}-\mathbf{a}I$.

Problem 14.15 Using the expansion in Problem 14.14 show that $R(\theta)=e^{i\theta L_3/\hbar}$ satisfies Eqs. (14.33)–(14.35).

Solution: Using $[Q_i,P_j]=i\hbar\delta_{ij}$ we have the commutation relations

$$\begin{aligned}&[L_3,Q_1]=[Q_1P_2-Q_2P_1,Q_1]=Q_1[P_2,Q_1]-Q_2[P_1,Q_1]=i\hbar Q_2,\\&[L_3,Q_2]=[Q_1P_2-Q_2P_1,Q_2]=Q_1[P_2,Q_2]-Q_2[P_1,Q_2]=-i\hbar Q_1,\\&[L_3,Q_3]=0.\end{aligned}$$

Setting $J=L_3/\hbar$ we have

$$[J,Q_1]=iQ_2,\qquad [J,Q_2]=-iQ_1,\qquad [J,Q_3]=0$$

and, using the expansion in Problem 14.14 with $B=-iJ$,

$$\begin{aligned}R^*(\theta)Q_iR^\theta&=e^{-i\theta J}Q_ie^{i\theta J}\\&=Q_i-i\theta[J,Q_i]-\frac{\theta^2}{2!}[J,[J,Q_i]]+i\frac{\theta^3}{3!}[J,[J,[J,Q_i]]]+\frac{\theta^4}{4!}\dots\end{aligned}$$

Now, using

$$\begin{aligned}&[J,Q_1]=iQ_2,\quad [J,[J,Q_1]]=i[J,Q_2]=Q_1,\quad [J,[J,[J,Q_1]]]=iQ_2,\ \text{etc.}\\&[J,Q_2]=-iQ_1,\quad [J,[J,Q_2]]=-i[J,Q_1]=Q_2,\quad [J,[J,[J,Q_2]]]=-iQ_1,\ \text{etc.}\end{aligned}$$

we have

$$\begin{aligned}R^*(\theta)Q_1R^\theta&=Q_1+\theta Q_2-\frac{\theta^2}{2!}Q_1-\frac{\theta^3}{3!}Q_2+\frac{\theta^4}{4!}Q_1+\dots\\&=Q_1\Big(1-\frac{\theta^2}{2!}-\frac{\theta^4}{4!}+\dots\Big)+Q_2\Big(\theta-\frac{\theta^3}{3!}+\dots\Big)\\&=Q_1\cos\theta+Q_2\sin\theta,\end{aligned}$$

and

$$\begin{aligned}
R^*(\theta) Q_2 R^\theta &= Q_2 - \theta Q_1 - \frac{\theta^2}{2!} Q_2 + \frac{\theta^3}{3!} Q_1 + \dots \\
&= Q_1 \Big(-\theta + \frac{\theta^3}{3!} - \dots \Big) + Q_2 \Big(1 - \frac{\theta^2}{2!} + \dots \Big) \\
&= -Q_1 \sin\theta + Q_2 \cos\theta .
\end{aligned}$$

The final identity (14.35), $R^*(\theta) Q_3 R^\theta = 0$, follows trivially from $[J, Q_3] = 0$.

Problem 14.16 **Show that the time reversal of angular momentum $\mathbf{L} = \mathbf{Q} \times \mathbf{P}$ is $\Theta^* L_i \Theta = -L_i$, and from the commutation relations $[L_i, L_j] = i\hbar \epsilon_{ijk} L_k$ show that Θ must anti-unitary.**

$**$ **NOTE: rewording of last part of question.**

Solution: For the time reversal operator Θ we have from Example 14.10,

$$\Theta^* Q_i \Theta = Q_i, \qquad \Theta^* P_j \Theta = -P_j, \qquad \Theta^* \Theta = \Theta \Theta^* = I.$$

Hence

$$\begin{aligned}
L'_i = \Theta^* L_i \Theta &= \Theta^* \epsilon_{ijk} Q_j P_k \Theta \\
&= \epsilon_{ijk} \Theta^* Q_j \Theta \Theta^* P_k \Theta \\
&= -\epsilon_{ijk} Q_j P_k = -L_i .
\end{aligned}$$

From this transformation law it follows that

$$[L'_i, L'_j] = [-L_i, -L_j] = [L_i, L_j] = i\hbar \epsilon_{ijk} L_k = -i\hbar \epsilon_{ijk} L'_k .$$

Thus the commutation relations are *not* invariant under time reversal. However

$$\begin{aligned}
[L'_i, L'_j] &= [\Theta^* L_i \Theta, \Theta^* L_j \Theta] \\
&= \Theta^* [L_i, L_j] \Theta \\
&= \Theta^* i\hbar \epsilon_{ijk} L_k \Theta \\
&= \epsilon_{ijk} \Theta^* i I \Theta \Theta^* L_k \Theta \\
&= \epsilon_{ijk} \Theta^* i I \Theta L'_k .
\end{aligned}$$

Comparing this with the previously derived commutation relation, we have

$$\Theta^* i I \Theta = -iI,$$

i.e. Θ is anti-unitary.

Problem 14.17 **Show that the correctly normalized fermion states are**

$$\frac{1}{\sqrt{N!}} \sum_P (-1)^P P |\varphi_{a_1}\rangle |\varphi_{a_2}\rangle \dots |\varphi_{a_N}\rangle$$

and normalized boson states are

$$\frac{1}{\sqrt{N!}\sqrt{n_0!}\sqrt{n_1!}\dots}\sum_P P|\varphi_{a_1}\rangle|\varphi_{a_2}\rangle\dots|\varphi_{a_N}\rangle.$$

Solution: The states $|\varphi_a\rangle$ are defined as the one-particle energy eigenstates,

$$h_i|\varphi_a\rangle = \varepsilon_a|\varphi_a\rangle$$

and states $|\Phi_k\rangle$ which are linear combinations of $|\varphi_{a_1}\rangle|\varphi_{a_2}\rangle\dots|\varphi_{a_N}\rangle$ are eigenstates of the entire system

$$H|\Phi_k\rangle = E_k|\Phi_k\rangle, \qquad \text{where} \qquad E_k = \sum_{a=0}^{\infty} n_a\varepsilon_a, \quad N = \sum_{a=0}^{\infty} n_a.$$

Since the occupation numbers n_a can only take values 0 or 1 for fermions we must have

$$|\Phi_k\rangle = \alpha A|\varphi_{a_1}\rangle|\varphi_{a_2}\rangle\dots|\varphi_{a_N}\rangle$$

and

$$\langle\varphi_{a_i}\,|\,\varphi_{a_j}\rangle = \delta_{ij}.$$

Hence

$$\begin{aligned}1 = \langle\Phi_k\,|\,\Phi_k\rangle &= \frac{|\alpha|^2}{(N!)^2}\sum_P\sum_{P'}(-1)^P(-1)^{P'}P\langle\varphi_{a_1}|\dots\langle\varphi_{a_N}|\,P'|\varphi_{a_1}\rangle\dots|\varphi_{a_N}\rangle\\ &= \frac{|\alpha|^2}{(N!)^2}N! = \frac{|\alpha|^2}{N!}\end{aligned}$$

since P and P' must be identical permutations of $1, 2, \dots, N$ for a non-zero result in the sum, and there are precisely $N!$ such permutations. We can therefore take $\alpha = \sqrt{N!}$, so that

$$|\Phi_k\rangle = \sqrt{N!}\frac{1}{N!}\sum_P(-1)^P P|\varphi_{a_1}\rangle|\varphi_{a_2}\rangle\dots|\varphi_{a_N}\rangle,$$

as required. For bosons we take

$$|\Phi_k\rangle = \beta S|\varphi_{a_1}\rangle|\varphi_{a_2}\rangle\dots|\varphi_{a_N}\rangle$$

and

$$1 = \langle\Phi_k\,|\,\Phi_k\rangle = \frac{|\beta|^2}{(N!)^2}\sum_P\sum_{P'}P\langle\varphi_{a_1}|\dots\langle\varphi_{a_N}|\,P'|\varphi_{a_1}\rangle\dots|\varphi_{a_N}\rangle.$$

For each permutation P of $1, 2, \dots, N$ the identical permutation P' provides a value 1 in the sum on the right-hand side. If we fix $a_1, \dots, a_N$ all further permutations P' which permute the n_0 states $|\phi_{b_i}\rangle$ having energy ε_0 among each other are permitted and contribute a value 1 to the sum. Similarly permutations among the n_1 states $|\phi_{b_i}\rangle$ having energy ε_1 have a similar effect, etc. This gives rise to an additional $n_0!n_1!n_2!\ldots$ values of 1 for each permutation P among the $\langle\phi_{a_i}|$. Hence

$$1 = \frac{|\beta|^2}{(N!)^2} N!n_0!n_1!n_2!\ldots$$

so that we may take

$$\beta = \frac{\sqrt{N!}}{\sqrt{n_0!}\sqrt{n_1!}\sqrt{n_2!}\ldots}$$

whence

$$|\Phi_k\rangle = \frac{1}{\sqrt{N!}\sqrt{n_0!}\sqrt{n_1!}\ldots} \sum_P P|\varphi_{a_1}\rangle|\varphi_{a_2}\rangle\ldots|\varphi_{a_N}\rangle.$$

Problem 14.18 Calculate the canonical partition function, mean energy U and entropy S, for a system having just two energy levels 0 and E. If $E = E(a)$ for a parameter a, calculate the force A and verify the thermodynamic relation $\mathrm{d}S = \frac{1}{T}(\mathrm{d}U + A\mathrm{d}a)$.

Solution: The canonical partition function is

$$Z = 1 + \mathrm{e}^{-\beta E}$$

and the mean energy is

$$U = -\frac{\partial Z}{\partial \beta} = \frac{E\mathrm{e}^{-\beta E}}{1+\mathrm{e}^{-\beta E}} = \frac{E}{\mathrm{e}^{\beta E}+1}.$$

Setting

$$w_0 = \frac{1}{Z}, \quad w_1 = \frac{\mathrm{e}^{-\beta E}}{Z}, \qquad w_0 + w_1 = 1,$$

we have

$$\begin{aligned} S &= -k(w_0 \ln w_0 + w_1 \ln w_1) \\ &= -k\Big(-\frac{\ln Z}{Z} + \frac{\mathrm{e}^{-\beta E}}{Z}(-\ln Z - \beta E)\Big) \\ &= k\Big(\ln Z + \frac{\beta E\mathrm{e}^{-\beta E}}{1+\mathrm{e}^{-\beta E}}\Big) \\ &= k\Big(\ln(1+\mathrm{e}^{-\beta E}) + \frac{\beta E\mathrm{e}^{-\beta E}}{1+\mathrm{e}^{-\beta E}}\Big). \end{aligned}$$

This result is also simple to obtain from the displayed equation after Eq. (14.50), namely $S = k(\ln Z + \beta U)$.
If $E = E(a)$ the force can be obtained from Eq. (50), or more directly from

$$A = -\sum_k w_k \frac{\partial E_k}{\partial a} = -w_1 \frac{\mathrm{d}E}{\mathrm{d}a} = -\frac{\mathrm{e}^{-\beta E}}{1+\mathrm{e}^{-\beta E}}\frac{\mathrm{d}E}{\mathrm{d}a}.$$

We now compute

$$\begin{aligned}\mathrm{d}S &= k\Big(\frac{-\beta \mathrm{e}^{-\beta E}\mathrm{d}E}{1+\mathrm{e}^{-\beta E}} + \frac{\beta \mathrm{d}E \mathrm{e}^{-\beta E}}{1+\mathrm{e}^{-\beta E}} - \frac{\beta^2 E \mathrm{e}^{-\beta E}\mathrm{d}E}{(1+\mathrm{e}^{-\beta E})^2}\Big)\\ &= -\frac{k\beta^2 E\mathrm{e}^{-\beta E}\mathrm{d}E}{(1+\mathrm{e}^{-\beta E})^2}\end{aligned}$$

while

$$\begin{aligned}\frac{1}{T}\mathrm{d}U &= k\beta\Big(\frac{\mathrm{d}E\mathrm{e}^{-\beta E}}{1+\mathrm{e}^{-\beta E}} - \frac{E\beta \mathrm{e}^{-\beta E}\mathrm{d}E}{1+\mathrm{e}^{-\beta E}} + \frac{E\beta\mathrm{e}^{-2\beta E}\mathrm{d}E}{(1+\mathrm{e}^{-\beta E})^2}\Big)\\ &= k\beta\Big(\frac{\mathrm{e}^{-\beta E}}{1+\mathrm{e}^{-\beta E}} - \frac{E\beta\mathrm{e}^{-\beta E}}{(1+\mathrm{e}^{-\beta E})^2}\Big)\mathrm{d}E\end{aligned}$$

and

$$\frac{1}{T}A\mathrm{d}a = k\beta\frac{-\mathrm{e}^{-\beta E}}{1+\mathrm{e}^{-\beta E}}\frac{\mathrm{d}E}{\mathrm{d}a}\mathrm{d}a = -k\beta\frac{\mathrm{e}^{-\beta E}}{1+\mathrm{e}^{-\beta E}}\mathrm{d}E.$$

Hence

$$\mathrm{d}S = \frac{1}{T}(\mathrm{d}U + A\mathrm{d}a) = -\frac{k\beta^2 E\mathrm{e}^{-\beta E}}{(1+\mathrm{e}^{-\beta E})^2}\mathrm{d}E.$$

Problem 14.19 **Let $\rho = \mathrm{e}^{-\beta H}$ be the unnormalized canonical distribution. For a free particle of mass m in one-dimension show that its position representation form $\rho(x,x';\beta) = \langle x|\rho|x'\rangle$ satisfies the diffusion equation**

$$\frac{\partial\rho(x,x';\beta)}{\partial\beta} = \frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}\rho(x,x';\beta)$$

with 'initial' condition $\rho(x,x';0) = \delta(x-x')$. Verify that the solution is

$$\rho(x,x';\beta) = \Big(\frac{m}{2\pi\hbar^2\beta}\Big)^{1/2}\mathrm{e}^{-m(x-x')^2/2\hbar^2\beta}.$$

Solution: If $\rho = \mathrm{e}^{-\beta H}$ then

$$\frac{\partial\rho}{\partial\beta} = -H\mathrm{e}^{-\beta H} = -H\rho.$$

Hence the position representation function $\rho(x,x';\beta) = \langle x|\rho|x'\rangle$ satisfies

$$\frac{\partial}{\partial\beta}\rho(x,x';\beta) = \langle x|\frac{\partial\rho}{\partial\beta}|x'\rangle = \langle x|-H\rho|x'\rangle.$$

For a free particle the position representation wave function $\psi(x) = \langle x\,|\,\psi\rangle$ satisfies

$$\langle x|H|\psi\rangle = -\frac{\hbar^2}{2m}\frac{\partial^2\psi(x)}{\partial x^2}$$

and setting $|\psi\rangle = \rho|x'\rangle$ we arrive at the diffusion equation

$$\frac{\partial\rho(x,x';\beta)}{\partial\beta} = \frac{\hbar^2}{2m}\frac{\partial^2\langle x|\rho|x'\rangle}{\partial x^2} = \frac{\hbar^2}{2m}\frac{\partial^2\rho(x,x';\beta)}{\partial x^2}.$$

The function

$$\rho(x,x';\beta) = \Big(\frac{m}{2\pi\hbar^2\beta}\Big)^{1/2}\mathrm{e}^{-m(x-x')^2/2\hbar^2\beta}.$$

is easily verified to be a solution by direct substitution (see also Problem 12.16, where this is essentially the solution which gives rise to the Greens' function for the diffusion equation). For this solution

$$\int_{-\infty}^{\infty}\rho(x,x';\beta)\mathrm{d}x = 1$$

and subsituting $n = \sqrt{m/2\beta\hbar^2}$ in the functions $g_n(x)$ of Problem 12.4 we have

$$\rho(x,x';\beta) \to \delta(x-x') \qquad \text{as } \beta \to 0.$$

Problem 14.20 A solid can be regarded as being made up of $3N$ independent quantum oscillators of angular frequency ω. Show that the canonical partition function is given by

$$Z = \Big(\frac{\mathrm{e}^{-\beta\hbar\omega/2}}{1-\mathrm{e}^{-\beta\hbar\omega}}\Big)^{3N},$$

and the specific heat is given by

$$C_V = \frac{\mathrm{d}U}{\mathrm{d}T} = 3Nk\Big(\frac{T_0}{T}\Big)^2\frac{\mathrm{e}^{T_0/T}}{(\mathrm{e}^{T_0/T}-1)^2} \qquad \textbf{where} \qquad kT_0 = \hbar\omega.$$

Show that the high temperature limit $T \gg T_0$ is the classical value $C_V = 3Nk$.

Solution: The hamiltonian of a harmonic oscillator is

$$h_i = \frac{p_i^2}{2m} + \frac{m\omega^2}{2}q_i^2$$

and the energy levels are $(n_i + \frac{1}{2})\hbar\omega$,

$$h_i|n_i\rangle = (n_i + \tfrac{1}{2})\hbar\omega|n_i\rangle.$$

The energy levels of the $3N$ oscillator system $H = \sum_{i=1}^{3N} h_i$ are

$$H|n_1\rangle\ldots|n_{3N}\rangle = E_n|n_1\rangle\ldots|n_{3N}\rangle$$

where

$$E_n = \hbar\omega \sum_{i=1}^{3N} (n_i + \tfrac{1}{2}).$$

Hence, the canonical partition function is

$$\begin{aligned} Z &= \sum_n \mathrm{e}^{-\beta E_n} \\ &= \sum_{n_1=0}^{\infty} \cdots \sum_{n_{3N}=0}^{\infty} \mathrm{e}^{-\beta\hbar\omega \sum_{i=1}^{3N}(n_i+\frac{1}{2})} \\ &= \prod_{i=1}^{3N} \sum_{n=0}^{\infty} \mathrm{e}^{-\beta\hbar\omega(n+\frac{1}{2})} \\ &= \Big(\frac{\mathrm{e}^{-\beta\hbar\omega/2}}{1-\mathrm{e}{-}\beta\hbar\omega}\Big)^{3N}. \end{aligned}$$

The average energy is

$$U = -\frac{\partial \ln Z}{\partial \beta} = 3N\hbar\omega\Big[-\frac{1}{2} + \frac{1}{1-\mathrm{e}^{-\beta\hbar\omega}}\Big]$$

and the specific heat at constant volume is, using $kT = 1/\beta$,

$$\begin{aligned} C_V = \frac{\mathrm{d}U}{\mathrm{d}T} &= -\frac{1}{kT^2}\frac{\mathrm{d}U}{\mathrm{d}\beta} \\ &= -\frac{3N\hbar\omega}{kT^2}\Big[\frac{-\mathrm{e}^{-\beta\hbar\omega}\hbar\omega}{1-\mathrm{e}^{-\beta\hbar\omega}}\Big] \\ &= 3Nk\Big(\frac{T_0}{T}\Big)^2 \frac{\mathrm{e}^{-T_0/T}}{(1-\mathrm{e}^{-T_0/T})^2} \end{aligned}$$

where $kT_0 = \hbar\omega$.
For $T \gg T_0$, we have $T_0/T \ll 1$ and

$$\begin{aligned} C_V &= 3Nk\Big(\frac{T_0}{T}\Big)^2 \frac{1-(T_0/T)+\frac{1}{2}(T_0/T)^2+\ldots}{(T_0/T)-\frac{1}{2}(T_0/T)^2+\ldots)^2} \\ &\approx 3Nk\frac{1-T_0/T+\ldots}{(1-\frac{1}{2}(T_0/T)^2+\ldots)^2} \\ &\approx 3Nk\Big(1+O\Big(\frac{T_0}{T}\Big)^2\Big) \end{aligned}$$

Problem 14.21 **Show that the average occupation numbers for the classical distribution, $Z_{\mathbf{Boltzmann}}$ are given by**

$$\langle n_a \rangle = -\frac{1}{\beta}\frac{\partial}{\partial \varepsilon_a} \ln Z_{\mathbf{Boltzmann}} = \lambda \mathrm{e}^{-\beta\varepsilon_a}.$$

Hence show that

$$\langle n_a\rangle_{\mathbf{Fermi}} < \langle n_a\rangle_{\mathbf{Boltzmann}} < \langle n_a\rangle_{\mathbf{Bose}}$$

and that all three types agree approximately for low occupation numbers $\langle n_a\rangle \ll 1$.

Solution: Dividing by the standard Gibbs factor $N!$ we take

$$w(n_0, n_1, \dots) = \frac{1}{Z} e^{\alpha \sum_a n_a - \beta \sum_a \varepsilon_a n_a} \frac{1}{n_0!} \frac{1}{n_1!} \dots$$

in the Botzmann grand canonical ensemble,

$$\begin{aligned}
Z_{\text{Boltzmann}} &= \sum_{n_0}^{\infty} \sum_{n_1}^{\infty} \dots e^{\alpha \sum_a n_a - \beta \sum_a \varepsilon_a n_a} \frac{1}{n_0!} \frac{1}{n_1!} \dots \\
&= \sum_{n_0}^{\infty} \frac{e^{(\alpha - \beta\varepsilon_0) n_0}}{n_0!} \sum_{n_1}^{\infty} \frac{e^{(\alpha - \beta\varepsilon_1) n_1}}{n_1!} \dots \\
&= \exp\left(e^{\alpha - \beta\varepsilon_0}\right) \exp\left(e^{\alpha - \beta\varepsilon_1}\right) \dots \\
&= \exp\left(e^{\alpha - \beta\varepsilon_0} + e^{\alpha - \beta\varepsilon_1} \dots\right)
\end{aligned}$$

so that

$$\ln Z_{\text{Boltzmann}} = e^{\alpha} \sum_a^{\infty} e^{-\beta\varepsilon_a}.$$

Hence

$$\langle n_a\rangle_{\text{Boltzmann}} = -\frac{1}{\beta} \frac{\partial \ln Z_{\text{Boltzmann}}}{\partial \varepsilon_a} = \lambda e^{-\beta\varepsilon_a}$$

where $\lambda = e^{\alpha}$. From the text we have

$$\langle n_a\rangle_{\text{Fermi}} = \frac{1}{\lambda^{-1} e^{\beta\varepsilon_a} + 1} = \frac{\lambda e^{-\beta\varepsilon_a}}{1 + \lambda e^{-\beta\varepsilon_a}}$$

and

$$\langle n_a\rangle_{\text{Bose}} = \frac{1}{\lambda^{-1} e^{\beta\varepsilon_a} - 1} = \frac{\lambda e^{-\beta\varepsilon_a}}{1 - \lambda e^{-\beta\varepsilon_a}}.$$

Since $\lambda e^{-\beta\varepsilon_a} > 0$ it follows immediately that

$$\langle n_a\rangle_{\text{Fermi}} < \langle n_a\rangle_{\text{Boltzmann}} < \langle n_a\rangle_{\text{Bose}}.$$

In the limit $\langle n_a\rangle_{\text{Boltzmann}} \to 0$, we have that $\lambda e^{-\beta\varepsilon_a} \ll 1$, in wheich case

$$\langle n_a\rangle_{\text{Fermi}} \approx \langle n_a\rangle_{\text{Bose}} \approx \lambda e^{-\beta\varepsilon_a} = \langle n_a\rangle_{\text{Boltzmann}}.$$

Problem 14.22 **A *spin system* consists of N particles of magnetic moment μ in a magnetic field B. When n particles have spin up, $N - n$**

spin down, the energy is $E_n = n\mu B - (N-n)\mu B = (2n-N)\mu B$. Show that the canonical partition function is

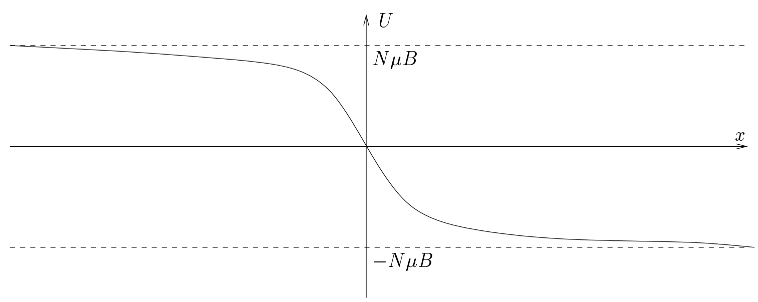
$$Z = \frac{\sinh\big((N+1)\beta\mu B\big)}{\sinh\beta\mu B}.$$

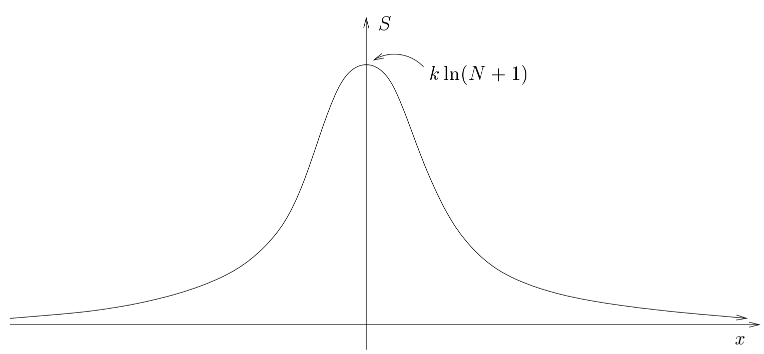
Evaluate the mean energy U and entropy S, sketching their dependence on the variable $x = \beta\mu B$.

Solution: The canonical partition function is, setting $x = \beta\mu B$,

$$\begin{aligned} Z = \sum_{n=0}^{N} e^{-\beta E_n} &= e^{Nx} + e^{(N-2)x} + \cdots + e^{-Nx} \\ &= e^{-Nx}\big(1 + e^{2x} + \cdots + e^{2Nx}\big) \\ &= e^{-Nx}\frac{e^{2(N+1)x} - 1}{e^{2x} - 1} \\ &= \frac{e^{(N+1)x} - e^{-(N+1)x}}{e^{x} - e^{-x}} \\ &= \frac{\sinh(N+1)x}{\sinh x}. \end{aligned}$$

The mean energy and entropy are given by





$$U = -\frac{\partial \ln Z}{\partial B} = -\frac{\partial \ln Z}{\partial x}\mu\beta = \big[-(N+1)\coth(N+1)x + \coth x\big]\mu B$$

and

$$S = k(\ln Z + U\beta) = k\big[\ln\sinh(N+1)x - \ln\sinh x + x(-(N+1)\coth(N+1)x + \coth x)\big].$$

At $x = 0$ (i.e. $\beta = 0$ or temperature $T \to \infty$, or $B = 0$) we have $U = 0$, while as $x \to \pm\infty$ the mean energy has limits $\mp N\mu B$. The entropy S has maximum value $k\ln(N+1)$ at $x = 0$, and approaches 0 rapidly as $x \to \pm\infty$. The latter is deduced from the limits

$$\ln\sinh y \approx y - \ln 2 + \mathrm{e}^{-2y}, \quad \coth y \approx 1 + 2\mathrm{e}^{-2y} \qquad \text{as } y \to \infty$$

which gives that the dominant term in the behaviour as $|x| \to \infty$ arises from the $x\coth x$ term and results in $S \approx 2|x|\mathrm{e}^{-2|x|}$ as $|x| \to \infty$.

Chapter 16

Problem 15.1 **Show that the group of unimodular matrices $SL(n,\mathbb{R}) = \{A \in GL(n,\mathbb{R}) \mid \det A = 1\}$ is a differentiable manifold.**

Solution: By Example 15.5 the group $GL(n,\mathbb{R})$ is an open submanifold of $\mathbb{R}^{n^2}$ and therefore has dimension n^2. Given any matrix $\mathsf{A} = [a_{ij}]$ we may expand the determinant by the first row,

$$\det \mathsf{A} = a_{11} \det \mathsf{A}_1 - a_{12} \det \mathsf{A}_2 + \dots \pm a_{1n} \det \mathsf{A}_n \qquad (*)$$

where A_i is the minor corresponding to the element a_{1i}; i.e. the matrix formed by removing the first row and ith column from A. For each $i = 1, 2, \dots, n$ define the set

$$U_i = \{\mathsf{A} \in SL(n,\mathbb{R}) \mid \det \mathsf{A}_i \neq 0\} \subset SL(n,\mathbb{R}).$$

If $SL(n,\mathbb{R})$ is given the relative topology induced by $GL(n,\mathbb{R})$ (necessarily Hausdorff by Corollary 10.6), then each subset U_i is open in this topology since it is the inverse image $\psi_i^{-1}(\dot{\mathbb{R}}$ of the open set $\dot{\mathbb{R}} = \mathbb{R} - \{0\}$ under the continuous map $\psi_i : SL(n,\mathbb{R}) \to \mathbb{R}$ defined by $\psi_i(\mathsf{A}) = \det(\mathsf{A}_i)$. Define the map $\phi_i : U_i \to \mathbb{R}^{n^2-1}$ by

$$\phi_i(\mathsf{A}) = (a_{i1}, \dots, a_{ii-1}, a_{1,i+1}, \dots, a_{1n}, a_{21}, a_{22}, \dots, a_{nn}).$$

These maps are all continuous and one-to-one onto the image $\phi_i(U_i) \subset \mathbb{R}^{n^2-1}$, since a_{1i} is uniquely determined by Eq. $(*)$ since $\det \mathsf{A}_i \neq 0$. Every element belongs to at least one U_i for if $\det \mathsf{A}_i = 0$ for all $i = 1, \dots, n$ then by Eq. $(*)$ $\det \mathsf{A} = 0 \neq 1$. The transformation functions $\phi_j \circ \phi_i^{-1}$ are trivially continuous. Hence this is a differentiable structure, making $SL(n,\mathbb{R})$ into a differentiable manifold of dimension $n^2 - 1$.

Problem 15.2 **On the n-sphere S^n find coordinates corresponding to (i) stereographic projection, (ii) spherical polars.**

Solution: The unit n-sphere is defined as

$$S^n = \{(x^1, x^2, \dots, x^{n+1}) \in \mathbb{R}^{n+1} \mid (x^1)^2 + (x^2)^2 + \dots + (x^n)^2 = 1\}.$$

(i) Let $N = (0, 0, \dots, 1)$, and define the map $\mathrm{St}_N : S^n - \{N\} \to \mathbb{R}^n$ by $\mathrm{St}_N : (x^1, \dots, x^{n+1}) \mapsto (X^1, \dots, X^n)$ where

$$X^i = \frac{x^i}{1 - x^{n+1}} \qquad (i = 1, \dots, n).$$

This map is one-to-one since it is reversible, i.e. the the inverse $\mathrm{St}_N^{-1} : \mathbb{R}^n \to S^n - \{N\}$ exists, since

$$\sum_{i=1}^{n} (X^i)^2 = \frac{1 + x^{n+1}}{1 - xn + 1}$$

so that

$$x^{n+1} = \frac{R-1}{R+1} \qquad \text{where} \qquad R = \sqrt{\textstyle\sum_{i=1}^{n}(X^i)^2}.$$

Hence

$$x^i = X^i(1 - x^{n+1}) = \frac{2X^i}{1+R}.$$

(ii) There are more than one logical ways of extending polar coordinates to S^n, but the most natural is the following: Since $|x^{n+1}| \leq 1$ there exists an angle θ_1 such that $x^{n+1} = \cos\theta_1$ $(0 \leq \theta_1 \leq \pi)$. As $|x^n| \leq \sqrt{1-(x^{n+1})^2} = |\sin\theta_1|$ there exists an angle θ_2 $(0 \leq \theta_2 \leq \pi)$ such that $x^n = \sin\theta_1 \cos\theta_2$. Continuing in this way we find

$$\begin{aligned}
x^{n+1} &= \cos\theta_1 \\
x^n &= \sin\theta_1 \cos\theta_2 \\
x^{n-1} &= \sin\theta_1 \sin\theta_2 \cos\theta_3 \\
&\dots \\
x^2 &= \sin\theta_1 \sin\theta_2 \dots \sin\theta_{n-1} \sin\theta_n \\
x^1 &= \sin\theta_1 \sin\theta_2 \dots \sin\theta_{n-1} \cos\theta_n.
\end{aligned}$$

The change in the last step, where $\sin\theta_n$ and $\cos\theta_n$ are reversed, is not compulsory but agrees with the usual convention whereby the final angle θ_n is measured from the x^1-axis. For example in three-dimensional spherical polars take $\theta = \theta_1$ and $\phi = \theta_2$ (otherwise we could set $\phi = \theta_2 + \pi/2$ and retain the same order of cos and sin in the above sequence). The ranges of angles are $0 \leq \theta_i \leq \pi$ for $i = 1, \dots, n-1$ and $0 \leq \theta_n \leq 2\pi$.

Problem 15.3 Show that the real projective n-space P^n defined in Example 10.15 as the set of straight lines through the origin in $\mathbb{R}^{n+1}$ is a differentiable manifold of dimension n, by finding an atlas of compatible charts that cover it.

Solution: On $\mathbb{R}^{n+1}$ define the equivalence relation

$$(x_1, x_2, \dots, x_{n+1}) \equiv (x'_1, x'_2, \dots, x'_{n+1})$$

iff there exists $\lambda \neq 0$ such that $x_i = \lambda x'_i$ for $i = 1, 2, \dots, n+1$. The projective space P^n consists of the equivalence classes, $[(x_1, x_2, \dots, x_{n+1})]$, which can be thought of as the set of all lines through the origin.
Set U_i to be the set of lines with $x_i \neq 0$, $(i = 1, 2, \dots, n+1)$. These are clearly open sets in the topology induced on P^n from $\mathbb{R}^{n+1}$ by this equiavlence relation (see Section 10.4). Define the maps $\phi_i : U_i \to \mathbb{R}^n$ by

$$\phi_i([(x_1, x_2, \dots, x_{n+1})]) = (y^1_{(i)}, y^2_{(i)}, \dots, y^n_{(i)})$$

where

$$y^1_{(i)} = \frac{x_1}{x_i}, \dots, y^{i-1}_{(i)} = \frac{x_{i-1}}{x_i}, y^i_{(i)} = \frac{x_{i+1}}{x_i}, \dots, y^n_{(i)} = \frac{x_{n+1}}{x_i}.$$

In other words, the ratio $x_i/x_i = 1$ is omitted. The coordinates y^i are clearly independent of the choice of representive in $[(x_1, x_2, \ldots, x_{n+1})]$, for if $x_i = \lambda x'_i$ for $i = 1, 2, \ldots, n+1$ then $x'_j/x'_i = \lambda x_j/\lambda x_i = x_j/x_i$ for all j. Since at least one $x_i \neq 0$ on every line every representative lies in at least one U_i and the charst $(U_i, \phi_i; y^j_{(i)})$ cover all of P^n. If $[(x_1, x_2, \ldots, x_{n+1})] \in U_i \cap U_j$ (so that $x_i \neq 0$ and $x_j \neq 0$) the coordinate transformation functions $\phi_j \circ \phi_i^{-1}$ are given by

$$y^k_{(j)} = \frac{x_k}{x_j} = \frac{x_k}{x_i}\frac{x_i}{x_j} = y^k_{(i)} y^i_{(j)}$$

which are clearly C^∞ and invertible, since

$$y^k_{(i)} = y^k_{(j)} y^j_{(i)}.$$

This is therefore a differentiable structure making P^n into an n-dimensional differentiable manifold.

Problem 15.4 Define the complex projective n-space CP^n in a similar way to Example 10.15 as lines in $\mathbb{C}^{n+1}$ of the form $\lambda(z^0, z^1, \ldots, z^n)$ where $\lambda, z^0, \ldots, z^n \in \mathbb{C}$. Show that CP^n is a differentiable (real) manifold of dimension $2n$.

Solution: Since $\mathbb{C}$ is identical with the manifold $\mathbb{R}^2$ (the complex plane $(x, y) \equiv x + iy$, it is clear that $\mathbb{C}^n$ is a manifold of dimension $2n$. The complex projective plane CP^n is defined similarly to the real case in Problem 15.3, with some differences of notation. Points of CP^n are denoted $[(z^0, z^1, \ldots, z^n)]$, and again we define open sets U_i by $z^i \neq 0$ which cover all of CP^n. Set $\phi_i : U_i \to \mathbb{C}^n \equiv \mathbb{R}^{2n}$ to be the maps $\phi_i : [(z^0, z^1, \ldots, z^n)] \mapsto (\zeta^1_{(i)}, \zeta^2_{(i)}, \ldots, \zeta^n_{(i)})$ where

$$\zeta^1_{(i)} = \frac{z^0}{z^i}, \ldots, \zeta^i_{(i)} = \frac{z^{i-1}}{z^i}, \zeta^{i+1}_{(i)} = \frac{z^{i+1}}{z^i}, \zeta^n_{(i)} = \frac{z^n}{z^i}.$$

The coordinate transformation maps $\zeta^k_{(j)} = \zeta^k_{(i)}\zeta^i_{(j)}$ are again C^∞ and invertible, making CP^n into a manifold of dimension $2n$.

Problem 15.5 Let $\mathbb{R}'$ be the manifold consisting of $\mathbb{R}$ with differentiable structure generated by the chart $(\mathbb{R}; y = x^3)$. Show that the identity map $\mathrm{id}_{\mathbb{R}} : \mathbb{R}' \to \mathbb{R}$ is a homeomorphism, which is not a diffeomorphism.
[$**$ NOTE: In the text, the word 'differentiable' before 'homeomorphism' should be omitted. Alternatively we could define the chart as $\mathbb{R}; y = x^{1/3}$ and the sentence stands.]

Solution: The chart defining the manifold $\mathbb{R}'$ is $(\mathbb{R}, \phi; y)$ where $y = x^3$. The map $f = \mathrm{id}_{\mathbb{R}} : \mathbb{R}' \to \mathbb{R}$ has coordinate representation

$$\hat{f}(y) = f \circ \phi^{-1}(y) = f \circ y^{1/3} = y^{1/3}.$$

This map is one-to-one with inverse $\hat{f}^{-1}(y) = y^3$. Both $\hat{f}$ and $\hat{f}^{-1}(y)$ are clearly continuous and therefore f is a homeomorphism. However f is not differentiable at the origin, for

$$\frac{\mathrm{d}\hat{f}(y)}{\mathrm{d}y} = \frac{1}{3}y^{-2/3} \to \infty \quad \text{as } y \to 0.$$

Had the original differentiable structure been defined as $y = x^{1/3}$ then f would have coordinate representation y^3 and it would have been a differentiable homeomorphism. However in that case its inverse f^{-1} would not be differentiable at $y = 0$, so it again would not be a diffeomorphism.

Problem 15.6 **Show that the set of real $m \times n$ matrices $M(m, n; \mathbb{R})$ is a manifold of dimension mn. Show that the matrix multiplication map $M(m, k; \mathbb{R}) \times M(k, n; \mathbb{R}) \to M(m, n; \mathbb{R})$ is differentiable.**

Solution: Just as in Example 15.5 the manifold structure is defined by the one-to-one surjective map $\phi : M(m, n; \mathbb{R}) \to \mathbb{R}^{mn}$ defined by

$$\phi(A) = (a_{11}, \dots, a_{1n}, a_{22}, \dots, a_{2n}, a_{31}, \dots, a_{mn}).$$

The product map isdefined by

$$\pi(AB) = [\textstyle\sum_{r=1}^{k} a_{ir} b_{ru}]$$

where we take indices $i, j = 1, \dots, m$, indices $r, s = 1, \dots, k$ and indices $u, v = 1, \dots n$. The product manifold $M(m, k; \mathbb{R}) \times M(k, n; \mathbb{R})$ is a manifold which can be identified with $\mathbb{R}^{mk+kn}$, using the standard coordinates

$$(A, B) \longrightarrow (a_{11}, \dots, a_{mk}, b_{11}, \dots, b_{kn}).$$

in these coordinates, the product map π is differentiable for

$$\hat{\pi}_{jv}(A, B) = \textstyle\sum_{s=1}^{k} a_{js} b_{sv}$$

are continuous functions which are differentiable, since

$$\frac{\partial \hat{\pi}_{jv}}{\partial a_{ir}} = \textstyle\sum_{s=1}^{k} \delta_{ij} \delta_{sr} b_{sv} = \delta_{ij} b_{rv}$$
$$\frac{\partial \hat{\pi}_{jv}}{\partial b_{ru}} = \textstyle\sum_{s=1}^{k} a_{js} \delta_{sr} \delta_{uv} = a_{jr} \delta_{uv}.$$

Problem 15.7 **Let $\gamma : \mathbb{R} \to \mathbb{R}^2$ be the curve $x = 2t + 1$, $y = t^2 - 3t$. Show that at an arbitrary parameter value t the tangent vector to the curve is $X_{\gamma(t)} = \dot{\gamma} = 2\partial_x + (2t - 3)\partial_y = 2\partial_x + (x - 4)\partial_y$. If $f : \mathbb{R}^2 \to \mathbb{R}$ is the**

function $f = x^2 - y^2$, write f as a function of t along the curve and verify the identities

$$X_{\gamma(t)} f = \frac{\mathrm{d}f(t)}{\mathrm{d}t} = \langle (\mathrm{d}f)_{\gamma(t)}, X_{\gamma(t)} \rangle = C_1^1 (\mathrm{d}f)_{\gamma(t)} \otimes X_{\gamma(t)}.$$

Solution: If $f : \mathbb{R}^2 \to \mathbb{R}$ is any differentiable function then

$$\begin{aligned} X_{\gamma(t)}(f) = \dot\gamma(f) &= \frac{\mathrm{d}f \circ \gamma}{\mathrm{d}t} \\ &= \frac{\mathrm{d}f(2t+1, t^2 - 3t)}{\mathrm{d}t} \\ &= 2\frac{\partial f}{\partial x} + (2t-3)\frac{\partial f}{\partial y}. \end{aligned}$$

Hence

$$X_{\gamma(t)} = 2\partial_x + (2t-3)\partial_y = 2\partial_x + (x-4)\partial_y$$

since $2t = x - 1$.

If $f = x^2 - y^2$ then

$$\begin{aligned} X_{\gamma(t)}(f) &= 2 \cdot 2x + (2t-3)(-2y) \\ &= 2(4t+2) - 2(2t-3)(t^2-3t) \\ &= -4t^3 + 18t^2 - 10t + 4. \end{aligned}$$

The bracket product is

$$\begin{aligned} \langle (\mathrm{d}f)_{\gamma(t)}, X_{\gamma(t)} \rangle &= \langle 2x\mathrm{d}x - 2y\mathrm{d}y, \partial_x + (x-4)\partial_y \rangle_{\gamma(t)} \\ &= \big(4x - 2y(x-4)\big)_{\gamma(t)} \\ &= 4(2t+1) - 2(t^2-3t)(2t-3) \\ &= -4t^3 + 18t^2 - 10t + 4, \end{aligned}$$

as above. The tensor $T = (\mathrm{d}f)_{\gamma(t)} \otimes X_{\gamma(t)}$ has basis decomposition

$$\begin{aligned} T &= \big[(2x\mathrm{d}x - 2y\mathrm{d}y) \otimes (2\partial_x + (x-4)\partial_y)\big]_{\gamma(t)} \\ &= 4(2t+1)\mathrm{d}x \otimes \partial_x - 4(t^2-3t)\mathrm{d}y \otimes \partial_x \\ &\quad + 2(2t+1)(2t-3)\mathrm{d}x \otimes \partial_y - 2(t^2-3t)(2t-3)\mathrm{d}y \otimes \partial_y. \end{aligned}$$

Hence

$$C_1^1 (\mathrm{d}f)_{\gamma(t)} \otimes X_{\gamma(t)} = T_1^1 + T_2^2 = 4(2t+1) - 2(t^2-3t)(2t-1) = X_{\gamma(t)} f.$$

Problem 15.8 **Let $x^1 = x$, $x^2 = y$, $x^3 = z$ be ordinary rectangular cartesian coordinates in $\mathbb{R}^3$, and let $x'^1 = r$, $x'^2 = \theta$, $x'^3 = \phi$ be the usual**

transformation to polar coordinates.

(a) Calculate the Jacobian matrices $[\partial x^i/\partial x'^j]$ **and** $[\partial x'^i/\partial x^j]$**.**

(b) In polar coordinates, work out the components of the covariant vector fields having components in rectangular coordinates (i) $(0,\,0,\,1)$**, (ii)** $(1,\,0,\,0)$**, (iii)** $(x,\,y,\,z)$**.**

(c) In polar coordinates, what are the components of the contravariant vector fields whose components in rectangular coordinates are (i) $(x,\,y,\,z)$**, (ii)** $(0,\,0,\,1)$**, (iii)** $(-y,\,x\,,0)$**.**

(d) If g_{ij} **is the covariant tensor field whose components in rectangular coordinates are** δ_{ij}**, what are its components** g'_{ij} **in polar coordinates?**

Solution: We have

$$\begin{aligned} x^1 x &= r\sin\theta\cos\phi & x'^1 &= r = \sqrt{x^2+y^2+z^2} \\ x^2 y &= r\sin\theta\sin\phi & x'^2 &= \theta = \arccos\frac{z}{r} \\ x^3 z &= r\cos\theta & x'^3 &= \phi = \arctan\frac{y}{x}. \end{aligned}$$

(a) The Jacobian matrices are easily computed:

$$\left[\frac{\partial x^i}{\partial x'^j}\right] = \begin{pmatrix} \sin\theta\cos\phi & r\cos\theta\cos\phi & -r\sin\theta\sin\phi \\ \sin\theta\sin\phi & r\cos\theta\sin\phi & r\sin\theta\cos\phi \\ \cos\theta & -r\sin\theta & 0 \end{pmatrix}$$

and

$$\begin{aligned} \left[\frac{\partial x'^i}{\partial x^j}\right] &= \begin{pmatrix} x/r & y/r & z/r \\ xz/r^2\sqrt{x^2+y^2} & yz/r^2\sqrt{x^2+y^2} & -\sqrt{x^2+y^2}/r^2 \\ -y/(x^2+y^2) & x/(x^2+y^2) & 0 \end{pmatrix} \\ &= \begin{pmatrix} \sin\theta\cos\phi & \sin\theta\sin\phi & \cos\theta \\ (\cos\theta\cos\phi)/r & (\cos\theta\sin\phi)/r & -\sin\theta/r \\ -\sin\phi/r\sin\theta & \cos\phi/r\sin\theta & 0 \end{pmatrix} \end{aligned}$$

(b) (i) If $\alpha_i = (0,0,1)$ then

$$\alpha'_i = \alpha_j \frac{\partial x^j}{\partial x'^i} = \frac{\partial x^3}{\partial x'^i} = (\cos\theta,\, -r\sin\theta,\, 0).$$

Alternatively, $\alpha = \mathrm{d}z = \cos\theta \mathrm{d}r - r\sin\theta\mathrm{d}\theta$.
(ii) If $\alpha_i = (1,0,0)$ then

$$\alpha'_i = \frac{\partial x^1}{\partial x'^i} = (\sin\theta\cos\phi,\, r\cos\theta\cos\phi,\, -r\sin\theta\sin\phi).$$

(iii) If $\alpha_i = (x,y,z)$ then

$$\alpha = x\mathrm{d}x + y\mathrm{d}y + z\mathrm{d}z = \tfrac{1}{2}\mathrm{d}(x^2+y^2+z^2) = \tfrac{1}{2}\mathrm{d}r^2 = r\mathrm{d}r.$$

Hence $\alpha'_i = (r,\, 0,\, 0)$. Of course, these components could also be found from the covariant transformation law of components as in (i) and (ii).

(c) (i) If $X^i = (x,\, y,\, z)$ then

$$\begin{aligned} X'^1 &= \frac{\partial x'^1}{\partial x^j} X^j = \frac{x^2}{r} + \frac{y^2}{r} + \frac{z^2}{r} = r, \\ X'^2 &= \frac{\partial x'^2}{\partial x^j} X^j = \frac{x^2 z}{r^2\sqrt{x^2+y^2}} + \frac{y^2 z}{r^2\sqrt{x^2+y^2}} - \frac{z\sqrt{x^2+y^2}}{r^2} = 0, \\ X'^3 &= \frac{\partial x'^3}{\partial x^j} X^j = -\frac{yx}{x^2+y^2} + \frac{xy}{x^2+y^2} = 0. \end{aligned}$$

(ii) If $X^i = (0, 0, 1)$then

$$X'^i = \frac{\partial x'^i}{\partial x^3} = (\cos\theta,\, -\frac{\sin\theta}{r},\, 0).$$

(iii) If $X^i = (-y,\, x,\, 0)$ then a similar calculation to part (i) results in $X'^i = (0,\, 0,\, 1)$.

(d) If $g_{ij} = \delta_{ij}$ then

$$g'_{i'j'} = g_{kl} \frac{\partial x^i}{\partial x'^{i'}} \frac{\partial x^j}{\partial x'^{j'}} = \sum_{i=1}^{3} \frac{\partial x^i}{\partial x'^{i'}} \frac{\partial x^i}{\partial x'^{j'}}.$$

Hence

$$\begin{aligned} g'_{11} &= \sin^2\theta\cos^2\phi + \sin^2\theta\sin^2\phi + \cos^2\theta &&= 1 \\ g'_{22} &= r^2\cos^2\theta\cos^2\phi + r^2\cos^2\theta\sin^2\phi + r^2\sin^2\theta &&= r^2 \\ g'_{33} &= r^2\sin^2\theta\sin^2\phi + r^2\sin^2\theta\cos^2\phi &&= r^2\sin^2\theta \end{aligned}$$

and all non-diagonal elements vanish,

$$g'_{12} = g'_{21} = g'_{13} = g'_{31} = g'_{23} = g'_{32} = 0.$$

Problem 15.9 **Show that the curve**

$$2x^2 + 2y^2 + 2xy = 1$$

can be converted by a rotation of axes to the standard form for an ellipse

$$x'^2 + 3y'^2 = 1.$$

If $x' = \cos\psi$**,** $y' = \frac{1}{\sqrt{3}}\sin\psi$ **is used as a parametrization of this curve, show that**

$$x = \frac{1}{\sqrt{2}}\Big(\cos\psi + \frac{1}{\sqrt{3}}\sin\psi\Big), \qquad y = \frac{1}{\sqrt{2}}\Big(-\cos\psi + \frac{1}{\sqrt{3}}\sin\psi\Big).$$

Compute the components of the tangent vector

$$X = \frac{\mathrm{d}x}{\mathrm{d}\psi}\partial_x + \frac{\mathrm{d}y}{\mathrm{d}\psi}\partial_y.$$

Show that if $f = xy$ then $X(f) = (2/\sqrt{3})\big(x^2 - y^2\big)$.

[** **NOTE: 'If $f = xy$' was accidentally omitted in the text.**]

Solution: A rotation

$$\begin{aligned} x &= x' \cos\alpha + y' \sin\alpha \\ y &= -x' \sin\alpha + y' \cos\alpha \end{aligned}$$

converts the equation $2x^2 + 2y^2 + 2xy = 1$ to

$$2x'^2(1 - \cos\alpha\sin\alpha) + 2y'^2(1 + \cos\alpha\sin\alpha) + 2x'y'(\cos^2\alpha - \sin^2\alpha) = 1.$$

Setting $\alpha = \pi/4$ we have $\cos\alpha = \sin\alpha = 1/\sqrt{2}$, and the equation becomes

$$x'^2 + 3y'^2 = 1.$$

In parametric form $x' = \cos\psi$, $y' = \frac{1}{\sqrt{3}}\sin\psi$ substitution in the transformation equations with $\cos\alpha = \sin\alpha = 1/\sqrt{2}$ results in

$$\begin{aligned} x &= \frac{1}{\sqrt{2}}\Big(\cos\psi + \frac{1}{\sqrt{3}}\sin\psi\Big), \\ y &= \frac{1}{\sqrt{2}}\Big(-\cos\psi + \frac{1}{\sqrt{3}}\sin\psi\Big), \end{aligned}$$

and

$$\begin{aligned} X &= \frac{\mathrm{d}x}{\mathrm{d}\psi}\partial_x + \frac{\mathrm{d}y}{\mathrm{d}\psi}\partial_y \\ &= \frac{1}{\sqrt{2}}\Big(-\sin\psi + \frac{1}{\sqrt{3}}\cos\psi\Big)\partial_x + \frac{1}{\sqrt{2}}\Big(\sin\psi + \frac{1}{\sqrt{3}}\cos\psi\Big)\partial_y \end{aligned}$$

so that, if $f = xy$ then

$$\begin{aligned} Xf &= \frac{1}{\sqrt{2}}\big(-\sin\psi + \frac{1}{\sqrt{3}}\cos\psi\big)y + \frac{1}{\sqrt{2}}\big(\sin\psi + \frac{1}{\sqrt{3}}\cos\psi\big)x \\ &= \frac{4}{3}\sin\psi\cos\psi \end{aligned}$$

on sustituting for x and y. A simple calculation gives that $x^2 - y^2 = (2/\sqrt{3})\cos\psi\sin\psi$, so that

$$X(f) = (2/\sqrt{3})\big(x^2 - y^2\big).$$

Problem 15.10 **Show that the tangent space $T_{(p,q)}(M \times N)$ at any point (p,q) of a product manifold $M \times N$ is naturally isomorphic to the direct sum of tangent spaces $T_p(M) \oplus T_q(N)$.**

Solution: The tangent space $T_p(M)$ is the set of all maps $X_p : \mathcal{F}_p(M) \to \mathbb{R}$ which are linear and satisfy the Leibnitz rule, Eqs, (15.3) and (15.4). In particular $T_{(p,q)}(M \times N)$ is the set of all maps $A_{(p,q)} : \mathcal{F}_{(p,q)}(M \times N) \to \mathbb{R}$ such that if $f, g : M \times N \to \mathbb{R}$ are differentiable functions at (p,q) then

$$A_{(p,q)}(af + bg) = aA_{(p,q)}f + bA_{(p,q)}g \tag{1}$$

$$A_{(p,q)}(fg) = f(p,q)A_{(p,q)}g + g(p,q)A_{(p,q)}f. \tag{2}$$

We wish to show that $T_{(p,q)}(M \times N) \cong T_p(M) \oplus T_q(N)$. From Sec. 3.4 $T_p(M) \oplus T_q(N)$ consists of all pairs (X_p, Y_q) where $X_p \in T_p(M)$, $Y_q \in T_q(N)$ with the addition laws

$$(X_p, Y_q) + a(X'_p, Y'_q) = (X_p + aX'_p, Y_q + aY'_q).$$

For each $(X_p, Y_q) \in T_p(M) \oplus T_q(N)$ define the action $\mathcal{F}_{(p,q)}(M \times N) \to \mathbb{R}$ by

$$(X_p, Y_q)f = X_p f_q + Y_q f_p$$

where functions $f_q : M \to \mathbb{R}$ and $f_p : N \to \mathbb{R}$ are defined by

$$f_q(p) = f_p(q) = f(p,q).$$

It is straightforward to show that these functions are differentiable at p and q respectively. Furthermore, if f and g belong to $\mathcal{F}_{(p,q)}(M \times N)$ then

$$(af + bg)_q = af_q + fg_q \qquad \text{and} \qquad (fg)_q = f_q g_q,$$

for

$$\begin{aligned}(af + bg)_q(p) &= (af + bg)(q,p) \\ &= af(p,g) + bg(p,q) \\ &= af_q(p) + bg_q(p)\end{aligned}$$

and

$$(fg)_q(p) = (fg)(p,q) = f(p,q)g(p,q) = f_q(p)g_q(p).$$

Similarly,

$$(af + bg)_p = af_p + fg_p \qquad \text{and} \qquad (fg)_p = f_p g_p.$$

Hence

$$\begin{aligned}(X_p, Y_q)(af + bg) &= X_p(af + bg)_q + Y_q(af + bg)_p \\ &= X_p(af_q + bg_q) + Y_q(af_p + bg_p) \\ &= aX_p f_q + bX_p g_q + aY_q f_p + bX_q g_p \\ &= a(X_p f_q + Y_q f_p) + b(X_p g_q + X_q g_p) \\ &= a(X_p, Y_q)f + b(X_p, Y_q)g\end{aligned}$$

and

$$\begin{aligned}(X_p, Y_q)(fg) &= X_p(fg)_q + Y_q(fg)_p \\ &= X_p(f_q g_q) + Y_q(f_p g_p) \\ &= f_q(p) X_p g_q + g_q(p) X_p f_q + f_p(q) Y_q g_p + g_p(q) Y_q f_p \\ &= f(p,q)(X_p g_q + Y_q g_p) + g(q,p)(X_p f_q + Y_q f_p) \\ &= f(p,q)(X_p, Y_q) g + g(p,q)(X_p, Y_q) f\end{aligned}$$

Hence $A_{(p,q)}$ satisfies (1) and (2). Furthermore the correspondence between $T_p(M) \oplus T_q(N)$ and $T_{(p,q)}(M \times N)$ is a vector space isomorphism, for

$$\begin{aligned}\big((X_p, Y_q) + a(X'_p, Y'_q)\big) f &= X_p f_q + Y_q f_p + a X'_p f_q + a Y'_q f_p \\ &= (X_p + a X'_p) f_q + (Y_q + a Y'_q) f_p \\ &= (X_p + a X'_p, Y_q + a Y'_q) f.\end{aligned}$$

It is a one-to-one correspondence since $(0,0)f = 0$, and must therefore be onto, since

$$\dim T_{(p,q)}(M \times N) = \dim(M \times N) = m + n = \dim(T_p(M) \oplus T_q(N)).$$

To construct the X_p, Y_p corresponding to any $A_{(p.q)} \in T_{(p,q)}(M \times N)$, let $f \in \mathcal{F}_p(M)$ and $g \in \mathcal{F}_q(N)$ be any differentiable functions at $p \in M$ and $q \in N$ resp. Set $\bar{f}, \bar{g} \in \mathcal{F}_{(q,p)}(M \times N)$ be defined by

$$\bar{f}(p,q) = f(p) \quad \forall q, \qquad \bar{g}(p,q) = g(q) \quad \forall p$$

and set X_p and Y_q to be tangent vectors at $p \in M$ and $q \in N$ defined by

$$X_p f = A_{(p,q)} \bar{f}, \qquad Y_q g = A_{(p,q)} \bar{g}.$$

These operators are linear and satisfy the Leibnitz rule and must of necessity be the tangent vectors such that $A_{(p,q)} = (X_p, Y_q)$ in the above definition.

Problem 15.11 On the unit 2-sphere express the vector fields ∂_x and ∂_y in terms of the polar coordinate basis ∂_θ and ∂_ϕ. Again in polar coordinates, what are the dual forms to these vector fields?

Solution: On the unit sphere

$$x = \sin\theta\cos\phi, \qquad y = \sin\theta\sin\phi, \qquad z = \cos\theta.$$

Expressing ∂_x and ∂_y in polar coordinates

$$\partial_x = \frac{\partial\theta}{\partial x}\partial_\theta + \frac{\partial\phi}{\partial x}\partial_\phi, \qquad \partial_y = \frac{\partial\theta}{\partial y}\partial_\theta + \frac{\partial\phi}{\partial y}\partial_\phi,$$

Using Problem 15.8 with $r = 1$ we have

$$\frac{\partial\theta}{\partial x} = \cos\theta\cos\phi \qquad \frac{\partial\theta}{\partial y} = \cos\theta\sin\phi$$
$$\frac{\partial\phi}{\partial x} = -\frac{\sin\phi}{\sin\theta} \qquad \frac{\partial\phi}{\partial y} = \frac{\cos\phi}{\sin\theta},$$

so that

$$\partial_x = \cos\theta\cos\phi\partial_\theta - \frac{\sin\phi}{\sin\theta}\partial_\phi$$
$$\partial_y = \cos\theta\sin\phi\partial_\theta + \frac{\cos\phi}{\sin\theta}\partial_\phi.$$

The dual basis $\varepsilon^1 = a_1\mathrm{d}\theta + a_2\mathrm{d}\phi$, $\varepsilon^2 = b_1\mathrm{d}\theta + b_2\mathrm{d}\phi$ satisfy

$$\langle\varepsilon^1, \partial_x\rangle = 1,\ \langle\varepsilon^1, \partial_y\rangle = 0, \qquad \langle\varepsilon^2, \partial_x\rangle = 0,\ \langle\varepsilon^2, \partial_y\rangle = 1.$$

Hence

$$a_1\cos\theta\cos\phi - a_2\frac{\sin\phi}{\sin\theta} = 1$$
$$a_1\cos\theta\sin\phi + a_2\frac{\cos\phi}{\sin\theta} = 0$$

which solves for

$$a_1 = \frac{\cos\phi}{\cos\theta}, \qquad a_2 = -\sin\theta\sin\phi.$$

Similarly,

$$b_1\cos\theta\cos\phi - b_2\frac{\sin\phi}{\sin\theta} = 0$$
$$b_1\cos\theta\sin\phi + b_2\frac{\cos\phi}{\sin\theta} = 1$$

has solution

$$b_1 = \frac{\sin\phi}{\cos\theta}, \qquad b_2 = \sin\theta\cos\phi.$$

Thus the dual basis is

$$\varepsilon^1 = \frac{\cos\phi}{\cos\theta}\mathrm{d}\theta - \sin\theta\sin\phi\mathrm{d}\phi,$$
$$\varepsilon^2 = \frac{\sin\phi}{\cos\theta}\mathrm{d}\theta + \sin\theta\cos\phi\mathrm{d}\phi.$$

Problem 15.12 **Express the vector field ∂_ϕ in polar coordinates (θ, ϕ) on the unit 2-sphere in terms of stereographic coordinates X and Y.**

Solution: On the unit sphere

$$X = \frac{x}{1-z} = \frac{\sin\theta\cos\phi}{1-\cos\theta} = \frac{2\sin\frac{1}{2}\theta\cos\frac{1}{2}\theta}{2\sin^2\frac{1}{2}\theta} = \cot\tfrac{1}{2}\theta\cos\phi$$
$$Y = \frac{y}{1-z} = \cdots = \cot\tfrac{1}{2}\theta\sin\phi$$

Hence

$$\begin{aligned}\partial_\phi &= \frac{\partial X}{\partial\phi}\partial_X + \frac{\partial Y}{\partial\phi}\partial_Y \\ &= -\cot\tfrac{1}{2}\theta\sin\phi\partial_X + \cot\tfrac{1}{2}\theta\sin\phi\partial_Y \\ &= -Y\partial_X + X\partial_Y \qquad .\end{aligned}$$

Problem 15.13 **Show that if $\rho_p = r_i(\mathrm{d}x^i)_p = \alpha^*\omega_{\alpha(p)}$ then the components are given by**

$$r_i = \frac{\partial y^a}{\partial x^i}\Big|_{\phi(p)} w_a$$

where $\omega_{\alpha(p)} = w_a(\mathrm{d}y^a)_{\alpha(p)}$.

If α is a diffeomorphism, define a map $\alpha_* : T_p^{(1,1)}(M) \to T_{\alpha(p)}^{(1,1)}(N)$ by setting

$$\alpha_* T\big(\omega_{\alpha(p)}, X_{\alpha(p)}\big) = T\big(\alpha^*\omega_{\alpha(p)}, \alpha_*^{-1}X_{\alpha(p)}\big)$$

and show that the components transform as

$$(\alpha_* T)^a{}_b = T^i{}_j \frac{\partial y^a}{\partial x^i}\frac{\partial x^j}{\partial y^b}.$$

Solution: From Eq. (15.18) we have that ρ_p is defined by

$$\langle\rho_p, X_p\rangle = \langle\omega_{\alpha(p)}, \alpha_* X_p\rangle.$$

Setting $\rho_p = r_i(\mathrm{d}x^i)_p$ and

$$X_p = X^i\Big(\frac{\partial}{\partial x^i}\Big)_p$$

we have, by Eq. (15.16)

$$\alpha_* X_p = X^i\frac{\partial y^a}{\partial x^i}\Big|_{\phi(p)}\Big(\frac{\partial}{\partial y^a}\Big)_{\alpha(p)}.$$

Hence

$$r_i X^i = w_a X^i \frac{\partial y^a}{\partial x^i}\Big|_{\phi(p)}$$

and, as X^i are arbitrary, it follows that

$$r_i = \frac{\partial y^a}{\partial x^i}\bigg|_{\phi(p)} w_a.$$

If α is a homeomorphism then the inverse map α^{-1} is represented by differentiable functions $x^i = x^i(ya)$. Set

$$\alpha_* T = (\alpha_* T)^a{}_b \Big(\frac{\partial}{\partial y^a}\Big)_{\alpha(p)} \otimes \big(\mathrm{d}y^b\big)_{\alpha(p)}$$

and we have

$$\alpha_* T\big(\omega_{\alpha(p)}, X_{\alpha(p)}\big) = (\alpha_* T)^a{}_b w_a X^b$$

where

$$X_{\alpha(p)} = X^b \Big(\frac{\partial}{\partial y^b}\Big)_{\alpha(p)}.$$

Using the inverse form of Eq. (15.16) we have

$$\big(\alpha^{-1}\big)_* X_{\alpha(p)} = X^b \frac{\partial x^i}{\partial y^b} \Big(\frac{\partial}{\partial x^i}\Big)_p.$$

Using the transformation of components of ω derived above, the defining equation for $\alpha_* T$ is thus represented in components by

$$(\alpha_* T)^a{}_b w_a X^b = T^i{}_j w_a \frac{\partial y^a}{\partial x^i} X^b \frac{\partial x^j}{\partial y^b}$$

and since w_a and X^b are arbitrary, we have the required formula

$$(\alpha_* T)^a{}_b = T^i{}_j \frac{\partial y^a}{\partial x^i} \frac{\partial x^j}{\partial y^b}.$$

Problem 15.14 **If $\gamma : \mathbb{R} \to M$ is a curve on M and $p = \gamma(t_0)$ and $\alpha : M \to N$ is a differentiable map show that**

$$\alpha_* \dot{\gamma}_p = \dot{\sigma}_{\alpha(p)} \qquad \textbf{where} \qquad \sigma = \alpha \circ \gamma : \mathbb{R} \to N.$$

Solution: Let $f : N \to \mathbb{R}$ be an arbitrary differentiable function. Then

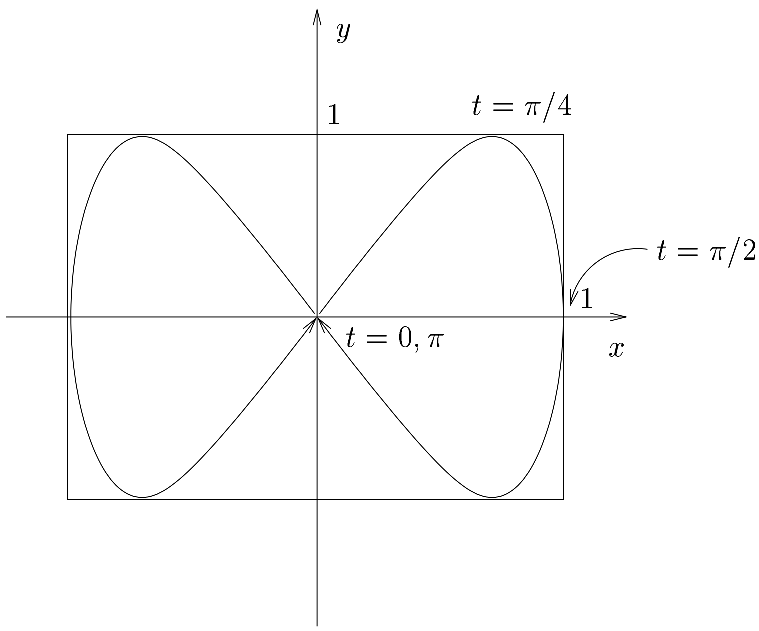
$$\begin{aligned}
\dot{\sigma}_{\alpha(p)}(f) &= \frac{\mathrm{d}}{\mathrm{d}t} f(\sigma(t))\Big|_{t=t_0} \\
&= \frac{\mathrm{d}}{\mathrm{d}t} f(\alpha \circ \gamma(t))\Big|_{t=t_0} \\
&= \frac{\mathrm{d}}{\mathrm{d}t} (f \circ \alpha)\big(\gamma(t)\big)\Big|_{t=t_0} \\
&= \dot{\gamma}(f \circ \alpha)\Big|_{t=t_0} \\
&= \dot{\gamma}_p(f \circ \alpha) \\
&= \alpha_* \dot{\gamma}_p(f).
\end{aligned}$$

Hence

$$\alpha_* \dot{\gamma}_p = \dot{\sigma}_{\alpha(p)}.$$

Problem 15.15 Is the map $\alpha : \mathbb{R} \to \mathbb{R}^2$ given by $x = \sin t$, $y = \sin 2t$ (i) an immersion, (ii) an embedded submanifold?

Solution: The map $x = \sin t$, $y = \sin 2t$ can be represented graphically as a figure of eight, shown below. (i) It is an immersion, for



$$\alpha_* \frac{\mathrm{d}}{\mathrm{d}t} = \cos t \partial_x + 2 \cos 2t \partial_y$$

is equal to the zero vector if and only if $cost = cos2t = 0$, i.e. $t = (n + \frac{1}{2})\pi$ and $2t = (m + \frac{1}{2})\pi$ for some integers n and m. This is clearly impossible for it would imply $2n + 1 = m + \frac{1}{2}$.

(ii) The map is not an embedding, as it is not injective for $\alpha(0) = \alpha(\pi) = (0, 0)$.

Problem 15.16 Show that the map $\alpha : \dot{\mathbb{R}}^2 \to \mathbb{R}^3$ defined by

$$u = x^2 + y^2, \qquad v = 2xy, \qquad w = x^2 - y^2$$

is an immersion. Is it an embedded submanifold?

Evaluate $\alpha^*(u\mathrm{d}u + v\mathrm{d}v + w\mathrm{d}w)$ and $\alpha_*(\partial_x)_{(a,b)}$. Find a vector field X on $\dot{\mathbb{R}}^2$ for which $\alpha_* X$ is not a well-defined vector field.

Solution: The Jacobian of this map is

$$\begin{pmatrix} \partial u/\partial x & \partial u/\partial x \\ \partial v/\partial x & \partial v/\partial x \\ \partial w/\partial x & \partial w/\partial x \end{pmatrix} = \begin{pmatrix} 2x & 2y \\ 2y & 2x \\ 2x & -2y \end{pmatrix}.$$

Thus $\alpha_* : T_{(x,y)}(\dot{\mathbb{R}}^2) \to T_{(u,v,w)}(\mathbb{R}^3)$ is not injective at (x, y) only if the two columns of this equation are linearly dependent, i.e. there exists a, b not both zero such that

$$ax + by = 0 \tag{1}$$
$$ay + bx = 0 \tag{2}$$
$$ax - by = 0 \tag{3}$$

If $y \neq 0$ then taking Eq. (3) from Eq. (1) gives $b = 0$. From Eq. (2) we then have $ay = 0$, whence $a = 0$, so that a and b are not both zero. If $y = 0$ then $x \neq 0$ (else $(x, y) = (0, 0) \notin \dot{\mathbb{R}}^2$), and Eqns. (1)and (2) give that $a = b = 0$. Hence the map α is everywhere injective on $\dot{\mathbb{R}}^2$, and α is an immersion. It is not however an embedded submanifold since $\alpha(-x, -y) = \alpha(x, y)$ for all x, y.

$$\begin{aligned} \alpha^*(u\mathrm{d}u + v\mathrm{d}v + w\mathrm{d}w) &= (x^2 + y^2)(2x\mathrm{d}x + 2y\mathrm{d}y) \\ &\quad + 2xy(2y\mathrm{d}x + 2x\mathrm{d}y) \\ &\quad + (x^2 - y^2)(2x\mathrm{d}x - 2y\mathrm{d}y) \\ &= 4x(x^2 + y^2)\mathrm{d}x + 4y(x^2 + y^2)\mathrm{d}y. \end{aligned}$$

$$\begin{aligned} \alpha_*(\partial_x)_{(a,b)} &= \left(\frac{\partial u}{\partial x}\right)_{(a,b)} \partial_u + \left(\frac{\partial v}{\partial x}\right)_{(a,b)} \partial_v + \left(\frac{\partial w}{\partial x}\right)_{(a,b)} \partial_w \\ &= 2a\partial_u + 2b\partial_v + 2a\partial_w \end{aligned}$$

The vector field $\alpha_* \partial_x$ is not well defined, for $\alpha(a, b) = \alpha(-a, -b)$ but

$$\alpha_*(\partial_x)_{(a,b)} \neq \alpha_*(\partial_x)_{(-a,-b)}.$$

Problem 15.17 Show that the components of the Lie product $[X, Y]^k$ given by Eq. (15.25) transform as a contravariant vector field under a coordinate transformation $x'^j(x^i)$.

Solution: The formula for the components of the Lie product $[X, Y]^k$ given by Eq. (15.25) is

$$[X, Y]^k = X^i Y^k{}_{,i} - Y^i X^k{}_{,i}.$$

The contravariant transformation law of components gives

$$X'^{i'} = X^i \frac{\partial x'^{i'}}{\partial x^i}, \quad Y'^{k'} = Y^k \frac{\partial x'^{k'}}{\partial x^k}$$

so that

$$\begin{aligned}
[X,Y]'^{k'} &= X'^{i'} Y'^{k'}{}_{,i'} - Y'^{i'} X'^{k'}{}_{,i'} \\
&= X^i \frac{\partial x'^{i'}}{\partial x^i} \Big(Y^k \frac{\partial x'^{k'}}{\partial x^k} \Big)_{,j} \frac{\partial x^j}{\partial x'^{i'}} - Y^i \frac{\partial x'^{i'}}{\partial x^i} \Big(X^k \frac{\partial x'^{k'}}{\partial x^k} \Big)_{,j} \frac{\partial x^j}{\partial x'^{i'}} \\
&= X^i \Big(Y^k{}_j \frac{\partial x'^{k'}}{\partial x^k} + Y^k \frac{\partial^2 x'^{k'}}{\partial x^k \partial x^j} \Big) \delta^j_j - Y^i \Big(X^k{}_j \frac{\partial x'^{k'}}{\partial x^k} + X^k \frac{\partial^2 x'^{k'}}{\partial x^k \partial x^j} \Big) \delta^j_j \\
&= \big(X^i Y^k{}_{,i} - Y^i X^k{}_{,i} \big) \frac{\partial x'^{k'}}{\partial x^k} \\
&= [X,Y]^k \frac{\partial x'^{k'}}{\partial x^k},
\end{aligned}$$

where we have used

$$\frac{\partial x'^{i'}}{\partial x^i} \frac{\partial x^j}{\partial x'^{i'}} = \delta^j_j$$

and

$$\big(X^i Y^k - Y^i X^k \big) \frac{\partial^2 x'^{k'}}{\partial x^k \partial x^j} = X^i Y^k \Big(\frac{\partial^2 x'^{k'}}{\partial x^k \partial x^j} - \frac{\partial^2 x'^{k'}}{\partial x^j \partial x^k} \Big) = 0.$$

Problem 15.18 Show that the Jacobi identity can be written

$$\mathcal{L}_{[X,Y]} Z = \mathcal{L}_X \mathcal{L}_Y Z - \mathcal{L}_Y \mathcal{L}_X Z,$$

and this property extends to all tensors T:

$$\mathcal{L}_{[X,Y]} T = \mathcal{L}_X \mathcal{L}_Y T - \mathcal{L}_Y \mathcal{L}_X T.$$

Solution: Using the fromula (15.31) for Lie derivative, $\mathcal{L}_X Y = [X,Y]$, the Jacobi identity

$$[[X,Y],Z] + [[Y,Z],X] + [[Z,X],Y] = 0$$

reads

$$\begin{aligned}
\mathcal{L}_{[X,Y]} Z &= -[[Y,Z],X] - [[Z,X],Y] \\
&= [X,[Y,Z]] - [Y,[X,Z]] \\
&= \mathcal{L}_X(\mathcal{L}_Y Z) - \mathcal{L}_Y(\mathcal{L}_X Z) \\
&= \mathcal{L}_X \mathcal{L}_Y Z - \mathcal{L}_Y \mathcal{L}_X Z.
\end{aligned}$$

(i) For tensors of type (0,0), $T = f : M \to \mathbb{R}$, the definition $\mathcal{L}_X f = Xf$, $\mathcal{L}_Y f = Yf$, gives

$$(\mathcal{L}_X \mathcal{L}_Y - \mathcal{L}_Y \mathcal{L}_X) f = X(Yf) - Y(Xf) = [X, Y] f$$

by definition of Lie bracket, Eq. (15.20).
(ii) For tensors of type $(1, 0)$, the result follows from the above on setting $T = Z$.
(iii) For tensors of type $(0, 1)$, $T = \omega$, a covector field on M, we use Eq. (15.38):

$$\begin{aligned}\langle(\mathcal{L}_X \mathcal{L}_Y - \mathcal{L}_Y \mathcal{L}_X)\omega, Z\rangle &= X\langle\mathcal{L}_Y \omega, Z\rangle - \langle\mathcal{L}_Y \omega, \mathcal{L}_X Z\rangle \\ &\quad - Y\langle\mathcal{L}_X \omega, Z\rangle - \langle\mathcal{L}_X \omega, \mathcal{L}_Y Z\rangle \\ &= X(Y\langle\omega, Z\rangle - \langle\omega, \mathcal{L}_Y Z\rangle) - Y\langle\omega, \mathcal{L}_X Z\rangle + \langle\omega, \mathcal{L}_Y \mathcal{L}_X Z\rangle \\ &\quad - Y(X\langle\omega, Z\rangle - \langle\omega, \mathcal{L}_X Z\rangle) + X\langle\omega, \mathcal{L}_Y Z\rangle - \langle\omega, \mathcal{L}_X \mathcal{L}_Y Z\rangle \\ &= [X, Y]\langle\omega, Z\rangle - \langle\omega, \mathcal{L}_{[X,Y]} Z\rangle \\ &= \langle\mathcal{L}_{[X,Y]}\omega, Z\rangle,\end{aligned}$$

by Eq. (15.38). Since Z is an arbitrary vector field we have

$$\mathcal{L}_{[X,Y]}\omega = (\mathcal{L}_X \mathcal{L}_Y - \mathcal{L}_Y \mathcal{L}_X)\omega.$$

(iv) Every tensor T of type (r, s) with 'order' $r + s > 1$ can be made up of a sum of tensor products $A \otimes B$ where A and B have lower orders. Assume that the relation holds for A and B, and we prove the result for $A \times B$, using Eq. (15.35):

$$\mathcal{L}_X(A \otimes B) = A \otimes (\mathcal{L}_X B) + (\mathcal{L}_X A) \otimes B.$$

The inductive prove proceeds as follows:

$$\begin{aligned}\mathcal{L}_{[X,Y]}(A \otimes B) &= A \otimes \mathcal{L}_{[X,Y]} B + \mathcal{L}_{[X,Y]} A \otimes B \\ &= A \otimes (\mathcal{L}_X \mathcal{L}_Y - \mathcal{L}_Y \mathcal{L}_X) B + (\mathcal{L}_X \mathcal{L}_Y - \mathcal{L}_Y \mathcal{L}_X) A \otimes B\end{aligned}$$

by the induction hypothesis

$$\begin{aligned}&= \mathcal{L}_X(A \otimes \mathcal{L}_Y B) - \mathcal{L}_X A \otimes \mathcal{L}_Y B - \mathcal{L}_Y(A \otimes \mathcal{L}_X B) + \mathcal{L}_Y A \otimes \mathcal{L}_X B \\ &\quad + \mathcal{L}_X(\mathcal{L}_Y A \otimes B) - \mathcal{L}_Y A \otimes \mathcal{L}_X B - \mathcal{L}_Y(\mathcal{L}_X A \otimes B) + \mathcal{L}_X A \otimes \mathcal{L}_Y B\end{aligned}$$

using Eq. (15.35)

$$\begin{aligned}&= \mathcal{L}_X(A \otimes \mathcal{L}_Y B) - \mathcal{L}_Y(A \otimes \mathcal{L}_X B) + \mathcal{L}_X(\mathcal{L}_Y A \otimes B) - \mathcal{L}_Y(\mathcal{L}_X A \otimes B) \\ &= \mathcal{L}_X \mathcal{L}_Y(A \otimes B) - \mathcal{L}_X(\mathcal{L}_Y A \otimes B) - \mathcal{L}_Y \mathcal{L}_X(A \otimes B) + \mathcal{L}_Y(\mathcal{L}_X A \otimes B) \\ &\quad + \mathcal{L}_X \mathcal{L}_Y(A \otimes B) - \mathcal{L}_X(A \otimes \mathcal{L}_Y B) - \mathcal{L}_Y \mathcal{L}_X(A \otimes B) + \mathcal{L}_Y(A \otimes \mathcal{L}_X B) \\ &= 2\mathcal{L}_X \mathcal{L}_Y(A \otimes B) - 2\mathcal{L}_Y \mathcal{L}_X(A \otimes B) - (\mathcal{L}_X \mathcal{L}_Y A - \mathcal{L}_Y \mathcal{L}_X A) \otimes B \\ &\quad - A \otimes (\mathcal{L}_X \mathcal{L}_Y B - \mathcal{L}_Y \mathcal{L}_X B) \\ &= 2\mathcal{L}_X \mathcal{L}_Y(A \otimes B) - 2\mathcal{L}_Y \mathcal{L}_X(A \otimes B) - (\mathcal{L}_{[X,Y]} A) \otimes B - A \otimes (\mathcal{L}_{[X,Y]} B)\end{aligned}$$

by the induction hypothesis. Hence, using Eq. (15.35) with X replaced by $[X,Y]$

$$2\mathcal{L}_{[X,Y]}(A\otimes B) = 2\mathcal{L}_X\mathcal{L}_Y(A\otimes B) - 2\mathcal{L}_Y\mathcal{L}_X(A\otimes B)$$

which gives the result on dividing by 2.

Problem 15.19 **Let $\alpha : M \to N$ be a diffeomorphism between manifolds M and N and X a vector field on M that generates a local one-parameter group of transformations σ_t on M. Show that the vector field $X' = \alpha_* X$ on N generates the local flow $\sigma'_t = \alpha \circ \sigma_t \circ \alpha^{-1}$.**

Solution: First of all, we must show the local flow conditions (i') and (ii') hold for the transformations $\sigma'_t = \alpha \circ \sigma_t \circ \alpha^{-1}$:
(i) Since σ_t is a diffeomorphism of an open neighbourhood $U \to \sigma_t(U)$ and α is a diffeomorphism, it follows that the composition $\sigma'_t = \alpha\circ\sigma_t\circ\alpha^{-1}$ is a diffeomorphism $\alpha(U) \to \alpha\circ\sigma_t(U)$.
(ii') For all $t, s, t+s \in I_\epsilon = (-\epsilon, \epsilon)$, from $\sigma_{t+s}(p) = \sigma_t \circ \sigma_s(p)$ it follows that

$$\begin{aligned}
\sigma'_{t+s}(\alpha(p)) &= \alpha\circ\sigma_t\circ\alpha^{-1}(\alpha(p)) \\
&= \alpha\circ\sigma_t\circ\sigma_s(p) \\
&= \alpha\circ\sigma_t\circ\alpha^{-1}\circ\alpha\circ\sigma_s\alpha^{-1}(\alpha(p)) \\
&= \sigma'_t\circ\sigma'_s(\alpha(p)).
\end{aligned}$$

Finally this local one-parameter group generates X' in a neighbourhood of any point $\alpha(p)$ where $p \in U$, for

$$\begin{aligned}
X'_{\alpha(p)}f &= (\alpha_* X)_{\alpha(p)}f \\
&= X_p(f\circ\alpha) \\
&= \frac{\mathrm{d}}{\mathrm{d}t}\big(f\circ\alpha(\sigma_t(p))\big)\Big|_{t=0} \\
&= \frac{\mathrm{d}}{\mathrm{d}t}\big(f\circ\alpha\circ\sigma_t\circ\alpha^{-1}(\alpha(p))\big)\Big|_{t=0} \\
&= \frac{\mathrm{d}}{\mathrm{d}t}\big(f\circ\sigma'_t(\alpha(p))\big)\Big|_{t=0} \qquad .
\end{aligned}$$

Problem 15.20 **For any real positive number n show that the vector field $X = x^n\partial_x$ is differentiable on the manifold $\mathbb{R}^+$ consisting of the positive real line $\{x \in \mathbb{R} \,|\, x > 0\}$. Why is this not true in general on the entire real line $\mathbb{R}$? As done for the case $n = 2$ in Example 15.13, find the maximal one-parameter subgroup σ_t generated by this vector field at any point $x > 0$.**

Solution: $X = x^n\partial_x$ is C^∞ on $\mathbb{R}^+$ for if f is any differentiable function on $\mathbb{R}^+$ then $Xf(x) = x^n f'(x)$ is C^∞ since it is a product of two C^∞ functions on $\mathbb{R}^+$, since all derivatives of x^n exist

$$\frac{\mathrm{d}^m}{\mathrm{d}x^m}x^n = n(n-1)\dots(n-m+1)x^{n-m}.$$

If n is not an integer then x^n is not differentiable at $x = 0$. For example, if $n = \frac{1}{2}$ then

$$\frac{\mathrm{d}}{\mathrm{d}x}\big(x^{1/2}f'(x)\big) = \frac{1}{2x^{1/2}}f'(x) + x^{1/2}f''(x)$$

which approaches $\pm\infty$ as $x \to 0$ if $f'(x) \neq 0$. Henc X is not smooth on all of $\mathbb{R}$.

Solving the differential equation

$$\frac{\mathrm{d}x}{\mathrm{d}t} = x^n \quad \Longrightarrow \quad x^{-n}\frac{\mathrm{d}x}{\mathrm{d}t} = 1$$

can be achieved by integrating both sides from 0 to t,

$$\int_0^t \frac{1}{1-n}\frac{\mathrm{d}x^{-n+1}}{\mathrm{d}t}\mathrm{d}t = t,$$

i.e.

$$\left.\frac{x^{1-n}}{1-n}\right]_0^t = t$$

and for any point x_0, writing $x = \sigma_t(x_0)$ we have

$$\frac{1}{1-n}\big[\big(\sigma_t(x_0)\big)^{1-n} - x_0^{1-n}\big] = t.$$

Hence

$$\sigma_t(x) = \big((1-n)t + x^{1-n}\big)^{1/(1-n)}.$$

It is straightforward to check that

$$\frac{\mathrm{d}}{\mathrm{d}t}f(\sigma_t(x))\Big|_{t=0} = f'(x)\frac{\mathrm{d}\sigma_t}{\mathrm{d}t}\Big|_{t=0} = x^n f'(x).$$

The case $n = 1$ is slightly different, and has the solution

$$\sigma_t(x) = x\mathrm{e}^t.$$

Problem 15.21 **On the manifold $\mathbb{R}^2$ with coordinates (x, y), let X be the vector field $X = -y\partial_x + x\partial_y$. Determine the integral curve through any point (x, y), and the one-parameter group generated by X.**

Find coordinates (x', y') such that $X = \partial_{x'}$.

Solution: The integral curves of X are defined by the differential equations

$$\frac{\mathrm{d}x}{\mathrm{d}t} = -y, \qquad \frac{\mathrm{d}y}{\mathrm{d}t} = x.$$

From these equations $\mathrm{d}^2x/\mathrm{d}t^2 = -x$ and the solution which passes through (x_0, y_0) at $t = 0$ is

$$\begin{aligned} x &= x_0 \cos t - y_0 \sin t \\ y &= x_0 \sin t + y_0 \cos t. \end{aligned}$$

The one-parameter group of transformations generating X is therefore

$$\sigma_t(x, y) = (x \cos t - y \sin t, x \sin t + y \cos t).$$

It is simple to check that

$$\frac{\mathrm{d}}{\mathrm{d}t} f(\sigma_t(x, y))\Big|_{t=0} = -y\frac{\partial f}{\partial x} + x\frac{\partial f}{\partial y}.$$

The streamlines tangent to X with $y_0 = 0$ have the form

$$x = x_0 \cos t, \qquad y = x_0 \sin t.$$

Hence $x^2 + y^2 = r^2 = x_0^2$, i.e. $r = x_0$ is constant along the integral curves, and t is the angular coordinate θ. The required coordinates are therefore $x' = \theta$, $y' = \theta$, and

$$X = -r \sin\theta\Big(\frac{\partial r}{\partial x}\partial_r + \frac{\partial \theta}{\partial x}\partial_\theta\Big) + r\cos\theta\Big(\frac{\partial r}{\partial y}\partial_r + \frac{\partial \theta}{\partial y}\partial_\theta\Big).$$

Since

$$J = \frac{\partial(x, y)}{\partial(r, \theta)} = \begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix}$$

we have

$$\frac{\partial(r, \theta)}{\partial(x, y)} = J^{-1} = \begin{pmatrix} \cos\theta & \sin\theta \\ -(1/r)\sin\theta) & (1/r)\cos\theta \end{pmatrix}$$

and

$$\begin{aligned} X &= -r\sin\theta\big(\cos\theta\partial_r - \frac{\sin\theta}{r}\partial_\theta\big) + r\cos\theta\big(\sin\theta\partial_r + \frac{\cos\theta}{r}\partial_\theta\big) \\ &= \partial_\theta = \partial_{x'}. \end{aligned}$$

Problem 15.22 Repeat the previous problem for the vector fields, $X = y\partial_x + x\partial_y$ and $X = x\partial_x + y\partial_y$.

Solution: The equations for integral curves of $X = y\partial_x + x\partial_y$ are

$$\frac{\mathrm{d}x}{\mathrm{d}t} = y, \qquad \frac{\mathrm{d}y}{\mathrm{d}t} = x,$$

having solution

$$x = x_0 \cosh t + y_0 \sinh t$$
$$y = x_0 \sinh t + y_0 \cosh t.$$

Hence the one-parameter group of transformations generating X is

$$\sigma_t(x, y) = (x \cosh t + y \sinh t, x \sinh t + y \cosh t.$$

The streamlines tangent to X with $y_0 = 0$ have the form

$$x = x_0 \cosh t, \qquad y = x_0 \sinh t.$$

If $|x| > |y|$ set $x' = t$, $y' = \sqrt{x^2 - y^2}$ so that

$$x = y' \cosh x', \qquad y = y' \sinh x'.$$

and

$$\begin{aligned} X &= y' \sinh x' \Big(\frac{\partial x'}{\partial x} \partial_{x'} + \frac{\partial y'}{\partial x} \partial_{y'} \Big) + y' \cosh x' \Big(\frac{\partial x'}{\partial y} \partial_{x'} + \frac{\partial y'}{\partial y} \partial_{y'} \Big) \\ &= y' \sinh x' \Big(-\frac{\sinh x'}{y'} \partial_{x'} + \cosh x' \partial_{y'} \Big) + y' \cosh x' \Big(\frac{\cosh x'}{y'} \partial_{x'} - \sinh x' \partial_{y'} \Big) \\ &= \partial_{x'}. \end{aligned}$$

For $|y| > |x|$ the transformation is $x = y' \sinh x'$, $y = y' \cosh x'$. These transformations still do not cover the lines $y = \pm x$. It is possible, for example to cover $y = x$, $x > y$, by setting $x_0 - y_0 = 2$, $x_0 + y_0 = 2y'$ and $x' = t$ in the above one-parameter group of transfs.,

$$x = y' \mathrm{e}^{x'} + \mathrm{e}^{-x'}$$
$$y = y' \mathrm{e}^{x'} - \mathrm{e}^{-x'}$$

etc.

For $X = x\partial_x + y\partial_y$ the integral curves are given by

$$\frac{\mathrm{d}x}{\mathrm{d}t} = x, \qquad \frac{\mathrm{d}y}{\mathrm{d}t} = y,$$

having solution $x = x_0 \mathrm{e}^t$, $y = y_0 \mathrm{e}^t$. Hence

$$\sigma_t(x, y) = (x\mathrm{e}^t, y\mathrm{e}^t).$$

In this case $x/y = x_0/y_0 = \text{const.}$, and setting initial conditions $x_0^2 + y_0^2 = 1$ we may set

$$y_0 = \cos y', \qquad x' = y = \tfrac{1}{2} \ln(x^2 + y^2)$$

i.e.

$$x = e^{x'} \sin y', \qquad y = e^{x'} \cos y'.$$

Using

$$\frac{\partial(x',y')}{\partial(x,y)} = \begin{pmatrix} \mathrm{e}^{-x'}\sin y' & \mathrm{e}^{-x'}\cos y' \\ \mathrm{e}^{-x'}\cos y' & -\mathrm{e}^{-x'}\sin y' \end{pmatrix}$$

we have

$$\begin{aligned} X &= e^{x'}\sin y'\big(\mathrm{e}^{-x'}\sin y'\partial_{x'} + \mathrm{e}^{-x'}\cos y'\partial_{y'}\big) + e^{x'}\cos y'\big(\mathrm{e}^{-x'}\cos y'\partial_{x'} - \mathrm{e}^{-x'}\sin y'\partial_{y'}\big) \\ &= \partial_{x'}. \end{aligned}$$

Problem 15.23 On a compact manifold show that every vector field X is complete. [*Hint*: Let σ_t be a local flow generating X, and let ϵ be the least bound required on a finite open covering. Set $\sigma_t = (\sigma_{t/N})^N$ for N large enough that $|t| < \epsilon N$.]

Solution: By Theorem 15.2, for each $p \in M$ there exists a neighbourhood U_p and real number ϵ_p such that there is a local one-parameter group of diffeomorphisms $\sigma_{(p)t}$ on U_p which generates $X\big|_U$. By the uniqueness theorem of differential equations

$$\frac{\mathrm{d}x^i}{\mathrm{d}t} = X^i(x^1,\dots,x^n)$$

the action of the groups $\sigma_{(p)t}$ and $\sigma_{(q)t}$ on any intersection $U_p \cap U_q$ is identical. Since M is compact there is a finite subcovering which may be labelled $U_1, U_2, \dots, U_k$, with corresponding positive numbers $\epsilon_1, \dots, \epsilon_k$. Let $\epsilon = \min\{\epsilon_a \mid 1 \le a \le k\}$. Define the map $\sigma : (-\epsilon, \epsilon) \times M \to M$ by

$$\sigma_t(p) \equiv \sigma(t,p) = \sigma_{(a)t}(p) \qquad \text{for} \qquad p \in U_a, \quad |t| < \epsilon.$$

For any real number t let N be any positive integer such that $|t|/N < \epsilon$, and set

$$\sigma_t(p) = \big(\sigma_{t/N}\big)^N(p)$$

where

$$\big(\sigma_{t/N}\big)^N = \underbrace{\sigma_{t/N} \circ \sigma_{t/N} \circ \cdots \circ \sigma_{t/N}}_{N \text{ times}}.$$

This extends σ to all values of t is independent of the choice of N. For any t and s pick N such both $|t|/N < \epsilon$ and $|s|/N < \epsilon$; then

$$\sigma_t \circ \sigma_s = \big(\sigma_{t/N}\big)^N \big(\sigma_{s/N}\big)^N = \big(\sigma_{t/N+s/N}\big)^N = \sigma_{t+s}.$$

Hence $\sigma : \mathbb{R}\times M \to M$ is a global one-parameter group of diffeomorphisms generating X, which is therefore a complete vector field.

Problem 15.24 **Show that the Lie derivative $\mathcal{L}_X$ commutes with all operations of contraction C^i_j on a tensor field T,**

$$\mathcal{L}_X C^i_j T = C^i_j \mathcal{L}_X T.$$

Solution: If ω and X are arbitrary covector and vector fields then, since $\widetilde{\varphi}$ is an invertible map, the fields $\widetilde{\varphi}\omega, \widetilde{\varphi}X$ ranges over all possible covector and vector fields resp. Now by the equation before Eq. (15.33)

$$\begin{aligned}
&\widetilde{\varphi}(C^i_j T)\big(\widetilde{\varphi}\omega^1, \dots, \widetilde{\varphi}\omega^{r-1}, \widetilde{\varphi}X_1, \dots, \widetilde{\varphi}X_{s-1}\big)\\
&\quad = \widetilde{\varphi}\big(C^i_j T(\omega^1, \dots, \omega^{r-1}, X_1, \dots, X_{s-1})\big)\\
&\quad = \widetilde{\varphi}\Big(\sum_{k=1}^n T(\omega^1, \dots, \omega^{i-1}\varepsilon^k, \omega^i, \dots, \omega^{r-1}, X_1, \dots, X_{j-1}, e_k, X_j, \dots, X_{s-1})\Big)\\
&\quad = \Big(\sum_{k=1}^n \widetilde{\varphi}T(\widetilde{\varphi}\omega^1, \dots, \widetilde{\varphi}\varepsilon^k, \dots, \widetilde{\varphi}\omega^{r-1}, \widetilde{\varphi}X_1, \dots, \widetilde{\varphi}e_k, \dots, \widetilde{\varphi}X_{s-1})\\
&\quad = (C^i_j \widetilde{\varphi}T)(\widetilde{\varphi}\omega^1, \dots, \widetilde{\varphi}\omega^{r-1}, \widetilde{\varphi}X_1, \dots, \widetilde{\varphi}X_{s-1})\Big)
\end{aligned}$$

since $\{\widetilde{\varphi}e_i\}$, $\{\widetilde{\varphi}\varepsilon^j\}$ are a basis and dual basis at each point of M by (15.32):

$$\langle \widetilde{\varphi}\varepsilon^j, \widetilde{\varphi}e_i\rangle = \widetilde{\varphi}\langle \varepsilon^j, e_i\rangle = \delta^j_i.$$

Hence, as the $\widetilde{\varphi}\omega^a$ and $\widetilde{\varphi}X_b$ are arbitrary covector and vector fields,

$$\widetilde{\varphi}(C^i_j T) = C^i_j \widetilde{\varphi}T.$$

Using Eq. (15.34)

$$\begin{aligned}
\mathcal{L}_X C^i_j T &= -\frac{\mathrm{d}\widetilde{\sigma}_t C^i_j T}{\mathrm{d}t}\Big|_{t=0}\\
&= -\frac{\mathrm{d}C^i_j \widetilde{\sigma}_t T}{\mathrm{d}t}\Big|_{t=0}\\
&= C^i_j \mathcal{L}_X T.
\end{aligned}$$

Problem 15.25 **Prove the formula (15.39) for the Lie derivative of a general tensor.**

Solution: Let

$$T = T^{ij\dots}{}_{kl\dots} \frac{\partial}{\partial x^i} \otimes \frac{\partial}{\partial x^i} \otimes \dots \otimes \mathrm{d}x^k \otimes \mathrm{d}x^l \otimes \dots$$

By eq. (15.35)

$$\begin{aligned}\mathcal{L}_X T = \mathcal{L}_X(T^{ij\cdots}{}_{kl\ldots})\frac{\partial}{\partial x^i} \otimes \cdots \otimes \mathrm{d}x^k \otimes \ldots \\ + T^{ij\cdots}{}_{kl\ldots}\mathcal{L}_X\Big(\frac{\partial}{\partial x^i}\Big) \otimes \cdots \otimes \mathrm{d}x^k \otimes \ldots \quad + \ldots \\ + T^{ij\cdots}{}_{kl\ldots}\frac{\partial}{\partial x^i} \otimes \cdots \otimes \mathcal{L}_X\big(\mathrm{d}x^k\big) \otimes \ldots \quad + \ldots\end{aligned}$$

Now

$$\mathcal{L}_X(T^{ij\cdots}{}_{kl\ldots}) = X(T^{ij\cdots}{}_{kl\ldots}) = T^{ij\cdots}{}_{kl\ldots,m}X^m.$$

As seen in the text

$$\mathcal{L}_X\Big(\frac{\partial}{\partial x^i}\Big) = -X^j{}_{,i}\frac{\partial}{\partial x^j}$$

and

$$\begin{aligned}\langle\mathcal{L}_X\mathrm{d}x^k, \partial_{x^i}\rangle &= \mathcal{L}_X\langle\mathrm{d}x^k, \partial_i\rangle - \langle\mathrm{d}x^k, \mathcal{L}_X\partial_{x^i}\rangle \\ &= X\big(\delta^k_i\big) + \langle\mathrm{d}x^k, X^m{}_{,i}\partial_{x^m}\rangle \\ &= X^m{}_{,i}\delta^k_m = X^k{}_{,i} \qquad .\end{aligned}$$

Hence

$$\mathcal{L}_X\mathrm{d}x^k = X^k{}_{,i}\mathrm{d}x^i$$

and, after some relabelling of summation (dummy) indices,

$$\begin{aligned}\big(\mathcal{L}_X T\big)^{ij\cdots}{}_{kl\ldots} = T^{ij\cdots}{}_{kl\ldots,m}X^m - T^{mj\cdots}{}_{kl\ldots}X^i{}_{,m} - T^{im\cdots}{}_{kl\ldots}X^j{}_{,m} - \ldots \\ + T^{ij\cdots}{}_{ml\ldots}X^m{}_{,k} + T^{ij\cdots}{}_{km\ldots}X^m{}_{,l} + \ldots\end{aligned}$$

Problem 15.26 **Let D_k be an involutive distribution spanned locally by coordinate vector fields $e_\alpha = \partial/\partial x^\alpha$, where greek indices α, β, etc. all range from 1 to k. If $X_\alpha = A^\beta{}_\alpha e_\beta$ is any local basis spanning a distribution D^k, show that the matrix of functions $[A^\beta_{\ \alpha}]$ is non-singular everywhere on its region of definition, and that $[X_\alpha, X_\beta] = C^\gamma_{\alpha\beta}X_\gamma$ where**

$$C^\gamma_{\alpha\beta} = \big(A^\delta{}_\alpha A^\eta{}_{\beta,\delta} - A^\delta{}_\beta A^\eta{}_{\alpha,\delta}\big)\big(A^{-1}\big)^\gamma{}_\eta.$$

Solution: If the vector fields X_α span D^k then, since e_β lie in D^k, we must have a set of functions $B^\gamma{}_\beta$ such that

$$e_\beta = B^\gamma{}_\beta X_\gamma.$$

Multiplying both sides of this equation by $A^\beta_{\ \alpha}$ (summation convention), we have

$$X_\alpha = A^\beta{}_\alpha e_\beta = A^\beta{}_\alpha B^\gamma{}_\beta X_\gamma$$

whence

$$A^\beta{}_\alpha B^\gamma{}_\beta = \delta^\gamma_\alpha$$

and the matrix $\mathsf{A} = [A^\beta{}_\alpha]$ is non-singular, having inverse $\mathsf{B} = [B^\gamma{}_\beta] = \mathsf{A}^{-1}$.

For any $f \in \mathcal{F}(M)$

$$\begin{aligned}
[X_\alpha, X_\beta]f &= A^\delta{}_\alpha \partial_{x^\delta}\big(A^\eta{}_\beta f_{,\eta}\big) - A^\delta{}_\beta \partial_{x^\delta}\big(A^\eta{}_\alpha f_{,\eta}\big) \\
&= A^\delta{}_\alpha A^\eta{}_{\beta,\delta} f_{,\eta} + A^\delta{}_\alpha A^\eta{}_\beta f_{,\eta\delta} \\
&\quad - A^\delta{}_\beta A^\eta{}_{\alpha,\delta} f_{,\eta} - A^\delta{}_\beta A^\eta{}_\alpha f_{,\eta\delta} \\
&= \big(A^\delta{}_\alpha A^\eta{}_{\beta,\delta} - A^\delta{}_\beta A^\eta{}_{\alpha,\delta}\big) f_{,\eta}
\end{aligned}$$

since $f_{,\eta\delta} = f_{,\delta\eta}$. Hence

$$\begin{aligned}
[X_\alpha, X_\beta]f &= \big(A^\delta{}_\alpha A^\eta{}_{\beta,\delta} - A^\delta{}_\beta A^\eta{}_{\alpha,\delta}\big) e_\eta f \\
&= \big(A^\delta{}_\alpha A^\eta{}_{\beta,\delta} - A^\delta{}_\beta A^\eta{}_{\alpha,\delta}\big)\big(A^{-1}\big)^\gamma_\eta X_\gamma f,
\end{aligned}$$

which gives the desired result.

Problem 15.27 **There is a classical version of the Frobenius theorem stating that a system of partial differential equations of the form**

$$\frac{\partial f^\beta}{\partial x^j} = A^\beta_j\big(x^1, \dots, x^k, f^1(x), \dots, f^r(x)\big)$$

where $i, j = 1, \dots, k$ **and** $\alpha, \beta = 1, \dots, r$ **has a unique local solution through any point** $(a^1, \dots, a^k, b^1, \dots, b^r)$ **if and only if**

$$\frac{\partial A^\beta_j}{\partial x^i} - \frac{\partial A^\beta_i}{\partial x^j} + A^\alpha_i \frac{\partial A^\beta_j}{\partial y^\alpha} - A^\alpha_j \frac{\partial A^\beta_i}{\partial y^\alpha} = 0$$

where $A^\beta_j = A^\beta_j\big(x^1, \dots, x^k, y^1, \dots, y^r\big)$**. Show that this statement is equivalent to the version given in Theorem 15.4. [em Hint: On** $\mathbb{R}^n$ **where** $n = r + k$ **consider the distribution spanned by vectors**

$$Y_i = \frac{\partial}{\partial x^i} + A^\beta_i \frac{\partial}{\partial y^\beta} \qquad (i = 1, \dots, k)$$

and show that the integrability condition is precisely the involutive condition $[Y_i, Y_j] = 0$**, while the condition for an integral submanifold of the form** $y^\beta = f^\beta(x^1, \dots, x^k)$ **is** $A^\beta_j = f^\beta_{,j}$**.]**

Solution: On $\mathbb{R}^n$ let D^k be the distribution spanned by vectors

$$Y_i = \frac{\partial}{\partial x^i} + A^\beta_i \frac{\partial}{\partial y^\beta} \qquad (i = 1, \dots, k)$$

where $A^\beta_i = A^\beta_i(x,y)$. This distribution has local integral manifolds of the form $y^\beta = f^\beta(x^1, \ldots x^n)$, if there is a map $\psi : U \subset \mathbb{R}^k \to \mathbb{R}^n$ defined by

$$\psi(x^1, \ldots, x^k) = (x^1, \ldots, x^k, f^1(x), \ldots f^r(x)).$$

In this case the distribution D^k is spanned by $Y_i = \psi_* \partial_{x^i}$, having the property that for every differentiable function $g = g(x,y)$ on $\mathbb{R}^n$,

$$\begin{aligned} Y_i g = \Big(\psi_* \frac{\partial}{\partial x^i}\Big)_{\psi(p)} g &= \frac{\partial}{\partial x^i} g \circ \psi\Big|_{p=(x^i)} \\ &= \frac{\partial g}{\partial x^i} + f^\alpha{}_{,i} \frac{\partial g}{\partial y^\alpha}, \end{aligned}$$

i.e.

$$Y_i = \frac{\partial}{\partial x^i} + A^\alpha_i \frac{\partial}{\partial y^\alpha} \qquad \text{where} \qquad A^\alpha_j = f^\alpha_{,j}.$$

By the Frobenius theorem, Theorem 15.4, and the remarks preceding it, a necessary and sufficient condition for the distribution D^k to have such integral manifolds is that the Y_i are an involutive set of vector fields. however, because of the requirement $Y_i = \psi_* \partial_{x^i}$, they must also commute:

$$[Y_i, Y_j] = [\psi_* \frac{\partial}{\partial x^i}, \psi_* \frac{\partial}{\partial x^j}] = \psi_* [\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}] = 0.$$

the last step follows from

$$\begin{aligned} [\psi_* \frac{\partial}{\partial x^i}, \psi_* \frac{\partial}{\partial x^j}] f &= \frac{\partial}{\partial x^i}\Big(\frac{\partial f \circ \psi}{\partial x^j} \circ \psi\Big) - \frac{\partial}{\partial x^j}\Big(\frac{\partial f \circ \psi}{\partial x^i} \circ \psi\Big) \\ &= \psi_* \left(\frac{\partial}{\partial x^i}\Big(\frac{\partial f \circ \psi}{\partial x^j}\Big) - \frac{\partial}{\partial x^j}\Big(\frac{\partial f \circ \psi}{\partial x^i}\Big) \right) \\ &= 0. \end{aligned}$$

Thus the functions A^α_j satisfy the partial differential equations

$$\frac{\partial f^\alpha}{\partial x^j} = A^\alpha_j(x^1, \ldots, x^k, f^1(x), \ldots, f^r(x))$$

if and only if $[Y_i, Y_j] = 0$, i.e.

$$\begin{aligned} 0 = [Y_i, Y_j] &= \Big[\frac{\partial}{\partial x^i} + A^\alpha_i \frac{\partial}{\partial y^\alpha}, \frac{\partial}{\partial x^j} + A^\beta_i \frac{\partial}{\partial y^\beta}\Big] \\ &= \Big(\frac{\partial A^\beta_j}{\partial x^i}\Big) \frac{\partial}{\partial y^\beta} - \frac{\partial A^\alpha_i}{\partial x^j}\Big) \frac{\partial}{\partial y^\alpha} \\ &\quad + A^\alpha_i \frac{\partial A^\beta_j}{\partial y^\alpha} \frac{\partial}{\partial y^\beta} - A^\beta_j \frac{\partial A^\alpha_i}{\partial y^\alpha} \frac{\partial}{\partial y^\alpha} \\ &= \Big(\frac{\partial A^\beta_j}{\partial x^i} - \frac{\partial A^\beta_i}{\partial x^j} + A^\alpha_i \frac{\partial A^\beta_j}{\partial y^\alpha} - A^\alpha_j \frac{\partial A^\beta_i}{\partial y^\alpha}\Big) \frac{\partial}{\partial y^\beta} \end{aligned}$$

on relabelling summation indices α and β in some places. This is essentially a reformulation of the Frobenius theorem as an integration condition on first order partial d.e.'s.

Chapter 16

Problem 16.1 **Let $x^1 = x$, $x^2 = y$, $x^3 = z$ be coordinates on the manifold $\mathbb{R}^3$. Write out the components α_{ij} and $(\mathrm{d}\alpha)_{ijk}$ etc. for each of the following 2-forms:**

$$\begin{aligned}
\alpha &= \mathrm{d}y \wedge \mathrm{d}z + \mathrm{d}x \wedge \mathrm{d}y,\\
\beta &= x\,\mathrm{d}z \wedge \mathrm{d}y + y\,\mathrm{d}x \wedge \mathrm{d}z + z\,\mathrm{d}y \wedge \mathrm{d}x,\\
\gamma &= \mathrm{d}(r^2(x\,\mathrm{d}x + y\,\mathrm{d}y + z\,\mathrm{d}z)), \qquad \textbf{where } r^2 = x^2 + y^2 + z^2.
\end{aligned}$$

Solution: From Eq. (8.10) we have

$$\begin{aligned}
\alpha &= \mathrm{d}y \wedge \mathrm{d}z + \mathrm{d}x \wedge \mathrm{d}y,\\
&= \tfrac{1}{2}(\mathrm{d}y \otimes \mathrm{d}z - \mathrm{d}z \otimes \mathrm{d}y) + \tfrac{1}{2}(\mathrm{d}x \otimes \mathrm{d}y - \mathrm{d}y \otimes \mathrm{d}x).
\end{aligned}$$

Hence

$$[\alpha_{ij}] = \begin{pmatrix} 0 & \frac{1}{2} & 0 \\ -\frac{1}{2} & 0 & \frac{1}{2} \\ 0 & -\frac{1}{2} & 0 \end{pmatrix}.$$

Similarly $\beta = x\,\mathrm{d}z \wedge \mathrm{d}y + y\,\mathrm{d}x \wedge \mathrm{d}z + z\,\mathrm{d}y \wedge \mathrm{d}x$ has components

$$[\beta_{ij}] = \begin{pmatrix} 0 & -\frac{1}{2}z & \frac{1}{2}y \\ \frac{1}{2}z & 0 & -\frac{1}{2}x \\ -\frac{1}{2}y & \frac{1}{2}x & 0 \end{pmatrix},$$

and

$$\begin{aligned}
\gamma &= \mathrm{d}(r^2(x\,\mathrm{d}x + y\,\mathrm{d}y + z\,\mathrm{d}z))\\
&= \mathrm{d}(r^2)(x\,\mathrm{d}x + y\,\mathrm{d}y + z\,\mathrm{d}z)) + r^2(\mathrm{d}x \wedge \mathrm{d}x + \mathrm{d}y \wedge \mathrm{d}y + \mathrm{d}z \wedge \mathrm{d}z)\\
&= (2x\,\mathrm{d}x + 2y\,\mathrm{d}y + 2z\,\mathrm{d}z)) \wedge (x\,\mathrm{d}x + y\,\mathrm{d}y + z\,\mathrm{d}z))\\
&= 0.
\end{aligned}$$

For differentials of these 2-forms we have

$$\begin{aligned}
\mathrm{d}\alpha &= \mathrm{d}^2y \wedge \mathrm{d}z + \mathrm{d}y \wedge \mathrm{d}^2z + \mathrm{d}^2x \wedge \mathrm{d}y + \mathrm{d}x \wedge \mathrm{d}^2y = 0\\
\mathrm{d}\beta &= \mathrm{d}x \wedge \mathrm{d}z \wedge \mathrm{d}y + \mathrm{d}y \wedge \mathrm{d}x \wedge \mathrm{d}z + \mathrm{d}z \wedge \mathrm{d}y \wedge \mathrm{d}x\\
&= -3\mathrm{d}x \wedge \mathrm{d}z \wedge \mathrm{d}y\\
\mathrm{d}\gamma &= 0.
\end{aligned}$$

Hence

$$(\mathrm{d}\alpha)_{ijk} = (\mathrm{d}\gamma)_{ijk} = 0 \qquad \text{and} \qquad (\mathrm{d}\beta)_{ijk} = \frac{-3}{3!}\epsilon_{ijk} = -\tfrac{1}{2}\epsilon.$$

Problem 16.2 On the manifold $\mathbb{R}^n$ compute the exterior derivative d of the differential form

$$\alpha = \sum_{i=1}^{n} (-1)^{i-1} x^i \mathrm{d}x^1 \wedge \cdots \wedge \mathrm{d}x^{i-1} \wedge \mathrm{d}x^{i+1} \wedge \cdots \wedge \mathrm{d}x^n.$$

Do the same for $\beta = r^{-n}\alpha$ where $r^2 = (x^1)^2 + \cdots + (x^n)^2$.

Solution: Using the antiderivative law (ED4)

$$\begin{aligned}\mathrm{d}\alpha &= \sum_{i=1}^{n} (-1)^{i-1} \mathrm{d}x^i \wedge \mathrm{d}x^1 \wedge \cdots \wedge \mathrm{d}x^{i-1} \wedge \mathrm{d}x^{i+1} \wedge \cdots \wedge \mathrm{d}x^n \\ &= n \mathrm{d}x^1 \wedge \mathrm{d}x^2 \wedge \cdots \wedge \mathrm{d}x^n\end{aligned}$$

since $i-1$ interchanges are needed to bring the 1-form $\mathrm{d}x^i$ to the ith place, between $\mathrm{d}x^{i-1}$ and $\mathrm{d}x^{i+1}$.

Using

$$\mathrm{d}r = \frac{1}{r} \sum_{j=1}^{n} x^j \mathrm{d}x^j$$

and (ED4) we have

$$\begin{aligned}\mathrm{d}\beta = \mathrm{d}(r^{-n}\alpha) &= -nr^{-n-1} \mathrm{d}r \wedge \alpha + r^{-n} \mathrm{d}\alpha \\ &= -nr^{-n-2} \Big[\sum_{i=1}^{n} \sum_{j=1}^{n} x^j \mathrm{d}x^j \wedge (-1)^{i-1} x^i \mathrm{d}x^1 \wedge \cdots \wedge \mathrm{d}x^{i-1} \wedge \mathrm{d}x^{i+1} \wedge \cdots \wedge \mathrm{d}x^n \Big] \\ &\quad + r^{-n} n \mathrm{d}x^1 \wedge \mathrm{d}x^2 \wedge \cdots \wedge \mathrm{d}x^n \\ &= -nr^{-n-2} \sum_{i=1}^{n} (x^i)^2 \mathrm{d}x^1 \wedge \cdots \wedge \mathrm{d}x^n + r^{-n} n \mathrm{d}x^1 \wedge \cdots \wedge \mathrm{d}x^n \\ &= 0\end{aligned}$$

for in the double sum only terms with $j = i$ survive, and $i - 1$ interchanges are needed to bring $\mathrm{d}x^j$ to the ith slot in the wedge product.

Problem 16.3 Show that the right-hand side of Eq. (16.6) transforms as a tensor field of type $(0, 3)$. Generalize this result to the right-hand side of Eq. (16.7), to show that this equation could be used as a local definition of exterior derivative independent of the choice of coordinate system.

Solution: The transformation of the components $(\mathrm{d}\alpha)_{ijk} = \frac{1}{3}(\alpha_{ij,k} + \alpha_{jk,i} + \alpha_{ki,j})$

given in Eq. (16.6) is given by

$$
\begin{aligned}
(\mathrm{d}\alpha')_{i'j'k'} &= \frac{1}{3}\big(\alpha'_{i'j',k'} + \alpha'_{j'k',i'} + \alpha'_{k'i',j'}\big) \\
&= \frac{1}{3}\Big(\frac{\partial}{\partial x'^{k'}}\Big(\alpha_{ij}\frac{\partial x^i}{\partial x'^{i'}}\frac{\partial x^j}{\partial x'^{j'}}\Big) + \dots\Big) \\
&= \frac{1}{3}\Big(\frac{\partial x^k}{\partial x'^{k'}}\alpha_{ij,k}\frac{\partial x^i}{\partial x'^{i'}}\frac{\partial x^j}{\partial x'^{j'}} + \alpha_{ij}\frac{\partial^2 x^i}{\partial x'^{i'}\partial x'^{k'}}\frac{\partial x^j}{\partial x'^{j'}} \\
&\qquad + \alpha_{ij}\frac{\partial x^i}{\partial x'^{i'}}\frac{\partial^2 x^j}{\partial x'^{j'}\partial x'^{k'}} + \dots\Big) \\
&= \frac{1}{3}\Bigg[(\alpha_{ij,k} + \alpha_{jk,i} + \alpha_{ki,j})\frac{\partial x^i}{\partial x'^{i'}}\frac{\partial x^j}{\partial x'^{j'}}\frac{\partial x^k}{\partial x'^{k'}} \\
&\qquad + \alpha_{ij}\Big(\frac{\partial^2 x^i}{\partial x'^{i'}\partial x'^{k'}}\frac{\partial x^j}{\partial x'^{j'}} + \frac{\partial x^i}{\partial x'^{i'}}\frac{\partial^2 x^j}{\partial x'^{j'}\partial x'^{k'}}\Big) \\
&\qquad + \alpha_{jk}\Big(\frac{\partial^2 x^j}{\partial x'^{j'}\partial x'^{i'}}\frac{\partial x^k}{\partial x'^{k'}} + \frac{\partial x^j}{\partial x'^{j'}}\frac{\partial^2 x^k}{\partial x'^{k'}\partial x'^{i'}}\Big) \\
&\qquad + \alpha_{ki}\Big(\frac{\partial^2 x^k}{\partial x'^{k'}\partial x'^{j'}}\frac{\partial x^i}{\partial x'^{i'}} + \frac{\partial x^k}{\partial x'^{k'}}\frac{\partial^2 x^i}{\partial x'^{i'}\partial x'^{j'}}\Big)\Bigg] \\
&= (\mathrm{d}\alpha)_{ijk}\frac{\partial x^i}{\partial x'^{i'}}\frac{\partial x^j}{\partial x'^{j'}}\frac{\partial x^k}{\partial x'^{k'}}
\end{aligned}
$$

as the second derivative terms cancel in pairs (e.g. the first and fourth, second and fifth, third and sixth).

For an r-form we have Eq. (16.7), which transforms as

$$
(\mathrm{d}\alpha')_{i'_1\dots i'_{r+1}} = \frac{(-1)^r}{r+1}\sum_{\text{cyclic }\pi}(-1)^\pi \alpha'_{i'_{\pi(1)}\dots i'_{\pi(r)},i'_{\pi(r+1)}}.
$$

where all $(-1)^\pi = 1$ for r even ($r+1$ odd) and $(-1)^\pi = \pm 1$ alternately for r odd.

Thus

$$
\begin{aligned}
(-1)^r(r+1)(\mathrm{d}\alpha')_{i'_1\dots i'_{r+1}} &= \frac{\partial}{\partial x'^{i'_{r+1}}}\Big(\alpha_{i_1\dots i_r}\frac{\partial x^{i_1}}{\partial x'^{i'_1}}\dots\frac{\partial x^{i_r}}{\partial x'^{i'_r}}\Big) \\
&\quad + (-1)^r\frac{\partial}{\partial x'^{i'_1}}\Big(\alpha_{i_2 i_3\dots i_{r+1}}\frac{\partial x^{i_2}}{\partial x'^{i'_2}}\dots\frac{\partial x^{i_{r+1}}}{\partial x'^{i'_{r+1}}}\Big) \\
&\quad + (-1)^{2r}\frac{\partial}{\partial x'^{i'_2}}\Big(\alpha_{i_3\dots i_{r+1}i_1}\frac{\partial x^{i_3}}{\partial x'^{i'_3}}\dots\Big) \\
&\quad \dots \\
&\quad + (-1)^{r^2}\frac{\partial}{\partial x'^{i'_r}}\Big(\alpha_{i_{r+1}i_1\dots i_{r-1}}\frac{\partial x^{i_{r+1}}}{\partial x'^{i'_{r+1}}}\dots\Big) \\
&= \big(\alpha_{i_1\dots i_r,i_{r+1}} + (-1)^r\alpha_{i_2 i_3\dots i_{r+1},i_1} + \dots\big)\frac{\partial x^{i_1}}{\partial x'^{i'_1}}\dots\frac{\partial x^{i_r}}{\partial x'^{i'_r}}\frac{\partial x^{i_{r+1}}}{\partial x'^{i'_{r+1}}} \\
&\quad + \alpha_{i_1\dots i_r}\Big(\frac{\partial^2 x^{i_1}}{\partial x'^{i'_1}\partial x'^{i'_{r+1}}}\frac{\partial x^{i_2}}{\partial x'^{i'_2}}\dots\frac{\partial x^{i_r}}{\partial x'^{i'_r}} + \dots\Big) \\
&\quad + \dots
\end{aligned}
$$

Each second derivative term

$$\frac{\partial^2 x^j}{\partial x'^{i'_p}\partial x'^{i'_q}},$$

with say $p > q$, arises in two places as

$$
\begin{aligned}
&(-1)^{rp}\alpha_{i_{p+1}\dots i_q\dots i_{p-1}}\frac{\partial x^{i_{p+1}}}{\partial x'^{i'_{p+1}}}\dots\frac{\partial^2 x^{i_q}}{\partial x'^{i'_q}\partial x'^{i'_p}}\dots\frac{\partial x^{i_{p-1}}}{\partial x'^{i'_{p-1}}} \\
&+ (-1)^{rq}\alpha_{i_{q+1}\dots i_p\dots i_{q-1}}\frac{\partial x^{i_{q+1}}}{\partial x'^{i'_{q+1}}}\dots\frac{\partial^2 x^{i_p}}{\partial x'^{i'_p}\partial x'^{i'_q}}\dots\frac{\partial x^{i_{q-1}}}{\partial x'^{i'_{q-1}}}.
\end{aligned}
$$

Replacing the dummy summation i_p in the second expression with i_q the two terms are seen to differ in sign by $(-1)^{r(p+q)}$ multiplied by the sign of the permuation

$$\pi = \begin{pmatrix} i_{q+1} & \dots & i_{p-1} & i_q & i_{p+1} & \dots & i_{q-1} \\ i_{p+1} & \dots & i_{q-1} & i_q & i_{q+1} & \dots & i_{p-1} \end{pmatrix}.$$

$p-(q+1)$ interchanges will bring the i_q to the first place in the top line, then $(p-q)$ r-cycles will bring the top line to the bottom line, so that

$$(-1)^\pi = (-1)^{(p-q)-1+(r-1)(p-q)} = (-1)^{r(p-q)-1}.$$

The two terms therefore differ in sign by

$$(-1)r(p+q) + r(p-q) - 1 = (-1)^{2rp-1} = -1,$$

and all second order derivative terms cancel in pairs. Thus the components given in Eq. (16.7) transform as a covariant tensor of type $(0, r+1)$,

$$(\mathrm{d}\alpha')_{i'_1\dots i'_{r+1}} = (\mathrm{d}\alpha)_{i_1\dots i_{r+1}}\frac{\partial x^{i_1}}{\partial x'^{i'_1}}\dots\frac{\partial x^{i_{r+1}}}{\partial x'^{i'_{r+1}}}$$

and (16.7) can be used as a local definition of exterior derivative independent of the choice of coordinate system.

Problem 16.4 **Let $\varphi : \mathbb{R}^2 \to \mathbb{R}^3$ be the map**

$$(x, y) \to (u, v, w) \qquad \textbf{where } u = \sin(xy),\ v = x + y,\ w = 2.$$

For the 1-form $\omega = w_1 \mathrm{d}u + w_2 \mathrm{d}v + w_3 \mathrm{d}w$ on $\mathbb{R}^3$ evaluate $\varphi^* \omega$. For any function $f : \mathbb{R}^3 \to \mathbb{R}$ verify Theorem 16.2, that $\mathrm{d}(\varphi^* f) = \varphi^* \mathrm{d}f$.

Solution: For each $i = 1, 2, 3$ we have

$$\varphi^* w_i(x, y) = w_i \circ \varphi(x, y) = w_i(\sin(xy), x + y, 2).$$

Hence

$$\begin{aligned}
\varphi^* \omega &= \varphi^* w_1 \varphi^* \mathrm{d}u + \varphi^* w_2 \varphi^* \mathrm{d}v + \varphi^* w_3 \varphi^* \mathrm{d}w \\
&= \varphi^* w_1 \mathrm{d}\varphi^* u + \varphi^* w_2 \mathrm{d}\varphi^* v + \varphi^* w_3 \mathrm{d}\varphi^* w \\
&= \varphi^* w_1 \mathrm{d}\varphi^* u + \varphi^* w_2 \mathrm{d}\varphi^* v + \varphi^* w_3 \mathrm{d}\varphi^* w
\end{aligned}$$

Now

$$\begin{aligned}
\mathrm{d}\varphi^* u &= \mathrm{d}u \circ \varphi = \mathrm{d}\sin(xy) = y\cos(xy)\mathrm{d}x + x\cos(xy)\mathrm{d}y \\
\mathrm{d}\varphi^* v &= \mathrm{d}v \circ \varphi = \mathrm{d}(x + y) = \mathrm{d}x + \mathrm{d}y \\
\mathrm{d}\varphi^* w &= \mathrm{d}w \circ \varphi = \mathrm{d}2 = 0.
\end{aligned}$$

Hence

$$\begin{aligned}
\varphi^* \omega &= \varphi^* w_1(x, y)(y\cos(xy)\mathrm{d}x + x\cos(xy)\mathrm{d}y) + \varphi^* w_2(x, y)(\mathrm{d}x + \mathrm{d}y) \\
&= (w_1(\sin(xy), x + y, 2) y\cos(xy) + w_2(\sin(xy), x + y, 2))\mathrm{d}x \\
&\quad + (w_1(\sin(xy), x + y, 2) x\cos(xy) + w_2(\sin(xy), x + y, 2))\mathrm{d}y
\end{aligned}$$

For any differentiable function $f : \mathbb{R}^3 \to \mathbb{R}$,

$$\begin{aligned}
\mathrm{d}(\varphi^* f) = \mathrm{d}(f \circ \varphi) &= \mathrm{d}(f(\sin(xy), x + y, 2) \\
&= f_{,u}(\sin(xy), x + y, 2)(y\cos(xy)\mathrm{d}x + x\cos y\mathrm{d}y) \\
&\quad + f_{,v}(\sin(xy), x + y, 2)(\mathrm{d}x + \mathrm{d}y) \\
&= f_{,u} \circ \varphi\, \mathrm{d}(\varphi^* u) + f_{,v} \circ \varphi\, \mathrm{d}(\varphi^* v) + f_{,w} \circ \varphi\, \mathrm{d}(\varphi^* w),
\end{aligned}$$

where $f_{,u} = \partial f / \partial u$ etc. Hence

$$\begin{aligned}
\varphi^* \mathrm{d}f &= \varphi^*(f_{,u}\mathrm{d}u + f_{,v}\mathrm{d}v + f_{,w}\mathrm{d}w) \\
&= f_{,u} \circ \varphi\, \mathrm{d}(\varphi^* u) + f_{,v} \circ \varphi\, \mathrm{d}(\varphi^* v) + f_{,w} \circ \varphi\, \mathrm{d}(\varphi^* w) \\
&= \mathrm{d}(\varphi^* f)
\end{aligned}$$

by the above discussion.

Problem 16.5 **If α is an r-form on a differentiable manifold M, show that for any vector fields $X_1, X_2, \dots, X_{r+1}$**

$$\mathrm{d}\alpha(X_1, X_2, \dots, X_{r+1}) = \frac{1}{r+1}\Big[\sum_{i=1}^{r+1}(-1)^{i+1}X_i\alpha(X_1, X_2, \dots, \hat{X}_i, \dots, X_{r+1}) + \sum_{i=1}^{r}\sum_{j=i+1}^{r+1}(-1)^{i+j}\alpha([X_i, X_j], \dots, \hat{X}_i, \dots, \hat{X}_j, \dots, X_{r+1})\Big]$$

where $\hat{X}_i$ signifies that the argument X_i is to be omitted. The case $r = 0$ simply asserts that $\mathrm{d}f(X) = Xf$, while Eq. (16.14) is the case $r = 1$. Proceed by induction, assuming the identity is true for all $(r-1)$-forms, and use the fact that any r-form can be written locally as a sum of tensors of the type $\omega \wedge \beta$ where ω is a 1-form and β an $(r-1)$-form.
[NOTE: Correction in last sentence. In the text it was stated that β was an r-form.]**

Solution: For $r = 0$ there is no double sum, and the required expression reads

$$\mathrm{d}\alpha(X_1) = \frac{1}{1}(-1)^2 X_1\alpha$$

which reads $\mathrm{d}f(X) = Xf$ if we set $\alpha = f$ and $X_1 = X$.
For $r = 1$, the expression reads

$$\begin{aligned}\mathrm{d}\alpha(X_1, X_2) &= \tfrac{1}{2}\big(X_1\alpha(X_2) - X_2\alpha(X_1) + (-1)^{1+2}\alpha([X_1, X_2])\big) \\ &= \tfrac{1}{2}\big(X_1\langle X_2, \alpha\rangle - X_2\langle X_1, \alpha\rangle - \langle [X_1, X_2], \alpha\rangle\big)\end{aligned}$$

which is identical with Eq. (16.14) if we set $\alpha = \omega$, $X_1 = X$ and $X_2 = Y$.
To prove for general order r, assume it true for all $(r-1)$-forms β,

$$\mathrm{d}\beta(X_1, X_2, \dots, X_r) = \frac{1}{r}\Big[\sum_{i=1}^{r}(-1)^{i+1}X_i\beta(X_1, X_2, \dots, \hat{X}_i, \dots, X_r) + \sum_{i=1}^{r-1}\sum_{j=i+1}^{r}(-1)^{i+j}\beta([X_i, X_j], \dots, \hat{X}_i, \dots, \hat{X}_j, \dots, X_r)\Big]$$

We will prove that it now holds for any r-form α of the form $\alpha = \omega \wedge \beta$ where ω is a 1-form. This implies the result for arbitrary r-forms by linearity, since every α is a sum of forms of this type.
In what follows we make use of the forms version of Eqs. (8.2) and (8.9): If α is an r-form and β an s-form then

$$\begin{aligned}(\alpha \wedge \beta)(X_1, \dots X_{r+s}) &= \mathcal{A}(\alpha \otimes \beta)(X_1, \dots X_{r+s}) \\ &= \frac{1}{(r+s)!}\sum_{\pi}\alpha(X_{\pi(1)}, \dots, X_{\pi(r)})\beta(X_{\pi(r+1)}, \dots, X_{\pi(r+s)}).\end{aligned}$$

From $\alpha = \omega \wedge \beta$ we have

$$\mathrm{d}\alpha = \mathrm{d}\omega \wedge \beta - \omega \wedge \mathrm{d}\beta$$

and

$$\mathrm{d}\alpha(X_1, \ldots, X_{r+1}) = (\mathrm{d}\omega \wedge \beta)(X_1, \ldots, X_{r+1}) - (\omega \wedge \mathrm{d}\beta)(X_1, \ldots, X_{r+1}).$$

Now

$$\begin{aligned}
(\mathrm{d}\omega \wedge \beta)(X_1, \ldots, X_{r+1}) &= \frac{1}{(r+1)!} \sum_{\pi} \mathrm{d}\omega \otimes \beta\big(X_{\pi(1)}, \ldots, X_{\pi(r+1)}\big) \\
&= \frac{1}{r(r+1)} \sum_{i=1}^{r} \sum_{j=i+1}^{r+1} (-1)^{i+1+j} \mathrm{d}\omega(X_i, X_j)\beta(X_1, \ldots, \hat{X}_i, \ldots, \hat{X}_j, \ldots, X_{r+1}) \\
&\quad + \frac{1}{r(r+1)} \sum_{i=2}^{r+1} \sum_{j=1}^{i-1} (-1)^{j+1+i} \mathrm{d}\omega(X_j, X_i)\beta(X_1, \ldots, \hat{X}_j, \ldots, \hat{X}_i, \ldots, X_{r+1}) \\
&= \frac{2}{r(r+1)} \sum_{i=1}^{r} \sum_{j=i+1}^{r+1} (-1)^{i+1+j} \mathrm{d}\omega(X_i, X_j)\beta(X_1, \ldots, \hat{X}_i, \ldots, \hat{X}_j, \ldots, X_{r+1}) \\
&= \frac{1}{r(r+1)} \sum_{i=1}^{r} \sum_{j=i+1}^{r+1} (-1)^{i+j+1} \big[X_i\omega(X_j) - X_j\omega(X_i) - \omega([X_i, X_j])\big] \cdot \\
&\qquad \cdot \beta(X_1, \ldots, \hat{X}_i, \ldots, \hat{X}_j, \ldots, X_{r+1})
\end{aligned} \tag{1}$$

and, from the induction hypothesis for $\mathrm{d}\beta$

$$\begin{aligned}
-(\omega \wedge \mathrm{d}\beta)(X_1, \ldots, X_{r+1}) &= \frac{-1}{r+1} \sum_{i=1}^{r+1} (-1)^{i+1} \omega(X_i) \mathrm{d}\beta(X_1, \ldots, \hat{X}_i, \ldots, X_{r+1}) \\
&= \frac{-1}{r+1} \sum_{i=1}^{r+1} (-1)^{i+1} \omega(X_i) \Big[\sum_{j=1}^{i-1} (-1)^{j+1} X_j \beta(X_1, \ldots, \hat{X}_j, \ldots, \hat{X}_i, \ldots, X_{r+1}) \\
&\quad + \sum_{j=i+1}^{r+1} (-1)^{j} X_j \beta(X_1, \ldots, \hat{X}_i, \ldots, \hat{X}_j, \ldots, X_{r+1}) \\
&\quad + \sum_{j=1}^{i-1} \sum_{k=j+1}^{i-1} (-1)^{j+k} \beta([X_j, X_k], X_1, \ldots, \hat{X}_j, \ldots, \hat{X}_k, \ldots \hat{X}_i, \ldots, X_{r+1}) \\
&\quad + \sum_{j=1}^{i-1} \sum_{k=i+1}^{r+1} (-1)^{j+k+1} \beta([X_j, X_k], X_1, \ldots, \hat{X}_j, \ldots, \hat{X}_i, \ldots \hat{X}_k, \ldots, X_{r+1}) \\
&\quad + \sum_{j=1}^{r} \sum_{k=i+1}^{r+1} (-1)^{j+k} \beta([X_j, X_k], X_1, \ldots, \hat{X}_i, \ldots, \hat{X}_j, \ldots \hat{X}_k, \ldots, X_{r+1}) \Big].
\end{aligned} \tag{2}$$

Of course i cannot have the value 1 for the first, third or fourth line of this equation to make sense, or the value 2 for the third line to make any contribution. Using

$$\begin{aligned}\alpha(Y_1,\dots,Y_r) &= \omega\wedge\beta(Y_1,\dots,Y_r)\\ &= \frac{1}{r}\sum_{i=1}^{r}(-1)^{i+1}\omega(Y_i)\beta(Y_1,\dots,\hat{Y}_i,\dots,Y_r)\end{aligned}$$

we have that the RHS of the required equation for $\mathrm{d}\alpha$ is

$$\begin{aligned}&\frac{1}{r(r+1)}\sum_{i=1}^{r+1}(-1)^{i+1}\Big[X_i\Big(\sum_{j=1}^{i-1}(-1)^{j+1}\omega(X_j)\beta(X_1,\dots,\hat{X}_j,\dots,\hat{X}_i,\dots,X_{r+1})\Big)\\ &\quad X_i\Big(\sum_{j=i+1}^{r+1}(-1)^{j}\omega(X_j)\beta(X_1,\dots,\hat{X}_i,\dots,\hat{X}_j,\dots,X_{r+1})\Big)\Big]\\ &+\frac{1}{r(r+1)}\sum_{i=1}^{r}\sum_{j=i+1}^{r+1}(-1)^{i+j}\Big[\omega([X_i,X_j])\beta(X_1,\dots,\hat{X}_i,\dots,\hat{X}_j,\dots,X_{r+1})\\ &\quad+\sum_{k=1}^{i-1}(-1)^{k}\omega(X_k)\beta([X_i,X_j],X_1,\dots,\hat{X}_k,\dots,\hat{X}_i,\dots,\hat{X}_j,\dots,X_{r+1})\\ &\quad+\sum_{k=i+1}^{j-1}(-1)^{k+1}\omega(X_k)\beta([X_i,X_j],X_1,\dots,\hat{X}_i,\dots,\hat{X}_k,\dots,\hat{X}_j,\dots,X_{r+1})\\ &\quad+\sum_{k=j+1}^{r+1}(-1)^{k}\omega(X_k)\beta([X_i,X_j],X_1,\dots,\hat{X}_i,\dots,\hat{X}_j,\dots,\hat{X}_k,\dots,X_{r+1})\end{aligned}$$

To show that this is equal to $\mathrm{d}\alpha(X_1,\dots,X_{r+1})$ it is a straightforward matter to check that these terms match those in Eqs. (1) and (2) term by term, after expanding

$$X_i\big(\omega(X_j)\beta(X_1,\dots,X_{r+1})\big) = X_i\big(\omega(X_j)\big)\beta(X_1,\dots,X_{r+1})+\omega(X_j)X_i\big(\beta(X_1,\dots,X_{r+1})\big)$$

etc. Hence the result follows by induction.

Problem 16.6 Show that the Laplacian operator on $\mathbb{R}^3$ may be defined by

$$\mathrm{d}*\mathrm{d}\phi = \nabla^2\phi\,\mathrm{d}x\wedge\mathrm{d}y\wedge\mathrm{d}z = \Big(\frac{\partial^2\phi}{\partial x^2}+\frac{\partial^2\phi}{\partial y^2}+\frac{\partial^2\phi}{\partial z^2}\Big)\mathrm{d}x\wedge\mathrm{d}y\wedge\mathrm{d}z$$

where $*$ is the Hodge star operator of Section 8.6.
Use this to express the Laplacian operator in spherical polar coordinates $(r,\ \theta,\ \phi)$.

Solution: If $\alpha=\alpha_i\mathrm{d}x^i$ is a 1-form then, from Problem 8.12 we have

$$(*\alpha)_{jk} = \frac{\sqrt{|g|}}{2}\epsilon_{ijk}\alpha^i \qquad \text{where} \qquad \alpha^i = g^{ij}\alpha_i. \qquad (*)$$

In rectangular cartesian coordinates $\mathrm{d}x^1 = x$, $\mathrm{d}x^2 = y$, $\mathrm{d}x^3 = z$ we have $g_{ij} = g^{ij} = \delta_{ij}$. Hence, setting $\alpha = \mathrm{d}\phi = \phi_{,i}\mathrm{d}x^i$, we have

$$(*\mathrm{d}\phi)_{jk} = \tfrac{1}{2}\epsilon_{ijk}\phi_{,i}$$

so that

$$*\mathrm{d}\phi = \frac{\partial\phi}{\partial x}\mathrm{d}y\wedge\mathrm{d}z + \frac{\partial\phi}{\partial y}\mathrm{d}z\wedge\mathrm{d}x + \frac{\partial\phi}{\partial z}\mathrm{d}x\wedge\mathrm{d}y.$$

Thus

$$\begin{aligned}\mathrm{d}*\mathrm{d}\phi &= \frac{\partial^2\phi}{\partial x^2}\mathrm{d}x\wedge\mathrm{d}y\wedge\mathrm{d}z + \frac{\partial^2\phi}{\partial y^2}\mathrm{d}y\wedge\mathrm{d}z\wedge\mathrm{d}x + \frac{\partial^2\phi}{\partial z^2}\mathrm{d}z\wedge\mathrm{d}x\wedge\mathrm{d}y\\ &= \Big(\frac{\partial^2\phi}{\partial x^2} + \frac{\partial^2\phi}{\partial y^2} + \frac{\partial^2\phi}{\partial z^2}\Big)\mathrm{d}x\wedge\mathrm{d}y\wedge\mathrm{d}z\\ &= \nabla^2\phi\,\mathrm{d}x\wedge\mathrm{d}y\wedge\mathrm{d}z.\end{aligned}$$

In polar coordinates

$$x = r\sin\theta\cos\phi, \qquad y = r\sin\theta\sin\phi, \qquad z = r\cos\theta$$

it follows that

$$\begin{aligned}\mathrm{d}x &= \sin\theta\cos\phi\,\mathrm{d}r + r\cos\theta\cos\phi\,\mathrm{d}\theta - r\sin\theta\sin\phi\,\mathrm{d}\phi\\ \mathrm{d}y &= \sin\theta\sin\phi\,\mathrm{d}r + r\cos\theta\sin\phi\,\mathrm{d}\theta + r\sin\theta\cos\phi\,\mathrm{d}\phi\\ \mathrm{d}z &= \cos\theta\,\mathrm{d}r - r\sin\theta\,\mathrm{d}\theta.\end{aligned}$$

Hence a straightforward calculation gives

$$\begin{aligned}G &= \mathrm{d}x\otimes\mathrm{d}x + \mathrm{d}x\otimes\mathrm{d}x + \mathrm{d}x\otimes\mathrm{d}x\\ &= \mathrm{d}r\otimes\mathrm{d}r + r^2\mathrm{d}\theta\otimes\theta + r^2\sin^2\theta\mathrm{d}\phi\otimes\mathrm{d}\phi,\end{aligned}$$

and the components of the tensor $g_{ij}\mathrm{d}x^i\otimes\mathrm{d}x^j$ in polar coordinates are

$$[g_{ij}] = \begin{pmatrix}1 & 0 & 0\\ 0 & r^2 & 0\\ 0 & 0 & r^2\sin^2\theta\end{pmatrix}, \qquad [g^{ij}] = \begin{pmatrix}1 & 0 & 0\\ 0 & r^{-2} & 0\\ 0 & 0 & r^{-2}\sin^{-2}\theta\end{pmatrix}.$$

The determinant is $g = r^4\sin^2\theta$, so that for any function f

$$\mathrm{d}f = \frac{\partial f}{\partial r}\,\mathrm{d}r + \frac{\partial f}{\partial\theta}\,\mathrm{d}\theta + \frac{\partial f}{\partial\phi}\,\mathrm{d}\phi$$

and from Eq. $(*)$ above

$$*\mathrm{d}f = r^2\sin\theta\Big(\frac{\partial f}{\partial r}\mathrm{d}\theta\wedge\mathrm{d}\phi + \frac{1}{r^2}\frac{\partial f}{\partial\theta}\mathrm{d}\phi\wedge\mathrm{d}r + \frac{1}{r^2\sin^2\theta}\frac{\partial f}{\partial\phi}\mathrm{d}r\wedge\mathrm{d}\theta\Big).$$

Hence

$$\begin{aligned}
\mathrm{d} * \mathrm{d}f &= \Big[\frac{\partial}{\partial r}\big(r^2\sin\theta\frac{\partial f}{\partial r}\big) + \frac{\partial}{\partial\theta}\big(\sin\theta\frac{\partial f}{\partial\theta}\big) + \frac{\partial}{\partial\phi}\big(\frac{1}{\sin\theta}\frac{\partial f}{\partial\phi}\big)\Big]\mathrm{d}r\wedge\mathrm{d}\theta\wedge\mathrm{d}\phi \\
&= \nabla^2 f\,\mathrm{d}x\wedge\mathrm{d}y\wedge\mathrm{d}z
\end{aligned}$$

where

$$\begin{aligned}
\nabla^2 f &= \frac{1}{r^2\sin\theta}\Big[\frac{\partial}{\partial r}\big(r^2\sin\theta\frac{\partial f}{\partial r}\big) + \frac{\partial}{\partial\theta}\big(\sin\theta\frac{\partial f}{\partial\theta}\big) + \frac{\partial}{\partial\phi}\big(\frac{1}{\sin\theta}\frac{\partial f}{\partial\phi}\big)\Big] \\
&= \frac{1}{r^2}\frac{\partial}{\partial r}\big(r^2\frac{\partial f}{\partial r}\big) + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\big(\sin\theta\frac{\partial f}{\partial\theta}\big) + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial\phi}\big(\frac{1}{\sin\theta}\frac{\partial f}{\partial\phi}\big) \\
&= \frac{\partial^2 f}{\partial r^2} + \frac{2}{r}\frac{\partial f}{\partial r} + \frac{1}{r^2}\frac{\partial^2 f}{\partial\theta^2} + \frac{\cot\theta}{r^2}\frac{\partial f}{\partial\theta} + \frac{1}{r^2\sin^2\theta}\frac{\partial^2 f}{\partial\phi^2}.
\end{aligned}$$

Problem 16.7 **Let $\omega = yz\mathrm{d}x + xz\mathrm{d}y + 3\mathrm{d}z$. Show that the Pfaffian system $\omega = 0$ has integral surfaces $g = z^3\mathrm{e}^{xy} =$ const., and express ω in the form $f\mathrm{d}g$.**
Do the same for $\omega = yz\mathrm{d}x + xz\mathrm{d}y + z^2\mathrm{d}z$.
[NOTE misprint in first ω given in text. The second case is suggested as an additional or alternative exercise.]**

Solution: If $\omega = yz\mathrm{d}x + xz\mathrm{d}y + 3\mathrm{d}z$ then

$$\begin{aligned}
\mathrm{d}\omega &= y\mathrm{d}z\wedge\mathrm{d}x + z\mathrm{d}y\wedge\mathrm{d}x + x\mathrm{d}z\wedge\mathrm{d}y + z\mathrm{d}x\wedge\mathrm{d}y \\
&= y\mathrm{d}z\wedge\mathrm{d}x + x\mathrm{d}z\wedge\mathrm{d}y \qquad .
\end{aligned}$$

Hence

$$\mathrm{d}\omega\wedge\omega = xzy\mathrm{d}z\wedge\mathrm{d}x\wedge\mathrm{d}y + xzy\mathrm{d}z\wedge\mathrm{d}y\wedge\mathrm{d}x = 0$$

and by the Frobenius theorem there exist local functions f and g such that $\omega f\mathrm{d}g$. These functions satisfy the differential equations

$$\begin{aligned}
f\frac{\partial g}{\partial x} &= yz \\
f\frac{\partial g}{\partial y} &= xz \\
f\frac{\partial g}{\partial z} &= 3.
\end{aligned}$$

Hence

$$x\frac{\partial g}{\partial x} = y\frac{\partial g}{\partial y} = \frac{xyz}{f}$$

which has solution $g = G(\ln x + \ln y, z) = G(\ln(xy), z)$ for arbitrary differentiable funtions G. Thus

$$g = F(u, z) \qquad \text{where} \qquad u = xy$$

and substituting in the $\partial g/\partial z$ equation gives

$$z\frac{\partial F}{\partial z} = \frac{3z}{f} = \frac{3}{y}\frac{\partial g}{\partial x} = 3\frac{\partial F}{\partial u}.$$

Hence

$$F = H(\ln z + \tfrac{1}{3}u)$$

for any function H. Taking $H(v) = \mathrm{e}^{3v}$ we have

$$g = \mathrm{e}^{3\ln z + xy} = z^3\mathrm{e}^{xy}$$

which gives the desired integral surfaces. The function f is

$$f = \frac{3}{\partial g/\partial z} = \frac{1}{z^2}\mathrm{e}^{-xy}.$$

It is straightforward to check that $\omega = f\mathrm{d}g$.

The discussion of the 1-form $\omega = yz\mathrm{d}x + xz\mathrm{d}y + z^2\mathrm{d}z$ follows along very similar lines. Again $\mathrm{d}\omega \wedge \omega = 0$ and an argument similar to that above shows that $\omega = f\mathrm{d}g$ where

$$g = z^2 + 2xy, \qquad f = \frac{z}{2}.$$

Problem 16.8 **Given an $r \times r$ matrix of 1-forms Ω, show that the equation**

$$\mathrm{d}\mathsf{A} = \Omega\mathsf{A} - \mathsf{A}\Omega$$

is soluble for an $r \times r$ matrix of functions A only if

$$\Theta\mathsf{A} = \mathsf{A}\Theta$$

where $\Theta = \mathrm{d}\Omega - \Omega \wedge \Omega$.
If the equation has a solution for arbitrary initial values $\mathsf{A} = \mathsf{A}_0$ at any point $p \in M$, show that there exists a 2-form α such that $\Theta = \alpha\mathsf{I}$ and $\mathrm{d}\alpha = 0$.

Solution: The matrix equation $\mathrm{d}\mathsf{A} = \Omega\mathsf{A} - \mathsf{A}\Omega$ is shorthand for

$$\mathrm{d}A_{ij} = \Omega_{ik}A_{kj} - A_{ik}\Omega_{kj}$$

where summation convention over the repeated index k is adopted. Hence

$$\begin{aligned}\mathsf{O} = \mathrm{d}^2\mathsf{A} &= \mathrm{d}\Omega\mathsf{A} - \Omega \wedge \mathrm{d}A - \mathrm{d}A \wedge \Omega - \mathsf{A}\mathrm{d}\Omega \\ &= \mathrm{d}\Omega\mathsf{A} - \Omega \wedge (\Omega\mathsf{A} - \mathsf{A}\Omega) - (\Omega\mathsf{A} - \mathsf{A}\Omega) \wedge \Omega - \mathsf{A}\mathrm{d}\Omega \\ &= (\mathrm{d}\Omega - \Omega \wedge \Omega)\mathsf{A} + \mathsf{A}(\Omega \wedge \Omega - \mathrm{d}\Omega) \\ &= \Theta\mathsf{A} - \mathsf{A}\Theta\end{aligned}$$

where $\Theta = \mathrm{d}\Omega - \Omega \wedge \Omega$. [Note that $(\Omega \wedge \Omega)_{ij} = \Omega_{ik} \wedge \Omega_{kj} \neq 0$ in general.] Thus

$$\Theta \mathsf{A} = \mathsf{A} \Theta.$$

If the matrix of functions A can take arbitrary values A_0 at a given point p, then at this point

$$\Theta_{ik}(p)(\mathsf{A}_0)_{kj} = (\mathsf{A}_0)_{ik} \Theta_{kj}(p)$$

and setting $(\mathsf{A}_0)_{kj} = \delta_{ka}\delta_{jb}$ for fixed a and b gives

$$\Theta_{ik}(p)\delta_{ka}\delta_{jb} = \delta_{ia}\delta_{kb}\Theta_{kj}(p),$$

i.e.

$$\Theta_{ia}(p)\delta_{jb} = \delta_{ia}\Theta_{bj}(p).$$

Setting $b = j$ gives $\Theta_{ia}(p) = \delta_{ia}\Theta_{bb}(p)$ (no summation over b), whence $\Theta_{aa}(p) = \Theta_{bb}(p)$ for all $a, b = 1, \dots n$. Writing $\alpha = \Theta_{aa}(p)$ we have then $\Theta_{ij}(p) = \alpha \delta_{ij}$ and as p is an arbitrary point there exists a 2-form $\alpha : M \to \mathbb{R}$ such that $\Theta = \alpha \mathsf{I}$. Therefore

$$\mathrm{d}\Omega - \Omega \wedge \Omega = \alpha \mathsf{I}$$

and, since Ω is a matrix of 1-forms

$$\begin{aligned} \mathrm{d}\alpha \mathsf{I} &= \mathrm{d}^2\Omega - \mathrm{d}\Omega \wedge \Omega + \Omega \wedge \mathrm{d}\Omega \\ &= -\mathrm{d}\Omega \wedge \Omega + \Omega \wedge \mathrm{d}\Omega \\ &= -(\alpha \mathsf{I} + \Omega \wedge \Omega) \wedge \Omega + \Omega \wedge (\alpha \mathsf{I} + \Omega \wedge \Omega) \\ &= -\alpha \mathsf{I} \wedge \Omega + \Omega \wedge \alpha \mathsf{I} \\ &= \mathsf{O}. \end{aligned}$$

The last step follows from

$$\begin{aligned} (\alpha \mathsf{I} \wedge \Omega)_{ij} &= \alpha \delta_{ik} \wedge \Omega_{kj} \\ &= \alpha \wedge \Omega_{ij} \\ &= \Omega_{ij} \wedge \alpha \qquad \text{(since } \alpha \text{ is a 2-form)} \\ &= \Omega_{ik} \wedge \alpha \delta_{kj} \\ &= (\Omega \wedge \alpha \mathsf{I})_{ij}. \end{aligned}$$

Hence $\mathrm{d}\alpha = 0$.

Problem 16.9 **For a reversible process $\sigma : T \to K$, using absolute temperature T as the parameter, set**

$$\sigma^* \theta = c \, \mathrm{d}T$$

where c is known as the specific heat for the process. For a perfect gas show that for a process at constant volume, $V =$ const. the specific heat is given by

$$c_V = \left(\frac{\partial U}{\partial T}\right)_V.$$

For a process at constant pressure show that

$$c_p = c_V + R,$$

while for an adiabatic process, $\sigma^*\theta = 0$,

$$pV^\gamma = \textbf{const.} \qquad \textbf{where} \qquad \gamma = \frac{c_p}{c_V}.$$

Solution: If $U = U(T, V)$ then

$$\mathrm{d}U = \left(\frac{\partial U}{\partial T}\right)_V \mathrm{d}T + \left(\frac{\partial U}{\partial V}\right)_T \mathrm{d}V$$

and

$$\sigma * \theta = c\mathrm{d}T = \sigma^*(\mathrm{d}U + p\mathrm{d}V) = \left(\frac{\partial U}{\partial T}\right)_V \mathrm{d}T + \left(p + \left(\frac{\partial U}{\partial V}\right)_T\right)\mathrm{d}V$$

For a constant volume process $V =$const. we have

$$\sigma^*\mathrm{d}V = 0 \quad \Longrightarrow \quad \sigma^*\theta = c_V\mathrm{d}T \quad \Longrightarrow \quad c_V = \left(\frac{\partial U}{\partial T}\right)_V.$$

For a perfect gas,

$$\left(\frac{\partial U}{\partial V}\right)_T = 0,$$

i.e. $U = U(T)$, so that for a constant pressure process $p =$const.

$$\begin{aligned}\sigma * \theta = c_p\mathrm{d}T &= \left(\frac{\partial U}{\partial T}\right)_V \mathrm{d}T + p\mathrm{d}V \\ &= c_V\mathrm{d}T + p\left(\frac{\partial V}{\partial T}\right)_p \mathrm{d}T.\end{aligned}$$

Since $pV = RT$, where R is the perfect gas constant, we have

$$p\left(\frac{\partial V}{\partial T}\right)_p = R$$

whence

$$c_p\mathrm{d}T = (c_V + R)\mathrm{d}T$$

i.e.

$$c_p = (c_V + R).$$

For an adiabatic process $\sigma^*\theta = 0$, i.e. $c = 0$,

$$0 = c_V \mathrm{d}T + p\mathrm{d}V.$$

From $pV = RT$ we have

$$\mathrm{d}T = \frac{p\mathrm{d}V + V\mathrm{d}p}{R}$$

so that

$$\begin{aligned} 0 &= p\mathrm{d}V + \frac{c_V}{R}(p\mathrm{d}V + V\mathrm{d}p) \\ &= \frac{1}{R}\big(c_p p\mathrm{d}V + c_V V\mathrm{d}p\big) \qquad \text{since } R = c_p - c_V \\ &= \frac{V^{\gamma-1}}{\gamma R}\mathrm{d}\big(pV^\gamma\big) \qquad \text{where} \qquad \gamma = \frac{c_p}{c_V}. \end{aligned}$$

Hence $pV^\gamma = \text{const.}$ for adiabatic processes.

Chapter 17

Problem 17.1 **Show that in spherical polar coordinates**

$$\mathrm{d}x \wedge \mathrm{d}y \wedge \mathrm{d}z = r^2 \sin\theta \mathrm{d}r \wedge \mathrm{d}\theta \wedge \mathrm{d}\phi,$$

and that $a^2 \sin\theta \mathrm{d}\theta \wedge \mathrm{d}\phi$ **is a volume element on the 2-sphere** $x^2+y^2+z^2 = a^2$**.**

Solution: In polar coordinates

$$x = r\sin\theta\cos\phi, \qquad y = r\sin\theta\sin\phi, \qquad z = r\cos\theta$$

we have

$$\begin{aligned}
\mathrm{d}x &= \sin\theta\cos\phi\,\mathrm{d}r + r\cos\theta\cos\phi\,\mathrm{d}\theta - r\sin\theta\sin\phi\,\mathrm{d}\phi \\
\mathrm{d}y &= \sin\theta\sin\phi\,\mathrm{d}r + r\cos\theta\sin\phi\,\mathrm{d}\theta + r\sin\theta\cos\phi\,\mathrm{d}\phi \\
\mathrm{d}z &= \cos\theta\,\mathrm{d}r - r\sin\theta\,\mathrm{d}\theta.
\end{aligned}$$

Hence, after a straightforward calculation

$$\begin{aligned}
\mathrm{d}x \wedge \mathrm{d}y \wedge \mathrm{d}z &= \det\frac{\partial(x,y,z)}{\partial(r,\theta\phi)}\mathrm{d}r \wedge \mathrm{d}\theta \wedge \mathrm{d}\phi \\
&= r^2 \sin\theta \mathrm{d}r \wedge \mathrm{d}\theta \wedge \mathrm{d}\phi.
\end{aligned}$$

On the 2-sphere $x^2+y^2+z^2 = a^2$, polar coordinates θ, ϕ are defined everwhere except on the region $0 < \theta\pi$, $0 < \phi < 2\pi$. On the this region it is clear that the 2-form $\omega = a^2 \sin\theta \mathrm{d}\theta \wedge \mathrm{d}\phi \neq 0$. This 2-form is undefined at $\phi = 0$ (or equivalently $\phi = 2\pi$), but has a unique continuous limit at these points, for the coordinate transformation $\phi' = \phi - \pi$, $\theta' = \theta$ results in

$$\omega = a^2 \sin\theta \mathrm{d}\theta \wedge \mathrm{d}\phi' \neq 0 \qquad \text{at } \phi' = \pm\pi, \quad \theta' \neq 0, \pi.$$

To see the behaviour at $\theta = 0, \pi$, consider rectangluar coordinates x, y for $z > 0$ and $z < 0$ respectively. Since

$$x = a\sin\theta\cos\phi, \qquad y = a\sin\theta\sin\phi, \qquad z = a\cos\theta$$

we have $z\mathrm{d}z + x\mathrm{d}x + y\mathrm{d}y + z\mathrm{d}z = 0$, so that

$$-\sin\theta\mathrm{d}\theta = \frac{1}{a}\mathrm{d}z = -\frac{x\mathrm{d}x + y\mathrm{d}y}{az}$$

and, from $\tan\phi = y/x$,

$$\sec^2\phi\mathrm{d}\phi = \frac{x\mathrm{d}y - y\mathrm{d}x}{x^2}.$$

Hence

$$\mathrm{d}\phi = \cos^2\phi\frac{x\mathrm{d}y - y\mathrm{d}x}{x^2} = \frac{x\mathrm{d}y - y\mathrm{d}x}{a^2\sin^2\theta} = \frac{x\mathrm{d}y - y\mathrm{d}x}{a^2 - z^2}$$

and

$$\begin{aligned}\omega &= a\frac{x\mathrm{d}x + y\mathrm{d}y}{z} \wedge \frac{x\mathrm{d}y - y\mathrm{d}x}{a^2 - z^2} \\ &= \frac{a}{z}\frac{(x+y)^2\mathrm{d}x \wedge \mathrm{d}y}{x^2+y^2} \\ &= \frac{a\mathrm{d}x \wedge \mathrm{d}y}{\pm\sqrt{a^2 - x^2 - y^2}}.\end{aligned}$$

At $\theta = 0, \pi$ we have $z = a$ and $z = -a$ resp., i.e. $x = y = 0$. Hence the limiting values of ω at these points are $\pm\mathrm{d}x \wedge \mathrm{d}y \neq 0$. Thus ω defines a non-vanishing 2-form on the entire 2-sphere.

Problem 17.2 Show that the 2-sphere $x^2 + y^2 + z^2 = 1$ is an orientable manifold.

Solution: Steregraphic projection coordinates from the north and south pole are given in Example 15.4

$$\begin{aligned} X &= \frac{x}{1-z}, \qquad & Y &= \frac{y}{1-z} \\ X' &= \frac{x}{1+z}, \qquad & Y' &= \frac{y}{1+z}. \end{aligned}$$

These coordinates cover the 2-sphere. From

$$X^2 + Y^2 = \frac{x^2+y^2}{(1-z)^2} = \frac{1-z^2}{(1-z)^2} = \frac{1-z}{1+z}$$

it follows immediately that

$$X' = \frac{X}{X^2+Y^2}, \qquad Y' = \frac{Y}{X^2+Y^2}$$

and the jacobian matrix is

$$\mathsf{J} = \frac{\partial(X', Y')}{\partial(X, Y)} = \frac{1}{(X^2+Y^2)^2}\begin{pmatrix} (Y^2 - X^2 & -2YX \\ -2YX & X^2 - Y^2 \end{pmatrix}$$

and the jacobian determinant is

$$\det \mathsf{J} = -\frac{(X^2-Y^2)^2}{(X^2+Y^2)^4} - \frac{4X^2Y^2}{(X^2+Y^2)^4} = \frac{-1}{(X^2+Y^2)^2}$$

so that these coordinates are oppositely oriented throughout their region of intersection $(X, Y), (X', Y') \neq (0, 0)$. Choose $X'' = X'$ and $Y'' = -Y'$ and the coordinates are positively oriented throughout their region of intersection,

$$\det \frac{\partial(X'', Y'')}{\partial(X, Y)} = \frac{1}{(X^2+Y^2)^2} > 0.$$

Hence S^2 is an oriented manifold.

Problem 17.3 **Show that the definition of the integral of an n-form over a manifold M given in Eq. (17.1) is independent of the choice of partition of unity subordinate to $\{U_a\}$.**

Solution: If g_a and h_a are partitions of unity subordinate to $\{U_a\}$ and α any n-form on M, then

$$\begin{aligned}\sum_b \int_M h_b \alpha &= \sum_b \sum_a \int_M g_a h_b \alpha \\ &= \sum_a \Big(\sum_b \int_M h_b g_a \alpha \Big) \\ &= \sum_a \int_M g_a \alpha\end{aligned}$$

since the support of each n-form $h_b g_a \alpha$ is included in U_a. Hence the definition of integral of α over M is independent of the choice of partition

$$\int_M \alpha = \sum_a \int_M g_a \alpha = \sum_b \int_M h_b \alpha.$$

Problem 17.4 **Let $\alpha = y^2 \mathrm{d}x + x^2 \mathrm{d}y$. If γ_1 is the stretch of y-axis from $(x = 0, y = -1)$ to $(x = 0, y = 1)$, and γ_2 the unit right semicircle connecting these points, evaluate**

$$\int_{\gamma_1} \alpha, \qquad \int_{\gamma_2} \alpha, \qquad \textbf{and} \qquad \int_{S^1} \alpha.$$

Verify Stokes' theorem for the unit circle and the unit right semicircular region encompassed by γ_1 and γ_2.

Solution: If γ is a curve given parametrically by $x = x(t)$, $y = y(t)$ we have, by Example 17.2,

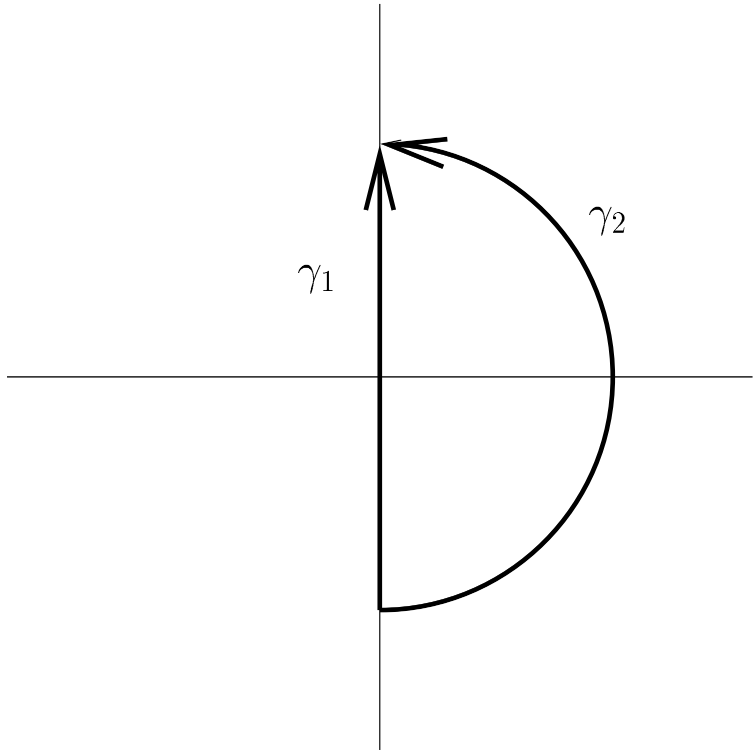
$$\int_\gamma \alpha = \int_{t_1}^{t_2} y^2 \frac{\mathrm{d}x}{\mathrm{d}t} + x^2 \frac{\mathrm{d}y}{\mathrm{d}t}\, \mathrm{d}t.$$

For γ_1 the parametric representation is $x = 0$, $y = t$, $t_1 = -1$, $t_2 = 1$, so that

$$\int_{\gamma_1} \alpha = \int_{-1}^{1} (t^2.0 + 0.1)\mathrm{d}t = 0.$$

The curve γ_2 has the parametric representation is $x = \cos t$, $y = \sin t$, $t_1 = -\pi/2$, $t_2 = \pi/2$:

$$\int_{\gamma_2} \alpha = \int_{-\pi/2}^{\pi/2} (-\sin^2 t. \sin t + \cos^2 t. \cos t)\mathrm{d}t.$$



Since

$$\begin{aligned}\cos^3 t - \sin^3 t &= (\cos t - \sin t)(\cos^2 t + \cos t \sin t + \sin^2 t)\\ &= (\cos t - \sin t)(1 + \cos t \sin t)\\ &= \cos t - \sin t + \cos^2 t \sin t - \sin^2 t \cos t\end{aligned}$$

we have

$$\int_{\gamma_2} \alpha = \left[\sin t + \cos t - \frac{\cos^3 t}{3} - \frac{\sin^3 t}{3}\right]_{-\pi/2}^{\pi/2} = \frac{4}{3}\,.$$

For $\gamma_3 = S^1$ the parametric representation is the same as for γ^2, but $t_1 = 0$, $t_2 = 2\pi$, giving

$$\int_{\gamma_3} \alpha = \left[\sin t + \cos t - \frac{\cos^3 t}{3} - \frac{\sin^3 t}{3}\right]_{0}^{2\pi} = 0.$$

The exterior derivative $\mathrm{d}\alpha$ is

$$\mathrm{d}\alpha = 2y\mathrm{d}y \wedge \mathrm{d}x + 2x\mathrm{d}x \wedge \mathrm{d}y = 2(x-y)\mathrm{d}x \wedge \mathrm{d}y.$$

Using polar coordinates, $x = r\cos\theta$, $y = r\sin\theta$,

$$\mathrm{d}\alpha = 2r^2(\cos\theta - \sin\theta)\mathrm{d}r \wedge \mathrm{d}\theta$$

and the integral over the the unit disc D is

$$\begin{aligned}\int_D r d\alpha &= \int_0^1 \int_0^{2\pi} 2r^2(\cos\theta - \sin\theta)\mathrm{d}r \wedge \mathrm{d}\theta \\ &= \left[\frac{2r^3}{3}\right]_0^1 \left[\sin\theta + \cos\theta\right]_0^{2\pi} \\ &= 0 \\ &= \int_{\gamma_3} \alpha.\end{aligned}$$

If D' is the right half disc then

$$\int_{D'} r d\alpha = \frac{2}{3}\left[\sin\theta + \cos\theta\right]_{-\pi/2}^{\pi/2} = \frac{4}{3} = \int_{\gamma_2} \alpha - \int_{\gamma_1} \alpha.$$

Since the boundary of D' is $\partial D' = \gamma_2 - \gamma_1$, we have Stokes' theorem satisfied

$$\int_{D'} r d\alpha = \int_{\partial D'} \alpha.$$

Problem 17.5 **If $\alpha = x\,\mathrm{d}y \wedge \mathrm{d}z + y\,\mathrm{d}z \wedge \mathrm{d}x + z\,\mathrm{d}x \wedge \mathrm{d}y$ compute $\int_{\partial\Omega} \alpha$ where Ω is (i) the unit cube, (ii) the unit ball in $\mathbb{R}^3$. In each case verify Stokes' theorem,**

$$\int_{\partial\Omega} \alpha = \int_{\Omega} \mathrm{d}\alpha.$$

Solution: (i) On the face $x = 0$ the outward normal is $\mathbf{n} = (-1, 0, 0)$, and a correctly ordered pair of tangent vectors is $(\mathbf{e} = (0, 0, 1), \mathbf{f} = (0, 1, 0)$, so that $(\mathbf{n}, \mathbf{e}, \mathbf{f})$ is a correctly oriented triple with respect to the standard basis $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ of $\mathbb{R}^3$. A parametric description of this face is $x = 0$, $y = \lambda_2$, $z = \lambda_1$, so that

$$\mathbf{e} = \frac{\partial x^i}{\partial \lambda_1}\mathbf{e}_i, \quad \mathbf{f} = \frac{\partial x^i}{\partial \lambda_2}\mathbf{e}_i,$$

and

$$\int_{x=0} \alpha = \int_0^1 \int_0^1 -x \frac{\partial y}{\partial \lambda_2}\frac{\partial z}{\partial \lambda_1} \mathrm{d}\lambda_1\,\mathrm{d}\lambda_2 = \int_0^1 \int_0^1 -0.1.1 \mathrm{d}\lambda_1 \mathrm{d}\,\lambda_2 = 0.$$

Similarly, the face $x = 1$ has parametric representation $y = \lambda_1$, $z = \lambda_2$ and

$$\int_{x=1} \alpha = \int_0^1 \int_0^1 x \frac{\partial y}{\partial \lambda_1}\frac{\partial z}{\partial \lambda_2} \mathrm{d}\lambda_1\,\mathrm{d}\lambda_2 = \int_0^1 \int_0^1 1 \mathrm{d}\lambda_1 \mathrm{d}\,\lambda_2 = 1\,.$$

The other faces are dealt with in a completely analogous fashion:

$$\int_{y=0} \alpha = \int_{z=0} \alpha = 0$$

and

$$\int_{y=1} \alpha = \int_0^1 \int_0^1 y \mathrm{d}\lambda_1 \, \mathrm{d}\lambda_2 = 1,$$

$$\int_{z=1} \alpha = \int_0^1 \int_0^1 z \mathrm{d}\lambda_1 \, \mathrm{d}\lambda_2 = 1.$$

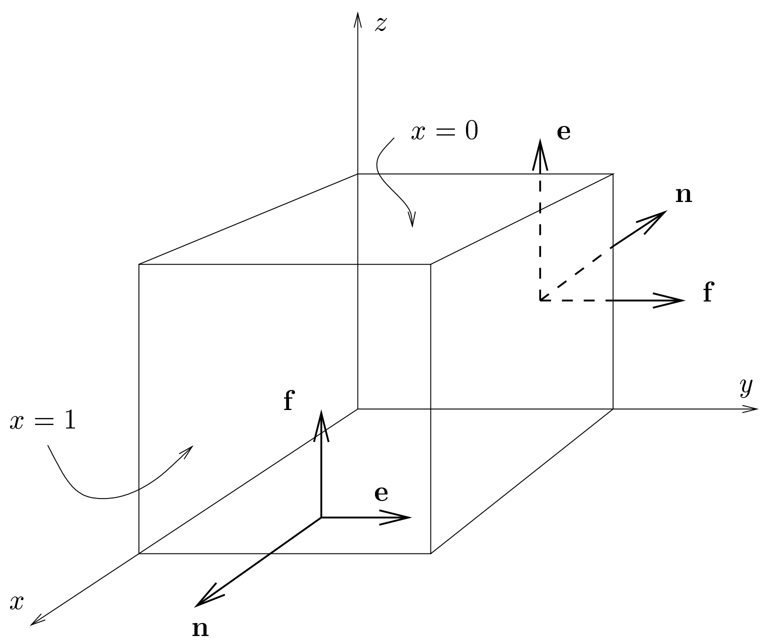
Hence

$$\int_{\partial\Omega} \alpha = 3.$$

To verify Stokes' theorem,

$$\int_\Omega \mathrm{d}\alpha = \int_\Omega 3 \mathrm{d}x \wedge \mathrm{d}y \wedge \mathrm{d}z = 3 \int_0^1 \int_0^1 \int_0^1 \mathrm{d}x \, \mathrm{d}y \, \mathrm{d}z = 3\,.$$

(ii) Setting r, θ, ϕ to be spherical polar coordinates



$$x = r \sin\theta \cos\phi, \qquad y = r \sin\theta \sin\phi, \qquad z = r \cos\theta$$

we may use θ, ϕ as parametric coordinates on the boundary $\partial\Omega = S^2$, positively oriented since ∂_r is an outward normal. From Example 17.3

$$\begin{aligned}
\int_{\partial\Omega}\alpha &= \iint_{r=1} \alpha_{ij}\frac{\partial x^i}{\partial\theta}\frac{\partial x^j}{\partial\phi}\mathrm{d}\theta\,\mathrm{d}\phi \\
&= \int_0^{\pi}\mathrm{d}\theta\int_0^{2\pi}\mathrm{d}\phi\; x\Big(\frac{\partial y}{\partial\theta}\frac{\partial z}{\partial\phi}-\frac{\partial z}{\partial\theta}\frac{\partial y}{\partial\phi}\Big)_{r=1}+y\Big(\frac{\partial z}{\partial\theta}\frac{\partial x}{\partial\phi}-\frac{\partial x}{\partial\theta}\frac{\partial z}{\partial\phi}\Big)_{r=1} \\
&\qquad + z\Big(\frac{\partial y}{\partial\theta}\frac{\partial z}{\partial\phi}-\frac{\partial z}{\partial\theta}\frac{\partial y}{\partial\phi}\Big)_{r=1} \\
&= \int_0^{\pi}\mathrm{d}\theta\int_0^{2\pi}\mathrm{d}\phi\;\sin^3\theta\cos^2\phi+\sin^3\theta\sin^2\phi+\sin^3\theta\sin^2\phi \\
&\qquad +\cos\theta(\cos\theta\sin\theta\cos^2\phi+\cos\theta\sin\theta\sin^2\phi) \\
&= \int_0^{\pi}\mathrm{d}\theta\int_0^{2\pi}\mathrm{d}\phi\;\sin^3\theta+\cos^2\theta\sin\theta \\
&= \int_0^{\pi}\sin\theta\,\mathrm{d}\theta\int_0^{2\pi}\mathrm{d}\phi \\
&= 4\pi.
\end{aligned}$$

In polar coordinates the exterior derivative is (see Problem 17.1)

$$\mathrm{d}\alpha = 3\mathrm{d}x\wedge\mathrm{d}y\wedge\mathrm{d}z = r^2\sin\theta\mathrm{d}r\wedge\mathrm{d}\theta\wedge\mathrm{d}\phi,$$

so that

$$\int_\Omega\mathrm{d}\alpha = 3\int_0^1 r^2\mathrm{d}r\int_0^{\pi}\sin\theta\,\mathrm{d}\theta\int_0^{2\pi}\mathrm{d}\phi = 4\pi$$

in agreement with Stokes' theorem.

Problem 17.6 **Let S be the surface of a cylinder of elliptical cross-section and height $2h$ given by**

$$x=a\cos\theta,\quad y=b\sin\theta,\ (0\le\theta<2\pi),\qquad -h\le z\le h.$$

(a) Compute $\int_S\alpha$ where $\alpha = x\,\mathrm{d}y\wedge\mathrm{d}z+y\,\mathrm{d}z\wedge\mathrm{d}x-2z\,\mathrm{d}x\wedge\mathrm{d}y$.

(b) Show $\mathrm{d}\alpha=0$, and find a 1-form ω such that $\alpha=\mathrm{d}\omega$.

(c) Verify Stokes' theorem $\int_S\alpha=\int_{\partial S}\omega$.

Solution: (a) Set $\lambda_1 = \theta$, $\lambda_2 = z$ on the surface S, and use Example 17.3

$$\begin{aligned}
\int_S \alpha &= \int\!\!\int \alpha_{ij} \frac{\partial x^i}{\partial \lambda_1}\frac{\partial x^j}{\partial \lambda_2} \mathrm{d}\lambda_1\, \mathrm{d}\lambda_2 \\
&\equiv \int\!\!\int \alpha_{ij} \Big(\frac{\partial x^i}{\partial \lambda_1}\frac{\partial x^j}{\partial \lambda_2} - \frac{\partial x^j}{\partial \lambda_1}\frac{\partial x^i}{\partial \lambda_2}\Big) \mathrm{d}\lambda_1\, \mathrm{d}\lambda_2 \\
&= \int_0^{2\pi} \mathrm{d}\theta \int_{-h}^{h} \mathrm{d}z \Big[x\Big(\frac{\partial y}{\partial \theta}\frac{\partial z}{\partial z} - \frac{\partial z}{\partial \theta}\frac{\partial y}{\partial z}\Big) + y\Big(\frac{\partial z}{\partial \theta}\frac{\partial x}{\partial z} - \frac{\partial x}{\partial \theta}\frac{\partial z}{\partial z}\Big) \\
&\qquad - 2z\Big(\frac{\partial x}{\partial \theta}\frac{\partial y}{\partial z} - \frac{\partial y}{\partial \theta}\frac{\partial x}{\partial z}\Big)\Big] \\
&= \int_0^{2\pi} \mathrm{d}\theta \int_{-h}^{h} \mathrm{d}z a \cos .b\cos\theta - b\sin\theta.(-a\sin\theta) \\
&= \int_0^{2\pi} \mathrm{d}\theta \int_{-h}^{h} \mathrm{d}z\; ab \\
&= 4\pi hab.
\end{aligned}$$

(b) $\mathrm{d}\alpha = \mathrm{d}x \wedge \mathrm{d}y \wedge \mathrm{d}z + \mathrm{d}y \wedge \mathrm{d}z \wedge \mathrm{d}x - 2\mathrm{d}z \wedge \mathrm{d}x \wedge \mathrm{d}y = 0$. To solve for $\alpha = \mathrm{d}\omega$ where $\omega = A\mathrm{d}x + B\mathrm{d}y + C\mathrm{d}z$, i.e.

$$\alpha = \mathrm{d}A \wedge \mathrm{d}x + \mathrm{d}B \wedge \mathrm{d}y + \mathrm{d}C \wedge \mathrm{d}z,$$

we read off the components of $\mathrm{d}y \wedge \mathrm{d}z$, $\mathrm{d}z \wedge \mathrm{d}x$ and $\mathrm{d}x \wedge \mathrm{d}y$ resp.:

$$x = -B_{,z} + C_{,y} \tag{1}$$

$$y = -C_{,x} + A_{,z} \tag{2}$$

$$-2z = -A_{,y} + B_{,x} \tag{3}$$

where $B_{,z} = \partial B/\partial z$ etc. Try $C = 0$, and (1) and (2) give

$$B = -zx + g(x, y)$$
$$A = zy + f(x, y)$$

for some functions f and g. Substituting in (3) results in

$$-2z = -z + f_{,y} - z + g_{,x}$$

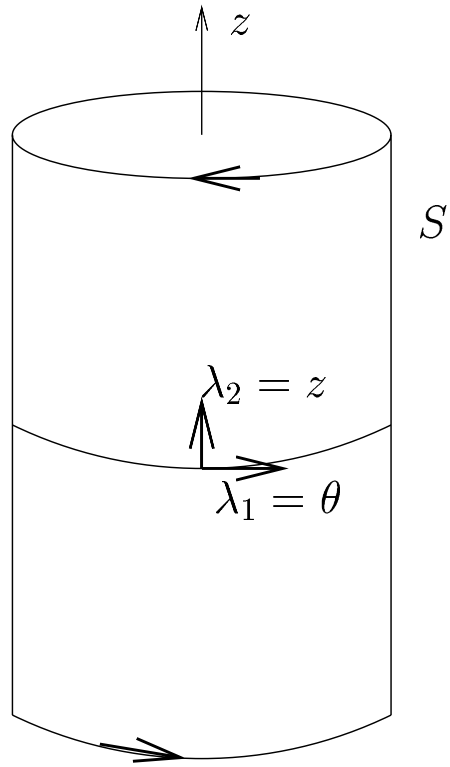
i.e.

$$f_{,y} + g_{,x} = 0.$$

The functions $f = g = 0$ satisfy this equation and a suitable 1-form ω is

$$\omega = zy\mathrm{d}x - zx\mathrm{d}y.$$

(c) The boundary of S consists of two ellipses at $z = \pm h$. We may take as parameter $\lambda = -\theta$ on $z = h$ (i.e. a clockwise direction), for $\mathbf{n} = \partial_z$ is the outward normal to the



boundary ellipse and $\mathbf{n}, \partial_\theta$ has negative orientation w.r.t. the positive orientation on S. Similarly, on $z = -h$ we take $\lambda = -\theta$ (i.e. anticlockwise direction). Hence

$$\begin{aligned}
\int_{\partial S} \omega &= \int_{z=-h} -h(y\mathrm{d}x - x\mathrm{d}y) + \int_{z=h} h(y\mathrm{d}x - x\mathrm{d}y) \\
&= \int_0^{2\pi} -h(b\sin\theta.(-a\sin\theta)\mathrm{d}\theta - a\cos\theta.b\cos\theta\mathrm{d}\theta) \\
&\quad - \int_0^{2\pi} h(b\sin\theta.(-a\sin\theta)\mathrm{d}\theta - a\cos\theta.b\cos\theta\mathrm{d}\theta) \\
&= 2\int_0^{2\pi} hab\mathrm{d}\theta \\
&= 4\pi hab = \int_S \alpha.
\end{aligned}$$

Problem 17.7 A torus in $\mathbb{R}^3$ may be represented parametrically by

$$x = \cos\phi(a + b\cos\psi) \qquad y = \sin\phi(a + b\cos\psi) \qquad z = b\sin\psi$$

where $0 \le \phi < 2\pi$, $0 \le \psi < 2\pi$. If b is replaced by a variable ρ that ranges

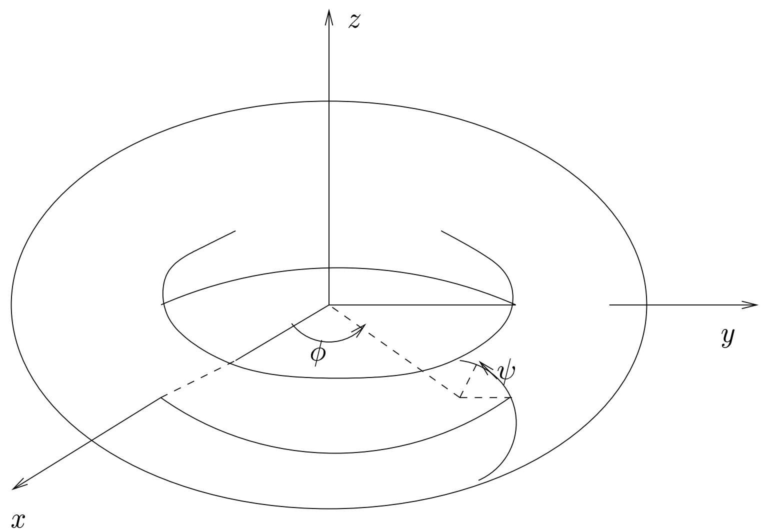
from 0 to b, show that

$$\mathrm{d}x \wedge \mathrm{d}y \wedge \mathrm{d}z = \rho(a + \rho\cos\psi)\mathrm{d}\phi \wedge \mathrm{d}\psi \wedge \mathrm{d}\rho.$$

By integrating this 3-form over the region enclosed by the torus, show that the volume of the solid torus is $2\pi^2 ab^2$. Can you see this by a simple geometrical argument?

Evaluate the volume by performing the integral of the 2-form $\alpha = x\,\mathrm{d}y \wedge \mathrm{d}z$ over the surface of the torus and using Stokes' theorem.

Solution: For each fixed ϕ the curve



$$\begin{aligned} x - a\cos\phi &= b\cos\phi\cos\psi \\ y - a\sin\phi &= b\sin\phi\cos\psi \\ z &= b\sin\psi \end{aligned}$$

describes a circle centre center $(a\cos\phi, a\sin\phi, 0)$, radius b,

$$(x - a\cos\phi)^2 + (y - a\sin\phi)^2 = b^2\cos^2\psi = b^2 - z^2.$$

As ϕ varies from 0 to 2π these circles describe the surface of a torus (donut). Replacing b by a variable ρ and forming differentials, we have

$$\begin{aligned} \mathrm{d}x &= -\sin\phi(a + \rho\cos\psi)\mathrm{d}\phi - \rho\cos\phi\sin\psi\mathrm{d}\psi + \cos\phi\cos\psi\mathrm{d}\rho \\ \mathrm{d}y &= \cos\phi(a + \rho\cos\psi)\mathrm{d}\phi - \rho\sin\phi\sin\psi\mathrm{d}\psi + \sin\phi\cos\psi\mathrm{d}\rho \\ \mathrm{d}z &= \rho\cos\psi\mathrm{d}\phi + \sin\psi\mathrm{d}\rho \end{aligned}$$

Hence

$$\begin{aligned}
\mathrm{d}x \wedge \mathrm{d}y \wedge \mathrm{d}z = \mathrm{d}\phi \wedge \mathrm{d}\psi \wedge \mathrm{d}\rho\big[&\rho \sin^2\phi(a+\rho\cos\psi)\sin^2\psi \\
&\sin^2\phi(a+\rho\cos\psi)\cos^2\psi\rho + \rho\cos^2\phi\sin^2\psi(a+\rho\cos\psi) \\
&\cos^2\phi\cos^2\psi(a+\rho\cos\psi)\rho\big] \\
= \mathrm{d}\phi \wedge \mathrm{d}\psi \wedge \mathrm{d}\rho&(a+\rho\cos\psi)\rho
\end{aligned}$$

The volume of the region D enclosed by the torus is therefore

$$\begin{aligned}
V = \int_D \mathrm{d}x \wedge \mathrm{d}y \wedge \mathrm{d}z &= \int_0^{2\pi} \mathrm{d}\phi \int_0^{2\pi} \mathrm{d}\psi \int_0^b \mathrm{d}\rho (a+\rho\cos\psi)\rho \\
&= 2\pi \int_0^{2\pi} \mathrm{d}\psi \Big(a\frac{b^2}{2} + \frac{b^3}{3}\cos\psi\Big) \\
&= 2\pi^2 ab^2.
\end{aligned}$$

As this volume is generated by a sequence of discs of area πb^2 sith centres strung along a circle of circumference $2\pi a$, the volume may be expeceted to be their product $2\pi a.\pi b^2 = 2\pi^2 ab^2$. Since $\mathrm{d}\alpha = \mathrm{d}x \wedge \mathrm{d}y \wedge \mathrm{d}z$, Stokes' theorem gives

$$V = \int_D \mathrm{d}x \wedge \mathrm{d}y \wedge \mathrm{d}z = \int_D \mathrm{d}\alpha = \int_{\partial D} \alpha.$$

Hence

$$\begin{aligned}
V &= \int_{\partial D} x\,\mathrm{d}y \wedge \mathrm{d}z \\
&= \int_{\partial D} \cos\phi(a+\rho\cos\psi)\Big(\frac{\partial y}{\partial\phi}\frac{\partial z}{\partial\psi} - \frac{\partial z}{\partial\phi}\frac{\partial y}{\partial\psi}\Big)\mathrm{d}\phi \wedge \mathrm{d}\psi \\
&= \int_0^{2\pi} \mathrm{d}\phi \int_0^{2\pi} \mathrm{d}\psi \cos\phi(a+\rho\cos\psi)\cos\phi(a+\rho\cos\psi)b\cos\psi \\
&= \int_0^{2\pi} \cos^2\phi \mathrm{d}\phi \int_0^{2\pi} (a+\rho\cos\psi)^2 b\cos\psi \mathrm{d}\psi \\
&= \int_0^{2\pi} \frac{\cos 2\phi + 1}{2}\mathrm{d}\phi \int_0^{2\pi} (a^2 + 2ab\cos\psi + b^2\cos^2\psi)b\cos\psi\mathrm{d}\psi \\
&= \pi \int_0^{2\pi} 2ab^2\cos^2\psi\,\mathrm{d}\psi \\
&= \pi.2\pi ab^2 = 2\pi^2 ab^2,
\end{aligned}$$

as required.

Problem 17.8 **Show that in n dimensions, if V is a regular n-domain with boundary $S = \partial V$,and we set α to be an $(n-1)$-form with components**

$$\alpha = \sum_{i=1}^{n} (-1)^{i+1} A^i \mathrm{d}x^1 \wedge \cdots \wedge \mathrm{d}x^{i-1} \wedge \mathrm{d}x^{i+1} \wedge \cdots \wedge \mathrm{d}x^n,$$

Stokes' theorem can be reduced to the *n-dimensional Gauss theorem*

$$\underset{V}{\int\cdots\int} A^i{}_{,i}\mathrm{d}x^1\ldots\mathrm{d}x^n = \underset{S}{\int\cdots\int} A^i\mathrm{d}S_i$$

where $\mathrm{d}S_i = (-1)^{i+1}\mathrm{d}x^1\ldots\mathrm{d}x^{i-1}\mathrm{d}x^{i+1}\ldots\mathrm{d}x^n$ is a 'vector volume element' normal to S.
[Note: Term $(-1)^{i+1}$ should be included in definition of S_i.]**

Solution: Suspending the summation convention

$$\begin{aligned}\mathrm{d}\alpha &= \sum_{i=1}^{n}(-1)^{i+1}\sum_{j=1}^{n} A^i{}_{,j}\mathrm{d}x^j\wedge\mathrm{d}x^1\wedge\cdots\wedge\mathrm{d}x^{i-1}\wedge\mathrm{d}x^{i+1}\wedge\cdots\wedge\mathrm{d}x^n,\\ &= \sum_{i=1}^{n}(-1)^{i+1}A^i{}_{,i}\mathrm{d}x^i\wedge\mathrm{d}x^1\wedge\cdots\wedge\mathrm{d}x^{i-1}\wedge\mathrm{d}x^{i+1}\wedge\cdots\wedge\mathrm{d}x^n,\\ &= \sum_{i=1}^{n}(-1)^{i+1}(-1)^{i-1}A^i{}_{,i}\mathrm{d}x^i\wedge\mathrm{d}x^1\wedge\cdots\wedge\mathrm{d}x^{i-1}\wedge\mathrm{d}x^i\wedge\mathrm{d}x^{i+1}\wedge\cdots\wedge\mathrm{d}x^n,\end{aligned}$$

after performing $i-1$ interchanges

$$= A^i{}_{,i}\wedge\mathrm{d}x^1\wedge\mathrm{d}x^2\wedge\cdots\wedge\mathrm{d}x^n,$$

on reinstating the summation convention on the index i. Hence, in a coordinate chart $(U,\varphi;x^i)$, by Stokes' theorem

$$\begin{aligned}\underset{\varphi(V)}{\int\cdots\int} A^i{}_{,i}\mathrm{d}x^1\ldots\mathrm{d}x^n &= \int_V \mathrm{d}\alpha\\ &= \int_{\partial V}\alpha\\ &= \sum_{i=1}^{n}\int_{\partial V}(-1)^{i+1}A^i\mathrm{d}x^1\wedge\cdots\wedge\mathrm{d}x^{i-1}\wedge\mathrm{d}x^{i+1}\wedge\cdots\wedge\mathrm{d}x^n\\ &= \underset{S}{\int\cdots\int} A^i\mathrm{d}S_i \qquad (S=\varphi(\partial V))\end{aligned}$$

where

$$\begin{aligned}\mathrm{d}S_i &= (-1)^{i+1}\mathrm{d}x^1\ldots\mathrm{d}x^{i-1}\mathrm{d}x^{i+1}\ldots\mathrm{d}x^n\\ &= (-1)^{i+1}\frac{\partial x^1}{\partial\lambda^{j_1}}\cdots\frac{\partial x^{i-1}}{\partial\lambda^{j_{i-1}}}\frac{\partial x^{i+1}}{\partial\lambda^{j_i}}\cdots\frac{\partial x^n}{\partial\lambda^{j_{n-1}}}\epsilon^{j_1\ldots j_{n-1}}\mathrm{d}\lambda^1\ldots\mathrm{d}\lambda^{n-1}\end{aligned}$$

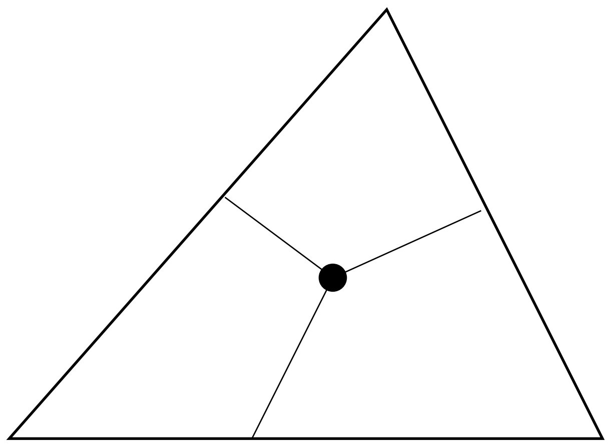
in parametric form. $\mathrm{d}S_i$ is "normal" to the surface ∂V in the sense that if

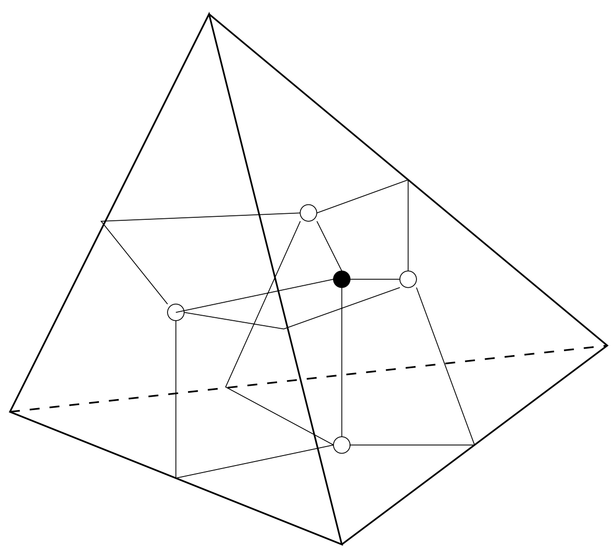
$$B^i = \sum_{j=1}^{n-1} b^j\frac{\partial x^i}{\partial\lambda^j}$$

is any vector tangent to the surface then it is straightforward to verify that $B^i \mathrm{d}S_i = 0$.

Problem 17.9 Show that any tetrahedron may be divided into 'cubical' regions.
Describe a procedure for achieving the same result for a general k-simplex.
[Note: the caption to Fig 17.4 should read *Dividing a cube into "triangular" cells*]

Solution: For a 2-simplex (triangle) pick any point inside the triangle, which we call its "centre" and join to the centres of the edges. This divides the 2-simples into three quadrilateral regions, or 2-cubes. For a 3-simplex or tetrahedron pick a





point or centre in the interior of the 3-simplex and join to centres of all four faces. Retaining the divisions of these faces relative these centres divides the 3-simplex into four 3-cubes as shown in the Figure. An n-simplex has $n+1$ faces. Join a central point p_0 to centres $p_1, \dots, p_{n+1}$ of these faces and cubulate each face with respect to these centres. This inductive process cubulates the n-simplex into $n+1$ n-cubes.

Problem 17.10 **For any pair of subspaces H and K of the exterior algebra $\Lambda^*(M)$, set $H \wedge K$ to be the vector subspace spanned by all $\alpha \wedge \beta$ where $\alpha \in H$, $\beta \in K$. Show that**

(a) $Z^p(M) \wedge Z^q(M) \subseteq Z^{p+q}(M)$,

(b) $Z^p(M) \wedge B^q(M) \subseteq B^{p+q}(M)$,

(c) $B^p(M) \wedge B^q(M) \subseteq B^{p+q}(M)$.

Solution: (a) Let $\alpha \in Z^p(M)$, $\beta \in Z^q(M)$, so that $\mathrm{d}\alpha = 0$ and $\mathrm{d}\beta = 0$. Then $\alpha \wedge \beta$ is a $(p+q)$-form and

$$\mathrm{d}(\alpha \wedge \beta) = \mathrm{d}\alpha \wedge \beta + (-1)^p \alpha \wedge \mathrm{d}\beta = 0.$$

Hence $\alpha \wedge \beta \in Z^{p+q}(M)$ and we have shown that

$$Z^p(M) \wedge Z^q(M) \subseteq Z^{p+q}(M).$$

(b) If $\alpha \in Z^p(M)$, $\beta \in B^q(M)$, then $\mathrm{d}\alpha = 0$ and $\beta = \mathrm{d}\gamma$ where $\gamma \in C^{q-1}(M)$. Hence

$$\mathrm{d}(\alpha \wedge \gamma) = \mathrm{d}\alpha \wedge \gamma + (-1)^p \alpha \wedge \mathrm{d}\gamma = (-1)^p \alpha \wedge \beta.$$

Thus $\alpha \wedge \beta = (-1)^p \mathrm{d}(\alpha \wedge \gamma)$, so that $\alpha \wedge \beta \in B^{p+q}(M)$, and it follows that

$$Z^p(M) \wedge B^q(M) \subseteq B^{p+q}(M).$$

(c) Since $B^p(M) \subseteq Z^p(M)$ it is immediate from (b) that

$$B^p(M) \wedge B^q(M) \subseteq Z^p(M) \wedge B^q(M) \subseteq B^{p+q}(M).$$

Problem 17.11 **Show that for any set of real numbers $a_1, \dots, a_k$ there exists a closed r-form α whose periods $\int_{C_i} \alpha = a_i$.**

Solution: This is a direct corollary of de Rham' theorem (Theorem 17.4). Let $\{[C_i]\}$ be a basis of the rth homology space $H_r(M)$ where $C_i \in Z_r(M)$ are r-cycles ($i = 1, \dots, b_r$). Theorem 17.4 implies that there exists a dual basis $\{[\varepsilon^j]\}$ ($j = 1, \dots, b_r$) of the cohomology space $H^r(M)$, where each ε^j is an r-cocycle (closed r-form), such that

$$\langle [C_i], [\varepsilon^j] \rangle = \int_{C_i} \varepsilon^j = \delta^j_i.$$

The closed r-form $\alpha = a_j \varepsilon^j$ then has periods

$$\int_{C_i} \alpha = \int_{C_i} a_j \varepsilon^j = a_j \delta^j_i = a_i.$$

Problem 17.12 **If S^1 is the unit circle, show that $b^0 = b^1 = 1$.**

Solution: Since S^1 is a connected topological space (see Problem 10.19) we have by Example 17.5 that $H^0(S^1) = \mathbb{R}$, i.e. $b^0 = 1$.

Let ω be any 1-form on S^1, expressed locally in angular coordinates as $\omega = f(\theta)\mathrm{d}\theta$, where $x = \cos\theta$, $y = \sin\theta$. Then ω is a closed 1-form ($\omega \in Z^1(S^1)$), since

$$\mathrm{d}\omega = f'(\theta)\mathrm{d}\theta \wedge \mathrm{d}\theta = 0.$$

ω is exact if there exists a function $F \in \mathcal{F}(S^1)$ such that $\omega = \mathrm{d}F$,

$$\omega = \frac{\mathrm{d}F}{\mathrm{d}\theta}\mathrm{d}\theta, \qquad \text{where} \qquad F(\theta) = \int_0^\theta f(\theta')\mathrm{d}\theta'.$$

While this equation is true locally for any 1-form, it is not true globally in general since the function must be continuous on S^1, i.e. we require peridocity $F(2\pi) \equiv \lim_{\theta\to 2\pi} F(\theta) = F(0) = 0$. This condition may be expressed

$$\int_{S^1} \omega = \int_0^{2\pi} f(\theta)\mathrm{d}\theta = 0.$$

If ω and ω' are any pair of closed forms $\mathrm{d}\omega = \mathrm{d}\omega' = 0$ and ω' is not exact, $\int_{S^1}\omega \neq 0$, then there exists a unique $a \in \mathbb{R}$ such that $\omega' - a\omega$ is exact, for

$$\int_{S^1} \omega' - a\omega = 0 \qquad \text{where} \qquad a = \frac{\int_{S^1}\omega'}{\int_{S^1}\omega}.$$

The a with this property is uniquely defined, for if $\omega' - a'\omega$ is exact then

$$(a - a')\int_{S^1} \omega = 0$$

where is only possible if $a = a'$. Hence $\omega' - a\omega = \mathrm{d}F$ and every closed 1-form $\omega' \sim a\omega$ for a unique $a \in \mathbb{R}$, i.e. $H^1(S^1) = \mathbb{R}$ or equivalently the first Betti number is $b^1 = 1$.

Problem 17.13 **Let** $$\alpha = \frac{x\mathrm{d}y - y\mathrm{d}x}{x^2 + y^2}$$

Show that α is a closed 1-form on $\mathbb{R}^2 - \{\mathbf{0}\}$. Compute its integral over the unit circle S^1 and show that it is not exact. What does this tell us of the de Rham cohomology of $\mathbb{R}^2 - \{\mathbf{0}\}$ and S^1?

Solution: The 1-form α is clearly defined on $\mathbb{R}^2 - \{\mathbf{0}\}$ and closed, for

$$\begin{aligned} \mathrm{d}\alpha &= \frac{\mathrm{d}x \wedge \mathrm{d}y - \mathrm{d}y \wedge \mathrm{d}x}{x^2 + y^2} - \frac{(2x\mathrm{d}x + 2y\mathrm{d}y) \wedge (x\mathrm{d}y - y\mathrm{d}x)}{(x^2+y^2)^2} \\ &= \frac{2\mathrm{d}x \wedge \mathrm{d}y}{x^2+y^2} - \frac{2(x^2+y^2)\mathrm{d}x \wedge \mathrm{d}y}{(x^2+y^2)^2} \\ &= 0\,. \end{aligned}$$

The 1-form α is not exact for if $\alpha = \mathrm{d}f$ then we would have

$$\int_{S^1} \alpha = f(2\pi) - f(0) = 0$$

where we parametrize the unit circle as in Problem 7.12 ($x = \cos\theta,\ y = \sin\theta$). However

$$\int_{S^1} \alpha = \int_0^{2\pi} \frac{\cos^2\theta\mathrm{d}\theta + \sin^2\theta\mathrm{d}\theta}{1} = \int_0^{2\pi} \mathrm{d}\theta = 2\pi \neq 0.$$

Hence α is not exact.

The map $\varphi : \mathbb{R}^2 - \{\mathbf{0}\} \to S^1$ defined by

$$\varphi : (r\cos\theta, r\sin\theta) \mapsto (\cos\theta, \sin\theta)$$

induces a homomorphism

$$\varphi^* : \Lambda^1(S^1) \to \Lambda^1(\mathbb{R}^2 - \{\mathbf{0}\})$$

which has the properties $\varphi^*(\omega \wedge \omega') = \varphi^*(\omega) \wedge \varphi^*(\omega')$ and $\varphi^*\mathrm{d}\omega = \mathrm{d}\varphi^*\omega$. It therefore induces a map

$$\varphi^* : Z^1(S^1) \to Z^1(\mathbb{R}^2 - \{\mathbf{0}\}) \qquad \text{and} \qquad \varphi^* : B^1(S^1) \to B^1(\mathbb{R}^2 - \{\mathbf{0}\}).$$

The map is non-trivial for $\alpha = \varphi^*(\mathrm{d}\theta)$ as $\theta = \arctan(y/x)$ so that

$$\mathrm{d}\theta = \frac{x\mathrm{d}y - y\mathrm{d}x}{x^2}\frac{1}{\sec^2\theta} = \frac{x\mathrm{d}y - y\mathrm{d}x}{x^2}\frac{x^2}{x^2+y^2} = \alpha.$$

Since $H^1(S^1)$ is one-dimensional we thus have

$$H^1(\mathbb{R}^2 - \{\mathbf{0}\}) = H^1(S^1) = \mathbb{R}$$

and by Problem 17.12, $b^1 = 1$.

Problem 17.14 **Prove that every closed 1-form on S^2 is exact. Show that this statement does not extend to 2-forms by showing that the 2-form**

$$\alpha = r^{-3}(x\,\mathrm{d}y\wedge\mathrm{d}z + y\,\mathrm{d}z\wedge\mathrm{d}x + z\,\mathrm{d}x\wedge\mathrm{d}y)$$

is closed, but has non-vanishing integral on S^2.
[NOTE: Coefficient r^{-3} in α incorrectly given in text.]**

Solution: Let $N = (0,0,1)$ and $S = (0,0,-1)$ be the north and south pole on the unit sphere S^2. By stereographic projections St_N and St_S (see Example 15.4) we have $S^2 - \{N\} \cong S^2 - \{S\} \cong \mathbb{R}^2$. If ω is a closed form on S^2, $\mathrm{d}\omega = 0$, then by the Poincaré lemma there exist functions $f : S^2 - \{N\} \to \mathbb{R}$ and $g : S^2 - \{S\} \to \mathbb{R}$ such that $\omega = \mathrm{d}f$ and $\omega = \mathrm{d}g$ on the domains of definition of these functions. On $U = S^2 - \{N, S\}$ we have $\mathrm{d}f\mathrm{d}g$ and since this is a connected open set we must have $f = g+\text{const.}$ on U. Since either function may be modified by an arbitrary constant we may pick f and g to be equal at some point $p \in U$. It then follows that $f = g$ on U. Let $h : S^2 \to \mathbb{R}$ be deined by $h = f$ on $S^2 - \{N\}$ and $h(S) = g(S)$. Then since h is identical with g on $S^2 - \{S\}$ it is differentiable at the south pole S and therefeore $f \in \mathcal{F}(S^2)$. It is clear that $\omega = \mathrm{d}h$ on all of S^2, and ω is an exact form.

Since $r^2 = x^2 + y^2 + z^2$, we have

$$2r\mathrm{d}r = 2x\,\mathrm{d}x + 2y\,\mathrm{d}y + 2z\,\mathrm{d}z$$

and

$$\begin{aligned}
\mathrm{d}\alpha &= -3r^{-4}r^{-1}(x\,\mathrm{d}x + y\,\mathrm{d}y + z\,\mathrm{d}z)\wedge(x\,\mathrm{d}y\wedge\mathrm{d}z + y\,\mathrm{d}z\wedge\mathrm{d}x + z\,\mathrm{d}x\wedge\mathrm{d}y)\\
&\quad + r^{-3}(\mathrm{d}x\wedge\mathrm{d}y\wedge\mathrm{d}z + \mathrm{d}y\wedge\mathrm{d}z\wedge\mathrm{d}x + \mathrm{d}z\wedge\mathrm{d}x\wedge\mathrm{d}y)\\
&= -3r^{-5}(x^2+y^2+z^2)\mathrm{d}x\wedge\mathrm{d}y\wedge\mathrm{d}z + 3r^{-3}\mathrm{d}x\wedge\mathrm{d}y\wedge\mathrm{d}z\\
&= 0.
\end{aligned}$$

Thus α is a closed 2-form. Evaluating its integral over S^2 in spherical polar coordinates (see Problem 17.1)

$$\begin{aligned}
&\int_{S^2}\alpha = \int_{S^2} r^{-3}(x\,\mathrm{d}y\wedge\mathrm{d}z + y\,\mathrm{d}z\wedge\mathrm{d}x + z\,\mathrm{d}x\wedge\mathrm{d}y)\\
&= \int_{r=1} \sin\theta\cos\phi\sin\theta\cos\phi\mathrm{d}\phi\wedge(-\sin\theta\mathrm{d}\theta) + \sin\theta\sin\phi(-\sin\theta\mathrm{d}\theta)\wedge(-\sin\theta\sin\phi\mathrm{d}\phi)\\
&\qquad + \cos\theta(\cos\theta\cos\phi\mathrm{d}\theta - \sin\theta\sin\phi\mathrm{d}\phi)\wedge(\cos\theta\sin\phi\mathrm{d}\theta + \sin\theta\cos\phi\mathrm{d}\phi)\\
&= \int_{r=1}\mathrm{d}\theta\wedge\mathrm{d}\phi(\sin^3\theta\cos^2\phi + \sin^3\theta\sin^2\phi + \cos^2\theta\sin\theta\cos^2\phi + \cos^2\theta\sin\theta\sin^2\phi)\\
&= \int_0^\pi \mathrm{d}\theta\sin\theta\int_0^{2\pi}\mathrm{d}\phi\\
&= 4\pi \neq 0.
\end{aligned}$$

Hence α is not exact, for if $\alpha = \mathrm{d}\omega$ then by Stokes' theorem

$$\int_{S^2}\alpha = \int_{S^2}\mathrm{d}\omega = \int_{\partial S^2}\omega = 0$$

since the compact manifold S^2 has no boundary.

Problem 17.15 Show that the Maxwell 2-form satisfies the identities

$$\varphi \wedge *\varphi = *\varphi \wedge \varphi = 4(\mathbf{B}^2 - \mathbf{E}^2)\Omega$$
$$\varphi \wedge \varphi = - * \varphi \wedge *\varphi = 8\mathbf{B} \cdot \mathbf{E}\Omega$$

where $\Omega = \mathrm{d}x^1 \wedge \mathrm{d}x^2 \wedge \mathrm{d}x^3 \wedge \mathrm{d}x^4$**.**

Solution: The Maxwell 2-form and its dual are given by (see text)

$$\varphi = 2\big(B_3 dx^1 \wedge dx^2 + B_2 dx^3 \wedge dx^1 + B_1 dx^2 \wedge dx^3 + E_1 dx^1 \wedge dx^4 + E_2 dx^2 \wedge dx^4 + E_3\big)$$

and

$$*\varphi = 2\big(-E_3 dx^1 \wedge dx^2 - E_2 dx^3 \wedge dx^1 - E_1 dx^2 \wedge dx^3 + B_1 dx^1 \wedge dx^4 + B_2 dx^2 \wedge dx^4 + B_3 dx^3 \wedge dx^4\big),$$

where $\mathbf{E} = (E_1, E_2, E_3)$ is the electric field, $\mathbf{B} = (B_1, B_2, B_3)$ the magnetic field. In calculating $\varphi \wedge *\varphi$ the term $dx^1 \wedge dx^2$ only couples in a non-trivial way with $dx^3 \wedge dx^4$, $dx^3 \wedge dx^1$ with $dx^2 \wedge dx^4$ etc. The result is

$$\begin{aligned}\varphi \wedge *\varphi &= 4\big((B_3)^2 + (B_2)^2 + (B_1)^2 - (E_1)^2 - (E_2)^2 - (E_3)^2\big)\mathrm{d}x^1 \wedge \mathrm{d}x^2 \wedge \mathrm{d}x^3 \wedge \mathrm{d}x^4 \\ &= 4(\mathbf{B}^2 - \mathbf{E}^2)\Omega.\end{aligned}$$

Similarly,

$$\begin{aligned}*\varphi \wedge \varphi &= 4\big(-(E_3)^2 - (E_2)^2 - (E_1)^2 + (B_1)^2 + (B_2)^2 + (B_3)^2\big)\mathrm{d}x^1 \wedge \mathrm{d}x^2 \wedge \mathrm{d}x^3 \wedge \mathrm{d}x^4 \\ &= 4(\mathbf{B}^2 - \mathbf{E}^2)\Omega.\end{aligned}$$

In the same way

$$\begin{aligned}\varphi \wedge \varphi &= 4\big(B_3E_3 + B_2E_2 + B_1E_1 + E_1B_1 + E_2B_2 + E_3B_3\big)\mathrm{d}x^1 \wedge \mathrm{d}x^2 \wedge \mathrm{d}x^3 \wedge \mathrm{d}x^4 \\ &= 8\mathbf{B} \cdot \mathbf{E}\Omega.\end{aligned}$$

and since $*\varphi$ can be formed form φ by replacing $\mathbf{B}$ by $-\mathbf{E}$ and $\mathbf{E}$ by $\mathbf{B}$, we have

$$*\varphi \wedge *\varphi = 8(-\mathbf{E}) \cdot \mathbf{B}\Omega = -8\mathbf{B} \cdot \mathbf{E}\Omega.$$

Chapter 18

Problem 18.1 **Show directly from the transformation laws (18.1) and (18.11) that the components of the covariant derivative (18.6) of a vector field transform as a tensor of type** $(1,1)$**.**

Solution: The components of covariant derivative (18.6) are

$$Y^i{}_{;k} = Y^i{}_{,k} + \Gamma^i_{jk} Y^j,$$

which in transformed coordinates read

$$Y^{i'}{}_{;k'} = Y^{i'}{}_{,k'} + \Gamma^{i'}_{j'k'} Y^{j'}.$$

Using the vector transformation law we find Eq. (18.1) for the first term on the RHS,

$$Y'^{i'}_{,k'} = \frac{\partial x'^{i'}}{\partial x^i}\frac{\partial x^k}{\partial x'^{k'}} Y^i_{,k} + Y^i \frac{\partial x^k}{\partial x'^{k'}}\frac{\partial^2 x'^{i'}}{\partial x^i \partial x^k}, \tag{1}$$

and from the transformation law of Γ^i_{jk} given in Eq. (18.11) the second term is

$$\begin{aligned}
\Gamma^{i'}_{j'k'} Y^{j'} &= \Big(\frac{\partial x^k}{\partial x'^{k'}}\frac{\partial x^j}{\partial x'^{j'}}\frac{\partial x'^{i'}}{\partial x^i}\Gamma^i_{jk} + \frac{\partial^2 x^i}{\partial x'^{k'}\partial x'^{j'}}\frac{\partial x'^{i'}}{\partial x^i}\Big)\frac{\partial x'^{j'}}{\partial x^l} Y^l \\
&= \frac{\partial x^k}{\partial x'^{k'}}\delta^j{}_l \frac{\partial x'^{i'}}{\partial x^i}\Gamma^i_{jk} Y^l + \frac{\partial}{\partial x'^{j'}}\Big(\frac{\partial x^i}{\partial x'^{k'}}\frac{\partial x'^{i'}}{\partial x^i}\Big)\frac{\partial x'^{j'}}{\partial x^l} Y^l - \frac{\partial x^i}{\partial x'^{k'}}\frac{\partial}{\partial x'^{j'}}\Big(\frac{\partial x'^{i'}}{\partial x^i}\Big)\frac{\partial x'^{j'}}{\partial x^l} Y^l \\
&= \frac{\partial x^k}{\partial x'^{k'}}\frac{\partial x'^{i'}}{\partial x^i}\Gamma^i_{jk} Y^j + \frac{\partial}{\partial x^l}(\delta^{i'}_{k'}) Y^l - \frac{\partial x^i}{\partial x'^{k'}}\frac{\partial x'^{j'}}{\partial x^l}\frac{\partial}{\partial x'^{j'}}\Big(\frac{\partial x'^{i'}}{\partial x^i}\Big) Y^l \\
&= \frac{\partial x^k}{\partial x'^{k'}}\frac{\partial x'^{i'}}{\partial x^i}\Gamma^i_{jk} Y^j - \frac{\partial x^i}{\partial x'^{k'}}\frac{\partial^2 x'^{i'}}{\partial x^l \partial x^i} Y^l.
\end{aligned}$$

After a relabelling of dummy summation indices the second term on the right hand side cancels the second term on the RHS of (1), and we have

$$Y^{i'}{}_{;k'} = \frac{\partial x'^{i'}}{\partial x^i}\frac{\partial x^k}{\partial x'^{k'}}\big(Y^i_{,k} + \Gamma^i_{jk} Y^j\big) = \frac{\partial x'^{i'}}{\partial x^i}\frac{\partial x^k}{\partial x'^{k'}} Y^i_{;k},$$

as required.

Problem 18.2 **Show that the transformation law (18.11) can be written in the form**

$$\Gamma'^{i'}_{j'k'} = \frac{\partial x^k}{\partial x'^{k'}}\frac{\partial x^j}{\partial x'^{j'}}\frac{\partial x'^{i'}}{\partial x^i}\Gamma^i_{jk} - \frac{\partial^2 x'^{i'}}{\partial x^j \partial x^k}\frac{\partial x^j}{\partial x'^{j'}}\frac{\partial x^k}{\partial x'^{k'}}.$$

[NOTE: the second term on the RHS is incorrect in the text.]**

Solution: The argument is essentially the same as in the previous problem. In Eq.(18.11),

$$\Gamma'^{i'}_{j'k'} = \frac{\partial x^k}{\partial x'^{k'}}\frac{\partial x^j}{\partial x'^{j'}}\frac{\partial x'^{i'}}{\partial x^i}\Gamma^i_{jk} + \frac{\partial^2 x^i}{\partial x'^{k'}\partial x'^{j'}}\frac{\partial x'^{i'}}{\partial x^i},$$

the second term may be rewritten

$$\begin{aligned}\frac{\partial^2 x^i}{\partial x'^{k'}\partial x'^{j'}}\frac{\partial x'^{i'}}{\partial x^i} &= \frac{\partial}{\partial x'^{k'}}\Big(\frac{\partial x^i}{\partial x'^{j'}}\frac{\partial x'^{i'}}{\partial x^i}\Big) - \frac{\partial x^i}{\partial x'^{j'}}\frac{\partial}{\partial x'^{k'}}\Big(\frac{\partial x'^{i'}}{\partial x^i}\Big)\\ &= \frac{\partial}{\partial x'^{k'}}(\delta^{i'}_{j'}) - \frac{\partial x^j}{\partial x'^{j'}}\frac{\partial x^k}{\partial x'^{k'}}\frac{\partial}{\partial x^k}\Big(\frac{\partial x'^{i'}}{\partial x^j}\Big)\\ &= -\frac{\partial^2 x'^{i'}}{\partial x^j \partial x^k}\frac{\partial x^j}{\partial x'^{j'}}\frac{\partial x^k}{\partial x'^{k'}}.\end{aligned}$$

Problem 18.3 Show directly from (Cov1)–(Cov4) that $D_{fX}T = fD_XT$ for all vector fields X, tensor fields T and scalar functions $f : M \to \mathbb{R}$.

Solution: For T a scalar field, $T = g$ we have

$$D_{fX}g = (fX)g = f(Xg) = fD_Xg.$$

For T a vector field, $T = Y$ the desired condition is $D_{fX}Y = fD_XY$ is identical with the requirement(Con2) for a connection. If T is a covector field, $T = \omega$ then by (Cov2),

$$\begin{aligned}\langle D_{fX}\omega, Y\rangle &= D_{fX}\langle\omega, Y\rangle - \langle\omega, D_{fX}Y\rangle\\ &= fD_X\langle\omega, Y\rangle - \langle\omega, fD_XY\rangle\\ &= fD_X\langle\omega, Y\rangle - f\langle\omega, D_XY\rangle\\ &= f\langle D_X\omega, Y\rangle \qquad \text{by (Cov2)}\\ &= \langle fD_X\omega, Y\rangle.\end{aligned}$$

Since Y is an arbitrary vector field we have $D_{fX}\omega = fD_X\omega$.
Let T be a tensor of type (r, s) with total order $m = r+s > 1$ of the form $T = A\otimes S$, where A is a tensor of either type $(1, 0)$ or type $(0, 1)$ and the total order of S is $m - 1$. Assume the statement is true of all tensor fields having total order less than m, then by (Cov4)

$$\begin{aligned}D_{fX}(A\otimes S) &= D_{fX}A\otimes S + A\otimes D_{fX}S\\ &= fD_XA\otimes S + fA\otimes D_XS\\ &= fD_X(A\otimes S).\end{aligned}$$

Since every tensor T of with total order $m = r + s > 1$ can be written locally (in any coordinate neighbourhood) as a sum of tensors of this form,

$$T = \sum_i A_i\otimes S_i,$$

it follows from (Cov3) and the above that

$$\begin{aligned}
D_{fX}T &= \sum_i D_{fX}(A_i \otimes S_i) \\
&= \sum_i f D_X(A_i \otimes S_i) \\
&= f \sum_i D_X(A_i \otimes S_i) \\
&= f D_X T.
\end{aligned}$$

The result now follows for all tensors by induction on the total order.

Problem 18.4 Verify from the coordinate transformation rule (18.11) for Γ^i_{jk} , that the components of the covariant derivative of an arbitrary tensor field, defined in Eq. (18.14), transform as components of a tensor field.

Solution: Writing Eq. (18.14) so as to only include one index of each type

$$T^{ij\cdots}{}_{kl\ldots;p} = T^{ij\cdots}{}_{kl\ldots,p} + \Gamma^i_{ap} T^{aj\cdots}{}_{kl\ldots} + \cdots - \Gamma^a_{kp} T^{ij\cdots}{}_{al\ldots}$$

we have

$$\begin{aligned}
T^{i'j'\cdots}{}_{k'l'\ldots;p'} &= T^{i'j'\cdots}{}_{k'l'\ldots,p'} + \Gamma'^{i'}_{a'p'} T'^{a'j'\cdots}{}_{k'l'\ldots} + \cdots - \Gamma'^{a'}_{k'p'} T'^{i'j'\cdots}{}_{a'l'\ldots} - \ldots \\
&= \Big(T^{ij\cdots}{}_{kl\ldots} \frac{\partial x'^{i'}}{\partial x^i} \frac{\partial x'^{j'}}{\partial x^j} \cdots \frac{\partial x^k}{\partial x'^{k'}} \frac{\partial x^l}{\partial x'^{l'}} \Big)_{,p} \frac{\partial x^p}{\partial x'^{p'}} \\
&\qquad + \Gamma'^{i'}_{a'p'} T^{aj\cdots}{}_{kl\ldots} \frac{\partial x'^{a'}}{\partial x^a} \frac{\partial x'^{j'}}{\partial x^j} \cdots \frac{\partial x^k}{\partial x'^{k'}} \frac{\partial x^l}{\partial x'^{l'}} + \ldots \\
&\qquad - \Gamma'^{a'}_{k'p'} T^{ij\cdots}{}_{al\ldots} \frac{\partial x'^{i'}}{\partial x^i} \frac{\partial x'^{j'}}{\partial x^j} \cdots \frac{\partial x^a}{\partial x'^{a'}} \frac{\partial x^l}{\partial x'^{l'}} - \ldots
\end{aligned}$$

Expand the the first term on the RHS by the Leibnitz product rule of derivatives, substitute the transformation law derived in problem 18.2 for the Γ terms associated with contravariant indices,

$$\Gamma'^{i'}_{a'p'} = \frac{\partial x^a}{\partial x'^{a'}} \frac{\partial x^p}{\partial x'^{p'}} \frac{\partial x'^{i'}}{\partial x^i} \Gamma^i_{ap} - \frac{\partial^2 x'^{i'}}{\partial x^a \partial x^p} \frac{\partial x^a}{\partial x'^{a'}} \frac{\partial x^p}{\partial x'^{p'}}$$

etc., and the transformation Eq. (18.11) for the Γ terms associated with covariant indices,

$$\Gamma'^{a'}_{k'p'} = \frac{\partial x^p}{\partial x'^{p'}} \frac{\partial x^k}{\partial x'^{k'}} \frac{\partial x'^{a'}}{\partial x^a} \Gamma^a_{kp} + \frac{\partial^2 x^a}{\partial x'^{p'} \partial x'^{k'}} \frac{\partial x'^{a'}}{\partial x^a}.$$

All second order derivative then cancel and we arrive at

$$T^{i'j'\cdots}{}_{k'l'\ldots;p'} = \Big(T^{ij\cdots}{}_{kl\ldots,p} + \Gamma^i_{ap}T^{aj\cdots}{}_{kl\ldots} + \cdots - \Gamma^a_{kp}T^{ij\cdots}{}_{al\ldots}\Big)\cdot$$
$$\frac{\partial x'^{i'}}{\partial x^i}\frac{\partial x'^{j'}}{\partial x^j}\cdots\frac{\partial x^k}{\partial x'^{k'}}\frac{\partial x^l}{\partial x'^{l'}}\frac{\partial x^p}{\partial x'^{p'}}$$
$$= T^{ij\cdots}{}_{kl\ldots;p}\frac{\partial x'^{i'}}{\partial x^i}\frac{\partial x'^{j'}}{\partial x^j}\cdots\frac{\partial x^k}{\partial x'^{k'}}\frac{\partial x^l}{\partial x'^{l'}}\frac{\partial x^p}{\partial x'^{p'}}.$$

Problem 18.5 Show that the identity (18.18) follows from Eq. (18.16).

Solution: The definition of covariant derivative DT of a tensor field T is the tensor field of type $(r, s+1)$ defined by

$$DT(\omega^1, \omega^2, \ldots, \omega^r, Y_1, Y_2, \ldots, Y_s, X) = D_X T(\omega^1, \omega^2, \ldots, \omega^r, Y_1, Y_2, \ldots, Y_s).$$

If S is any tensor field of type (r, s) we may write $S(\omega^1, \ldots, \omega^r, Y_1, \ldots, Y_s)$ as a multiple contraction of a tensor of type $(r+s, r+s)$:

$$S(\omega^1, \ldots, \omega^r, Y_1, \ldots, Y_s) = C\, T \otimes \omega^1 \otimes \cdots \otimes \omega^r \otimes Y_1 \otimes \cdots \otimes Y_s$$

where

$$C \equiv C^1_{s+1} \ldots C^r_{s+r} C^{r+1}_1 \ldots C^{r+s}_s.$$

Assuming the result in Eq. (18.16) that covariant derivative commutes with contractions,

$$D\big(C^k_l S\big) = C^k_l DS,$$

we have, using (Cov4), that

$$\begin{aligned}
&X(T(\omega^1, \ldots, \omega^r, Y_1, \ldots, Y_s))\\
&\quad = D_X(C\, T \otimes \omega^1 \otimes \cdots \otimes \omega^r \otimes Y_1 \otimes \cdots \otimes Y_s)\\
&\quad = C\, D_X(T \otimes \omega^1 \otimes \cdots \otimes \omega^r \otimes Y_1 \otimes \cdots \otimes Y_s)\\
&\quad = C\,\big[(D_X T) \otimes \omega^1 \otimes \cdots \otimes \omega^r \otimes Y_1 \otimes \cdots \otimes Y_s\\
&\qquad\quad + T \otimes (D_X \omega^1) \otimes \cdots \otimes \omega^r \otimes Y_1 \otimes \cdots \otimes Y_s\\
&\qquad\quad + \ldots\\
&\qquad\quad + T \otimes \omega^1 \otimes \cdots \otimes \omega^r \otimes Y_1 \otimes \cdots \otimes (D_X Y_s)\big]\\
&\quad = D_X T(\omega^1, \ldots, \omega^r, Y_1, \ldots, Y_s)\\
&\quad + T(D_X\omega^1, \ldots, \omega^r, Y_1, \ldots, Y_s) + \cdots + T(\omega^1, \ldots, D_X\omega^r, Y_1, \ldots, Y_s)\\
&\quad + T(\omega^1, \ldots, \omega^r, D_X Y_1, \ldots, Y_s) + \cdots + T(\omega^1, \ldots, \omega^r, Y_1, \ldots, D_X Y_s).
\end{aligned}$$

This proves Eq. (18.18).

Problem 18.6 **Let f be a smooth function, $X = X^i\partial_{x^i}$ a smooth vector field and $\omega = w_i \mathrm{d}x^i$ a differential 1-form. Show that**

$$(D_jD_i - D_iD_j)f = 0,$$
$$(D_jD_i - D_iD_j)X = X^a R^k{}_{aji}\partial_{x^k},$$
$$(D_jD_i - D_iD_j)\omega = w_a R^a{}_{kij}\mathrm{d}x^k.$$

Why does the torsion tensor not appear in these formulae, in contrast with the Ricci identities (18.27)–(18.29)?
[**NOTE: interchange i and j in RHS term of last (**ω equation given in text.]

Solution: If T is a tensor field of type (r, s) then in a local coordinate chart

$$D_pT \equiv D_{\partial_{x^p}}\big(T^{ij\cdots}{}_{kl\ldots}\partial_{x^i}\otimes\partial_{x^j}\otimes\cdots\otimes \mathrm{d}x^k\otimes\mathrm{d}x^l\otimes\ldots\big)$$
$$= T^{ij\cdots}{}_{kl\ldots;p}\partial_{x^i}\otimes\partial_{x^j}\otimes\cdots\otimes \mathrm{d}x^k\otimes\mathrm{d}x^l\otimes\ldots$$

Note that the RHS is a tensor of type (r, s) for each fixed value $p = 1, \ldots n$. Thus for any scalar field f (type $(0, 0)$)

$$D_jD_if = D_j(f_{,i}) = (f_{,i})_{,j} = f_{ij}$$

and

$$(D_jD_i - D_iD_j)f = f_{,ij} - f_{,ji} = 0.$$

There is no torsion tensor in this result since components $f_{,i}$ are treated as scalars here, not as components of a covector as in Eq. (18.27).

If $X = X^k\partial_{x^k}$ is a vector field then

$$\begin{aligned}(D_jD_i - D_iD_j)X &= D_j\big(X^k{}_{;i}\partial_{x^k}\big) - D_i\big(X^k{}_{;j}\partial_{x^k}\big)\\
&= D_j\big((X^k{}_{,i} + \Gamma^k_{ai}X^a)\partial_{x^k}\big) - D_i\big((X^k{}_{,j} + \Gamma^k_{aj}X^a)\partial_{x^k}\big)\\
&= \big((X^k{}_{,i} + \Gamma^k_{ai}X^a))_{,j} + \Gamma^k_{bj}(X^b{}_{,i} + \Gamma^b_{ai}X^a)\big)\partial_{x^k}\\
&\quad - \big((X^k{}_{,j} + \Gamma^k_{aj}X^a))_{,i} + \Gamma^k_{bi}(X^b{}_{,j} + \Gamma^b_{aj}X^a)\big)\partial_{x^k}\\
&= \big(X^k{}_{,ij} + \Gamma^k_{ai,j}X^a + \Gamma^k_{ai}X^a{}_{,j} + \Gamma^k_{bj}X^b{}_{,i} + \Gamma^k_{bj}\Gamma^b_{ai}X^a\\
&\quad - X^k{}_{,ji} - \Gamma^k_{aj,i}X^a - \Gamma^k_{aj}X^a{}_{,i} - \Gamma^k_{bi}X^b{}_{,j} - \Gamma^k_{bi}\Gamma^b_{aj}X^a\big)\partial_{x^k}\\
&= \big(\Gamma^k_{ai,j} - \Gamma^k_{aj,i} + \Gamma^k_{bj}\Gamma^b_{ai} - \Gamma^k_{bi}\Gamma^b_{aj}\big)X^a\partial_{x^k}\\
&= X^aR^k{}_{aji}\partial_{x^k}.\end{aligned}$$

In Eq. (18.28) the components $X^k{}_{;i} = X^k{}_{,i} + \Gamma^k_{ai}X^a$ are treated as components of a tensor field of type $(1, 1)$, whereas in the above they are components of a vector field. It is this difference which leads to the lack of torsion tensor in the result.

They case of a covector field $\omega = w_k \mathrm{d}x^k$ is copmpletely analogous:

$$\begin{aligned}
(D_j D_i - D_i D_j)\omega &= D_j\big(w_{k;i}\mathrm{d}x^k\big) - D_i\big(w_{k;j}\mathrm{d}x^k\big) \\
&= D_j\big((w_{k,i} - \Gamma^a_{ki}w_a)\mathrm{d}x^k\big) - D_i\big((w_{k,j} - \Gamma^a_{kj}w_a)\mathrm{d}x^k\big) \\
&= \big((w_{k,i} - \Gamma^a_{ki}w_a))_{,j} - \Gamma^b_{kj}(w_{b,i} - \Gamma^a_{bi}w_a)\big)\mathrm{d}x^k \\
&\quad - \big((w_{k,j} - \Gamma^a_{kj}w_a))_{,i} + \Gamma^b_{ki}(w_{b,j} - \Gamma^a_{bj}w_a)\big)\mathrm{d}x^k \\
&= \big(X^k{}_{,ij} + \Gamma^k_{ai,j}X^a + \Gamma^k_{ai}X^a{}_{,j} + \Gamma^k_{bj}X^b{}_{,i} + \Gamma^k_{bj}\Gamma^b_{ai}X^a \\
&\quad - X^k{}_{,ji} - \Gamma^k_{aj,i}X^a - \Gamma^k_{aj}X^a{}_{,i} - \Gamma^k_{bi}X^b{}_{,j} - \Gamma^k_{bi}\Gamma^b_{aj}X^a\big)\partial_{x^k} \\
&= \big(\Gamma^a_{kj,i} - \Gamma^a_{ki,j} + \Gamma^a_{bi}\Gamma^b_{kj} - \Gamma^a_{bj}\Gamma^b_{ki}\big)w_a\partial_{x^k} \\
&= w_a R^a{}_{kij}\mathrm{d}x^k.
\end{aligned}$$

Comments regarding torsion are similar to those above.

Problem 18.7 Show that the coordinate expression for the Lie derivative of a vector field may be written

$$(\mathcal{L}_X Y)^i = [X, Y]^i = Y^i{}_{;j}X^j - X^i{}_{;j}Y^j + T^i_{jk}X^kY^j. \tag{18.30}$$

For a torsion-free connection show that the Lie derivative (15.39) of a general tensor field S of type (r, s) may be expressed by

$$\begin{aligned}
\big(\mathcal{L}_X S\big)^{ij\dots}{}_{kl\dots} &= S^{ij\dots}{}_{kl\dots;m}X^m - S^{mj\dots}{}_{kl\dots}X^i{}_{;m} - S^{im\dots}{}_{kl\dots}X^j{}_{;m} - \dots \\
&\quad + S^{ij\dots}{}_{ml\dots}X^m{}_{;k} + S^{ij\dots}{}_{km\dots}X^m{}_{;l} + \dots\,.
\end{aligned} \tag{18.31}$$

Write down the full version of this equation for a general connection with torsion.

Solution: From Eqs. (15.25) and (15.37) and Eq. (18.6) we have

$$\begin{aligned}
(\mathcal{L}_X Y)^i = [X, Y]^i &= Y^i{}_{,j}X^j - X^i{}_{,j}Y^j + T^i_{jk}X^kY^j \\
&= \big(Y^i{}_{;j} - \Gamma^i_{kj}Y^k\big)X^j - \big(X^i{}_{;j} - \Gamma^i_{kj}X^k\big)Y^j \\
&= Y^i{}_{;j}X^j - X^i{}_{;j}Y^j + T^i_{jk}Y^jX^k
\end{aligned}$$

where (see Eq. (18.21))

$$T^i_{jk} = \Gamma^i_{kj} - \Gamma^i_{jk}.$$

From Eq. (15.39)

$$\begin{aligned}
\big(\mathcal{L}_X S\big)^{ij\dots}{}_{kl\dots} &= S^{ij\dots}{}_{kl\dots,m}X^m - S^{mj\dots}{}_{kl\dots,m}X^i{}_{,m} - S^{im\dots}{}_{kl\dots,m}X^j{}_{,m} - \dots \\
&\quad + S^{ij\dots}{}_{ml\dots}X^m{}_{,k} + S^{ij\dots}{}_{km\dots}X^m{}_{,l} + \dots\,.
\end{aligned}$$

Substituting Eq. (18.14) for the tensor field S,

$$\begin{aligned}
S^{ij\dots}{}_{kl\dots,m} &= T^{ij\dots}{}_{kl\dots;m} - \Gamma^i_{am}S^{aj\dots}{}_{kl\dots} - \Gamma^j_{am}T^{ia\dots}{}_{kl\dots} - \dots \\
&\quad + \Gamma^a_{km}T^{ij\dots}{}_{al\dots} + \Gamma^a_{lm}T^{ij\dots}{}_{ka\dots} + \dots
\end{aligned}$$

and

$$X^i{}_{,m} = X^i{}_{;m} - \Gamma^i_{am} X^a \quad \text{etc.}$$

gives

$$\begin{aligned}
\big(\mathcal{L}_X S\big)^{ij\dots}{}_{kl\dots} &= S^{ij\dots}{}_{kl\dots;m} X^m - S^{mj\dots}{}_{kl\dots} X^i{}_{;m} - S^{aj\dots}{}_{kl\dots} \Gamma^i_{am} X^m + S^{mj\dots}{}_{kl\dots,m} \Gamma^i_{am} X^a + \dots \\
&\quad + S^{ij\dots}{}_{ml\dots} X^m{}_{;k} + S^{ij\dots}{}_{al\dots} \Gamma^a_{km} X^m - S^{ij\dots}{}_{ml\dots} \Gamma^m_{ak} X^a + \dots \\
&= S^{ij\dots}{}_{kl\dots;m} X^m - S^{mj\dots}{}_{kl\dots} X^i{}_{;m} - S^{im\dots}{}_{kl\dots} X^j{}_{;m} - \dots \\
&\quad + S^{ij\dots}{}_{ml\dots} X^m{}_{;k} + S^{ij\dots}{}_{km\dots} X^m{}_{;l} + \dots \\
&\quad + S^{aj\dots}{}_{kl\dots} T^i_{am} X^m + \dots + S^{ij\dots}{}_{ml\dots} T^m_{ak} X^a + \dots
\end{aligned}$$

Eq. (18.31) follows for a torsion-free connection, $T^i_{jk} = 0$.

Problem 18.8 Prove the Ricci identities (18.28) and (18.29).

Solution: Since $X^k{}_{;i}$ are components of a tensor field of type $(1,1)$,

$$\begin{aligned}
X^k{}_{;ij} - X^k{}_{;ji} &= (X^k{}_{;i})_{,j} + \Gamma^k_{aj} X^a{}_{;i} - \Gamma^a_{ij} X^k{}_{;a} \\
&\quad - (X^k{}_{;j})_{,i} - \Gamma^k_{ai} X^a{}_{;j} + \Gamma^a_{ji} X^k{}_{;a} \\
&= X^k{}_{,ij} + \Gamma^k_{ai,j} X^a + \Gamma^k_{ai} X^a{}_{,j} + \Gamma^k_{aj}\big(X^a{}_{,i} + \Gamma^a_{bi} X^b\big) - \Gamma^a_{ij} X^k{}_{;a} \\
&\quad - X^k{}_{,ji} - \Gamma^k_{aj,i} X^a - \Gamma^k_{aj} X^a{}_{,i} - \Gamma^k_{ai}\big(X^a{}_{,j} - \Gamma^a_{bj} X^b\big) + \Gamma^a_{ji} X^k{}_{;a} \\
&= \big(\Gamma^k_{ai,j} - \Gamma^k_{aj,i} + \Gamma^k_{mj}\Gamma^m_{ai} - \Gamma^k_{mi}\Gamma^m_{aj}\big) X^a + \big(\Gamma^a_{ji} - \Gamma^a_{ij}\big) X^k{}_{;a} \\
&= X^a R^k{}_{aji} + T^a_{ij} X^k{}_{;a}.
\end{aligned}$$

Similarly, for a 1-form $\omega = w_i dx^i$,

$$\begin{aligned}
w_{k;ij} - w_{k;ji} &= (w_{k;i})_{,j} - \Gamma^a_{kj} w_{a;i} - \Gamma^a_{ij} w_{k;a} \\
&\quad - (w_{k;j})_{,i} + \Gamma^a_{ki} w_{a;j} + \Gamma^a_{ji} w_{k;a} \\
&= w_{k,ij} - \Gamma^a_{ki,j} w_a - \Gamma^a_{ki} w_{a,j} - \Gamma^a_{kj}(w_{a,i} - \Gamma^b_{ai} w_b) - \Gamma^a_{ij} w_{k;a} \\
&\quad - w_{k,ij} + \Gamma^a_{kj,i} w_a + \Gamma^a_{kj} w_{a,i} + \Gamma^a_{ki}(w_{a,j} + \Gamma^b_{aj} w_b) + \Gamma^a_{ji} w_{k;a} \\
&= \big(\Gamma^a_{kj,i} - \Gamma^a_{ki,j} + \Gamma^a_{mi}\Gamma^m_{kj} - \Gamma^a_{mj}\Gamma^m_{ki}\big) w_a + \big(\Gamma^a_{ji} - \Gamma^a_{ij}\big) w_{k;a} \\
&= w_a R^a{}_{kij} + T^a_{ij} w_{k;a}.
\end{aligned}$$

Problem 18.9 For a torsion-free connection prove the generalized Ricci identities

$$\begin{aligned}
S^{kl\dots}{}_{mn\dots;ij} - S^{kl\dots}{}_{mn\dots;ji} &= S^{al\dots}{}_{mn\dots} R^k{}_{aji} + S^{ka\dots}{}_{mn\dots} R^l{}_{aji} + \dots \\
&\quad + S^{kl\dots}{}_{an\dots} R^a{}_{mij} + S^{kl\dots}{}_{ma\dots} R^a{}_{nij} + \dots
\end{aligned}$$

How is this equation modified in the case of torsion?

Solution: Consider firstly a *decomposable* tensor of type (r, s), i.e. one of the form

$$S = X \otimes Y \otimes \cdots \otimes \omega \otimes \rho \otimes \dots,$$

having components

$$S^{kl\dots}{}_{mn\dots} = X^k Y^l \dots w_m r_n \dots$$

Using the generalized Leibnitz rule (Cov4),

$$\begin{aligned} S^{kl\dots}{}_{mn\dots;ij} - S^{kl\dots}{}_{mn\dots;ji} &= (X^k Y^l \dots w_m r_n \dots)_{;ij} - (X^k Y^l \dots w_m r_n \dots)_{;ji} \\ &= (X^k{}_{;i} Y^l \dots w_m r_n \dots)_{;j} + (X^k Y^l{}_{;i} \dots w_m r_n \dots)_{;j} + \dots \\ &\quad + (X^k Y^l \dots w_{m;i} r_n \dots)_{;j} + (X^k Y^l \dots w_m r_{n;i} \dots)_{;j} + \dots \\ &\quad - (X^k{}_{;j} Y^l \dots w_m r_n \dots)_{;i} - (X^k Y^l{}_{;j} \dots w_m r_n \dots)_{;i} - \dots \\ &\quad - (X^k Y^l \dots w_{m;j} r_n \dots)_{;i} - (X^k Y^l \dots w_m r_{n;j} \dots)_{;i} - \dots \end{aligned}$$

Contributions from products of first covariant derivatives of components of X and Y cancel:

$$X^k{}_{;i} Y^l{}_{;j} \dots w_m r_n \dots + X^k{}_{;j} Y^l{}_{;i} \dots w_m r_n \dots - X^k{}_{;j} Y^l{}_{;i} \dots w_m r_n \dots - X^k{}_{;i} Y^l{}_{;j} \dots w_m r_n \dots = 0$$

leaving, by Eqs. (18.28) and (18.29),

$$\begin{aligned} &(X^k{}_{;ij} - X^k{}_{;ji}) Y^l \dots w_m r_n \dots + X^k (Y^l{}_{;ij} - Y^l{}_{;ji}) \dots w_m r_n \dots + \dots \\ &+ X^k Y^l \dots (w_{m:ij} - w_{m:ji} r_n \dots + X^k Y^l \dots w_m (r_{n:ij} - r_{n:ji} \dots + \dots \\ = {}& R^k{}_{aji} X^a Y^l \dots w_m r_n \dots + R^l{}_{aji} X^k Y^a \dots w_m r_n \dots + \dots + R^a{}_{mij} X^k Y^a \dots w_a r_n \dots + \dots \\ &+ T^a_{ij} X^k{}_{;a} Y^l \dots w_m r_n \dots + T^a_{ij} X^k Y^l{}_{;a} \dots w_m r_n + \dots + T^a_{ij} X^k Y^l \dots w_{m;a} r_n + \dots \\ = {}& R^k{}_{aji} X^a Y^l \dots w_m r_n \dots + \dots + R^a{}_{mij} X^k Y^a \dots w_a r_n \dots + \dots + T^a_{ij} (X^k Y^l \dots w_m r_n)_{;a} . \end{aligned}$$

Since every tensor field S of type (r, s) is (locally) a sum of decomposable tensor fields of this type, we have

$$\begin{aligned} S^{kl\dots}{}_{mn\dots;ij} - S^{kl\dots}{}_{mn\dots;ji} &= S^{al\dots}{}_{mn\dots} R^k{}_{aji} + S^{ka\dots}{}_{mn\dots} R^l{}_{aji} + \dots \\ &\quad + S^{kl\dots}{}_{an\dots} R^a{}_{mij} + S^{kl\dots}{}_{ma\dots} R^a{}_{nij} + \dots + T^a_{ij} S^{kl\dots}{}_{mn\dots;a}, \end{aligned}$$

which reduces to the desired expression if the connection is torsion-free, $T^a_{ij} = 0$.

Problem 18.10 **For arbitrary vector fields Y, Z and W show that the operator $\Sigma_{Y,Z,W} : \mathcal{T}(M) \to \mathcal{T}(M)$ defined by**

$$\Sigma_{Y,Z,W} X = D_W\big(\rho_{Y,Z} X\big) - \rho_{Z,[Y,W]} X - \rho_{Y,Z}\big(D_W X\big)$$

has the cyclic symmetry

$$\Sigma_{Y,Z,W} X + \Sigma_{Z,W,Y} X + \Sigma_{W,Y,Z} X = 0.$$

Express this equation in components with respect to a local coordinate chart and show that it is equivalent to the (second) Bianchi identity

$$R^i{}_{jkl;m} + R^i{}_{jlm;k} + R^i{}_{jmk;l} = R^i{}_{jpk}T^p_{ml} + R^i{}_{jpl}T^p_{km} + R^i{}_{jpm}T^p_{lk}. \tag{18.32}$$

Solution: Expanding each term in the cyclic sum we have, on substituting Eq. (18.22),

$$\begin{aligned}
\big(\Sigma_{Y,Z,W} + \Sigma_{Z,W,Y} + \Sigma_{W,Y,Z}\big)X &= \\
&= \big(D_W\rho_{Y,Z} + D_Y\rho_{Z,W} + D_Z\rho_{W,Y} \\
&\quad - \rho_{Y,Z}D_W - \rho_{Z,W}D_Y - \rho_{W,Y}D_Z \\
&\quad - \rho_{Z,[Y,W]} - \rho_{W,[Z,Y]} - \rho_{Y,[W,Z]}\big)\,X \\
&= \big(D_WD_YD_Z - D_WD_ZD_Y - D_WD_{[Y,Z]} \\
&\quad + D_YD_ZD_W - D_YD_WD_Z - D_YD_{[Z,W]} \\
&\quad + D_ZD_WD_Y - D_ZD_YD_W - D_ZD_{[W,Y]} \\
&\quad - D_YD_ZD_W + D_ZD_YD_W + D_{[Y,Z]}D_W \\
&\quad - D_ZD_WD_Y + D_WD_ZD_Y + D_{[Z,W]}D_Y \\
&\quad - D_WD_YD_Z + D_YD_WD_Z + D_{[W,Y]}D_Z \\
&\quad - D_ZD_{[Y,W]} + D_{[Y,W]}D_Z + D_{[Z,[Y,W]]} \\
&\quad - D_WD_{[Z,Y]} + D_{[Z,Y]}D_W + D_{[W,[Z,Y]]} \\
&\quad - D_YD_{[W,Z]} + D_{[W,Z]}D_Y + D_{[Y,[W,Z]]}\big)X \\
&= \big(D_{[Z,[Y,W]]} + D_{[W,[Z,Y]]} + D_{[Y,[W,Z]]}\big)X \\
&= D_{[Z,[Y,W]]+[W,[Z,Y]]+[Y,[W,Z]]}X \\
&= 0
\end{aligned}$$

on using the Jacobi identity. We have also used $[Y,Z] = -[Z,Y]$ etc. for cancellation of terms.

To express this identity in coordinates, set $\rho_{ij} \equiv \rho_{\partial_i,\partial_j}$ where $\partial_i \equiv \partial_{x^i}$. Then using Problem 18.6, and the fact that the commutators of coordinate vector fields vanish, $[\partial_i, \partial_j] = 0$ for all i, j,

$$\begin{aligned}
\rho_{ij}X &= \big(D_iD_j - D_jD_i - D_{[\partial_i,\partial_j]}\big)X \\
&= (D_iD_j - D_jD_i)X \\
&= X^pR^k{}_{pij}\partial_k.
\end{aligned}$$

In particular

$$\rho_{ij}\partial_k = R^b{}_{kij}\partial_b.$$

Set $X = \partial_j$, $Y = \partial_k$, $Z = \partial_l$ and $W = \partial_m$. Then

$$\begin{aligned}\Sigma_{Y,Z,W}X &= D_m\big(\rho_{kl}\partial_j\big) - \rho_{\partial_l,[\partial_k,\partial_m]}\partial_j - \rho_{kl}\big(D_m\partial_j\big)\\ &= D_m\big(R^i{}_{jkl}\partial_i\big) - \rho_{kl}\big(\Gamma^i_{jm}\partial_i\big)\\ &= R^i{}_{jkl,m}\partial_i + R^i{}_{jkl}\Gamma^p_{mi}\partial_p - \Gamma^p_{jm}R^i{}_{pkl}\partial_i\\ &= \big(R^i{}_{jkl,m} + R^p{}_{jkl}\Gamma^i_{mp} - \Gamma^p_{jm}R^i{}_{pkl}\big)\partial_i\\ &= \big(R^i{}_{jkl;m} + \Gamma^p_{km}R^i{}_{jpl} + \Gamma^p_{lm}R^i{}_{jkp}\big)\partial_i\end{aligned}$$

on using the appropriate expression for the covariant derivative $R^i{}_{jkl;m}$ from Eq. (18.14). Expanding the cyclical equation

$$(\Sigma_{Y,Z,W} + \Sigma_{Z,W,Y} + \Sigma_{W,Y,Z})X = 0$$

for our choice of coordinate basis vectors results in

$$\begin{aligned}0 = R^i{}_{jkl;m} &+ R^i{}_{jlm;k} + R^i{}_{jmk;l} + \Gamma^p_{km}R^i{}_{jpl} + \Gamma^p_{lm}R^i{}_{jkp}\\ &+ \Gamma^p_{lk}R^i{}_{jpm} + \Gamma^p_{mk}R^i{}_{jlp} + \Gamma^p_{ml}R^i{}_{jpk} + \Gamma^p_{kl}R^i{}_{jmp}\end{aligned}$$

Using the anti-symmetry on the last pair of indices of $R^i{}_{jkp} = -R^i{}_{jpk}$ we have

$$\begin{aligned}R^i{}_{jkl;m} + R^i{}_{jlm;k} + R^i{}_{jmk;l} &= (\Gamma^p_{lm} - \Gamma^p_{ml})R^i{}_{jpk} + (\Gamma^p_{mk} - \Gamma^p_{km})R^i{}_{jpl} + (\Gamma^p_{kl} - \Gamma^p_{lk})R^i{}_{jpm}\\ &= -R^i{}_{jpk}T^p_{ml} - R^i{}_{jpl}T^p_{km} - R^i{}_{jpm}T^p_{lk}.\end{aligned}$$

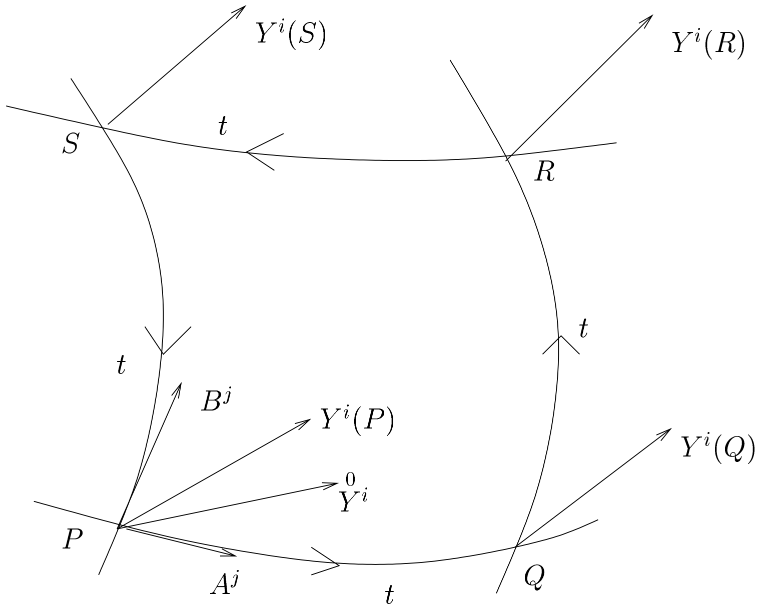
Problem 18.11 **Let $Y^i(t)$ be a vector that is parallel propagated along a curve having coordinate representation $x^j = \overset{0}{x}{}^j + A^j t$. Show that for $t \ll 1$**

$$Y^i(t) = \overset{0}{Y}{}^i - \overset{0}{\Gamma}{}^i_{ja}\overset{0}{Y}{}^j A^a t + \frac{t^2}{2}\big(\overset{0}{\Gamma}{}^i_{ka}\overset{0}{\Gamma}{}^k_{jb} - \overset{0}{\Gamma}{}^i_{ja,b}\big)A^a A^b \overset{0}{Y}{}^j + O(t^3)$$

where $\overset{0}{\Gamma}{}^i_{jk} = \Gamma^i_{jk}(\overset{0}{x}{}^a)$ and $\overset{0}{Y}{}^i = Y^i(0)$. From the point P, having coordinates $\overset{0}{x}{}^i$, parallel transport the tangent vector $\overset{0}{Y}{}^i$ around a coordinate rectangle $PQRSP$ whose sides are each of parameter length t and are along the a- and b-axes successively through these points. For example, the a-axis through P is the curve $x^j = \overset{0}{x}{}^j + \delta^j_a t$. Show that to order t^2, the final vector at P has components

$$Y^i = \overset{0}{Y}{}^i + t^2 \overset{0}{R}{}^i{}_{jba}\overset{0}{Y}{}^j$$

where $\overset{0}{R}{}^i{}_{jba}$ are the curvature tensor components at P.



[NOTE: In the term in parentheses in the first displayed equation in text one index a in the first term should be replaced by b and the superscript j in the second term replaced by i.]

Solution: Parallel propagating $Y(t)$ along a curve with tangent vector $x^j = \overset{0}{x}{}^j + A^j t$ ($A^j =$ const.), we have

$$\frac{\mathrm{d}Y^i}{\mathrm{d}t} + \Gamma^i_{ja} Y^j \frac{\mathrm{d}x^a}{\mathrm{d}t} = 0,$$

i.e.

$$\frac{\mathrm{d}Y^i}{\mathrm{d}t} = -\Gamma^i_{ja} Y^j A^a.$$

Differentiating again gives

$$\begin{aligned} \frac{\mathrm{d}^2 Y^i}{\mathrm{d}t^2} &= -\Gamma^i_{ja} \frac{\mathrm{d}Y^j}{\mathrm{d}t} A^a - \frac{\mathrm{d}}{\mathrm{d}t} \Gamma^i_{ja} Y^j A^a \\ &= \Gamma^i_{ja} \Gamma^j_{kb} Y^k A^b - \Gamma^i_{ja,b} A^b Y^j A^a. \end{aligned}$$

A Taylor expansion at $x^j = \overset{0}{x}{}^j$ gives

$$\begin{aligned} Y^i(t) &= \overset{0}{Y}{}^i + t\Big(\frac{\mathrm{d}Y^i}{\mathrm{d}t}\Big)_{x=\overset{0}{x}} + \frac{t^2}{2}\Big(\frac{\mathrm{d}^2 Y^i}{\mathrm{d}t^2}\Big)_{x=\overset{0}{x}} + O(t^3) \\ &= \overset{0}{Y}{}^i - t\overset{0}{\Gamma}{}^i_{ja} \overset{0}{Y}{}^j A^a + \frac{t^2}{2}\big(\overset{0}{\Gamma}{}^i_{ka} \overset{0}{\Gamma}{}^k_{jb} - \overset{0}{\Gamma}{}^i_{ja,b}\big) A^a A^b \overset{0}{Y}{}^j + O(t^3) \end{aligned}$$

On the a-axis connecting P to Q we have $A^j = \delta^j_a$ In the following we will treat a and b as constants and suspend summation convention over these indices (but maintain it on indices i, j, k). Parallel transporting Y from P to Q gives then, to order t^2

$$Y^i(Q) = \overset{0}{Y}{}^i - t\overset{0}{\Gamma}{}^i_{ja}\overset{0}{Y}{}^j + \tfrac{1}{2}t^2\big(\overset{0}{\Gamma}{}^i_{ka}\overset{0}{\Gamma}{}^k_{ja} - \overset{0}{\Gamma}{}^j_{ja,a}\big)\overset{0}{Y}{}^j + O(t^3).$$

Continuing by parallel transport from Q to R along the line $x^i(s) = x^i(Q) + s\delta^i_b$ to $s = t$ we have

$$\begin{aligned} Y^i(R) &= Y^i(Q) - t\Gamma^i_{jb}(Q)Y^i(Q) + \tfrac{1}{2}t^2\big(\Gamma^i_{kb}\Gamma^k_{jb} - \Gamma^i_{ja,b}\big)_Q Y_j(Q) + O(t^3) \\ &= \overset{0}{Y}{}^i - t\big(\overset{0}{\Gamma}{}^i_{ja}\overset{0}{Y}{}^j + \overset{0}{\Gamma}{}^i_{jb}\overset{0}{Y}{}^j\big) + \tfrac{1}{2}t^2\big(\overset{0}{\Gamma}{}^i_{ka}\overset{0}{\Gamma}{}^k_{ja} - \overset{0}{\Gamma}{}^i_{ja,a} \\ &\quad - 2\overset{0}{\Gamma}{}^i_{jb,a} + 2\overset{0}{\Gamma}{}^i_{kb}\overset{0}{\Gamma}{}^k_{ja} + \overset{0}{\Gamma}{}^i_{kb}\overset{0}{\Gamma}{}^k_{jb} - \overset{0}{\Gamma}{}^i_{jb,b}\big)\overset{0}{Y}{}^j + O(t^3) \end{aligned}$$

on substituting for $Y^i(Q)$ from the previous equation and $\Gamma^i_{jb}(Q) = \overset{0}{\Gamma}{}^i_{jb} + t\overset{0}{\Gamma}{}^i_{jb,a}$. Continuing in the same way (parameter t changed now to $-t$) from R to S gives

$$\begin{aligned} Y^i(S) &= Y^i(R) + t\Gamma^i_{jb}(R)Y^i(R) + \tfrac{1}{2}t^2\big(\Gamma^i_{kb}\Gamma^k_{jb} - \Gamma^i_{ja,b}\big)_R Y_j(R) + O(t^3) \\ &= \overset{0}{Y}{}^i - t\overset{0}{\Gamma}{}^i_{jb}\overset{0}{Y}{}^j + \tfrac{1}{2}t^2\big(-\overset{0}{\Gamma}{}^i_{jb,b} - 2\overset{0}{\Gamma}{}^i_{jb,a} + 2\overset{0}{\Gamma}{}^i_{kb}\overset{0}{\Gamma}{}^k_{ja} \\ &\quad + 2\overset{0}{\Gamma}{}^i_{ja,b} - 2\overset{0}{\Gamma}{}^i_{ka}\overset{0}{\Gamma}{}^k_{jb} + \overset{0}{\Gamma}{}^i_{kb}\overset{0}{\Gamma}{}^k_{jb}\big)\overset{0}{Y}{}^j + O(t^3) \end{aligned}$$

and finally from S back to P a parameter distance $-t$ along the b-axis

$$\begin{aligned} Y^i(P) &= Y^i(S) + t\Gamma^i_{jb}(S)Y^i(S) + \tfrac{1}{2}t^2\big(\Gamma^i_{kb}\Gamma^k_{jb} - \Gamma^i_{ja,b}\big)_S Y_j(S) + O(t^3) \\ &= \overset{0}{Y}{}^i + t^2\big(\overset{0}{\Gamma}{}^i_{ja,b} - \overset{0}{\Gamma}{}^i_{jb,a} + \overset{0}{\Gamma}{}^i_{kb}\overset{0}{\Gamma}{}^k_{ja} - \overset{0}{\Gamma}{}^i_{ka}\overset{0}{\Gamma}{}^k_{jb}\big)\overset{0}{Y}{}^j + O(t^3) \\ &= \overset{0}{Y}{}^i + t^2\overset{0}{R}{}^i_{jba}\overset{0}{Y}{}^j + O(t^3). \end{aligned}$$

Problem 18.12 **(a) Show that in a pseudo-Riemannian space the action principle**

$$\delta\int_{t_1}^{t_2} L\mathrm{d}t = 0$$

where $L = g_{ij}\dot{x}^i\dot{x}^j$ gives rise to geodesic equations with affine parameter t.

(b) For the sphere of radius a in polar coordinates,

$$\mathrm{d}s^2 = a^2(\mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\phi^2),$$

use this variation principle to write out the equations of geodesics, and read off from them the Christoffel symbols Γ^i_{jk}.

(c) Verify by direct substitution in the geodesic equations that $L = \dot{\theta}^2 +$

$\sin^2\theta\dot\phi^2$ **is a constant along the geodesics and use this to show that the general solution of the geodesic equations is given by**

$$b\cot\theta = -\cos(\phi-\phi_0) \quad \textbf{where } b, \phi_0 = \textbf{consts.}$$

(d) Show that these curves are great circles on the sphere.
[NOTE: Although the greek indices used in the text are acceptable, it is more consistent with the conventions of this book to adop roman indices.]

Solution: (a) The Euler-Lagrange equations (16.25) read

$$\begin{aligned} 0 &= \frac{\mathrm{d}}{\mathrm{d}t}\Big(\frac{\partial L}{\partial \dot x^i}\Big) - \frac{\partial L}{\partial x^i} \\ &= \frac{\mathrm{d}}{\mathrm{d}t}(2g_{ij}\dot x^j) - g_{jk,i}\dot x^j \dot x^k \\ &= 2g_{ij}\ddot x^j + 2g_{ij,k}\dot x^k \dot x^j - g_{jk,i}\dot x^j\dot x^k \end{aligned}$$

Hence

$$g_{ij}\ddot x^j + \tfrac{1}{2}(g_{ij,k}+g_{ik,j}-g_{jk,i})\dot x^j\dot x^k = 0$$

and multiplying by g^{im} results in

$$\ddot x^m + \Gamma^m_{jk}\dot x^j \dot x^k = 0,$$

the geodesic equation with affine parameter t. The Γ^m_{jk} are here the Christoffel symbols, Eq. (18.40).

(b) For $\mathrm{d}s^2 = a^2(\mathrm{d}\theta^2 + \sin^2\theta\mathrm{d}\phi^2)$, the Lagrangian is

$$L = a^2(\dot\theta^2 + \sin^2\theta\dot\phi^2).$$

The θ equation is

$$0 = \frac{\mathrm{d}}{\mathrm{d}t}\Big(\frac{\partial L}{\partial\dot\theta}\Big) - \frac{\partial L}{\partial\theta} = \frac{\mathrm{d}}{\mathrm{d}t}(2a^2\theta) - 2a^2\sin\theta\cos\theta\dot\phi^2,$$

which reads

$$\ddot\theta - \sin\theta\cos\theta\dot\phi^2 = 0. \tag{1}$$

Similarly, the ϕ equation is

$$0 = \frac{\mathrm{d}}{\mathrm{d}t}\Big(\frac{\partial L}{\partial\dot\phi}\Big) - \frac{\partial L}{\partial\phi} = 2a^2\frac{\mathrm{d}}{\mathrm{d}t}(\sin^2\theta\dot\phi), \tag{2}$$

i.e.

$$\ddot\phi + 2\cot\theta\dot\theta\dot\phi = 0 \tag{3}$$

Setting $x^1 = \theta$, $x^2 = \phi$, the Christoffel symbols can now be read off from these equations:

$$\Gamma^1_{22} = -\sin\theta\cos\theta, \quad \Gamma^2_{12} = \Gamma^2_{21} = \cot\theta$$

and all other $\Gamma^i_{jk} = 0$.

(c) From $L = a^2(\dot\theta^2 + \sin^2\theta\dot\phi^2)$ and Eqs. (1) and (3) we have

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}L &= a^2(2\dot\theta\ddot\theta + 2\sin^2\theta\dot\phi\ddot\phi + 2\sin\theta\cos\theta\dot\theta\dot\phi^2)\\ &= 2a^2(\dot\theta\sin\theta\cos\theta\dot\phi^2 - 2\sin^2\theta\cot\theta\dot\theta\dot\phi^2 + \sin\theta\cos\theta\dot\theta\dot\phi^2)\\ &= 2a^2(2\sin\theta\cos\theta - 2\sin\theta\cos\theta)\dot\theta\dot\phi^2\\ &= 0.\end{aligned}$$

From the ϕ equation in the form given in Eq. (2) we instantly have

$$\dot\phi = A\sin^{-2}\theta$$

for some constant A. From the constancy of L,

$$L = a^2\left(\left(\frac{\mathrm{d}\theta}{\mathrm{d}t}\right)^2 + \sin^2\theta\frac{\mathrm{d}\phi}{\mathrm{d}t}\right)^2\Big) = \text{const.}$$

we may change the affine parameter t to $\sqrt{L}t/a$ so that

$$\dot\theta^2 + \sin^2\dot\phi^2 = 1.$$

Then

$$\dot\theta^2 = 1 - \frac{A^2}{\sin^2\theta}$$

and

$$\frac{\mathrm{d}\theta}{\mathrm{d}\phi} = \frac{\dot\theta}{\dot\phi} = \frac{\sin\theta}{A}\sqrt{\sin^2\theta - A^2}.$$

Integration gives

$$\begin{aligned}\phi &= \int\frac{A\mathrm{d}\theta}{\sin\theta\sqrt{\sin^2\theta - A^2}} + \phi_0\\ &= \int\frac{A\mathrm{d}\theta}{\sin^2\theta\sqrt{1 - A^2/\sin^2\theta}} + \phi_0\\ &= \int\frac{-b\mathrm{d}\psi}{\sqrt{1 - b^2\psi^2}} + \phi_0\end{aligned}$$

where

$$\psi = \cot\theta, \qquad b = \frac{A}{\sqrt{1 - A^2}}.$$

Hence

$$\phi = -\cos^{-1}(b\psi) + \phi_0$$

or equivalently

$$b\cot\theta = -\cos(\phi - \phi_0).$$

(d) Great circles are the intercepts of planes $Ax + By + Cz = 0$ with the sphere $x^2 + y^2 + z^2 = a^2$, which in spherical polar coordinates reads

$$A\sin\theta\cos\phi + B\sin\theta\sin\phi + C\cos\theta = 0.$$

This equation may be rewritten as

$$\sin\theta\cos(\phi - \phi_0) + b\cos\theta = 0$$

where

$$\tan\phi_0 = \frac{B}{A}, \qquad b = \frac{C}{\sqrt{A^2 + B^2}},$$

in agreement with the above.

Problem 18.13 **Show directly from the tensor transformation laws of g_{ij} and g^{ij} that the Christoffel symbols**

$$\Gamma^i_{jk} = \tfrac{1}{2} g^{ia}(g_{aj,k} + g_{ak,j} - g_{jk,a})$$

transform as components of an affine connection.

Solution: Let

$$[a, jk] = \tfrac{1}{2}(g_{aj,k} + g_{ak,j} - g_{jk,a})$$

so that

$$\Gamma^i_{jk} = \tfrac{1}{2} g^{ia}[a, jk].$$

Under a coordinate transformation $x^j \to y^{j'}(x^1, \dots, x^n)$ we have

$$g'_{j'k'} = g_{jk}\frac{\partial x^j}{\partial y^{j'}}\frac{\partial x^k}{\partial y^{k'}}, \qquad g'^{j'k'} = g^{jk}\frac{\partial y^{j'}}{\partial x^j}\frac{\partial y^{k'}}{\partial x^k},$$

whence

$$\begin{aligned}
[a',j'k']' &= \tfrac{1}{2}(g'_{a'j',k'} + g'_{a'k',j'} - g'_{j'k',a'}) \\
&= \frac{1}{2}\Big(g_{aj,k}\frac{\partial x^a}{\partial y^{a'}}\frac{\partial^2 x^j}{\partial y^{k'}\partial y^{j'}} + g_{aj,k}\frac{\partial^2 x^a}{\partial y^{k'}\partial y^{a'}}\frac{\partial x^j}{\partial y^{j'}} + g_{aj,k}\frac{\partial x^k}{\partial y^{k'}}\frac{\partial x^a}{\partial y^{a'}}\frac{\partial x^j}{\partial y^{j'}} \\
&\quad + g_{ak,j}\frac{\partial x^a}{\partial y^{a'}}\frac{\partial^2 x^k}{\partial y^{j'}\partial y^{k'}} + g_{ak,j}\frac{\partial^2 x^a}{\partial y^{j'}\partial y^{a'}}\frac{\partial x^k}{\partial y^{k'}} + g_{ak,j}\frac{\partial x^j}{\partial y^{j'}}\frac{\partial x^a}{\partial y^{a'}}\frac{\partial x^k}{\partial y^{k'}} \\
&\quad - g_{jk,a}\frac{\partial x^j}{\partial y^{j'}}\frac{\partial^2 x^k}{\partial y^{a'}\partial y^{k'}} - g_{jk,a}\frac{\partial^2 x^j}{\partial y^{a'}\partial y^{j'}}\frac{\partial x^k}{\partial y^{k'}} + g_{jk,a}\frac{\partial x^a}{\partial y^{a'}}\frac{\partial x^j}{\partial y^{j'}}\frac{\partial x^k}{\partial y^{k'}}\Big) \\
&= [a, jk]\frac{\partial x^a}{\partial y^{a'}}\frac{\partial x^j}{\partial y^{j'}}\frac{\partial x^k}{\partial y^{k'}} + g_{ak}\frac{\partial^2 x^k}{\partial y^{j'}\partial y^{k'}}\frac{\partial x^a}{\partial y^{a'}}.
\end{aligned}$$

Hence

$$\begin{aligned}
\Gamma'^{i'}_{j'k'} &= g'^{i'a'}[a', j'k'] \\
&= g^{ib}\frac{\partial y^{i'}}{\partial x^i}\frac{\partial y^{a'}}{\partial x^b}\Big\{[a, jk]\frac{\partial x^a}{\partial y^{a'}}\frac{\partial x^j}{\partial y^{j'}}\frac{\partial x^k}{\partial y^{k'}} + g_{ak}\frac{\partial^2 x^k}{\partial y^{j'}\partial y^{k'}}\frac{\partial x^a}{\partial y^{a'}}\Big\}
\end{aligned}$$

and using the identities

$$\frac{\partial y^{a'}}{\partial x^b}\frac{\partial x^a}{\partial y^{a'}} = \delta^a_b, \qquad g^{ib}g_{bk} = \delta^i_{\ k}$$

there results

$$\Gamma'^{i'}_{\ j'k'} = \Gamma^i_{jk}\frac{\partial y^{i'}}{\partial x^i}\frac{\partial x^j}{\partial y^{j'}}\frac{\partial x^k}{\partial y^{k'}} + \frac{\partial y^{i'}}{\partial x^i}\frac{\partial^2 x^i}{\partial y^{j'}\partial y^{k'}}$$

as in Eq. (18.11).

Problem 18.14 Equation (18.46) has strictly only been proved for a vector field A. Show that it holds equally for a vector field whose components $A^i(t,\lambda)$ are only defined on the one-parameter family of curves γ.

Solution: If $A^i = A^i(t,\lambda)$ is only defined on the 2-surface swept out by the curves γ then we cannot use the Ricci identities since the covariant derivative $A^i_{\ ;j}$ can not in general be defined. However

$$\frac{DA^i}{\partial\lambda} = \frac{\partial A^i}{\partial\lambda} + \Gamma^i_{jk}A^j\frac{\partial\gamma^k}{\partial\lambda} = \frac{\partial A^i}{\partial\lambda} + \Gamma^i_{jk}A^jY^k$$
$$\frac{DA^i}{\partial t} = \frac{\partial A^i}{\partial t} + \Gamma^i_{jk}A^j\frac{\partial\gamma^k}{\partial t} = \frac{\partial A^i}{\partial t} + \Gamma^i_{jk}A^jX^k$$

Hence

$$\frac{D}{\partial t}\frac{D}{\partial\lambda}A^i = \frac{\partial}{\partial t}\Big(\frac{\partial A^i}{\partial\lambda} + \Gamma^i_{jk}A^jY^k\Big) + \Gamma^i_{jk}\Big(\frac{\partial A^j}{\partial\lambda} + \Gamma^j_{lm}A^lY^m\Big)X^k$$
$$\frac{D}{\partial t}\frac{D}{\partial\lambda}A^i = \frac{\partial}{\partial\lambda}\Big(\frac{\partial A^i}{\partial t} + \Gamma^i_{jk}A^jX^k\Big) + \Gamma^i_{jk}\Big(\frac{\partial A^j}{\partial t} + \Gamma^j_{lm}A^lX^m\Big)Y^k$$

and using

$$\frac{\partial}{\partial t}\Gamma^i_{jk} = \Gamma^i_{jk,l}X^l, \qquad \frac{\partial}{\partial\lambda}\Gamma^i_{jk} = \Gamma^i_{jk,l}Y^l,$$

we have

$$\begin{aligned}\frac{D}{\partial t}\frac{D}{\partial\lambda}A^i - \frac{D}{\partial\lambda}\frac{D}{\partial t}A^i &= \Gamma^i_{jk,l}X^lA^jY^k - \Gamma^i_{jk,l}Y^lA^jX^k \\ &\quad + \Gamma^i_{jk}\Big(\frac{\partial A^j}{\partial t}Y^k + A^j\frac{\partial Y^k}{\partial t} - \frac{\partial A^j}{\partial\lambda}X^k - A^j\frac{\partial X^k}{\partial\lambda} \\ &\quad + \frac{\partial A^j}{\partial\lambda}X^k + \Gamma^j_{lm}A^lY^mX^k - \frac{\partial A^j}{\partial t}Y^k - \Gamma^j_{lm}A^lX^mY^k\Big) \\ &= \big(\Gamma^i_{ak,j} - \Gamma^i_{aj,k} + \Gamma^i_{bj}\Gamma^b_{ak} - \Gamma^i_{bk}\Gamma^b_{aj}\big)A^aX^jY^k \\ &= R^i_{\ ajk}A^aX^jY^k,\end{aligned}$$

by Eq. (18.25). It is worth noting that this result is independent of the torsion-free condition (assumed in Eq. (18.44)) and holds even if $T^i_{jk} \neq 0$.

Problem 18.15 **Let $e_i = \partial_{x^i}$ be a coordinate basis.**

(a) Show that the first Bianchi identity reads

$$R^i{}_{[jkl]} = T^i_{[jk;l]} - T^a_{[jk}T^i_{l]a},$$

and reduces to the cyclic identity (18.26) in the case of a torsion-free connection.

(b) Show that the second Bianchi identity becomes

$$R^i{}_{j[kl;m]} = R^i{}_{ja[k}T^a_{ml]},$$

which is identical with Eq. (18.32) of Problem 18.10.

Solution: (a) Assuming $\varepsilon^k = \mathrm{d}x^k$ the first Bianchi identity (18.70) reads

$$\mathrm{d}\tau^i = \rho^i{}_k \wedge \mathrm{d}x^k - \omega^i{}_k \wedge \tau^k.$$

From Eqs. (18.69) and (18.66), we have

$$\begin{aligned} \mathrm{d}\tau^i &= \mathrm{d}(T^i_{jk}\mathrm{d}x^j \wedge \mathrm{d}x^k) \\ &= T^i_{jk,l}\mathrm{d}x^l \wedge \mathrm{d}x^j \wedge \mathrm{d}x^k \\ &= T^i_{[jk,l]}\mathrm{d}x^j \wedge \mathrm{d}x^k \wedge \mathrm{d}x^l \end{aligned}$$

and

$$\begin{aligned} \rho^i{}_k \wedge \mathrm{d}x^k &= \rho^i{}_j \wedge \mathrm{d}x^j \\ &= R^i{}_{jkl}\mathrm{d}x^k \wedge \mathrm{d}x^l \wedge \mathrm{d}x^j \\ &= R^i{}_{[jkl]}\mathrm{d}x^j \wedge \mathrm{d}x^k \wedge \mathrm{d}x^l \end{aligned}$$

since cyclic permutations ofthree indices are even, and from Eq.(18.68)

$$\begin{aligned} -\omega^i{}_k \wedge \tau^k &= -\omega^i{}_a \wedge \tau^a \\ &= -\Gamma^i_{al}\mathrm{d}x^l \wedge T^a_{jk}\mathrm{d}x^j \wedge \mathrm{d}x^k \\ &= -\Gamma^i_{a[l}T^a_{jk]}\mathrm{d}x^j \wedge \mathrm{d}x^k \wedge \mathrm{d}x^l. \end{aligned}$$

Hence the totally antisymmetrized coefficient of $\mathrm{d}x^j \wedge \mathrm{d}x^k \wedge \mathrm{d}x^l$ in the first Bianchi identity gives

$$T^i_{[jk,l]} = R^i{}_{[jkl]} - \Gamma^i_{a[l}T^a_{jk]}.$$

Now from the general covariant derivative formula (18.14)

$$T^i_{jk;l} = T^i_{jk,l} + \Gamma^i_{al}T^a_{jk} - \Gamma^a_{jl}T^i_{ak} - \Gamma^a_{kl}T^i_{ja}$$

so that, using $T^i_{ak} = -T^i_{ka}$,

$$\begin{aligned} T^i_{[jk;l]} &= T^i_{[jk,l]} + \Gamma^i_{a[l}T^a_{jk]} + \Gamma^a_{[jl}T^i_{k]a} - \Gamma^a_{[kl}T^i_{j]a} \\ &= R^i{}_{[jkl]} + \Gamma^a_{[jl}T^i_{k]a} - \Gamma^a_{[kl}T^i_{j]a} \\ &= R^i{}_{[jkl]} + T^a_{[kl}T^i_{j]a} \end{aligned}$$

using $T^a_{kl} = \Gamma^a_{lk} - \Gamma^a_{kl}$ and performing a cyclic permutation on the middle term in the penultimate line. After another cyclic permutation we obtain the desired

$$R^i{}_{[jkl]} = T^i_{[jk;l]} - T^a_{[jk}T^i_{l]a}.$$

If $T^i_{jk} = 0$ this reduces to the identity $R^i{}_{[jkl]} = 0$, which on using skew symmetry in the last pair of indices can be written in the cyclic form

$$2(R^i{}_{jkl} + R^i{}_{klj} + R^i{}_{ljk}) = 0.$$

(b) In the second Bianchi identity (18.71)

$$\mathrm{d}\rho^i{}_j = \rho^i{}_k \wedge \omega^k_j - \omega^i{}_k \wedge \rho^k_j$$

we have

$$\mathrm{d}\rho^i{}_j = R^i{}_{jkl,m}\mathrm{d}x^k \wedge \mathrm{d}x^l \wedge \mathrm{d}x^m = R^i{}_{j[kl,m]}\mathrm{d}x^k \wedge \mathrm{d}x^l \wedge \mathrm{d}x^m$$

and

$$\begin{aligned}
\rho^i{}_k \wedge \omega^k_j - \omega^i{}_k \wedge \rho^k_j \\
&= R^i{}_{klm}\Gamma^k_{ja}\mathrm{d}x^l \wedge \mathrm{d}x^m \wedge \mathrm{d}x^a - \Gamma^i_{ka}R^k{}_{jlm}\mathrm{d}x^a \wedge \mathrm{d}x^l \wedge \mathrm{d}x^m \\
&= \big(R^i{}_{alm}\Gamma^a_{jk} - R^a{}_{jlm}\Gamma^i_{ak}\big)\mathrm{d}x^k \wedge \mathrm{d}x^l \wedge \mathrm{d}x^m
\end{aligned}$$

after some interchanges of dummy summation indices. Hence

$$R^i{}_{j[kl,m]} = R^i{}_{a[lm}\Gamma^a_{|j|k]} - R^a{}_{j[lm}\Gamma^i_{|a|k]}$$

where a pair of vertical lines sorrounding an index or group of indices indicates that those indices are not to be included in the antisymmetrization bracket $[\dots]$. Expanding the covariant derivative of $R^i{}_{jkl}$, we have

$$\begin{aligned}
R^i{}_{j[kl;m]} &= R^i{}_{j[kl,m]} + \Gamma^i_{a[m}R^a{}_{|j|kl]} - \Gamma^a_{j[m}R^i{}_{|a|kl]} \\
&\quad - \Gamma^a_{[km}R^a{}_{|ja|l]} + \Gamma^i_{[lm}R^i{}_{|j|k]a} \\
&= \Gamma^a_{j[k}R^i{}_{|a|lm]} - \Gamma^i_{a[k}R^a{}_{|j|lm]} + \Gamma^i_{a[m}R^a{}_{|j|kl]} - \Gamma^a_{j[m}R^i{}_{|a|kl]} \\
&\quad - \Gamma^a_{[km}R^i{}_{|ja|l]} + \Gamma^i_{[lm}R^i{}_{|j|k]a} \\
&= \Gamma^a_{[km}R^i{}_{|j|l]a} - \Gamma^i_{[ml}R^i{}_{|j|k]a} \\
&= R^i{}_{ja[l}(-\Gamma^i_{km]} + \Gamma^i_{mk]}) \\
&= R^i{}_{ja[l}T^a_{km]} \\
&= R^i{}_{ja[k}T^a_{ml]}.
\end{aligned}$$

Problem 18.16 **In a Riemannian manifold (M, g) show that the sectional curvature $K(X, Y)$ at a point p, defined in Example 18.4, is independent of the choice of basis of the 2-space; i.e., $K(X', Y') = K(X, Y)$ if**

$X' = aX + bY$**,** $Y' = cX + dY$ **where** $ad - bc \neq 0$**.**
The space is said to be *isotropic* at $p \in M$ **if** $K(X, Y)$ **is independent of the choice of tangent vectors** X **and** Y **at** p**. If the space is isotropic at each point** p **show that**

$$R_{ijkl} = f(g_{ik}g_{jl} - g_{il}g_{jk})$$

where f **is a scalar field on** M**. If the dimension of the manifold is greater than 2, show *Schur's theorem*: a Riemannian manifold that is everywhere isotropic is a space of constant curvature,** f **= const. [*Hint*: Use the contracted Bianchi identity (18.59).]**

Solution: If $X' = aX + bY$, $Y' = cX + dY$ then, using the symmetries

$$R(X, Y, W, Z) = -R(Y, X, W, Z) = -R(X, Y, Z, W)$$

we have

$$\begin{aligned}
R(X', Y', X', Y') &= R(aX + bY, cX + dY, aX + bY, cX + dY) \\
&= R(aX, dY, aX, dY) + R(aX, dY, bY, cX) + R(bY, cX, aX, dY) + R(bY, cX, bY, cX) \\
&= ad(ad - bc)R(X, Y, Z, W) - bc(ad - bc)R(X, Y, Z, W) \\
&= (ad - bc)^2 R(X, Y, Z, W)
\end{aligned}$$

We can set $A(X, Y) = -G(X, Y, X, Y)$ where the tensor G defined by

$$\begin{aligned}
G(X, Y, Z, W) &= g(X, Z)g(Y, W) - g(X, W)g(Y, Z) \\
&= -G(Y, X, Z, W) = -G(X, Y, W, Z)
\end{aligned}$$

then, by the same argument to that above

$$A(X', Y') = (ad - bc)^2 A(X, Y).$$

Hence

$$\begin{aligned}
K(X', Y') = \frac{R(X', Y', X', Y')}{A(X', Y')} &= \frac{(ad - bc)^2 R(X, Y, Z, W)}{(ad - bc)^2 A(X, Y)} \\
&= \frac{R(X, Y, Z, W)}{A(X, Y)} = K(X, Y).
\end{aligned}$$

If $K(X, Y)$ is independent of X, Y at p then there exists a constant f at p such that

$$R(X, Y, X, Y) = -fA(X, Y) = fG(X, Y, X, Y).$$

Set S to the tensor of type $(0, 4)$, $S = R - fG$, having the following symmetries (true of both R and G separately):

$$S(X, Y, W, Z) = -S(Y, X, W, Z) = -S(X, Y, Z, W) = S(W, Z, X, Y)$$

and the cyclic symmetry

$$S(X, Y, W, Z) + S(X, W, Z, Y) + S(X, Z, Y, W) = 0.$$

Then, assuming $S(X,Y,X,Y) = 0$ for all tangent vectors X, Y at p (automatically true if $X \propto Y$), then

$$\begin{aligned}
0 &= S(X+W, Y+Z, X+W, Y+Z) \\
&= S(X,Y,X,Y) + S(X,Y,X,Z) + S(X,Y,W,Y) + S(X,Y,W,Z) \\
&\quad + S(X,Z,X,Y) + S(X,Z,X,Z) + S(X,Z,W,Y) + S(X,Z,W,Z) \\
&\quad + S(W,Y,X,Y) + S(W,Y,X,Z) + S(W,Y,W,Y) + S(W,Y,W,Z) \\
&\quad + S(W,Z,X,Y) + S(W,Z,X,Z) + S(W,Z,W,Y) + S(W,Z,W,Z) \\
&= 2\big(S(X,Y,X,Z) + S(X,Y,W,Y) + S(X,Z,W,Z) + S(W,Y,W,Z) \\
&\qquad + S(X,Y,W,Z) + S(X,Z,W,Z)\big)
\end{aligned}$$

Setting $Z = 0$ gives $S(X,Y,W,Y) = 0$, and similarly setting $X = 0$ gives $S(W,Y,W,Z) = 0$ etc. Hence

$$S(X,Y,W,Z) + S(X,Z,W,Y) = 0.$$

Using this identity together with the cyclic and other symmetries we have

$$\begin{aligned}
S(X,Y,W,Z) &= -S(X,W,Z,Y) - S(X,Z,Y,W) \\
&= S(X,W,Z,Y) + S(X,W,Y,Z) \\
&= -S(X,Y,W,Z) + S(X,W,Y,Z).
\end{aligned}$$

Hence

$$S(X,Y,W,Z) = \tfrac{1}{2} S(X,W,Y,Z)$$

and repeating this identity gives

$$S(X,W,Y,Z) = \tfrac{1}{2} S(X,Y,W,Z)$$

so that

$$S(X,Y,W,Z) = \tfrac{1}{4} S(X,Y,W,Z)$$

i.e. $S(X,Y,W,Z) = 0$. This shows that $R = fG$ where f is a scalar field on M, i.e. in components

$$R_{ijkl} = f(g_{ik}g_{jl} - g_{il}g_{jk}).$$

The Ricci tensor is

$$R_{ij} = g^{ab}R_{aibj} = fg^{ab}(g_{ab}g_{ij} - g_{aj}g_{ib} = (n-1)f\, g_{ab}$$

and Ricci scalar is

$$R = g^{ij}R_{ij} = n(n-1)f.$$

From the contracted Bianchi identities (18.59),

$$R^i{}_{j;i} = \tfrac{1}{2} R_{,i}$$

we have

$$(n-1)(f\delta^i{}_j)_{;i} = (n-1)f_{,i} = \tfrac{1}{2} n(n-1)f_{,i}$$

i.e.

$$(n-2)f_{,i} = 0.$$

Hence if the dimension $n > 2$ then $f = \text{const.}$, the space if a space of constant curvature.

Problem 18.17 **Show that a space with symmetric connection is locally flat if and only if there exists a local basis of vector fields $\{e_i\}$ that are *absolutely parallel*, $De_i = 0$.**
[NOTE: must assume space has symmetric connection]

Solution: If there exists a basis $\{e_i\}$ such that $De_i = 0$ then $D_X e_i = 0$ for all vector fields X. Hence

$$\rho_{X,Y} e_i = (D_X D_Y - D_Y D_X - D_{[X,Y]})e_i = 0$$

and the curvature 2-forms $\rho^i_{\ j} = 0$, since for any pair of vector fields X, Y

$$\rho^i_{\ j}(X,Y) = \langle \varepsilon^i, \rho_{X,Y} e_i \rangle = 0.$$

By Theorem 18.1 it follows that the space is locally flat.

Conversely, if the space is locally flat then there exists local coordinates such that $\Gamma^i jk = 0$, i.e.

$$D_i \partial_{x^j} = \Gamma^k_{ji} \partial_{x^k} = 0.$$

Setting $e_j = \partial_{x^j}$ we have $D_i e_j = 0$, i.e. for all vector fields $X = X^i e_i$

$$D_X e_j = X^i D_i e_j = 0.$$

In other words, $De_i = 0$.

Problem 18.18 **Let (M, φ) be a *surface of revolution* defined as a submanifold of $\mathbb{E}^3$ of the form**

$$x = g(u)\cos\theta, \quad y = g(u)\sin\theta, \quad z = h(u).$$

Show that the induced metric (see Example 18.1) is

$$\mathrm{d}s^2 = \big(g'(u)^2 + h'(u)^2\big)\mathrm{d}u^2 + g^2(u)\mathrm{d}\theta^2.$$

Picking the parameter u such that $g'(u)^2 + h'(u)^2 = 1$ (interpret this choice!), and setting the basis 1-forms to be $\varepsilon^1 = \mathrm{d}u$, $\varepsilon^2 = g\,\mathrm{d}\theta$, calculate the connection 1-forms $\omega^i_{\ j}$, the curvature 1-forms $\rho^i_{\ j}$, and the curvature tensor component R_{1212}.

Solution: The induced metric $g = \varphi^*(\mathrm{d}x \otimes \mathrm{d}x + \mathrm{d}y \otimes \mathrm{d}y + \mathrm{d}z \otimes \mathrm{d}z)$ is found from

$$\begin{aligned}
\mathrm{d}x &= g'(u)\cos\theta \mathrm{d}u - g(u)\sin\theta \mathrm{d}\theta \\
\mathrm{d}y &= g'(u)\sin\theta \mathrm{d}u + g(u)\cos\theta \mathrm{d}\theta \\
\mathrm{d}z &= h'(u)
\end{aligned}$$

so that

$$\mathrm{d}s^2 = \big(g'(u)^2 + h'(u)^2\big)\mathrm{d}u^2 + g^2(u)\mathrm{d}\theta^2.$$

If we choose the parameter u such that $g'(u)^2 + h'(u)^2 = 1$ then

$$\mathrm{d}s^2 = \mathrm{d}u^2 + g^2(u)\mathrm{d}\theta^2,$$

and u is the "radial distance prameter" along lines $\theta = \text{const}$. Let e_i be the orthonormal basis

$$e_1 = \partial_u, \qquad e_2 \frac{1}{g}\partial_\theta$$

with dual basis

$$\varepsilon^1 = \mathrm{d}u, \qquad \varepsilon^2 = g\,\mathrm{d}\theta$$

Setting $\omega_{ik} = -\omega_{ki} = A_{ikj}\varepsilon^j$, the first Cartan structural formulae give

$$\mathrm{d}\varepsilon^i = -\omega_{ik} \wedge \varepsilon^k = -A_{ikj}\varepsilon^j \wedge \varepsilon^k = A_{ijk}\varepsilon^j \wedge \varepsilon^k,$$

where $A_{ikj} = -A_{kij}$. In particular $A_{iij} = A_{22j} = 0$. Hence

$$\mathrm{d}\varepsilon^1 = 0 \quad \Longrightarrow \quad A_{121}\varepsilon^2 \wedge \varepsilon^1 = 0 \quad \Longrightarrow \quad A_{121} = A_{211} = 0.$$

and

$$\mathrm{d}\varepsilon^2 = g'(u)\mathrm{d}u \wedge \mathrm{d}\theta = \frac{g'}{g}\varepsilon^1 \wedge \varepsilon^2 \quad \Longrightarrow \quad A_{212} = -A_{122} = \frac{g'}{g}.$$

Thus $\omega_{12} = -\omega_{21} = -\frac{g'}{g}\varepsilon^2$ and

$$\omega^1_{\ 2} = -\omega^2_{\ 1} = -\frac{g'}{g}\varepsilon^2.$$

The second Cartan structural formulae give

$$\rho_{ij} = 2(\mathrm{d}\omega_{ij} + \omega_{ik} \wedge \omega^k_{\ j})$$

so that

$$\begin{aligned}
\rho_{12} &= 2\big(\mathrm{d}\big(-\frac{g'}{g}\varepsilon^2\big) + \omega^{12} \wedge \omega^2_{\ 2}\big) \\
&= 2\Big(-\frac{g''}{g} + \frac{(g')^2}{g^2}\Big)\varepsilon^1\varepsilon^2
\end{aligned}$$

and using

$$\rho_{12} = R_{12kl}\varepsilon^k \wedge \varepsilon^l = R_{1212}\varepsilon^1 \wedge \varepsilon^2 + R_{1221}\varepsilon^2 \wedge \varepsilon^1 = 2R_{1212}\varepsilon^1 \wedge \varepsilon^2$$

we have

$$R_{1212} = -\frac{g''(u)}{g(u)} + \frac{(g'(u))^2}{g(u)^2}.$$

Problem 18.19 **For the ellipsoid**

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$$

show that the sectional curvature is given by

$$K = \Big(\frac{x^2bc}{a^3} + \frac{y^2ac}{b^3} + \frac{z^2ab}{c^3}\Big)^{-2}.$$

Solution: There seems no easy way of doing this problem, and any reader who can come up with a simpler calculation will be gratefully acknowledged in the next update of this file.

Parametrically, we can define the ellipsoid by

$$x = a\sin\theta\cos\phi, \qquad y = b\sin\theta\sin\phi, \qquad z = c\cos\theta$$

which results in a metric

$$\begin{aligned} \mathrm{d}s^2 &= \mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2 \\ &= (a^2\cos^2\theta\cos^2\phi + b^2\cos^2\theta\sin^2\phi + c^2\sin^2\theta)\mathrm{d}\theta^2 \\ &\quad + 2(b^2 - a^2)\cos\theta\sin\theta\cos\phi\sin\phi\mathrm{d}\theta\mathrm{d}\phi \\ &= \sin^2\theta(a^2\sin^2\phi + b^2\cos^2\phi)\mathrm{d}\phi^2. \end{aligned}$$

Set, for example

$$\varepsilon^1 = A\mathrm{d}\theta, \qquad \varepsilon^2 = B\mathrm{d}\theta + C\mathrm{d}\phi,$$

then we have

$$g = \varepsilon^1 \otimes \varepsilon^1 + \varepsilon^2 \otimes \varepsilon^2$$

if and only if

$$\begin{aligned} C^2 &= \sin^2\theta(a^2\sin^2\phi + b^2\cos^2\phi) \\ B &= C^{-1}(b^2 - a^2)\cos\theta\sin\theta\cos\phi\sin\phi \\ A^2 &= a^2\cos^2\theta\cos^2\phi + b^2\cos^2\theta\sin^2\phi + c^2\sin^2\theta - B^2. \end{aligned}$$

Then in this basis $g_{ij} = $ consts. and $\omega_{ij} = -\omega_{ji}$. In the first Cartan structural formulae

$$\mathrm{d}\varepsilon^1 = -\omega_{12} \wedge \varepsilon^2, \qquad \mathrm{d}\varepsilon^2 = -\omega_{21} \wedge \varepsilon^1 = \omega_{12} \wedge \varepsilon^1$$

we have

$$\mathrm{d}\varepsilon^1 = \frac{\partial A}{\partial \phi}\mathrm{d}\phi \wedge \mathrm{d}\theta = -\frac{1}{CA}\frac{\partial A}{\partial \phi}\varepsilon^1 \wedge \varepsilon^2$$
$$\mathrm{d}\varepsilon^2 = \frac{\partial B}{\partial \phi}\mathrm{d}\phi \wedge \mathrm{d}\theta + \frac{\partial C}{\partial \theta}\mathrm{d}\theta \wedge \mathrm{d}\phi = \frac{1}{CA}\Big(\frac{\partial C}{\partial \theta} - \frac{\partial B}{\partial \phi}\Big)\varepsilon^1 \wedge \varepsilon^2$$

whence

$$\omega_{12} = -\omega_{21} = \frac{1}{CA}\frac{\partial A}{\partial \phi}\varepsilon^1 + \frac{1}{CA}\Big(\frac{\partial B}{\partial \phi} - \frac{\partial C}{\partial \theta}\Big)\varepsilon^2.$$

From the second Cartan structural formula

$$\rho = 2\mathrm{d}\omega_{12} - \omega_{1k} \wedge \omega^k{}_2 = 2\mathrm{d}\omega_{12}$$

since both terms in the k-sum vanish as $\omega_{11} = \omega^2{}_2 = \omega_{22} = 0$. A massive calculation should result in

$$\rho_{12} = K\varepsilon^1 \wedge \varepsilon 2$$

where

$$K = \frac{1}{\Big((bc/a)\sin^2\theta\cos^2\phi + (ac/b)\sin^2\theta\sin^2\phi + (ab/c)\cos^2\theta\Big)^2}$$
$$= \Big(\frac{bc}{a^3}x^2 + \frac{ac}{b^3}y^2 + \frac{ab}{c^3}z^2\Big)^{-2}$$

An easier case which definitely comes out correctly is the symmetric or spheroidal case, $a = b$, in which the metric is diagonal in these coordinates and

$$B = 0, \qquad C^2 = a^2\sin^2\theta, \qquad A^2 = a^2\cos^2\theta + c^2\sin^2\theta.$$

Then

$$\omega_{12} = -\frac{\cot\theta}{\big(a^2\cos^2\theta + c^2\sin^2\theta\big)^{1/2}}\varepsilon^2$$

and after a relatively straightforward calculation

$$\rho_{12} = \mathrm{d}\omega_{12} = \frac{c^2\varepsilon^1 \wedge \varepsilon^2}{\big(a^2\cos^2\theta + c^2\sin^2\theta\big)^2}$$

so that

$$K = \frac{c^2}{\big(a^2\cos^2\theta + c^2\sin^2\theta\big)^2} = \Big(\frac{c}{a^2}(x^2 + y^2) + \frac{a^2}{c^3}z^2\Big)^{-2}.$$

Problem 18.20 **A linearized plane gravitational wave is a solution of the linearized Einstein equations $\Box h_{\mu\nu} = 0$ of the form $h_{\mu\nu} = h_{\mu\nu}(u)$ where $u = x^3 - x^4 = z - ct$. Show that the harmonic gauge condition (18.89) implies that, up to undefined constants,**

$$h_{14} + h_{13} = h_{24} + h_{23} = h_{11} + h_{22} = 0, \qquad h_{34} = -\tfrac{1}{2}(h_{33} + h_{44}).$$

Use the remaining gauge freedom $\xi_\mu = \xi_\mu(u)$ to show that it is possible to transform $h_{\mu\nu}$ to the form

$$[h_{\mu\nu}] = \begin{pmatrix} \mathsf{H} & \mathsf{O} \\ \mathsf{O} & \mathsf{O} \end{pmatrix} \qquad \textbf{where} \qquad \mathsf{H} = \begin{pmatrix} h_{11} & h_{12} \\ h_{12} & -h_{11} \end{pmatrix}.$$

Setting $h_{11} = \alpha(u)$ and $h_{12} = \beta(u)$, show that the equation of geodesic deviation has the form

$$\ddot\eta_1 = \frac{\epsilon}{2}c^2\left(\alpha''\eta_1 + \beta''\eta_2\right) \qquad \ddot\eta_2 = \frac{\epsilon}{2}c^2\left(\beta''\eta_1 - \alpha''\eta_2\right)$$

and $\ddot\eta_3 = 0$. Make a sketch of the distribution of neighbouring accelerations of freely falling particles about a geodesic observer in the two cases $\beta = 0$ and $\alpha = 0$. These results are central to the observational search for gravity waves.

Solution: If $h_{\mu\nu} = h_{\mu\nu}(u)$ where $u = x^3 - x^4 = z - ct$ the harmonic condition $h^\nu{}_{\mu,\nu} = \frac{1}{2}h_{,\mu}$, reads

$$h'_{3\mu} + h'_{4\mu} = \tfrac{1}{2}h'\delta_{3\mu} - \tfrac{1}{2}h'\delta_{4\mu}$$

where $' \equiv \mathrm{d}/\mathrm{d}u$ and $h = h_{11} + h_{22} + h_{33} - h_{44}$. Hence, setting $\mu = 1, 2, 3, 4$ successively results in the following equations:

$$h'_{31} + h'_{41} = 0 \tag{1}$$
$$h'_{32} + h'_{42} = 0 \tag{2}$$
$$h'_{33} + h'_{43} = \tfrac{1}{2}h' \tag{3}$$
$$h'_{34} + h'_{44} = -\tfrac{1}{2}h'. \tag{4}$$

Up to arbitrary constants the first two equations imply

$$h_{14} = -h_{13}, \qquad h_{24} = -h_{23}.$$

Adding (3) and (4) gives

$$h_{34} = -\tfrac{1}{2}(h_{33} + h_{44}),$$

while subtracting (4) from (3) provides

$$h_{33} - h_{44} = h = h_{11} + h_{22} + h_{33} - h_{44}$$

i.e.

$$h_{11} + h_{22} = 0.$$

Still available are gauge transformations

$$h_{\mu\nu} \to h_{\mu\nu} + \xi_{\mu,\nu} + \xi_{\nu,\mu}$$

provided $\Box\xi_\mu = 0$. In particular we may set $\xi = \xi(u)$, having the effect

$$\begin{aligned} h_{44} &\to h_{44} - 2\xi_4' \\ h_{33} &\to h_{33} + 2\xi_3' \\ h_{14} &\to h_{14} - 2\xi_1' \\ h_{24} &\to h_{24} - 2\xi_2'. \end{aligned}$$

The ξ_μ may be chosen such that

$$h_{44} = h_{33} = h_{14} = h_{24} = 0$$

and using the above identities we also have $h_{34} = h_{13} = h_{23} = 0$. Hence the only non-zero components of $h_{\mu\nu}$ are $h_{11} = -h_{22}$ and $h_{12} = h_{21}$:

$$[h_{\mu\nu} = \begin{pmatrix} h_{11} & h_{12} & 0 & 0 \\ h_{12} & -h_{11} & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

The components of curvature tensor are

$$R_{\mu\nu\rho\sigma} = \frac{\epsilon}{2}(h_{\mu\sigma,\nu\rho} + h_{\nu\rho,\mu\sigma} - h_{\nu\sigma,\mu\rho} - h_{\mu\rho,\nu\sigma})$$

and the equation of geodesic equation is (summation convention over cartesian components $i, j = 1, 2, 3$)

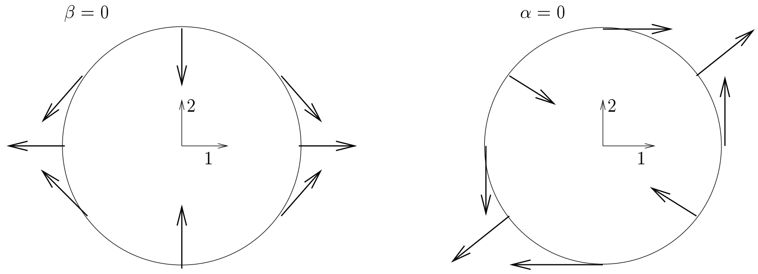
$$\ddot{\eta}_i = -R_{i4j4}\eta_j.$$

The components of the form R_{i4j4} are

$$\begin{aligned} R_{i434} &= R_{34i4} = 0, \\ R_{1414} &= -R2424 = -\frac{\epsilon}{2}h_{11}'' = -\frac{\epsilon}{2}\alpha'' \\ R_{1424} &= R2414 = -\frac{\epsilon}{2}h_{12}'' = -\frac{\epsilon}{2}\beta'' \end{aligned}$$

and the equations of geodesic equation read

$$\begin{aligned} \ddot{\eta}_3 &= 0 \\ \ddot{\eta}_1 &= -c^2\big(R_{1414}\eta_1 + R_{1424}\eta_2\big) = \frac{\epsilon}{2}c^2\big(\alpha''\eta_1 + \beta''\eta_2\big) \\ \ddot{\eta}_2 &= -c^2\big(R_{2414}\eta_1 + R_{2424}\eta_2\big) = \frac{\epsilon}{2}c^2\big(\beta''\eta_1 - \alpha''\eta_2\big) \end{aligned}$$



Problem 18.21 **Show that every two-dimensional space-time metric (signature 0) can be expressed locally in *conformal coordinates***

$$ds^2 = e^{2\phi}\big(dx^2 - dt^2\big) \qquad \textbf{where} \qquad \phi = \phi(x,t).$$

Calculate the Riemann curvature tensor component R_{1212}, and write out the two-dimensional Einstein vacuum equations $R_{ij} = 0$. What is their general solution?

Solution: It is no straightforward matter to show this by seeking a direct coordinate transformation, since it involves solving a coupled pair of simultaneous partial differential equations, and it is non-trivial to see that a local solution always exists. An alternative approach is to show that every two dimensional metric tensor g_{ij} $(i,j = 1,2)$ is *conformal* to a flat metric: there exists a scalar field ϕ such that the metric $\hat{g}_{ij} = e^{-2\phi} g_{ij}$ has vanishing Riemann tensor, $\hat{R}_{ijkl} = 0$. We calculate this Riemann tensor by first computing the Christoffel symbols

$$\begin{aligned}
\hat{\Gamma}^i_{jk} &= \tfrac{1}{2}\hat{g}^{ia}(\hat{g}_{aj,k} + \hat{g}_{ak,j} - \hat{g}_{jk,a}) \\
&= \tfrac{1}{2}g^{ia}(g_{aj,k} + g_{ak,j} - g_{jk,a} - 2g_{aj}\phi_{,k} - 2g_{ak}\phi_{,j} + 2g_{jk}\phi_{,a}) \\
&= \Gamma^i_{jk} - \delta^i_{\ j}\phi_{,k} - \delta^i_{\ k}\phi_{,j} + g^{ia}g_{jk}\phi_{,a}) \qquad (*)
\end{aligned}$$

and

$$\begin{aligned}
\hat{R}_{ijkl} &= \hat{g}_{ia}\hat{R}^a_{\ jkl} \\
&= e^{-2\phi}g_{ia}(\hat{\Gamma}^a_{jl,k} - \hat{\Gamma}^a_{jk,l} + \hat{\Gamma}^b_{jl}\hat{\Gamma}^a_{bk} - \hat{\Gamma}^b_{jk}\hat{\Gamma}^a_{bl} \\
&= e^{-2\phi}\big[R_{ijkl} + g_{ia}(-\delta^a_j\phi_{,l} - \delta^a_l\phi_{,j} + g^{ab}g_{jl}\phi_{,b})_{,k} \\
&\qquad - g_{ia}(-\delta^a_j\phi_{,k} - \delta^a_k\phi_{,j} + g^{ab}g_{jk}\phi_{,b})_{,l} + \dots\big] \\
&= e^{-2\phi}\big[R_{ijkl} - g_{il}\phi_{,jk} + g_{ik}\phi_{,jl} + g_{jl}\phi_{,ik} - g_{jk}\phi_{,il} + F_{ijkl}(g_{ab}, g_{ab,c}, \phi_{,a})\big].
\end{aligned}$$

Hence

$$\hat{R}_{1212} = e^{-2\phi}\big[R_{1212} + g_{11}\phi_{,22} + g_{22}\phi_{,11} - 2g_{12}\phi_{,12} + F(x^a, \phi_{,1}, \phi_{,2})\big]$$

and the equation $\hat{R}_{1212} = 0$ is a single partial differential equation of second order for ϕ which always has a local solution. Since all this is the only independent component of $\hat{R}_{ijkl}$ in two dimensions, the metric $\hat{g}_{ij}$ is a flat metric and by Theorem 18.1 there exist local coordinates such that

$$d\hat{s}^2 = \hat{g}_{ij}dx^i dx^j = dx^2 - dt^2$$

where we have set $x = x^1$, $t = x^2$. Hence $ds^2 = e^{2\phi}(dx^2 - dt^2)$, i.e.

$$g_{11} = -g_{22}e^{2\phi}, \quad g_{12} = 0.$$

Setting $\hat{\Gamma}^i_{jk} = 0$ in Eq. $(*)$, we have

$$\begin{aligned} \Gamma^1_{11} = \Gamma^1_{22} = \phi_{,x} \qquad & \Gamma^1_{12} = \phi_{,t}, \\ \Gamma^2_{11} = \Gamma^2_{22} = \phi_{,t} \qquad & \Gamma^2_{12} = \phi_{,x}, \end{aligned}$$

Hence

$$\begin{aligned} R_{1212} = e^{2\phi}R^1{}_{212} &= e^{2\phi}\big(\Gamma^1_{22,1} - \Gamma^1_{21,2} + \Gamma^a_{22}\Gamma^1_{a1} - \Gamma^a_{21}\Gamma^1_{a2}\big) \\ &= e^{2\phi}\big(\phi_{,xx} - \phi_{,tt} + (\phi_{,x})^2 + (\phi_{,t})^2 - (\phi_{,x})^2 - (\phi_{,y})^2\big) \\ &= e^{2\phi}\big(\phi_{,xx} - \phi_{,tt}\big) \end{aligned}$$

and

$$R_{11} = R^1{}_{1a1} = -e^{-2\phi}R_{2121} = -e^{-2\phi}R_{1212} = \phi_{,tt} - \phi_{xx}.$$

Similarly

$$R_{22} = R^1{}_{1a1} = e^{-2\phi}R_{1212} = \phi_{,xx} - \phi_{tt}$$

and $R_{12} = R^a_{1a2} = 0$. The equation $R_{ij} = 0$ therefeore is equivalent to the two-dimensional wave equation

$$\phi_{,tt} - \phi_{xx} = 0$$

which has general solution

$$\phi = f(x - t) + g(x + t).$$

Problem 18.22 **(a) For a perfect fluid in general relativity,**

$$T_{\mu\nu} = (\rho c^2 + P)U_\mu U_\nu + Pg_{\mu\nu}, \qquad (U^\mu U_\mu = -1)$$

show that the conservation identities $T^{\mu\nu}{}_{;\nu} = 0$ imply

$$\rho_{,\nu}U^\nu + \big(\rho + \frac{P}{c^2}\big)U^\nu_{;\nu} = 0,$$

$$(\rho c^2 + P)U^\mu_{;\nu}U^\nu + P_{,\nu}\big(g^{\mu\nu} + U^\mu U^\nu\big) = 0.$$

(b) For a pressure-free fluid show that the streamlines of the fluid (i.e. the curves $x^\mu(s)$ satisfying $\mathrm{d}x^\mu/\mathrm{d}s = U^\mu$) are geodesics, and ρU^μ is a covariant 4-current, $(\rho U^\mu)_{,\mu} = 0$.

(c) In the Newtonian approximation where

$$U_\mu = \Big(\frac{v_i}{c}, -1\Big) + O(\beta^2), \qquad P = O(\beta^2)\rho c^2, \qquad \Big(\beta = \frac{v}{c}\Big)$$

where $|\beta| \ll 1$ and $g_{\mu\nu} = \eta_{\mu\nu} + \epsilon h_{\mu\nu}$ with $\epsilon \ll 1$, show that

$$\epsilon h_{44} \approx -\frac{2\phi}{c^2}, \quad \epsilon h_{ij} \approx -\frac{2\phi}{c^2}\delta_{ij}, \qquad \textbf{where} \qquad \nabla^2\phi = 4\pi G\rho$$

and $h_{i4} = O(\beta)h_{44}$. Show in this approximation that the equations $T^{\mu\nu}{}_{;\nu} = 0$ approximate to

$$\frac{\partial\rho}{\partial t} + \nabla\cdot(\rho\mathbf{v}) = 0, \qquad \rho\frac{\mathrm{d}\mathbf{v}}{\mathrm{d}t} = -\nabla P - \rho\nabla\phi.$$

[NOTE: Corrections to the second, third and fifth displayed equations in this problem. Also note that the second displayed equation after Eq. (18.89) in the text should read

$$\epsilon\Box\varphi_{\mu\nu} = -2\kappa T_{\mu\nu} = -\frac{16\pi G}{c^4}T_{\mu\nu}. \qquad]$$

Solution: (a) The equation $T^{\mu\nu}{}_{;\nu} = 0$ reads

$$(\rho c^2 + P)_{,\nu}U^\mu U^\nu + (\rho c^2 + P)(U^\mu{}_{;\nu}U^\nu + U^\mu U^\nu{}_{;\nu}) + P_{,\nu}g^{\mu\nu} = 0. \tag{1}$$

Multiplying through by U_μ (summation convention throughout), and using $U_\mu U^\mu{}_{;\nu} = \frac{1}{2}(U_\mu U^\mu)_{,\nu} = 0$, we have

$$-(\rho c^2 + P)_{,\nu}U^\nu - (\rho c^2 + P)U^\nu{}_{;\nu} + P_{,\nu}U^\nu = 0,$$

i.e.

$$\rho_{,\nu}U^\nu + \big(\rho + \frac{P}{c^2}\big)U^\nu_{;\nu} = 0. \tag{2}$$

Substituting back into (1) gives

$$(\rho c^2 + P)U^\mu_{;\nu}U^\nu + P_{,\nu}\big(g^{\mu\nu} + U^\mu U^\nu\big) = 0. \tag{3}$$

(b) If $P = 0$ then Eq. (3) reads

$$\rho U^\mu_{;\nu}U^\nu = 0$$

and setting $U^\mu = \mathrm{d}x^\mu/\mathrm{d}s$ we have

$$\begin{aligned} 0 &= U^\mu_{;\nu} U^\nu \\ &= \frac{D}{\mathrm{d}s}\Big(\frac{\mathrm{d}x^\mu}{\mathrm{d}s}\Big) \\ &= \frac{\mathrm{d}^2 x^\mu}{\mathrm{d}s^2} + \Gamma^\mu_{\nu\rho}\frac{\mathrm{d}x^\nu}{\mathrm{d}s}\frac{\mathrm{d}x^\rho}{\mathrm{d}s} \end{aligned}$$

which means that the streamline curves curves $x^\mu(s)$ are geodesics. Eq. (2) can be written in 4-current form

$$\rho_{,\nu}U^\nu + \rho U^\nu_{;\nu} = \big(\rho U^\nu\big)_{;\nu} = 0.$$

(c) In the Newtonian approximation

$$U_\mu = \Big(\frac{v_i}{c}, -1\Big) + O(\beta^2), \qquad P = O(\beta^2)\rho c^2, \qquad \Big(\beta = \frac{v}{c}\Big)$$

and the tensor potential $\varphi_{\mu\nu} = h_{\mu\nu} - \frac{1}{2}h\eta_{\mu\nu}$ satisfies $\varphi^\alpha{}_{,\nu,\alpha} = 0$, and

$$\epsilon\Box\varphi_{\mu\nu} \approx \epsilon\nabla^2\varphi_{\mu\nu} = -2\kappa T_{\mu\nu} = -\frac{16\pi G}{c^4}\big[(\rho c^2 + P)U_\mu U_\nu + P\eta_{\mu\nu}\big]$$

i.e.,

$$\begin{aligned} \epsilon\nabla^2\varphi_{44} &= -\frac{16\pi G}{c^2}\rho \\ \epsilon\nabla^2\varphi_{i4} &= \frac{16\pi G}{c^3}\rho v_i \ll \epsilon\nabla^2\varphi_{44} \\ \epsilon\nabla^2\varphi_{ij} &= -\frac{16\pi G}{c^4}\big(\rho v_i v_j + P\delta_{ij}\big) \ll \epsilon\nabla^2\varphi_{44}. \end{aligned}$$

Now we can solve for $h_{\mu\nu}$,

$$h_{\mu\nu} = \varphi_{\mu\nu} - \tfrac{1}{2}\varphi\eta_{\mu\nu}, \qquad \varphi = \varphi^\mu{}_\mu = \varphi_{ii} - \varphi_{44}.$$

Since

$$\nabla^2\varphi = \nabla^2(\varphi_{ii} - \varphi_{44}) \approx -\nabla^2\varphi_{44} \approx \frac{16\pi G}{c^2}\rho$$

it follows that

$$\begin{aligned} \epsilon\nabla^2 h_{44} &\approx -\frac{8\pi G}{c^2}\rho \\ \epsilon\nabla^2 h_{i4} &\approx \frac{16\pi G}{c^3}\rho v_i \\ \epsilon\nabla^2 h_{ij} &\approx -\frac{16\pi G}{c^4}\rho\delta_{ij}. \end{aligned}$$

From Eq. (18.85), $\epsilon h_{44} = -2\phi/c^2$ where $\nabla^2\phi = 4\pi G\rho$, so that

$$\epsilon h_{44} \approx -\frac{2\phi}{c^2}, \qquad \epsilon h_{ij} \approx -\frac{2\phi}{c^2}\delta_{ij},$$

and $h_{i4} = O(\beta)h_{44}$.

For $P \ll \rho c^2$ $T^{4\nu}{}_{;\nu} = 0$ reduces

$$\rho_{,\nu}U^\nu + \rho U^\nu_{;\nu} \approx 0$$

and substituting the approximate components of U^μ,

$$\rho_{,i}\frac{v_i}{c} + \rho_{,4} + \rho\Big(\frac{v_{i,i}}{c} + \Gamma^\rho_{i\rho}\frac{v_i}{c}\Big) = 0.$$

Now

$$\Gamma^i_{jk} \approx O(h_{ij,k}) = O\Big(\frac{\phi_{,k}}{c^2}\Big)$$

and

$$\Gamma^4_{i4} \approx \tfrac{1}{2}g^{44}h_{4i,4} = -\tfrac{1}{2}h_{4i,4} \ll \frac{\phi_{,i}}{c^2}$$

so that

$$\Gamma^\rho_{i\rho}\frac{v_i}{c} \ll v_{i,i}.$$

Hence

$$\rho_{,i}v_i + \frac{\partial\rho}{\partial t} + \rho v_{i,i} = 0,$$

which can be written in standard 3-vector form

$$\frac{\partial\rho}{\partial t} + \nabla\cdot(\rho\mathbf{v}) = 0.$$

On neglecting the term P compared with ρc^2 and setting $\mu = i$ in Eq. (3), we have

$$\rho c^2\Big(\frac{v_{i,j}}{c}\frac{v_j}{c} + \Gamma^i_{44}\Big) + P_{,i} = 0.$$

Substituting

$$\Gamma^i_{44} = -\tfrac{1}{2}\epsilon h_{44,i} = \frac{\phi_{,i}}{c^2}$$

results in

$$\rho v_{i,j}v_j + \frac{\partial v_i}{\partial t} = -P_{,i} - \rho\phi_{,i}$$

which, in 3-vector notation, is

$$\rho\frac{\mathrm{d}\mathbf{v}}{\mathrm{d}t} = -\nabla P - \rho\nabla\phi.$$

Problem 18.23 **(a) Compute the components of the Ricci tensor $R_{\mu\nu}$ for a space-time that has a metric of the form**

$$\mathrm{d}s^2 = \mathrm{d}x^2 + \mathrm{d}y^2 - 2\,\mathrm{d}u\,\mathrm{d}v + 2H\,\mathrm{d}v^2 \qquad (H = H(x,y,u,v)).$$

(b) Show that the space-time is a vacuum if and only if $H = \alpha(x, y, v) + f(v)u$ where $f(v)$ is an arbitrary function and α satisfies the two-dimensional Laplace equation

$$\frac{\partial^2 \alpha}{\partial x^2} + \frac{\partial^2 \alpha}{\partial y^2} = 0,$$

and show that it is possible to set $f(v) = 0$ by a coordinate transformation $u' = ug(v),\ v' = h(v)$.

(c) Show that $R_{i4j4} = -H_{,ij}$ for $i, j = 1, 2$.

Solution: (a) There are two basic approaches to this problem. The Cartan method is to write

$$g = \varepsilon^1 \otimes \varepsilon^1 + \varepsilon^2 \otimes \varepsilon^2 + \varepsilon^3 \otimes \varepsilon^4 + \varepsilon^4 \otimes \varepsilon^3$$

where $\varepsilon^1 = \mathrm{d}x$, $\varepsilon^2 = \mathrm{d}y$, $\varepsilon^3 = \mathrm{d}v$, $\varepsilon^4 = -\mathrm{d}u + H\mathrm{d}v$, and use Cartan's structural formulae, with $g_{ij} =$ consts.
We will, however, adopt the more direct method here of calculating Christoffel symbols with coordinates set to be $x^1 = x$, $x^2 = y$, $x^3 = u$, $x^4 = v$,

$$[g_{\mu\nu}] = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 2H \end{pmatrix}, \qquad [g^{\mu\nu}] = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -2H & -1 \\ 0 & 0 & -1 & 0 \end{pmatrix}.$$

Then

$$\Gamma^1_{\mu\nu} = \tfrac{1}{2}(g_{1\mu,\nu} + g_{1\nu,\mu} - g_{\mu\nu,1}) \Longrightarrow \Gamma^1_{44} = -H_{,1}$$

and all other $\Gamma^1_{\mu\nu} = 0$. Similarly the only non-vanishing $\Gamma^2_{\mu\nu} = 0$ is

$$\Gamma^2_{44} = -H_{,2}.$$

From

$$\Gamma^3_{\mu\nu} = \tfrac{1}{2} g^{3\alpha}(g_{\alpha\mu,\nu} + g_{\alpha\nu,\mu} - g_{\mu\nu,\alpha})$$

we find

$$\Gamma^3_{44} = 2HH_{,3} - H_{,4}, \qquad \Gamma^3_{14} = -H_{,1}, \qquad \Gamma^3_{24} = -H_{,2}, \qquad \Gamma^3_{34} = -H_{,3},$$

and the only non-vanishing Christoffel symbol of the form $\Gamma^4_{\mu\nu}$ is

$$\Gamma^4_{44} = H_{,3}.$$

From these it is immediate that $\Gamma^\rho_{\mu\rho} = 0$ and the Ricci tensor is

$$R_{\mu\nu} = \Gamma^\rho_{\mu\nu,\rho} - \Gamma^\rho_{\mu\rho,\nu} + \Gamma^\alpha_{\mu\nu}\Gamma^\rho_{\alpha\rho} - \Gamma^\alpha_{\mu\rho}\Gamma^\rho_{\alpha\nu}$$
$$R_{\mu\nu} = \Gamma^\rho_{\mu\nu,\rho} - \Gamma^\alpha_{\mu\rho}\Gamma^\rho_{\alpha\nu}.$$

Thus

$$R_{11} = R_{22} = R_{33} = R_{12} = R_{13} = R_{23} = 0$$

and

$$R_{14} = \Gamma^3_{14,3} - \Gamma^3_{14}\Gamma^4_{34} = -H_{,13}$$
$$R_{24} = -H_{,23}$$
$$R_{34} = -H_{,33}$$
$$R_{44} = -H_{,11} - H_{,22} + 2HH_{,33}.$$

(b) The "two-dimensional vacuum equation" $R_{\mu\nu} = 0$ therefore gives

$$R_{14} = R_{24} = R_{34} = 0 \quad \Longrightarrow \quad H_{,3} \equiv \frac{\partial H}{\partial u} = f(x^4) = f(v).$$

Hence $H = f(v)u + \alpha(x, y, v)$. Substituting into the R_{44} equation, using $H_{33} \equiv \partial^2 H/\partial u^2 = 0$,

$$H_{,11} + H_{,22} = \frac{\partial^2 \alpha}{\partial x^2} + \frac{\partial^2 \alpha}{\partial y^2} = 0.$$

Try a coordinate transformation $u' = ug(v)$, $v' = h(v)$, or equivalently,

$$u = u'p(v'), \quad v = k(v').$$

the u, v terms in the metric transform as

$$-2\mathrm{d}u\mathrm{d}v + 2H\mathrm{d}v^2 = -2p\frac{\mathrm{d}k}{\mathrm{d}v'}\mathrm{d}u'\mathrm{d}v' + 2\left(H\frac{\mathrm{d}k}{\mathrm{d}v'} - u'\frac{\mathrm{d}p}{\mathrm{d}v'}\right)\frac{\mathrm{d}k}{\mathrm{d}v'}(\mathrm{d}v')^2.$$

Let us set

$$p\frac{\mathrm{d}k}{\mathrm{d}v'} = 1, \qquad \frac{\mathrm{d}p}{\mathrm{d}v'} = f(k(v')),$$

a pair of equations which may clearly be solved in principle since it amounts to solving the single second order d.e. for k:

$$\frac{\mathrm{d}^2 k}{\mathrm{d}v'^2} + f(k(v'))\left(\frac{\mathrm{d}k}{\mathrm{d}v'}\right)^2 = 0$$

followed by setting $p = 1/(\mathrm{d}k/\mathrm{d}v')$. We then have that the cofficient of $\mathrm{d}u\mathrm{d}v$ is still -2, while the linear term in u' in H has been eliminated. Thus $h \to H = \alpha(x, y, v')$ after this coordinate transformation.

(c) The Riemann tensor is

$$R^\mu{}_{\nu\rho\sigma} = \Gamma^\mu_{\nu\sigma,\rho} - \Gamma^\mu_{\nu\rho,\sigma} + \Gamma^\alpha_{\nu\sigma}\Gamma^\mu_{\alpha\rho} - \Gamma^\alpha_{\nu\rho}\Gamma^\mu_{\alpha\sigma}$$

so that

$$R^1{}_{414} = \mathbb{G}^1_{44,1} - \Gamma^1_{41,4} + \Gamma^\alpha_{44}\Gamma^1_{\alpha 1} - \Gamma^\alpha_{41}\Gamma^1_{\alpha 4} = -H_{,11}$$
$$R^1{}_{424} = \mathbb{G}^1_{44,2} - \Gamma^1_{42,4} + \Gamma^\alpha_{44}\Gamma^1_{\alpha 2} - \Gamma^\alpha_{42}\Gamma^1_{\alpha 4} = -H_{,12}$$
$$R^2{}_{424} = -H_{,22}$$

and for $i, j = 1, 2$

$$R_{i4j4} = g_{ik} R^k{}_{4j4} = -H_{,ij}.$$

Problem 18.24 **Show that a coordinate transformation $r = h(r')$ can be found such that the Schwarzschild solution has the form**

$$\mathrm{d}s^2 = -\mathrm{e}^{\mu(r')}\mathrm{d}t^2 + \mathrm{e}^{\nu(r')}(\mathrm{d}r'^2 + r'^2(\mathrm{d}\theta^2 + \sin^2\theta \mathrm{d}\phi^2)).$$

Evaluate the functions e^{μ} and e^{ν} explicitly.

Solution: Under a coordinate transformation $r = h(r')$ the Schwarzschild solution becomes

$$\mathrm{d}s^2 = -\Big(1 - \frac{2m}{r}\Big)c^2\mathrm{d}t^2 + \frac{1}{1 - 2m/h}\Big(\frac{\mathrm{d}h}{\mathrm{d}r'}\Big)^2 \mathrm{d}r'^2 + h^2(r')(\mathrm{d}\theta^2 + \sin^2\theta \mathrm{d}\phi^2)).$$

Fixing attention on the spatial parts of this metric the aim is to find a function h such that

$$\mathrm{e}^{\nu(r')} = \frac{h^2}{r'^2} = \frac{1}{1 - 2m/h}\Big(\frac{\mathrm{d}h}{\mathrm{d}r'}\Big)^2,$$

i.e.

$$\frac{1}{r'} = \frac{1}{\sqrt{h(h - 2m)}}\frac{\mathrm{d}h}{\mathrm{d}r'}.$$

Integrating, gives

$$\int \frac{\mathrm{d}r'}{r'} = \int \frac{\mathrm{d}h}{\sqrt{h(h-2m)}} = \int \frac{\mathrm{d}h}{\sqrt{(h-m)^2 - m^2}}$$

i.e.

$$\ln r' + a = \ln\big(\sqrt{(h-m)^2 - m^2}\big)$$

and exponentiating,

$$ar' = h - m + \sqrt{h(h-2m)}.$$

Hence

$$h = \frac{(ar' + m)^2}{2ar'}$$

and if we require that $h \to r'$ as $r' \to \infty$ then we must set $a = 2$, i.e.

$$r = h(r') = \frac{(2r' + m)^2}{4r'}$$

so that

$$\mathrm{e}^{\nu(r')} = \frac{h^2}{r'^2} = \Big(1 + \frac{m}{2r'}\Big)^4.$$

The function e^μ is found from

$$\begin{aligned} e^{\mu(r')} = -c^2 g_{44} &= -c^2\Big(1 - \frac{2m}{h}\Big) \\ &= -c^2\Big(1 - \frac{8mr'}{(r'+m)^2}\Big) \\ &= -c^2 \frac{(1 - m/2r')^2}{(1 + m/2r')^2} \end{aligned}$$

Problem 18.25 **Consider an oscillator at $r = r_0$ emitting a pulse of light (null geodesic) at $t = t_0$. If this is received by an observer at $r = r_1$ at $t = t_1$, show that**

$$t_1 = t_0 + \int_{r_0}^{r_1} \frac{dr}{c(1 - 2m/r)}.$$

By considering a signal emitted at $t_0 + \Delta t_0$, received at $t_1 + \Delta t_1$ (assuming the radial positions r_0 and r_1 to be constant) show that $\Delta t_0 = \Delta t_1$ and the gravitational redshift found by comparing *proper times* at emission and reception is given by

$$1 + z = \frac{\Delta\tau_1}{\Delta\tau_0} = \sqrt{\frac{1 - 2m/r_1}{1 - 2m/r_0}}.$$

Show that for two clocks at different heights h on the Earth's surface, this reduces to

$$z \approx \frac{GM}{c^2}\frac{h}{R^2}$$

where M and R are the mass and radius of the Earth.
[NOTE: The text statement $t_0 = t_1$ in line 5 should be replaced by $t_0 + \Delta t_0$, and the final displayed formula is incorrectly given in the text.]

Solution: For a radial null geodesic (light signal),

$$\frac{d\theta}{ds} = \frac{d\phi}{ds} = 0$$

we have

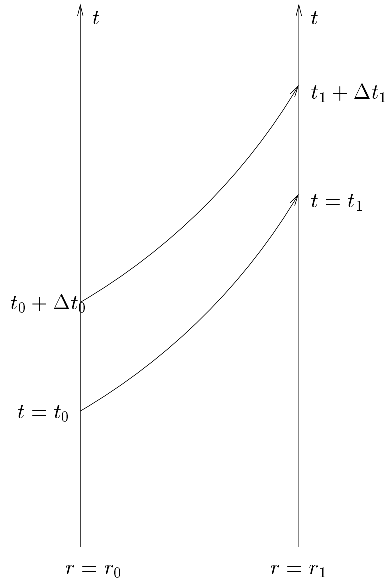
$$c^2\Big(\frac{dt}{ds}\Big)^2\Big(1 - \frac{2m}{r}\Big) - \frac{1}{1 - 2m/r}\Big(\frac{dr}{ds}\Big)^2 = 0.$$

Hence

$$\frac{dt}{dr} = \frac{1}{c(1 - 2m/r)}$$

which integrates to give

$$t = t_0 + \int_{r_0}^{r_1} \frac{dr}{c(1 - 2m/r)}.$$



For a second signal emitted at $t_0 + \Delta t_0$, the time of reception is

$$t_1 + \Delta t_1 = t_0 + \Delta t_0 + \int_{r_0}^{r_1} \frac{\mathrm{d}r}{c(1 - 2m/r)} = t_1 + \Delta t_0.$$

since $\Delta t_0 = \Delta t_1$. The proper time between the emission events at r_0 is given by

$$\left(\Delta\tau_0\right)^2 = \frac{1}{c^2}\Delta s^2 = \left(1 - \frac{2m}{r_0}\right)\left(\Delta t_0\right)^2,$$

while at reception the proper time difference $\Delta\tau_1$ is given by

$$\left(\Delta\tau_1\right)^2 = \left(1 - \frac{2m}{r_1}\right)\left(\Delta t_1\right)^2 = \left(1 - \frac{2m}{r_1}\right)\left(\Delta t_0\right)^2.$$

Hence the redshift is

$$1 + z = \frac{\Delta\tau_1}{\Delta\tau_0} = \sqrt{\frac{1 - 2m/r_1}{1 - 2m/r_0}}.$$

On the Earth's surface, set $r_0 = R$ and $r_1 = R + h$ where $h \ll R$. Noting that

$$m = \frac{GM}{c^2} \ll R$$

we have

$$\begin{aligned}1+z &= \sqrt{\frac{1-2m/r_1}{1-2m/r_0}} \\ &\approx -\frac{m}{r_1}+\frac{m}{r_0} \\ &= \frac{m}{R}-\frac{m}{R+h} \\ &\approx \frac{m}{R}\cdot\frac{h}{R} \\ &= \frac{MGh}{c^2R^2}.\end{aligned}$$

Problem 18.26 **In the Schwarzschild solution show the only possible closed photon path is a circular orbit at $r = 3m$, and show that it is unstable.**

Solution: From Problem 18.12 affinely parametrized geodesics may be calculated from the Lagrangian $L = g_{\mu\nu}\dot{x}^\mu\dot{x}^\nu$, the Euler-Lagrange equations reading

$$\frac{\mathrm{d}}{\mathrm{d}t}\Big(\frac{\partial L}{\partial \dot{x}^\mu}\Big) - \frac{\partial L}{\partial x^\mu} = \frac{\mathrm{d}}{\mathrm{d}t}(2g_{\mu\nu}\dot{x}^\nu) - g_{\nu\rho,\mu}\dot{x}^\nu\dot{x}^\rho = 0.$$

Since $g_{\mu\nu,4} = g_{\mu\nu,3} = 0$ (where $x^1 = r$, $x^2 = \theta$, $x^3 = \phi$) we have

$$\begin{aligned} g_{44}\frac{\mathrm{d}x^4}{\mathrm{d}s} &= \text{const.} \Longrightarrow \frac{\mathrm{d}x^4}{\mathrm{d}s} = \frac{k}{1-2m/r} \\ g_{33}\frac{\mathrm{d}x^3}{\mathrm{d}s} &= \text{const.} \Longrightarrow \frac{\mathrm{d}x^3}{\mathrm{d}s} = \frac{h}{r^2\sin\theta}, \end{aligned}$$

and the x^3-equation is

$$\frac{\mathrm{d}\theta}{\mathrm{d}s^2} + \frac{2}{r}\frac{\mathrm{d}\theta}{\mathrm{d}s}\frac{\mathrm{d}r}{\mathrm{d}s} - \sin\theta\cos\theta\Big(\frac{\mathrm{d}\phi}{\mathrm{d}s}\Big)^2 = 0 \qquad (1)$$

If the geodesic has initial conditions

$$\theta_0 = \frac{\pi}{2}, \qquad \frac{\mathrm{d}\theta}{\mathrm{d}s}\Big|_{s=0} = 0$$

then $\theta = \pi/2$ for all s is the unique solution of the differential equation (1) with these initial conditions. By a rotation of axis every geodesic may be assumed to be "equatorial", $\theta = \pi/2$.

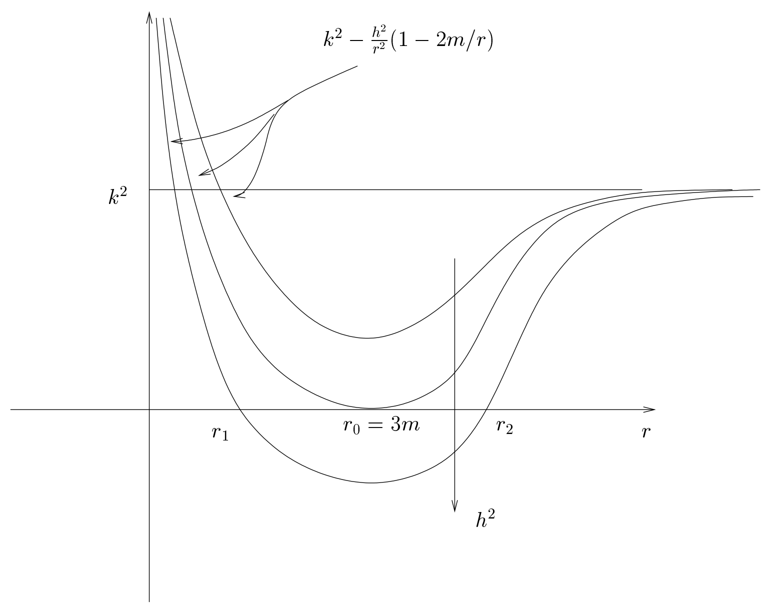
Instead of the r-equation, we may consider the null tangent condition:

$$\frac{1}{1-2m/r}\Big(\frac{\mathrm{d}r}{\mathrm{d}s}\Big)^2 + r^2\Big(\frac{\mathrm{d}\phi}{\mathrm{d}s}\Big)^2 - \Big(1-\frac{2m}{r}\Big)\Big(\frac{\mathrm{d}x^4}{\mathrm{d}s}\Big)^2 = 0,$$

i.e.

$$\left(\frac{\mathrm{d}r}{\mathrm{d}s}\right)^2 = k^2 - \frac{h^2}{r^2}\left(1 - \frac{2m}{r}\right) = r^{-3}f(r)$$

where $f(r)$ is a cubic, which must be positive throughout the orbit since $(\mathrm{d}r/\mathrm{d}s)^2 \geq 0$. The cubic can have at most two positive roots $r = r_1 > 0$, $r = r_2 > r_1$ as it has at least one negative root, for $f(0) = h^2 2m > 0$ and $f(-\infty) = -\infty$. Hence all



orbits either reach a maximum value $r = r_1$ and head towards $r = 0$, or lie between a minimum at $r = r_2$ and $r = \infty$. In both cases the orbits are clearly not closed. The only closed orbit occurs when there is a double root, $r_1 = r_2$, which occurs if

$$r^{-3}f(r) = \frac{\mathrm{d}}{\mathrm{d}r}(r^{-3}f((r)) = 0.$$

The second equation implies

$$\frac{2h^2}{r^3} - \frac{6mh^2}{r^4} = 0$$

i.e. $r = 3m$, clearly a circular orbit. As all other orbits (i.e. any small variations in parameters k and h) result in orbits which are not closed, this orbit is necessarily unstable.

Problem 18.27 **(a) A particle falls radially inwards from rest at infinity in a Schwarzschild solution. Show that it will arrive at $r = 2m$ in a finite *proper time* after crossing some fixed reference position r_0, but that coordinate time $t \to \infty$ as $r \to 2m$.**

(b) On an infalling extended body compute the tidal force in a radial

direction, by parallel propagating a tetrad (only the radial spacelike unit vector need be considered) and calculating R_{1414}.

(c) Estimate the total tidal force on a person of height 1.8 m, weighing 70 kg, falling headfirst into a solar mass black hole ($M_\odot = 2\times 10^{30}$ kg.), as he crosses $r = 2m$.

Solution: (a) As in Problem 18.26

$$\frac{\mathrm{d}x^4}{\mathrm{d}s} = \frac{k}{1-2m/r}$$

and for a radial timelike geodesic (freely falling particle with $\dot\theta = \dot\phi = 0$)

$$g_{\mu\nu}\frac{\mathrm{d}x^\mu}{\mathrm{d}s}\frac{\mathrm{d}x^\nu}{\mathrm{d}s} = -1 \quad\Longrightarrow\quad \left(\frac{\mathrm{d}r}{\mathrm{d}s}\right)^2 = k^2 - \left(1-\frac{2m}{r}\right).$$

If the particle is at rest at $r=\infty$ then $k=1$ (so that $\mathrm{d}r/\mathrm{d}s \to 0$ as $r\to\infty$),

$$\frac{\mathrm{d}r}{\mathrm{d}s} = \sqrt{\frac{2m}{r}}.$$

The proper time along the path is given by

$$\begin{aligned}\tau = \frac{s}{c} &= -\frac{1}{c}\int_{r_0}^{2m}\sqrt{\frac{2m}{r}}\mathrm{d}r\\ &= \frac{2}{3c\sqrt{2m}}\left(r_0^{3/2} - (2m)^{3/2}\right) < \infty.\end{aligned}$$

(b) The tangent vector to the infalling geodesic is

$$U^\mu = \left(-\sqrt{\frac{2m}{r}},\, 0,\, 0,\, \frac{1}{1-2m/r}\right), \qquad (U_\alpha U^\alpha = -1).$$

Let e^μ be the unit radial spacelike vector orthogonal to U^μ,

$$e^\mu = \left(-1,\, 0,\, 0,\, \frac{\sqrt{2m/r}}{1-2m/r}\right) \qquad (e_\alpha e^\alpha = 1, e_\alpha U^\alpha = 0).$$

This vector is parallel propagated along the geodesic, for

$$\begin{aligned} e_\mu e^\mu = 1 \quad &\Longrightarrow \quad 2e_\mu\frac{De^\mu}{\mathrm{d}s} = 0,\\ U_\mu e^\mu = 1 \quad &\Longrightarrow \quad U_\mu\frac{De^\mu}{\mathrm{d}s} = 0 \qquad \text{since } \frac{DU^\mu}{\mathrm{d}s} = 0,\end{aligned}$$

so that

$$\frac{De^1}{\mathrm{d}s} = \frac{De^4}{\mathrm{d}s} = 0.$$

Also

$$\frac{De^2}{\mathrm{d}s} = \frac{\mathrm{d}e^2}{\mathrm{d}s} + \Gamma^2_{\mu\nu} e^\mu U^\nu = 0$$

since $\Gamma^2_{11} = \Gamma^2_{14} = \Gamma^2_{44} = 0$, and similarly

$$\frac{De^3}{\mathrm{d}s} = 0.$$

Hence

$$\frac{De^\mu}{\mathrm{d}s} = \frac{\mathrm{d}e^\mu}{\mathrm{d}s} + \Gamma^\mu_{\nu\rho} e^\nu U^\rho = 0.$$

The tidal forces are found from Eq. (18.81),

$$\ddot{\eta}_i = K_{ij}\eta_i \qquad \text{where} \qquad K_{ij} = -c^2 R_{\mu\rho\nu\sigma} e^\mu_i e^\nu_j U^\rho U^\sigma$$

and $\dot{} \equiv \mathrm{d}/\mathrm{d}\tau$. In the radial direction

$$K_{11} = -c^2(e^1U^4 - e^4U^1)^2 R_{1414} = -c^2 R_{1414}$$

and from Eq, (18.97) and the field $R_{11} = 0$ we find from $\lambda = -\ln(1 - 2m/r)$

$$R_{1414} = \tfrac{1}{4}\mathrm{e}^{-\lambda}\left(4\frac{\lambda'}{r}\right) = -\frac{2m}{r^3}.$$

Hence $K_{11} = 2m/r^3$ and the tidal force per unit mass on an infalling body is

$$\ddot{\eta} = \frac{2mc^2}{r^3}\eta.$$

(c) As $r \to 2m$

$$\ddot{\eta} = \frac{c^2}{4m^2}\eta = \frac{c^6\eta}{4G^2M^2}.$$

If $\rho = m/l$ is the mass per unit lenght of the infalling person then the total tidal force is

$$F = \int_0^l \frac{\rho c^6 x}{4G^2M^2} dx = \frac{ml}{2}\frac{c^6}{4G^2M^2}.$$

Setting $m = 70$ kg, $l = 1.8$ m, $M = M_\odot = 2 \times 10^{30}$ kg, $G = 6.7 \times 10^{-11}$, $c = 3 \times 10^8 \mathrm{m\ s^{-1}}$, we have

$$F = 6 \times 10^1 1 \text{ Newtons} = 6 \times 10^7 \text{ tonnes}.$$

Problem 18.28 Show that for a closed Friedmann model of total mass M, the maximum radius is reached at $t = 2GM/3c^3$ where its value is $a_{\mathbf{max}} = 4GM/3\pi c^2$.

Solution: For a closed universe we have

$$a = \frac{8\pi G\rho_0}{3c^2}\alpha, \qquad t = \frac{8\pi G\rho_0}{3c^3}y$$

where

$$\alpha = \sin^2\eta, \qquad y = \eta - \sin\eta\cos\eta.$$

The maximum radius occurs at the maximum value of a, i.e. at at $\eta = \pi/2$ where $\alpha = 1$, $y = \pi/2$. Hence

$$t_{\max} = \frac{4\pi^2 G\rho_0}{3c^3}, \qquad a_{\max} = \frac{8\pi G\rho_0}{3c^2}$$

By Eq. (18.104) and (18.109) we have

$$M(t) = V(t)\rho(t) = 2\pi^2 a^3(t)\rho_0 a^{-3} = 2\pi^2\rho_0 = \text{const.}$$

Hence substituting $\rho_0 = M/2\pi^2$ into the above gives

$$t_{\max} = \frac{2GM}{3c^3}, \qquad a_{\max} = \frac{4GM}{3\pi c^3}.$$

Problem 18.29 **Show that the *radiation filled universe*, $P = \frac{1}{3}\rho c^2$ has $\rho \propto a^{-4}$ and the time evolution is for $k = 0$ is given by $a \propto t^{1/2}$. Assuming the radiation is black body, $\rho c^2 = a_S T^4$, where $a_S = 7.55 \times 10^{-15}$erg cm$^{-3}(^oK)^{-4}$, show that the temperature of the universe evolves with time as**

$$T = \left(\frac{3c^2}{32\pi G a_S}\right)^{1/4} t^{-1/2} = \frac{1.52\times 10^{10}}{\sqrt{t}}\ {}^oK \qquad (t \textbf{ in seconds}).$$

[NOTE: $P = \frac{1}{3}\rho c^2$. In text the factor c^2 was omitted. Similarly the factor c^2 was omitted in $\rho c^2 = a_S T^4$. A factor 10^10 was omitted on the RHS in the final equation.]

Solution: If $P = \frac{1}{3}\rho c^2$ then by Eq. (18.108)

$$\dot\rho + \frac{\dot a}{a}\rho = 0$$

which has solution $\rho = \rho_0 a^{-4}$ for some constant ρ_0. Wetting $k = 0$ in Eq. (18.106) gives

$$\frac{\dot a^2}{a^2} = \frac{8\pi G}{3}\rho_0 a^{-4}$$

whence

$$a\dot a = \frac{1}{2}\frac{\mathrm{d}a^2}{\mathrm{d}t} = \sqrt{\frac{8\pi G\rho_0}{3}}$$

so that, setting $a = 0$ at $t = 0$,

$$a^2 = \sqrt{\frac{8\pi G\rho_0}{3}}t$$

i.e.

$$a(t) = \left(\frac{8\pi G\rho_0}{3}\right)^{1/4} t^{1/2}.$$

Since

$$\begin{aligned} T^4 = \frac{\rho c^2}{a_S} &= \frac{\rho_0 c^2}{a_S} a^{-4} \\ &= \frac{3c^2}{32\pi G a_S t^2} \end{aligned}$$

we have

$$T = \left(\frac{3c^2}{32\pi G a_S}\right)^{1/4} t^{-1/2}$$

and setting, in c.g.s. units $c = 3 \times 10^{10}$, $G = 6.67 \times 10^{-8}$, and $a_S = 7.55 \times 10^{-15}$

$$T = \frac{1.52 \times 10^{10}}{\sqrt{t}}\ {}^oK.$$

where t is measured in seconds (i.e. after 1 second the temperature of the universe is $1.52 \times 10^{10}\ {}^o$ K.

Problem 18.30 Consider two radial light signals (null geodesics) received at the spatial origin of coordinates at times t_0 and $t_0 + \Delta t_0$, emitted from $\chi = \chi_1$ (or $r = r_1$ in the case of the flat models) at time $t = t_1 < t_0$. By comparing proper times between reception and emission show that the observer experiences a redshift in the case of an expanding universe ($a(t)$ increasing) given by

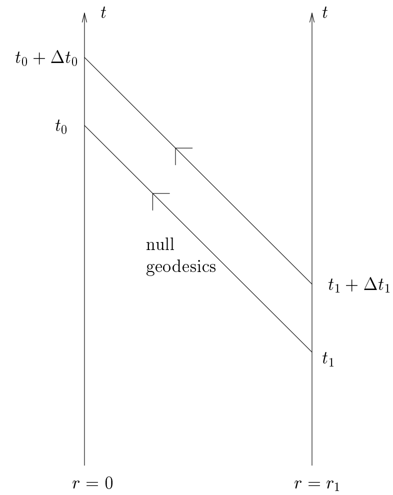
$$1 + z = \frac{\Delta t_0}{\Delta t_1} = \frac{a(t_0)}{a(t_1)}.$$

Solution: The Robertson-Walker line element is

$$\mathrm{d}s^2 = -c^2\mathrm{d}t^2 + a^2(t)(\mathrm{d}\chi^2 + f^2(\chi)(\mathrm{d}\theta^2 + \sin^2\theta\mathrm{d}\phi^2))$$

where $f(\chi) = \chi, \sin\chi, \sinh\chi$ depending on whether is $k = 0, 1, -1$ respectively. Along a null geodesic we have $\mathrm{d}s^2 = 0$, i.e.

$$c^2\left(\frac{\mathrm{d}t}{\mathrm{d}\chi}\right)^2 = a^2(t).$$



Integrating gives

$$\chi_1 = \int_0^{\chi_1} \mathrm{d}\chi = \int_{t_1}^{t_0} \frac{c\mathrm{d}t}{a(t)}$$

As this is independent of the emission and arrival times,

$$\begin{aligned}\int_{t_1}^{t_0} \frac{c\mathrm{d}t}{a(t)} &= \int_{t_1+\Delta t_1}^{t_0+\Delta t_0} \frac{c\mathrm{d}t}{a(t)} \\ &= \int_{t_1}^{t_0} \frac{c\mathrm{d}t}{a(t)} + \int_{t_0}^{t_0+\Delta t_0} \frac{c\mathrm{d}t}{a(t)} - \int_{t_1}^{t_1+\Delta t_1} \frac{c\mathrm{d}t}{a(t)}\end{aligned}$$

Hence

$$\int_{t_0}^{t_0+\Delta t_0} \frac{c\mathrm{d}t}{a(t)} = \int_{t_1}^{t_1+\Delta t_1} \frac{c\mathrm{d}t}{a(t)}$$

and in the limit $\Delta t_0, \Delta t_1 \to 0$,

$$\frac{c\Delta t_0}{a(t_0)} = \frac{c\Delta t_1}{a(t_1)}.$$

Thus, as Δt_0 and Δ_1 represent proper time differences at the respective events,

$$1 + z = \frac{\Delta t_0}{\Delta t_1} = \frac{a(t_0)}{a(t_1)}.$$

Problem 18.31 **By considering light signals as in the previous problem, show that an observer at $r = 0$, in the Einstein-de Sitter universe can at time $t = t_0$ see no events having radial distance $a(t_0)r > R_H = 3ct_0$. Show that the mass contained within this radius, called the particle horizon, is**

given by $M_H = 6c^3 t_0/G$.
[NOTE: The factor $a(t_0)$ in the definition of horizon distance R_H.]

Solution: In the Einstein-deSitter universe,

$$a(t) = At^{2/3} \qquad \text{where} \qquad A = (6\pi G\rho_0)^{1/3}$$

and the density is given by

$$\rho = \rho_0 a^{-3} = \frac{1}{6\pi G t^2}.$$

By Problem 18.30, if t_0 is the time of reception of a light signal emitted at r_1 at time t_1

$$\begin{aligned} r_1 = \int_{t_1}^{t_0} \frac{c\mathrm{d}t}{a(t)} &= \int_{t_1}^{t_0} \frac{c\mathrm{d}t}{At^{2/3}} \\ &= \frac{c}{A}\left[3t^{1/3}\right]_{t_1}^{t_0} \\ &= \frac{3c}{A}\left(t_0^{1/3} - t_1^{1/3}\right) \end{aligned}$$

Since $t_1 > 0$, it follows that $r_1 < 3ct_0^{1/3}/A$ and no signal can be received from events having coordinates

$$r > \frac{3ct_0^{1/3}}{A}.$$

i.e. having distance

$$a(t_0)r > R_H = a(t_0)\frac{3ct_0^{1/3}}{A} = 3Act_0^{1/3}\frac{3ct_0^{1/3}}{A} = 3ct_0.$$

The mass within this horizon distance is

$$\begin{aligned} M_H(t_0) &= \int_0^{r_H} \rho(t_0)a^3(t_0)4\pi r^2\mathrm{d}r \qquad \text{where } r_H = R_H/a(t_0) \\ &= \frac{4\pi}{3}\rho(t_0)a^3(t_0)(r_H)^3 \\ &= \frac{4\pi}{3}\rho(t_0)(R_H)^3 \\ &= \frac{4\pi}{3}\frac{27c^3t_0^3}{6\pi G t_0^2} \\ &= \frac{6c^3 t_0}{G}. \end{aligned}$$

Problem 18.32 **If a Lagrangian depends on second and higher order derivatives of the fields, $L = L(\Phi_A,\, \Phi_{A,\mu},\, \Phi_{A,\mu\nu}, \ldots)$ derive the generalized**

Euler-Lagrange equations

$$\frac{\delta L\sqrt{-g}}{\delta \Phi_A} \equiv \frac{\partial L\sqrt{-g}}{\partial \Phi_A} - \frac{\partial}{\partial x^\mu}\Big(\frac{\partial L\sqrt{-g}}{\partial \Phi_{A,\mu}}\Big) + \frac{\partial^2}{\partial x^\mu \partial x^\nu}\Big(\frac{\partial L\sqrt{-g}}{\partial \Phi_{A,\mu\nu}}\Big) - \cdots = 0.$$

Solution: If D is any region in a coordinate domain U,

$$\delta \int_D L(\Phi_A,\, \Phi_{A,\mu}, \Phi_{A,\mu\nu}, \ldots)\Omega$$
$$= \delta \int_{D\cap U} \frac{\partial L\sqrt{-g}}{\partial \Phi_A}\delta\Phi_A + \frac{\partial L\sqrt{-g}}{\partial \Phi_{A,\mu}}\delta\Phi_{A,\mu} + \frac{\partial L\sqrt{-g}}{\partial \Phi_{A,\mu\nu}}\delta\Phi_{A,\mu\nu} + \ldots \mathrm{d}^4x$$

Now

$$\frac{\partial L\sqrt{-g}}{\partial \Phi_{A,\mu}}\delta\Phi_{A,\mu} = \frac{\partial}{\partial x^\mu}\Big(\frac{\partial L\sqrt{-g}}{\partial \Phi_{A,\mu}}\delta\Phi_A\Big) - \frac{\partial}{\partial x^\mu}\Big(\frac{\partial L\sqrt{-g}}{\partial \Phi_{A,\mu}}\Big)\delta\Phi_A$$

and by the Gauss-Stokes theorem

$$\int_{D\cap U} \frac{\partial}{\partial x^\mu}\Big(\frac{\partial L\sqrt{-g}}{\partial \Phi_{A,\mu}}\delta\Phi_A\Big)\mathrm{d}^4x = \int_{\partial D\cap U} \frac{\partial L\sqrt{-g}}{\partial \Phi_{A,\mu}}\delta\Phi_A \mathrm{d}\sigma_\mu = 0$$

where $\delta\Phi_A = 0$ on the boundary of D and

$$\mathrm{d}\sigma_\mu = \frac{1}{3!}\epsilon_{\mu\nu\rho\sigma}\mathrm{d}x^\nu \wedge \mathrm{d}x^\rho \wedge \mathrm{d}x^\sigma.$$

Similarly

$$\begin{aligned}\frac{\partial L\sqrt{-g}}{\partial \Phi_{A,\mu\nu}}\delta\Phi_{A,\mu\nu} &= \frac{\partial}{\partial x^\nu}\Big(\frac{\partial L\sqrt{-g}}{\partial \Phi_{A,\mu\nu}}\delta\Phi_{A,\mu}\Big) - \frac{\partial}{\partial x^\nu}\Big(\frac{\partial L\sqrt{-g}}{\partial \Phi_{A,\mu\nu}}\Big)\delta\Phi_{A,\mu} \\ &= \frac{\partial}{\partial x^\nu}\Big(\frac{\partial L\sqrt{-g}}{\partial \Phi_{A,\mu\nu}}\delta\Phi_{A,\mu}\Big) - \frac{\partial^2}{\partial x^\mu\partial x^\nu}\Big(\frac{\partial L\sqrt{-g}}{\partial \Phi_{A,\mu\nu}}\delta\Phi_A\Big) \\ &\quad + \frac{\partial^2}{\partial x^\mu\partial x^\nu}\Big(\frac{\partial L\sqrt{-g}}{\partial \Phi_{A,\mu\nu}}\Big)\delta\Phi_A.\end{aligned}$$

By Stokes' theorem the first two terms have vanishing integrals over D, and similar arguments hold for higher derivatives, leaving

$$\delta \int_D L\Omega = \int_D \frac{\delta L\sqrt{-g}}{\delta \Phi_A}\delta\Phi_A\Omega$$

where

$$\frac{\delta L\sqrt{-g}}{\delta \Phi_A} = \frac{\partial L\sqrt{-g}}{\partial \Phi_A} - \frac{\partial}{\partial x^\mu}\Big(\frac{\partial L\sqrt{-g}}{\partial \Phi_{A,\mu}}\Big) + \frac{\partial^2}{\partial x^\mu \partial x^\nu}\Big(\frac{\partial L\sqrt{-g}}{\partial \Phi_{A,\mu\nu}}\Big) - \cdots.$$

Since $\delta\Phi_A$ is arbitrary over D we deduce that

$$\frac{\delta L\sqrt{-g}}{\delta \Phi_A} = 0.$$

Problem 18.33 **For a skew symmetric tensor $F^{\mu\nu}$ show that**

$$F^{\mu\nu}{}_{;\nu} = \frac{1}{\sqrt{-g}}\left(\sqrt{-g}F^{\mu\nu}\right)_{,\nu}.$$

Solution: Using the formula for general covariant derivative Eq. (18.14), and the fact that the connection is symmetric $\Gamma^{\mu}_{\nu\rho} = \Gamma^{\mu}_{\rho\nu}$

$$\begin{aligned}
F^{\mu\nu}{}_{;\nu} &= F^{\mu\nu}{}_{,\nu} + \Gamma^{\mu}_{\alpha\nu}F^{\alpha\nu} + \Gamma^{\nu}_{\alpha\nu}F^{\mu\alpha} \\
&= F^{\mu\nu}{}_{,\nu} + \Gamma^{\nu}_{\alpha\nu}F^{\mu\alpha} \\
&= F^{\mu\nu}{}_{,\nu} + \frac{1}{\sqrt{-g}}\left(\sqrt{-g}\right)_{,\alpha}F^{\mu\alpha} \qquad \text{using Eq. (18.119)} \\
&= \frac{1}{\sqrt{-g}}\left(\sqrt{-g}F^{\mu\nu}\right)_{,\nu}.
\end{aligned}$$

Problem 18.34 **Compute the Euler-Lagrange equations and energy-stress tensor for a scalar field Lagrangian in general relativity given by**

$$L_S = -\psi_{,\mu}\psi_{,\nu}g^{\mu\nu} - m^2\psi^2.$$

Verify $T^{\mu\nu}{}_{;\nu} = 0$.

Solution: We can work from the Euler-Lagrange equations (18.113) with $\Phi_A = \psi$ or directly from the total action principle:

$$\begin{aligned}
&\delta\int_D (-\psi_{,\mu}\psi_{,\nu}g^{\mu\nu} - m^2\psi^2)\Omega \\
&\qquad = \int_{D\cap U} (-2\psi_{,\mu}\delta\psi_{,\nu}g^{\mu\nu} - 2m^2\psi\delta\psi)\sqrt{-g} \\
&\qquad\qquad - \left[\psi_{,\mu}\psi_{,\nu}\delta g^{\mu\nu} + (\psi_{,\alpha}\psi_{,\alpha} + m^2\psi^2)\delta\sqrt{-g}\right]\mathrm{d}^4x.
\end{aligned}$$

The terms involving $\delta\psi$ can be written

$$\int_{D\cap U} -\left(2\psi_{,\mu}g^{\mu\nu}\sqrt{-g}\right)_{,\nu} + 2\left[\left(\psi_{,\mu}g^{\mu\nu}\sqrt{-g}\right)_{,\nu} - m^2\psi\sqrt{-g}\right]\delta\psi\mathrm{d}^4x.$$

The first term vanishes by Stokes' theorem and $\delta\psi = 0$ on ∂D, leaving, since $\delta\psi$ is arbitrary in D:

$$\left(\psi_{,\mu}g^{\mu\nu}\sqrt{-g}\right)_{,\nu} - m^2\psi\sqrt{-g} = 0.$$

Using

$$\Box\psi \equiv \psi^{;\mu}{}_{;\mu} = \frac{1}{\sqrt{-g}}\left(\sqrt{-g}g^{\mu\nu}\psi_{,\nu}\right)_{,\mu}$$

this equation can be written

$$\Box\psi - m^2\psi = 0,$$

known as the *covariant Klein-Gordon equation.*

Using Eq. (18.117)

$$\delta\sqrt{-g} = -\frac{\sqrt{-g}}{2} g_{\mu\nu}\delta g^{\mu\nu},$$

we have the coefficent of $-\frac{1}{2}T_{\mu\nu}\delta g^{\mu\nu}$ in the above expansion of $\delta\int L$,

$$T_{\mu\nu} = 2\psi_{,\mu}\psi_{,\nu} - g_{\mu\nu}\big(\psi_{,\alpha}\psi_{,\alpha} + m^2\psi^2\big).$$

Thus

$$\begin{aligned} T^{\mu\nu}{}_{;\nu} &= 2\psi^{;\mu}{}_{;\nu}\psi^{;\nu} + 2\psi^{;\mu}\psi^{;\nu}{}_{;\nu} - g^{\mu\nu}\big(2\psi^{;\alpha}{}_{;\nu}\psi_{;\alpha} + 2m^2\psi\psi_{,\nu}\big) \\ &= 2\psi^{;\mu}\big(\psi^{;\nu}{}_{;\nu} - m^2\psi\big) \end{aligned}$$

since

$$\psi^{;\alpha}{}_{;\nu} = g^{\alpha\beta}\psi_{;\beta\nu} = g^{\alpha\beta}\psi_{;\nu\beta} = \psi_{;\nu}{}^{;\alpha}.$$

Hence, using the covariant Klein-Gordon equation $T^{\mu\nu}{}_{;\nu} = 0$.

Problem 18.35 Prove the implication given in Eq. (18.127). Show that this equation and Eq. (18.126) imply $T^{\mu\nu}{}_{;\nu} = 0$ for the electromagnetic energy-stress tensor given in Eq. (18.125).

Solution: If

$$F_{\mu\nu,\rho} + F_{\nu\rho,\mu} + F_{\rho\mu,\nu} = 0$$

then, using $\Gamma^\alpha_{\mu\nu} = \Gamma^\alpha_{\nu\mu}$, we have

$$\begin{aligned} F_{\mu\nu;\rho} &+ F_{\nu\rho;\mu} + F_{\rho\mu;\nu} = \\ &F_{\mu\nu,\rho} - \Gamma^\alpha_{\mu\rho}F_{\alpha\nu} - \Gamma^\alpha_{\nu\rho}F_{\mu\alpha} \\ &+ F_{\nu\rho,\mu} - \Gamma^\alpha_{\rho\mu}F_{\nu\alpha} \\ &+ F_{\rho\mu,\nu} - \Gamma^\alpha_{\rho\nu}F_{\alpha\mu} - \Gamma^\alpha_{\mu\nu}F_{\rho\alpha} \\ &= -\Gamma^\alpha_{\mu\rho}(F_{\alpha\nu} + F_{\nu\alpha}) - \Gamma^\alpha_{\nu\rho}(F_{\mu\alpha} + F_{\alpha\mu}) - \Gamma^\alpha_{\nu\mu}(F_{\alpha\rho} + F_{\rho\alpha}) \\ &= 0 \end{aligned}$$

Eq. (18.125) should read

$$T_{\mu\nu} = \frac{1}{4\pi}\big(F_{\mu\alpha}F_\nu{}^\alpha - \frac{1}{4}F_{\alpha\beta}F^{\alpha\beta}g_{\mu\nu}\big),$$

and using $g^{\mu\nu}{}_{;\nu} = 0$ we have

$$
\begin{aligned}
T^{\mu\nu}{}_{;\nu} &= \frac{1}{4\pi}\big(F^{\mu}{}_{\alpha}F^{\nu\alpha} - \tfrac{1}{4}F_{\alpha\beta}F^{\alpha\beta}g^{\mu\nu}\big)_{;\nu} \\
&= \frac{1}{4\pi}\big(F^{\mu}{}_{\alpha;\nu}F^{\nu\alpha} + F^{\mu}{}_{\alpha}F^{\nu\alpha}{}_{;\nu} - \tfrac{1}{2}F_{\alpha\beta;\nu}F^{\alpha\beta}g^{\mu\nu}\big) \\
&= \frac{1}{4\pi}g^{\mu\nu}\big(F_{\nu\alpha;\beta}F^{\beta\alpha} - \frac{1}{2}F_{\alpha\beta;\nu}F^{\alpha\beta}\big)
\end{aligned}
$$

after a relabelling of summation index in first term and using Eq. (18.126),

$$
\begin{aligned}
&= \frac{1}{4\pi}g^{\mu\nu}F^{\beta\alpha}\big(\tfrac{1}{2}F_{\nu\alpha;\beta} - \tfrac{1}{2}F_{\nu\beta;\alpha} + \tfrac{1}{2}F_{\alpha\beta;\nu}\big) \\
&= \frac{1}{8\pi}g^{\mu\nu}F^{\beta\alpha}\big(F_{\nu\alpha;\beta} + F_{\beta\nu;\alpha} + F_{\alpha\beta;\nu}\big) \\
&= 0.
\end{aligned}
$$

Chapter 19

Problem 19.1 **Let $\mathsf{E}^i{}_j$ be the matrix whose (j,i)th component is 1 and all other components vanish. Show that these matrices form a basis of $\mathcal{GL}(n,\mathbb{R})$, and have the commutator relations**

$$[\mathsf{E}^i{}_j, \mathsf{E}^k{}_l] = \delta^i{}_l \mathsf{E}^k{}_j - \delta^k{}_j \mathsf{E}^i{}_l.$$

Write out the structure constants with respect to this algebra in this basis.
[NOTE: (j,i)th component is 1 , not (i,j)th as in text.]

Solution: From Example 19.4 we find $\mathcal{GL}(n,\mathbb{R}) \cong M_n(\mathbb{R})$, where commutators are matrix product commutators,

$$[\mathsf{A}, \mathsf{B}]^i{}_j = (\mathsf{AB} - \mathsf{BA})]^i{}_j = A^i{}_k B^k{}_j - B^i{}_k A^k{}_j.$$

Let $\mathsf{E}^i{}_j$ be the matrices whose (a,b)th component is

$$\left(\mathsf{E}^i{}_j\right)^a{}_b = \delta^i_b \delta^a_j,$$

i.e. the (j,i)th component is 1 $(a = j, b = i)$, and all others are 0. These are obviously linearly independent and n^2 in number. They therefore form a basis of $M_n(\mathbb{R})$, as every matrix A can be written uniquely as $\mathsf{A} = A^j{}_i \mathsf{E}^i{}_j$ for

$$A^j{}_i \left(\mathsf{E}^i{}_j\right)^a{}_b = A^j{}_i \delta^i_b \delta^a_j = A^a_b.$$

We have commutators

$$\begin{aligned}
[\mathsf{E}^i{}_j, \mathsf{E}^k_l]^a_b &= \left(\mathsf{E}^i_j \left(\mathsf{E}^k_l\right)^a{}_b - \left(\mathsf{E}^k_l \mathsf{E}^i{}_j\right)^a{}_b \right. \\
&= \left(\mathsf{E}^i{}_j\right)^a{}_c \left(\mathsf{E}^k_l\right)^c{}_b - \left(\mathsf{E}^k_l\right)^a{}_c \left(\mathsf{E}^i{}_j\right)^c{}_b \\
&= \delta^a_j \delta^i{}_c \delta^c_l \delta^k_b - \delta^k_c \delta^a_l \delta^c_j \delta^i_b \\
&= \delta^a_j \delta^i_l \delta^k_b - \delta^a_l \delta^k_j \delta^i_b \\
&= \delta^i_l \left(\mathsf{E}^k_j\right)^a{}_b - \delta^k_j \left(\mathsf{E}^i{}_l\right)^a{}_b
\end{aligned}$$

Hence we can write

$$[\mathsf{E}^i{}_j, \mathsf{E}^k_l] = \sum_{p,q} C^{pq}_{ij,kl} E^p{}_q$$

where

$$C^{pq}_{ij,kl} = \delta^i{}_l \delta^k{}_p \delta^q_j - \delta^k_j \delta^i{}_p \delta^q{}_l.$$

Problem 19.2 **Let $\mathsf{E}^i{}_j$ be the matrix defined as in the previous problem, and $\mathsf{F}^i{}_j = i\mathsf{E}^i{}_j$ where $i = \sqrt{-1}$. Show that these matrices form a basis of**

$\mathcal{GL}(n,\mathbb{C})$, and write all the commutator relations between these generators of $\mathcal{GL}(n,\mathbb{C})$.

Solution: The Lie group $GL(n,\mathbb{C})$ can be regarded as a real Lie group on an open submanifold of $\mathbb{R}^{2n}$, since the with coordinates real and imaginary parts of the components of the non-singular complex matrices. As in Example 19.4 all left invariant vector fields can be written

$$X = z^i{}_k A^k{}_j \frac{\partial}{\partial z^i{}_j} \qquad \text{where} \qquad z^i{}_k = x^i{}_k + i y^i{}_k$$

and $[X,Y] \longrightarrow [\mathsf{A},\mathsf{B}]$, so the Lie algebra is the complex Lie algebra $\mathcal{GL}(n,\mathbb{C}) \cong M_n(\mathbb{C})$, with matrix commutators. This may be considered to be a real Lie algebra $\mathcal{GL}(n,\mathbb{C})^R$ using the technique in Example 6.13 by regarded A and $i\mathsf{A}$ as linearly independent matrices and only allowing multiplication by real scalars. Setting $\mathsf{A} = \mathsf{A}_1 + i\mathsf{A}_2$ where A_1 and A_2 are real matrices, the bracket product is taken to be

$$[\mathsf{A},\mathsf{B}] = [\mathsf{A}_1 + i\mathsf{A}_2, \mathsf{B}_1 + i\mathsf{B}_2] = [\mathsf{A}_1,\mathsf{B}_1] - [\mathsf{A}_2,\mathsf{B}_2] + i([\mathsf{A}_1,\mathsf{B}_2 1] + [\mathsf{A}_2,\mathsf{B}_1]).$$

The elements of $GL(n,\mathbb{C})$ are spanned by $\mathsf{E}^i{}_j$ and $\mathsf{F}^i{}_j = i\mathsf{E}^i{}_j$ for

$$\mathsf{A} = (A_1{}^i{}_j + iA_2{}^i{}_j)\mathsf{E}^j{}_i = A_1{}^i{}_j \mathsf{E}^j{}_i + A_2{}^i{}_j \mathsf{F}^j{}_i$$

and the matrices $\mathsf{E}^i{}_j$ and $\mathsf{F}^i{}_j$ clearly form an l.i. set. They are therefore a basis of $GL(n,\mathbb{C})$. Commutators are

$$\begin{aligned}
[\mathsf{E}^i{}_j, \mathsf{E}^k_l] &= \delta^i_l \mathsf{E}^k_j - \delta^k_j \mathsf{E}^i{}_l \\
[\mathsf{F}^i{}_j, \mathsf{F}^k_l] &= -\delta^i_l \mathsf{E}^k_j + \delta^k_j \mathsf{E}^i{}_l \\
[\mathsf{E}^i{}_j, \mathsf{F}^k_l] &= \delta^i_l \mathsf{F}^k_j - \delta^k_j \mathsf{F}^i{}_l
\end{aligned}$$

Problem 19.3 **Define the $T_e(G)$-valued 1-form θ on a Lie group G, by setting**

$$\theta_g(X_g) = L_{g^{-1}*} X_g$$

for any vector field X on G (not necessarily left-invariant). Show that θ is left-invariant, $L_a^* \theta_g = \theta_{a^{-1}g}$ for all $a, g \in G$.

With respect to a basis E_i of left-invariant vector fields and its dual basis ε^i, show that

$$\theta = \sum_{i=1}^{n} \big(E_i\big)_e \varepsilon^i.$$

Solution: For any vector field X

$$\begin{aligned}\theta_{a^{-1}g}(X_{a^{-1}g}) &= L_{g^{-1}a*}X_{a^{-1}g} \\ &= (L_{g^{-1}} \circ L_a)_* X_{a^{-1}g} \\ &= L_{g^{-1}*}L_{a*}X_{a^{-1}g} \\ &= \theta_g(L_{a*}X_{a^{-1}g}) \\ &= L_{a*}\theta_g(X_{a^{-1}g})\end{aligned}$$

Since X is an arbitrary vector field $\theta_{a^{-1}g} = L_a^*\theta_g$, i.e. θ is left-invariant.

If E_i is a basis of left-invariant vector fields with dual basis ε^i, then setting $X_g = X^1 = (E_i)_g$ we have $X^i = \langle \varepsilon^i, X\rangle_g$. Hence (adopting summation convention)

$$\begin{aligned}\theta_g(X_g) &= L_{g^{-1}*}X_g \\ &= X^i L_{g^{-1}*}(E_i)_g \\ &= X^i (E_i)_e \\ &= \langle \varepsilon^i, X\rangle_g (E_i)_e \\ &= (E_i)_e \varepsilon^i(X_g)\end{aligned}$$

Since X is an arbitrary vector field, $\theta = \big(E_i\big)_e \varepsilon^i$.

Problem 19.4 **A function $f : G \to \mathbb{R}$ is said to be an analytic function on G if it can be expanded as a Taylor series at any point $g \in G$. Show that if X is a left-invariant vector field and f is an analytic function on G then**

$$f(g \exp tX) = \big(\mathrm{e}^{tX} f\big)(g)$$

where, for any vector field Y, we define

$$\mathrm{e}^Y f = f + Yf + \frac{1}{2!}Y^2 f + \frac{1}{3!}Y^3 f + \cdots = \sum_{i=0}^{\infty} \frac{Y^n}{n!} f.$$

The operator Y^n is defined inductively by $Y^n f = Y(Y^{n-1} f)$.

Solution: Set

$$F(t) = f(g \exp tX) = f(g\gamma(t))$$

where $\gamma(t) = \exp tX$ is the one-parameter group generated by X. Since the components of $\gamma(t)$ are analytic functions at $t = 0$ we have $F(t)$ is an analytic function of t,

$$F(t) = F(0) + t\dot{F}(0) + \tfrac{1}{2}t^2\ddot{F}(0) + \dots$$

The left invariant vector field $X = L_{g*}X$ is tangent to the curve $g\gamma(t)$ through g, since by Eq. (19.8)

$$Xf(\exp tX) = \frac{\mathrm{d}}{\mathrm{d}t} f(\exp tX)$$

so that for any differentiable function f at g

$$Xf(g\exp tX) = Xf \circ L_g(\exp tX) = \frac{\mathrm{d}}{\mathrm{d}t} f \circ L_g(\exp tX) = \frac{\mathrm{d}}{\mathrm{d}t} f(g\gamma(t)).$$

Hence

$$\begin{aligned}
\dot{F}(t) &= \frac{\mathrm{d}}{\mathrm{d}t}F(t) = (Xf)(g\gamma(t)), \\
\ddot{F}(t) &= \frac{\mathrm{d}^2}{\mathrm{d}t^2}F(t) = \frac{\mathrm{d}}{\mathrm{d}t}(Xf)(g\gamma(t)) = X^2 f(g\gamma(t)), \\
&\dots \\
F^{(n)}(t) &= \frac{\mathrm{d}^n}{\mathrm{d}t^n}F(t) = X^n f(g\gamma(t)).
\end{aligned}$$

In a neighbourhood of t then

$$\begin{aligned}
F(t) &= \sum_{n=0}^{\infty} \frac{t^n}{n!} \frac{\mathrm{d}^n}{\mathrm{d}t^n} F\Big|_{t=0} \\
&= f(g) + tXf(g) + \tfrac{1}{2}t^2 X^2 f(g) + \dots \\
&= \sum_{i=0}^{\infty} \Big[\frac{(tX)^n}{n!} f\Big](g),
\end{aligned}$$

i.e.

$$f(g\exp tX) = \big(\mathrm{e}^{tX} f\big)(g).$$

Problem 19.5 **Show that** $\exp tX \exp tY = \exp t(X+Y) + O(t^2)$

Solution: From Problem 19.4, for an arbitrary analytic function f,

$$\begin{aligned}
f(g\exp tX \exp tY) &= \mathrm{e}^{tY} f(g\exp tX) \\
&= \mathrm{e}^{tX}\mathrm{e}^{tY} f(g) \\
&= (1 + tX + O(t^2))(1 + tY + O(t^2))f(g) \\
&= (1 + tX + O(t^2))(f(g) + tYf(g) + O(t^2)) \\
&= f(g) + t(X+Y)f(g) + O(t^2) \\
&= \mathrm{e}^{t(X+Y)} f(g) + O(t^2) \\
&= f(g\exp t(X+Y)) + O(t^2).
\end{aligned}$$

It is possible to go further and find the $O(t^2)$ term as follows:

$$
\begin{aligned}
&f(g\exp tX\exp tY)\\
&\qquad = (1+tX+\tfrac{1}{2}t^2X^2+O(t^3))(f(g)+tYf(g)+\tfrac{1}{2}t^2Y^2f(g)+O(t^3))\\
&\qquad = f(g)+t(X+Y)f(g)+\tfrac{1}{2}t^2(X^2+2XY+Y^2)f(g)+O(t^3)\\
&\qquad = f(g)+t(X+Y)f(g)+\tfrac{1}{2}t^2((X+Y)^2+[X,Y])f(g)+O(t^3)\\
&\qquad = \mathrm{e}^{t(X+Y)+\frac{1}{2}t^2[X,Y]}f(g)+O(t^3)\\
&\qquad = f(g\exp(t(X+Y)+\tfrac{1}{2}t^2[X,Y]+O(t^3))
\end{aligned}
$$

Hence

$$\exp tX\exp tY = \exp\big(t(X+Y)+\tfrac{1}{2}t^2[X,Y]+O(t^3)\big).$$

Problem 19.6 **For any $n\times n$ matrix A, show that**

$$\frac{\mathrm{d}}{\mathrm{d}t}\det \mathrm{e}^{t\mathsf{A}}\Big|_{t=0} = \operatorname{tr}\mathsf{A}.$$

Solution: From Eq. (19.11)

$$\frac{\mathrm{d}}{\mathrm{d}t}\det \mathrm{e}^{t\mathsf{A}}\Big|_{t=0} = \frac{\mathrm{d}}{\mathrm{d}t}\mathrm{e}^{\operatorname{tr}t\mathsf{A}}\Big|_{t=0} = \frac{\mathrm{d}}{\mathrm{d}t}\mathrm{e}^{t\operatorname{tr}\mathsf{A}}\Big|_{t=0} = \operatorname{tr}\mathsf{A}$$

or more directly, by expansion of determinant at $t=0$:

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\det \mathrm{e}^{t\mathsf{A}}\Big|_{t=0} &= \frac{\mathrm{d}}{\mathrm{d}t}\big(\det(\mathsf{I}+t\mathsf{A}+\tfrac{1}{2}t^2\mathsf{A}^2+\dots)\big)\big|_{t=0}\\
&= \frac{\mathrm{d}}{\mathrm{d}t}\big(1+t\operatorname{tr}\mathsf{A}+O(t^2)\big)\big|_{t=0}\\
&= \operatorname{tr}\mathsf{A}.
\end{aligned}
$$

Problem 19.7 **Prove Eq. (19.11). One method is to find a matrix S that transforms A to upper-triangular Jordan form by a similarity transformation as in Section 4.2, and use the fact that both determinant and trace are invariant under such transformations.**

Solution: Under a similarity transformation $\mathsf{A}' = \mathsf{SAS}^{-1}$,

$$\operatorname{tr}\mathsf{A}' = \operatorname{tr}\mathsf{SAS}^{-1} = \operatorname{tr}\mathsf{S}^{-1}\mathsf{SA} == \operatorname{tr}\mathsf{A}.$$

and

$$\det\mathsf{A}' = \det\mathsf{S}\det\mathsf{A}(\det\mathsf{S})^{-1} = \det\mathsf{A}.$$

All powers of A and A' are related by the same similarity transformation,

$$\mathsf{A}'^2 = \mathsf{SAS}^{-1}\mathsf{SAS}^{-1} = \mathsf{SA}^2\mathsf{S}^{-1}$$

etc. Inductively we find that

$$\mathsf{A}'^n = \mathsf{SAS}^{-1}\mathsf{SA}^{n-1}\mathsf{S}^{-1} = \mathsf{SA}^n\mathsf{S}^{-1}.$$

Hence

$$e^{\mathsf{A}'} = \mathsf{I} + \mathsf{A}' + \tfrac{1}{2}\mathsf{A}'^2 + \cdots = \mathsf{S}e^{\mathsf{A}}\mathsf{S}^{-1}.$$

In Section 4.2 it was shown that we can find S (possibly a complex matrix) such that A' is an upper triangular matrix,

$$\mathsf{A}' = \begin{pmatrix} a'_{11} & \dots & & & \dots \\ 0 & a'_{22} & \dots & & \\ \dots & \dots & & & \\ 0 & 0 & 0 & \dots & a'_{nn} \end{pmatrix}.$$

Then A'^n is upper triangular for all n and $\left(\mathsf{A}'^n\right)_{ii} = (a'_{ii})^n$. Hence

$$\left(e^{\mathsf{A}'}\right)_{ii} = e^{a'_{ii}}.$$

Thus

$$\begin{aligned} \det e^{\mathsf{A}} &= \det\left(\mathsf{S}e^{\mathsf{A}}\mathsf{S}^{-1}\right) \\ &= \det e^{\mathsf{A}'} \\ &= e^{a'_{11}}e^{a'_{22}}\dots e^{a'_{nn}} \\ &= e^{a'_{11}} + a'_{22} + \cdots + a'_{nn} \\ &= e^{\operatorname{tr}\mathsf{A}'} \\ &= e^{\operatorname{tr}\mathsf{A}}. \end{aligned}$$

Problem 19.8 Show that $GL(n,\mathbb{C})$ and $SL(n,\mathbb{C})$ are connected Lie groups. Is $U(n)$ a connected group?

Solution: By the Jordan canonical form (see Section 4.2), for every $\mathsf{A} \in GL(n,\mathbb{C})$ there exists $\mathsf{S} \in GL(n,\mathbb{C})$ such that $\mathsf{B} = \mathsf{SAS}^{-1}$ only has non-zero components in the diagonal and superdiagonal. In fact, since $\det\mathsf{B}\det\mathsf{A} \neq 0$ we may assume $\mathsf{B} = [b_{ij}]$ where

$$b_{ii} \neq 0 \quad (i = 1,\dots,n), \qquad b_{i\,i+1} = 0 \text{ or } 1 \quad (i = 1,\dots,n-1).$$

Let $\gamma_1 : [0,1] \to GL(n,\mathbb{C})$ be the curve $\gamma_1(t) = \mathsf{G}(t) = [g_{ij}(t)]$ where

$$\begin{aligned} g_{ij}(t) &= b_{ij} && \text{if } j \neq i+1 \\ g_{i\,i+1}(t) &= (1-t)b_{i\,i+1} && \text{for } i = 1,\dots,n-1. \end{aligned}$$

The result is that $\mathsf{G}(1)$ is the diagonal matrix formed by setting all off-diagonal terms in B to zero. Now consider the curve $\gamma_2 : [0,1] \to GL(n,\mathbb{C})$ defined by $\gamma_2(t) = \mathsf{H}(t) = [h_{ij}(t)]$ where

$$\begin{aligned} h_{ij}(t) &= 0 && \text{if } j \neq i \\ h_{ii}(t) &= 1 + (1-t)b_{ii} && (i = 1, \dots, n). \end{aligned}$$

This curve continuously connects $\mathsf{G}(1)$ at $t = 0$ to I at $t = 1$. Let $\gamma : [0,2] \to GL(n,\mathbb{C})$ be the continuous curve connecting B to I, by

$$\gamma(t) = \begin{cases} \gamma_1(t) & 0 \leq t \leq 1 \\ \gamma_2(t-1) & 1 \leq t \leq 2. \end{cases}$$

Clearly we have $\gamma(0) = \mathsf{B}$ and $\gamma(2) = \mathsf{I}$. The continuous curve $\gamma' = \mathsf{S}^{-1}\gamma\mathsf{S}$ then connects $\gamma'(0) = \mathsf{A}$ and $\gamma'(2) = \mathsf{I}$. Suppose now A is not in G_0, the connected component of the identity. The set $C = \gamma'([0,2])$ is the image in $GL(n,\mathbb{C})$ under a continuous map of the connected interval $[0,2]$ in $\mathbb{R}$. By Theorem 10.17 it is a connected set, and by Theorem 10.16 the union $G_0 \cup C$ is also connected. This is a contradiction since G_0 is a maximal connected set and $\mathsf{A} \in G_0 \cup C \supset G_0$. Hence $G_0 = GL(n,\mathbb{C})$, and the group $GL(n,\mathbb{C})$ is connected.

While a similar argument may be used to show that $SL(n,\mathbb{C})$ is connected, an alternative argument is now to define the map $f : GL(n,\mathbb{C}) \to SL(n,\mathbb{C})$ by

$$f(\mathsf{A}) = (\det \mathsf{A})^{-1/n}\mathsf{A}.$$

Since det is a continuous map the image $SL(n,\mathbb{C})$ under f of the connected set $GL(n,\mathbb{C})$ must be connected (see Theorem 10.17).

The group $U(n)$ is connected. This is essentially dealt with in the concluding remarks of Example 10.22.

Problem 19.9 Show that the groups $SL(n,\mathbb{R})$ and $SO(n)$ are closed subgroups of $GL(N,\mathbb{R})$, and that $U(n)$ and $SU(n)$ are closed subgroups of $GL(n,\mathbb{C})$. Show furthermore that $SO(n)$ and $U(n)$ are compact Lie subgroups.

Solution: Let $\varphi : GL(n,\mathbb{R}) \to \mathbb{R}$ be the continuous map defined by $\varphi(\mathsf{A}) = \det \mathsf{A}$. Then $SL(n,\mathbb{R}) = \varphi^{-1}(\{1\})$ is a closed set since it is the inverse image of under a continuous map of the closed singleton set $\{1\} \subset \mathbb{R}$.

A matrix A is orthogonal iff $\mathsf{A}\mathsf{A}^T = \mathsf{I}$. Let $\psi : GL(n,\mathbb{R}) \to M_n(\mathbb{R}) \cong \mathbb{R}^{n^2}$ be the map defined by $\psi(\mathsf{A}) = \mathsf{A}\mathsf{A}^T$. The function ψ is continuous since the components are quadratic functions of the component of A. Hence $O(n) = \psi^{-1}(\{\mathsf{I}\})$ is a closed set. The group $SO(n) = O(n) \cap SL(n,\mathbb{R})$ is the intersection of two closed sets and is therefore closed.

The treatment for $U(n)$ is similar. Set $\rho : GL(n\mathbb{C}) \to M_n(\mathbb{C}) \cong \mathbb{R}^{2n^2}$ to be the continuous map defined by $\rho(\mathsf{A}) = \mathsf{A}\mathsf{A}^\dagger$. The group $U(n) = \rho^{-1}(\{\mathsf{I}\})$ is therefore closed, as is $SU(n) = U(n) \cap SL(n, \mathbb{C})$ since the group $SL(n, \mathbb{C}) = \det^{-1}(\{1\})$ is a closed as for the real case above.

If $\mathsf{A} \in SO(n)$ then

$$\mathsf{A}\mathsf{A}^T = \mathsf{I} \quad \Longrightarrow \quad \sum_j a_{ij} a_{kj} = \delta_{ik} \quad \Longrightarrow \quad \sum_{i,j} a_{ij} a_{ij} = n$$

i.e.

$$\operatorname{tr} \mathsf{A}\mathsf{A}^T = \sum_{i,j} (a_{ij})^2 = n.$$

Thus $SO(n)$ is a closed subset of the sphere of radius $\sqrt{n}$ in $M_n(\mathbb{R}) \cong \mathbb{R}^{n^2}$. By the Heine-Borel theorem (Theorem 10.8) and Theorem 10.13 this is a compact set.
The treatment of $U(n)$ is similar. If A is a unitary matrix, $\mathsf{A}\mathsf{A}^\dagger = \mathsf{I}$, then

$$\sum_j a_{ij} a_{kj} = \delta_{ik} \quad \Longrightarrow \quad \sum_{i,j} \left|a_{ij}\right|^2 = n$$

and using

$$\left|a_{ij}\right|^2 = \left(\text{Re } a_{ij}\right)^2 + \left(\text{Im } a_{ij}\right)^2$$

the unitary group is a closed subgroup of the sphere of radius $\sqrt{n}$ in $\mathbb{R}^{2n^2}$ and is therefore compact.

Problem 19.10 As in Example 2.13 let the symplectic group $Sp(n)$ consist of $2n \times 2n$ matrices S such that

$$\mathsf{S}^T \mathsf{J} \mathsf{S} = \mathsf{J}, \qquad \mathsf{J} = \begin{pmatrix} \mathsf{O} & \mathsf{I} \\ -\mathsf{I} & \mathsf{O} \end{pmatrix},$$

where O is the $n \times n$ zero matrix and I is the $n \times n$ unit matrix. Show that the Lie algebra $\mathcal{S}p(n)$ consists of matrices A satisfying

$$\mathsf{A}^T \mathsf{J} + \mathsf{J} \mathsf{A} = \mathsf{O}.$$

Verify that these matrices form a Lie algebra and generate the symplectic group. What is the dimension of the symplectic group? Is it a closed subgroup of $GL(2n, \mathbb{R})$? Is it compact?

Solution: Consider a one-parameter subgroup $\mathsf{S}(t) = \exp t\mathsf{A} = \mathrm{e}^{t\mathsf{A}}$. The symplectic condition is written

$$\left(\mathrm{e}^{t\mathsf{A}}\right)^T \mathsf{J} \mathrm{e}^{t\mathsf{A}} = \mathrm{e}^{t\mathsf{A}^T} \mathsf{J} \mathrm{e}^{t\mathsf{A}} = \mathsf{J}$$

and differentiating with respect to t gives

$$\mathsf{A}^T \mathrm{e}^{t\mathsf{A}^T} \mathsf{J} \mathrm{e}^{t\mathsf{A}} + \mathrm{e}^{t\mathsf{A}^T} \mathsf{J} \mathrm{e}^{t\mathsf{A}} \mathsf{A} = \mathsf{A}^T \mathsf{J} + \mathsf{J} \mathsf{A} = \mathsf{O}.$$

Matrices satisfying this relation form a Lie algebra, for if

$$\mathsf{A}^T\mathsf{J} + \mathsf{J}\mathsf{A} = \mathsf{B}^T\mathsf{J} + \mathsf{J}\mathsf{B} = \mathsf{O}$$

then the relation is clearly satisfied by linear combinations $\mathsf{A} + \lambda\mathsf{B}$ and the commutator satisfies

$$\begin{aligned}
[\mathsf{A},\mathsf{B}]^T\mathsf{J} + \mathsf{J}[\mathsf{A},\mathsf{B}] &= \big((\mathsf{AB})^T - (\mathsf{BA})^T\big)\mathsf{J} + \mathsf{J}[\mathsf{A},\mathsf{B}] \\
&= \mathsf{B}^T\mathsf{A}^T\mathsf{J} - \mathsf{A}^T\mathsf{B}^T\mathsf{J} + \mathsf{J}[\mathsf{A},\mathsf{B}] \\
&= -\mathsf{B}^T\mathsf{J}\mathsf{A} + \mathsf{A}^T\mathsf{J}\mathsf{B} + \mathsf{J}[\mathsf{A},\mathsf{B}] \\
&= \mathsf{JBA} - \mathsf{JAB} + \mathsf{J}(\mathsf{AB} - \mathsf{BA}) \\
&= \mathsf{O}.
\end{aligned}$$

If $\mathsf{A}^T\mathsf{J} + \mathsf{J}\mathsf{A} = \mathsf{O}$ then

$$(\mathsf{A}^T)^2\mathsf{J} = -\mathsf{A}^T\mathsf{J}\mathsf{A} = \mathsf{J}\mathsf{A}^2$$

etc. and in general

$$(\mathsf{A}^T)^n\mathsf{J} = (-1)^n\mathsf{J}\mathsf{A}$$

so that

$$\begin{aligned}
\mathrm{e}^{\mathsf{A}^T}\mathsf{J} &= (\mathsf{I} + \mathsf{A}^T + \tfrac{1}{2}t^2(\mathsf{A}^T)^2 + \dots)\mathsf{J} \\
&= \mathsf{J}(\mathsf{I} - \mathsf{A} + \tfrac{1}{2}t^2\mathsf{A}^2 + \dots) \\
&= \mathsf{J}\mathrm{e}^{-\mathsf{A}}.
\end{aligned}$$

Hence $\mathrm{e}^{\mathsf{A}} \in Sp(n)$ for

$$\mathrm{e}^{\mathsf{A}^T}\mathsf{J}\mathrm{e}^{\mathsf{A}} = \mathsf{J}$$

and this Lie algebra generates $Sp(n)$.

Let

$$\mathsf{A} = \begin{pmatrix} \mathsf{A}_1 & \mathsf{A}_2 \\ \mathsf{A}_3 & \mathsf{A}_4 \end{pmatrix}, \qquad \mathsf{A}^T = \begin{pmatrix} (\mathsf{A}_1)^T & (\mathsf{A}_3)^T \\ (\mathsf{A}_2)^T & (\mathsf{A}_4)^T \end{pmatrix}$$

then

$$\mathsf{A}^T\mathsf{J} + \mathsf{J}\mathsf{A} = \mathsf{O} \quad \Longrightarrow \quad (\mathsf{A}_3)^T = \mathsf{A}_3, \quad (\mathsf{A}_2)^T = \mathsf{A}_2, \quad \mathsf{A}_4 = -(\mathsf{A}_1)^T.$$

Thus the number of independent components associated with each of the submatrices A_2 and A_3 is $n(n+1)/2$, and A_1 has n^2 independent components. A_4 is completely defined once A_1 is known, so the total number of degrees of freedom is

$$n(n+1) + n^2 = 2n^2 + n.$$

This is the dimension of the Lie algebra $\mathcal{S}p(n)$ and the associated Lie group $Sp(n)$. The group is a closed subgroup of $SL(2n,\mathbb{R})$ since $Sp(n) = \varphi^{-1}(\{\mathsf{J}\})$ where $\varphi : Sp(n) \to M_{2n}(\mathbb{R})$ is the map (see Problem 19.9 for more details)

$$\varphi(\mathsf{S}) = \mathsf{S}^T\mathsf{J}\mathsf{S}.$$

It is not compact in general. For example $Sp(2) \cong SL(2,\mathbb{R})$ is a non-compact group since it contains all diagonal matrices of the form

$$\begin{pmatrix} a & 0 \\ 0 & a^{-1} \end{pmatrix} \qquad 0 < a < \infty,$$

clearly an unbounded subset of $M_2(\mathbb{R})$.

Problem 19.11 **Show that a group G acts effectively on G/H if and only if H contains no non-trivial normal subgroup of G. [*Hint*: The set of elements leaving all points of G/H fixed is $\bigcap_{a\in G} aHa^{-1}$]**
[NOTE: The words non-trivial must be inserted before "normal subgroup".**
The definition of *effective* action given on p.572 in text is inaccurate. It should read: The action of G on M is said to be effective if e is the only element that leaves every point $x \in M$ fixed,... The displayed equation which follows is correct.]

Solution: G acts effectively on $G/H = \{aH \mid a \in G\}$ iff $g(aH) = gaH = aH$ iff for all $h \in H$ there exists $h' \in H$ such that $gah = ah'$. Thus g must have have the form

$$g = ah'h^{-1}a^{-1} \in aHa^{-1}.$$

As this is required to hold for all $a \in G$ we have

$$g(aH) = aH \quad \forall a \in G \qquad \text{iff} \qquad g \in K = \bigcap_{a\in G} aHa^{-1}$$

Thus the action is effective iff $K = \{e\}$. If H has a normal subgroup $H' \neq \{e\}$ then $aH'a^{-1} = H'$ for all $a \in G$. Hence $K \supseteq H'$, i.e. $K \neq \{e\}$, which implies that the action is not effective. Conversely if the action is effective then $K = \{e\}$ and H has no normal subgroup other than $\{e\}$.

Problem 19.12 **Show that the special othogonal group $SO(n)$, the pseudo-orthogonal groups $O(p,q)$ and the symplectic group $Sp(n)$ are all closed subgroups of $GL(n,\mathbb{R})$.**

(a) Show that the complex groups $SL(n,\mathbb{C})$, $O(n,\mathbb{C})$, $U(n)$, $SU(n)$ are closed subgroups of $GL(n,\mathbb{C})$.

(b) Show that the unitary groups $U(n)$ and $SU(n)$ are compact groups.

Solution: This problem is, accidentally, almost a repeat of Problem 19.9. The only extra item is to show $O(p,q)$ is closed. Consider the map $\psi : GL(n,\mathbb{R}) \to M_n(\mathbb{R})$, where $n = p + q$, defined by

$$\psi(\mathsf{A}) = \mathsf{A}^T G \mathsf{A}$$

where $G = [g_{ij}]$ is the matrix

$$g_{ij} = \eta_i \delta_{ij} \qquad \text{where} \qquad \eta_1 = \cdots = \eta_p = 1, \quad \eta_{p+1} = \cdots = \eta_n = -1.$$

Then $O(p,q) = \psi^{-1}(\{\mathsf{I}\})$. The rest of the argument follows as in Problem 19.9. The symplectic group $Sp(n)$ is actually a closed subgroup of $GL(2n, C)$. The proof is given in Problem 19.10.

Problem 19.13 Show that the centre Z of a Lie group G, consisting of all elements which commute with every element $g \in G$, is a closed normal subgroup of G.

Show that the general complex linear group $GL(n+1, \mathbb{C})$ acts transitively but not effectively on complex projective n-space CP^n defined in Problem 15.4. Show that the centre of $GL(n+1, \mathbb{C})$ is isomorphic to $GL(1, \mathbb{C})$ and $GL(n+1, \mathbb{C})/GL(1, \mathbb{C})$ is a Lie group that acts effectively and transitively on CP^n.

Solution: The centre Z is defined as $Z = \{a \in G \mid ag = ga \quad \forall g \in G\}$. This is a subgroup, as shown in Example 2.20: (i) if $ag = ga$ and $bg = gb$ then $abg = agb = gab$, (ii) $eg = ge$, (iii) multiply $ag = ga$ on left and right by a^{-1} gives $ga^{-1} = a^{-1}g$. It is a normal subgroup since $hZ = Zh$ for all $h \in G$ since $ha = ah$ for all $a \in Z$. To show it is a closed subgroup, define for each $g \in G$ the map $\phi_g : G \to G$ by

$$\phi_g(a) = aga^{-1}.$$

Then $Z_g = \phi_g^{-1}(\{g\}) = \{a \in G \mid ag = ga\}$, known as the *centralizer of* g is a closed set. The centre is

$$Z = \bigcap_{g \in G} Z_g$$

is closed since it is an intersection of closed sets.

From Example 10.15 and Problem 15.4 each point of CP^n is an equivalence classe

$$[\mathbf{z}] = \{\lambda \mathbf{z} \mid \lambda \in \dot{\mathbb{C}}\}$$

where $\mathbf{z} \neq \mathbf{0}$ is a non-vanishing column vector:

$$\mathbf{z} = \begin{pmatrix} z^0 \\ z^1 \\ \vdots \\ z^n \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Any matrix of $\mathsf{A} \in GL(n+1, \mathbb{C})$ acts on CP^n in a natural way by setting $A[\mathbf{z}] = [A\mathbf{z}]$. This result of this action is well-defined since $\mathsf{A}\mathbf{z} \neq \mathbf{0}$ since A is non-singular, and it is independent of the choice of representative $\mathbf{z}$ for

$$\mathsf{A}[\lambda \mathbf{z}] = [\mathsf{A}(\lambda \mathbf{z})] = [\lambda \mathsf{A}\mathbf{z}] = [\mathsf{A}\mathbf{z}] = \mathsf{A}[\mathbf{z}].$$

The action is transitive for if $\mathbf{u}$ and $\mathbf{v}$ are any pair of complex $(n+1)$-vectors then there exists a non-singular matrix A such that $A\mathbf{u} = \mathbf{v}$. To show this, let $\mathbf{e}_0, \dots, \mathbf{e}_n$ be the natural basis of $\mathbb{C}^{n+1}$, having components $(\mathbf{e}_i)^j = \delta^j_i$. By the basis extension theorem, Theorem 3.7, there exists a basis $\mathbf{f}_0, \dots, \mathbf{f}_n$ such that $\mathbf{f}_0 = \mathbf{u}$. Let A be the non-singular operator defined by $B : \mathbf{e}_i \mapsto \mathbf{f}_i = B^j_i \mathbf{e}_j$. Then for any vector $\mathbf{z} = z^i \mathbf{e}_i$, we have

$$B : z^i \mathbf{e}_i \mapsto z^i \mathbf{f}_i = B^j_i z^i \mathbf{e}_j$$

which can be written in matrix notation $\mathbf{z} \mapsto \mathsf{B}\mathbf{z}$, where B is the non-singular matrix $[B^j_i]$. In this notation $\mathsf{B}\mathbf{e}_0 = \mathbf{f}_0 = \mathbf{u}$. Similarly, there exists a linear operator C with associated non-singular matrix C such that $\mathsf{C}\mathbf{e}_0 = \mathbf{v}$. The matrix $\mathsf{A} = \mathsf{C}\mathsf{B}^{-1}$ has the effect that $\mathsf{A}\mathbf{u} = \mathbf{v}$. Thus the action of A on CP^{n+1} is transitive. The action is not effective, however, for $\mu\mathsf{I}$ acts as the identity transformation for any $\mu \in \dot{\mathbb{C}}$:

$$\mu\mathsf{I}[\mathbf{z}] = [\mu\mathbf{z}] = [\mathbf{z}] = \mathsf{I}[\mathbf{z}].$$

The centre of $GL(n+1,\mathbb{C})$ consists of those non-singular matrices A that commute with all $\mathsf{B} \in GL(n+1,\mathbb{C})$, $\mathsf{AB} = \mathsf{BA}$. Taking $\mathsf{B} = \mathsf{I} + \mathsf{E}^i{}_j$ (see Problem 19.1 for the definition of $\mathsf{E}^i{}_j$), easily seen to be non-singular, we have

$$B^a_b = \delta^a_b + \delta^a_j \delta^i_b$$

and

$$\mathsf{AB} = \mathsf{BA} \quad \Longrightarrow \quad A^a_j \delta^i_b = \delta^a_j A^i_b.$$

Setting $b = i$ and $a = j$ we have $A^j_j = A^i_i$, while for $a \neq j$, $A^a_j = 0$. Since these relations hold for arbitrary i, j, a it is clear that $\mathsf{A} = \lambda\mathsf{I}$. Hence $Z = \{\lambda\mathsf{I}\} \cong GL(1,\mathbb{C})$.

Elements of $GL(n+1,\mathbb{C})/GL(1,\mathbb{C})$ may be written as cosets $\lambda\mathsf{A}$ which act on CP^n in the obvious way

$$\lambda\mathsf{A}[\mathbf{z}] = \mathsf{A}[\lambda\mathbf{z}] = \mathsf{A}[\mathbf{z}].$$

The action is transitive by the same argument as above. The action is effective for $\lambda\mathsf{A}$ acts as an identity on CP^n iff for every vector $\mathbf{z}$, $\lambda\mathsf{A}[\mathbf{z}] = [\mathbf{z}]$. This is equivalent to requiring that every vector $\mathbf{z}$ is an eigenvector, $\mathsf{A}\mathbf{z} = \alpha\mathbf{z}$ — clearly possible if and only if $\mathsf{A} = \mu\mathsf{I}$.

Problem 19.14 **Show that $SU(n+1)$ acts transitively on CP^n and the isotropy group of a typical point, taken for convenience to be the point whose equivalence class contains $(1,0,\dots,0,0)$, is $U(n)$. Hence show that the factor space $SU(n+1)/U(n)$ is homeomorphic to CP^n. Show similarly, that**

(a) $SO(n+1)/O(n)$ **is homeomorphic to real projective space** P^n**.**

(b) $U(n+1)/U(n) \cong SU(n+1)/SU(n)$ **is homeomorphic to** S^{2n+1}**.**

Solution: Let $\mathbb{C}^{n+1}$ have an inner product defined such that $\langle \mathbf{e}_i \mid \mathbf{e}_j \rangle = \delta_{ij}$, i.e. $\langle \mathbf{z} \,|\, \mathbf{w} \rangle = \sum_{i=0}^{n} \overline{z^i}\, w^i$. Define the action of $SU(n+1)$ on CP^n as in Problem 19.13 (i.e. treated as a subgroup of $GL(n+1, \mathbb{C})$). In each equivalence class $[\mathbf{u}] = \{\lambda \mathbf{u} \,|\, \lambda \in \dot{\mathbb{C}}\}$ there exists a representative $\mathbf{u}$ such that $\|\mathbf{u}\|^2 = \langle \mathbf{u} \mid \mathbf{u} \rangle = 1$. By Schmidt orthonormalization there exists an o.n. basis $\mathbf{f}_0, \dots \mathbf{f}_n$ such that $\mathbf{f}_0 = \mathbf{u}$. The linear map $A : \mathbf{e}_i \mapsto \mathbf{f}_i = A^j_{\ i} \mathbf{e}_j$ is has a unitary matrix $\mathsf{A} = [A^j_{\ i}]$ since

$$\delta_{ij} = \langle \mathbf{f}_i \,|\, \mathbf{f}_j \rangle = \overline{A^k_{\ i}} A^l_{\ j} \langle \mathbf{e}_k \,|\, \mathbf{e}_l \rangle = \sum_k \overline{A^k_{\ i}} A^k_{\ j},$$

i.e. $\mathsf{A}^\dagger \mathsf{A} = \mathsf{I}$. The matrix $\mathsf{A} \in U(n+1)$ and $\mathsf{A}\mathbf{e}_0 = \mathbf{u}$. We may multiply A by an arbitrary factor $\mathrm{e}^{i\lambda}$ without any effect on the the action on CP^n, so there is no loss of generality in assumed $\det \mathsf{A} = 1$, i.e. $\mathsf{A} \in SU(n+1)$. By an argument similar to that in Problem 19.14, it follows that for any points $[\mathbf{u}], [\mathbf{v}] \in CP^n$ there exists a $\mathsf{A} \in SU(n+1)$ such that $\mathsf{A}[\mathbf{u}] = [\mathbf{v}]$. Thus the action of $SU(n+1)$ on CP^n is transitive.

The isotropy group of $[\mathbf{e}_0]$ is the set of matrices $\mathsf{A} \in SU(n+1)$ such that $\mathsf{A}\mathbf{e}_0 = \lambda \mathbf{e}_0$. Since A is a unitary matrix the eigenvalue λ has unit modulus, $\lambda = \mathrm{e}^{i\phi}$. Furthermore $\mathsf{A}\mathbf{e}_0 = \lambda \mathbf{e}_0 \Longrightarrow A^i_{\ 0} = \lambda$ and $A^i_{\ 0} = 0$ for $i = 1, \dots n$. The unitary condition may also be written $\mathsf{A}\mathsf{A}^\dagger = \mathsf{I}$, i.e. $\sum_{k=0}^{n} \overline{A^i_{\ k}} A^j_{\ k} = \delta^{ij}$, and setting $i = j = 0$ we have

$$1 = \sum_{k=0}^{n} \overline{A^0_{\ k}} A^0_{\ k} = |A^0_{\ 0}|^2 + \sum_{k=1}^{n} \left|A^0_{\ k}\right|^2 \quad \Longrightarrow \quad \sum_{k=1}^{n} \left|A^0_{\ k}\right|^2 = 0.$$

Hence $A^0_{\ k} = 0$ for $k = 1, \dots, n$, and the matrices of the isotropy group are of the form

$$\mathsf{A} = \begin{pmatrix} \mathrm{e}^{i\phi} & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & \mathsf{A}' & \\ 0 & & & \end{pmatrix} \in SU(n+1).$$

The submatrix A' clearly belongs to $U(n)$ and has determinant $\mathrm{e}^{-i\phi}$. As in Example 19.10 the correspondence $AU(n) \longrightarrow A\mathbf{e}_0$ is one-to-one, continuous and has a continuous inverse. Hence

$$SU(n+1)/U(n) \cong CP^n.$$

The arguments for (a) and (b) are similar. We outline the arguments.

(a) Consider the action of $SO(n+1)$ on P^n, real projective space formed from $\mathbb{R}^{n+1}$ with a standard real inner product $\mathbf{e}_i \cdot \mathbf{e}_j = \delta_{ij}$. Again the action is transitive and the isotropy group of $\mathbf{e}_0$ consists of matrices having the form

$$\mathsf{A} = \begin{pmatrix} \pm 1 & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & \mathsf{A}' & \\ 0 & & & \end{pmatrix} \in SO(n+1),\ \mathsf{A}' \in O(n).$$

As above this implies that $SO(n+1)/O(n) \cong P^n$.

(b) Consider the action of $U(n+1)$ on $\mathbb{C}^{n+1}$ with standard inner product. The isotropy group of $[\mathbf{e}_0]$ consists of matrices

$$\mathsf{A} = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & & & \\ \vdots & & \mathsf{A}' & \\ 0 & & & \end{pmatrix} \in U(n+1), \ \mathsf{A}' \in U(n).$$

The correspondence $\mathsf{A}U(n) \longrightarrow \mathsf{A}\mathbf{e}_0$ is a homeomorphism, and setting

$$\mathsf{A}\mathbf{e}_0 = \alpha^0 \mathbf{e}_0 + \alpha^1 \mathbf{e}_1 + \dots + \alpha^n \mathbf{e}_n$$

we have

$$1 = \|\mathsf{A}\mathbf{e}_0\|^2 = \left|\alpha^0\right|^2 + \left|\alpha^1\right|^2 + \dots + \left|\alpha^n\right|^2.$$

Setting $\alpha^i = a^i + ib^i$, this condition is equivalent to

$$(a^0)^2 + (a^1)^2 + \dots + (a^n)^2 + (b^0)^2 + (b^1)^2 + \dots + (b^n)^2 = 1$$

so that

$$U(n+1)/U(n) \cong S^{2n+1}.$$

The discussion for $SU(n+1)/SU(n) \cong S^{2n+1}$ is essentially identical.

Problem 19.15 **As in Problem 9.2 every Lorentz transformation $\mathsf{L} = [L^i_{\ j}]$ has $\det \mathsf{L} = \pm 1$ and either $L^4_{\ 4} \geq 1$ or $L^4_{\ 4} \leq -1$. Hence show that the Lorentz group $G = O(3,1)$ has four connected components,**

$$G_0 = G^{++} : \ \det \mathsf{L} = 1, \ L^4_{\ 4} \geq 1 \qquad G^{+-} : \ \det \mathsf{L} = 1, \ L^4_{\ 4} \leq -1$$
$$G^{-+} : \ \det \mathsf{L} = -1, \ L^4_{\ 4} \geq 1 \qquad G^{--} : \ \det \mathsf{L} = -1, \ L^4_{\ 4} \leq -1.$$

Show that the group of components G/G_0 is isomorphic with the discrete abelian group $Z_2 \times Z_2$.

Solution: To show that the proper Lorentz group G_0 is connected, let $SO(n-1,1)$ be the special orthogonal group of dimension n, consisting of matrices L such that

$$\mathsf{L}^T \mathsf{G} \mathsf{L} = \mathsf{G} \qquad \text{where} \qquad \mathsf{G} = \begin{pmatrix} \mathsf{I} & \mathbf{0} \\ \mathbf{0}^T & -1 \end{pmatrix}$$

and $\det \mathsf{L} = 1$ and $\mathrm{L}^n_{\ n} \geq 1$. A similar argument to that in Example 19.10 can be used to show that

$$SO(n+1,1)/SO(n,1) \cong H^n$$

where $H^n \subset \mathbb{R}^{n+1}$ is the upper sheet of the hyperboloid

$$(x^1)^2 + \dots + (x^n)^2 - (x^{n+1})^2 = -1, \qquad x^{n+1} > 0.$$

This surface is homeomorphic to $\mathbb{R}^n$ by the continuous correspondence

$$\mathbf{x} = \begin{pmatrix} x^1 \\ \vdots \\ x^n \end{pmatrix} \longleftrightarrow \begin{pmatrix} x^1 \\ \vdots \\ x^n \\ x^{n+1} \end{pmatrix}$$

where

$$x^{n+1} = \sqrt{1 + \mathbf{x}^2} = \sqrt{1 + (x^1)^2 + \cdots + (x^n)^2}.$$

Hence H^n is connected and by the argument used in Example 10.22 using Theorem 10.22 it follows inductively that $SO(n,1)$ is a connected topological space for all n (the case $n = 0$ is trivial).

Let

$$\mathsf{L}_1 = \mathsf{I} = \begin{pmatrix} \mathsf{I} & \mathbf{0} \\ \mathbf{0}^T & 1 \end{pmatrix}, \quad \mathsf{L}_2 = \begin{pmatrix} \mathsf{I} & \mathbf{0} \\ \mathbf{0}^T & -1 \end{pmatrix}, \quad \mathsf{L}_3 = \begin{pmatrix} -\mathsf{I} & \mathbf{0} \\ \mathbf{0}^T & 1 \end{pmatrix}, \quad \mathsf{L}_4 = \begin{pmatrix} -\mathsf{I} & \mathbf{0} \\ \mathbf{0}^T & -1 \end{pmatrix}.$$

Every element of G^{+-} is of the form $\mathsf{L}_2\mathsf{L}$ where $\mathsf{L} \in G^{++}$, i.e. $G^{+-} = \mathsf{L}_2 G_0$, and similarly $G^{-+} = \mathsf{L}_3 G_0$, $G^{--} = \mathsf{L}_4 G_0$. Hence each of these sets are connected, since they are homeomorphic to G_0. Furthermore G_0 is the intersection of the inverse images of the positive real line $R^+ = \{x > 0 \,|\, x \in \mathbb{R}\}$ under the continuous maps $\det : O(3,1) \to \mathbb{R}$ and $L^4_4 : O(3,1) \to \mathbb{R}$. Hence G_0 is an open set, as are G^{+-}, G^{-+} and G^{--}. Thus $G_0 = G - (G^{+-} \cup G^{-+} \cup G^{--}$ is also a closed set. As G_0 is connected it is thus the connected component of the identity.

The multiplication table of the cosets G/G_0 is

G/G_0	L_1	L_2	L_3	L_4
L_1	L_1	L_2	L_3	L_4
L_2	L_2	L_1	L_4	L_3
L_3	L_3	L_4	L_1	L_2
L_4	L_4	L_3	L_2	L_1

Setting $Z_2 = \{1, -1\}$, the product group $Z_2 \times Z_2 = \{(\pm 1, \pm 1)\}$ has the same multiplication table if we make the correspondence

$$\mathsf{L}_1 \to (1,1), \qquad \mathsf{L}_2 \to (-1,1), \qquad \mathsf{L}_3 \to (1,-1), \qquad \mathsf{L}_4 \to (-1,-1).$$

Hence the group of components G/G_0 is isomorphic with $Z_2 \times Z_2$.

Problem 19.16 Show that the component of the identity G_0 of a locally connected group G is generated by any connected neighbourhood of the identity e: that is, every element of G_0 can be written as a product of elements from such a neighbourhood.

Hence show that every discrete normal subgroup N of a connected group G is contained in the centre Z of G.

Find an example of a discrete normal subgroup of the disconnected group $O(3)$ that is not in the centre of $O(3)$.

Solution: The first statement has been shown in Theorem 10.21.
If N is discrete then every $a \in N$ has an open neighbourhood $U \subset G$ such that $U \cap N = \{a\}$. Let $\varphi : G \to G$ be the map defined by

$$\varphi(g) = gag^{-1}.$$

Since $\varphi(e) = eae^{-1} = a \in U$ and φ is a continuous map, the set

$$V = \varphi^{-1}(U) = \{g \,|\, gag^{-1} \in U\}$$

is an open neighbourhood of e. As N is a normal subgroup $gag^{-1} \in N$ for all $g \in G$. Hence $gag^{-1} = a$ for all $g \in V$. In other words, $V \subset Z_a$ where Z_a is the centralizer of a, defined in Problem 19.13 as the set of elements which commute with a,

$$Z_a = \{g \,|\, ga = ag\} = \{g \in G \,|\, gag^{-1} = a\}.$$

As this is a subgroup of G containing a neighbourhood of the identity, and G is connected, it follows by Theorem 10.21 that $Z_a = G$. In other words, $ga = ag$ for all $g \in G$, i.e. $a \in Z$.

In $O(3)$ let N be the subset consisting of the two elements $\{\mathsf{I}, \mathsf{J}\}$ where

$$\mathsf{J} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

The reflection J does in commute with all rotations, for

$$\begin{pmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} -\cos\theta & \sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

and

$$\begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} -\cos\theta & -\sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Problem 19.17 **Let $\mathcal{A}$ be a Lie algebra, and for each element X of $\mathcal{A}$ define the linear operator $\mathrm{ad}_X : \mathcal{A} \to \mathcal{A}$ by $\mathrm{ad}_X(Y) = [X, Y]$.**
(a) Show that the map $X \mapsto \mathrm{ad}_X$ is a Lie algebra homomorphism of $\mathcal{A}$ into $\mathcal{GL}(\mathcal{A})$ (called the *adjoint representation*).

(b) For any Lie group G show that each inner automorphism $C_g : G \to G$ defined by $C_g(a) = gag^{-1}$ (see Section 2.4) is a Lie group automorphism, and the map $\mathrm{Ad} : G \to GL(\mathcal{G})$ defined by $\mathrm{Ad}(g) = C_{g*}$ is a Lie group homomorphism.

(c) Show that $\mathrm{Ad}_* = \mathrm{ad}$.
[NOTE: Rewording of first part of question].

Solution: The map ad_X is linear, for

$$\begin{aligned}\mathrm{ad}_X(aY+bZ) &= [X, aY+bZ]\\ &= a[X,Y]+b[X,Z]\\ &= a\,\mathrm{ad}_X(Y)+b\,\mathrm{ad}_X(Z).\end{aligned}$$

(a) The map $\varphi : \mathcal{A} \to L(\mathcal{A})$ defined by $\varphi(X) = \mathrm{ad}_X$ is a Lie algebra homomorphism, for on using the Jacobi identity,

$$\begin{aligned}\mathrm{ad}_{[X,Y]}(Z) &= [[X,Y],Z]\\ &= -[[Y,Z],X]-[[Z,X],Y]\\ &= [X,[Y,Z]]-[Y,[X,Z]]\\ &= (\mathrm{ad}_X\,\mathrm{ad}_Y-\mathrm{ad}_Y\,\mathrm{ad}_X)Z.\end{aligned}$$

That is,

$$\varphi([X,Y]) = [\varphi(X), \varphi(Y)]$$

where the bracket on the right is the operator bracket as defined in the Lie algebra $\mathcal{GL}(\mathcal{A})$.

(b) As seen in Chapter 2, for each $g \in G$ the map $C_g : G \to G$ is a homomorphism, since

$$C_g(ab) = gabg^{-1} = gag^{-1}gbg^{-1} = C_g(a)C_g(b),$$

and it is one-to-one, for

$$C_g(a) = b \Longrightarrow a = g^{-1}bg = C_{g^{-1}}(b),$$

i.e. $_g$ is invertible with $(C_g)^{-1} - C_{g^{-1}}$. Clearly it is a Lie group automorphism since it is differentiable with differentiable inverse. Further, since $C_{gh} = C_gC_h$, as

$$C_{gh}(a) = gha(gh)^{-1} = ghah^{-1}g^{-1} = C_gC_h(a),$$

it follows tht

$$\mathrm{Ad}(gh) = C_{gh*} = (C_gC_h)_* = C_{g*}\circ C_{h*} = \mathrm{Ad}(g)\,\mathrm{Ad}(h).$$

It remains to show that C_{g*} maps left invariant vector field to left invariant vector fields so that $C_{g*} : \mathcal{G} \to \mathcal{G}$. As $C_g = L_gR_{g^{-1}} = R_{g^{-1}}L_g$ we can write

$$C_{g*} = L_{g*}R_{g^{-1}*} = R_{g^{-1}*}L_{g*}$$

so that if X is a left invariant vector field, $L_{g*}X = X$, we have

$$C_{g*}X = R_{g^{-1}*}L_{g*}X = R_{g^{-1}*}X.$$

Hence

$$\begin{aligned} L_{h*}C_{g*}X &= L_{h*}R_{g^{-1}*}X \\ &= R_{g^{-1}*}L_{h*}X \\ &= R_{g^{-1}*}X \\ &= C_{g*}X. \end{aligned}$$

To show that the map $g \mapsto C_{g*}$ is differentiable is slightly technical (see Warner p.113 for details). Essentially it requires showing that for any fixed $X_e \in T_e(G) \cong \mathcal{G}$ the map $\alpha : G \to T_e(G)$ defined by $\alpha : g \mapsto C_{g*}X_e$ is C^∞.

(c) Since $Ad : G \to GL(\mathcal{G})$ is a Lie group homomorphism, the induced tangent map $\mathrm{Ad}_* : \mathcal{G} \to \mathcal{GL}(\mathcal{G})$ is a Lie algebra homomorphism. Hence for any $Y \in \mathcal{G}$, settin $\varphi = \mathrm{Ad}$ in Eq. (19.10) and using Example 19.5 we have

$$\begin{aligned} \mathrm{Ad}(\exp tX)Y &= \exp(\mathrm{Ad}_*\, tX)Y \\ &= \mathrm{e}^{\mathrm{Ad}_*\, tX}Y \\ &= (1 + t\,\mathrm{Ad}_*(X) + O(t^2))Y. \end{aligned}$$

Taking $\mathrm{d}/\mathrm{d}t$ at $t = 0$ gives

$$\begin{aligned} \mathrm{Ad}_*(X)Y &= \frac{\mathrm{d}}{\mathrm{d}t}\,\mathrm{Ad}(\exp tX)Y\Big|_{t=0} \\ &= \frac{\mathrm{d}}{\mathrm{d}t}C_{\exp tX*}Y\Big|_{t=0} \\ &= \frac{\mathrm{d}}{\mathrm{d}t}R_{\exp(-tX)*}L_{\exp tX*}Y\Big|_{t=0} \\ &= \frac{\mathrm{d}}{\mathrm{d}t}R_{\exp(-tX)*}Y\Big|_{t=0}. \end{aligned}$$

Now by Eq. (19.14) we have

$$[X,Y]_g = L_{g*}[X,Y]_e = -\frac{\mathrm{d}}{\mathrm{d}t}L_{g*}\big(R_{\exp tX*}Y\big)_e\Big|_{t=0} = \frac{\mathrm{d}}{\mathrm{d}t}\big(R_{\exp(-tX)*}Y\big)_g\Big|_{t=0}.$$

Hence

$$Ad_*(X)Y = [X,Y] = \mathrm{ad}_X(Y)$$

as required.

Problem 19.18 **(a) Show that the group of all Lie algebra automorphisms of a Lie algebra $\mathcal{A}$ form a Lie subgroup of Aut$(\mathcal{A}) \subseteq GL(\mathcal{A})$.**

(b) A linear operator $D : \mathcal{A} \to \mathcal{A}$ is called a *derivation* on $\mathcal{A}$ if $D[X,Y] =$

$[DX, Y] + [X, DY]$. **Prove that the set of all derivations of $\mathcal{A}$ form a Lie algebra, $\partial(\mathcal{A})$, which is the Lie algebra of Aut$(\mathcal{A})$.**

Solution: (a) A Lie algebra homomorphism is a map $\alpha : \mathcal{A} \to \mathcal{A}$ such that $\alpha \in GL(\mathcal{A})$ and

$$\alpha[X, Y] = [\alpha X, \alpha Y] \qquad \text{for all } X, Y \in \mathcal{A}.$$

If $\beta[X, Y] = [\beta X, \beta Y]$ is another Lie algebra homomorphism then so is the product $\alpha\beta$, for

$$\alpha\beta[X, Y] = \alpha[\beta X, \beta Y] = [\alpha\beta X, \alpha\beta Y].$$

If α is one-to-one and onto it is an automorphism , in which case its inverse is also a homorphism, for every $X \in \mathcal{A}$ may be written $X = \alpha X'$ for some $X' \in \mathcal{A}$, and

$$\alpha^{-1}[X, Y] = \alpha^{-1}[\alpha X', \alpha Y'] = \alpha^{-1}\alpha[X', Y'] = [X', Y'] = [\alpha^{-1}X, \alpha^{-1}Y].$$

To show that the group of all Lie algebra automorphisms, Aut$(\mathcal{A})$ is a matrix Lie group (Lie subgroup of $GL(\mathcal{A})$) is beyond the scope of this book. A more complete discussion of this rather technical question can be found in Helgason p.105.

(b) If D_1 D_2 are derivations then

$$\begin{aligned}
[D_1, D_2]([X, Y]) &= (D_1D_2 - D_2D_1)[X, Y] \\
&= D_1\big([D_2X, Y] + [X, D_2Y]\big) - D_2\big([D_1X, Y] + [X, D_1Y]\big) \\
&= [D_1D_2X, Y] + [D_2X, D_1Y] + [D_1X, D_2Y] + [X, D_1D_2Y] \\
&\quad - [D_2D_1X, Y] - [D_1X, D_2Y] - [D_2X, D_1Y] - [X, D_2D_1Y] \\
&= [(D_1D_2 - D_2D_1)X, Y] + [X, (D_1D_2 - D_2D_1)Y]
\end{aligned}$$

It is therefore a Lie algebra $\partial(\mathcal{A})$.
The Lie algebra of Aut$(\mathcal{A})$ contains all elements $D \in L(\mathcal{A})$ such that $e^{tD} \in$ Aut$(\mathcal{A})$, i.e.

$$e^{tD}[X, Y] = [e^{tD}X, e^{tD}Y.$$

Hence

$$\begin{aligned}
[X, Y] + tD[X, Y] + O(t^2) &= [X + tDX + O(t^2), Y + tDY + O(t^2)] \\
&= [X, Y] + t([DX, Y] + [X, DY]) + O(t^2)
\end{aligned}$$

and from the coefficient of t it follows that D is a derivation ,

$$D[X, Y] = [DX, Y] + [X, DY],$$

i.e. $D \in \partial(\mathcal{A})$. Conversely, if $D \in \partial(\mathcal{A})$ then

$$D^2[X, Y] = [D^2X, Y] + 2[DX, DY] + [X, D^2Y]$$

and continuing inductively we find

$$D^k[X, Y] = \sum_{i=0}^{k} \binom{k}{i} [D^iX, D^{k-i}Y]$$

whence a straightforward computation shows that

$$e^{tD}[X,Y] = [e^{tD}X, e^{tD}Y].$$

Thus $e^{tD} \in \mathrm{Aut}(\mathcal{A})$, which proves the result.

Problem 19.19 **Show that the non-translational Killing vectors of pseudo-Euclidean space with metric tensor $g_{ij} = \eta_{ij}$ are of the form**

$$X = A_j^{\ k} x^j \partial_{x^k} \qquad \textbf{where} \qquad A_{kl} = A_k^{\ j}\eta_{jl} = -A_{lk}.$$

Hence show that the Lie algebra of $SO(p,q)$ is generated by matrices I_{ij} with $i<j$, having matrix elements $(\mathsf{I}_{ij})_a^{\ b} = \eta_{ia}\delta_j^b - \delta_i^{\ b}\eta_{ja}$. Show that the commutators of these generators can be written (setting $\mathsf{I}_{ij} = -\mathsf{I}_{ji}$ if $i>j$)

$$[\mathsf{I}_{ij}, \mathsf{I}_{kl}] = \mathsf{I}_{il}\eta_{jk} + \mathsf{I}_{jk}\eta_{il} - \mathsf{I}_{ik}\eta_{jl} - \mathsf{I}_{jl}\eta_{ik}.$$

Solution: The Killing equations (19.18) with $g_{ij} = \eta_{ij} = \text{consts.}$ are

$$\xi_{i;j} + \xi_{j;i} = \xi_{i,j} + \xi_{j,i} = 0.$$

Hence, setting $i=j$ and with summation convention suspended for repeated indices on the same level,

$$\xi_{j,j} = 0.$$

Thus

$$\xi_{i,jj} = -\xi_{j,ij} = -\xi_{j,ji} = 0,$$

i.e.

$$\frac{\partial^2 \xi_i}{\partial (x^j)^2} = 0,.$$

so that ξ_i is linear in all x^j,

$$\xi_i = A_{ki}x^k + B_i.$$

Ignoring the translational part (see Example 19.3) amounts to setting $B_i = 0$, so we assume $\xi_i = A_{ki}x^k$. Substituting back into the Killing equations above,

$$\xi_{i,j} A_{ki}\delta_j^k = A_{ji} \quad \Longrightarrow \quad A_{ji} + A_{ji} = 0.$$

Let $(\mathsf{I}_{ij})_a^{\ b} = \eta_{ia}\delta_j^b - \delta_i^{\ b}\eta_{ja}$, then

$$\mathsf{A} = [A_a^{\ b}] = \sum_{i<j} A^{ij}\mathsf{I}_{ij} \tag{1}$$

for, if $a<b$ then

$$\sum_{i<j} A^{ij}(\mathsf{I}_{ij})_a^{\ b} = \sum_{i<j} A^{ij}\eta_{ia}\delta_j^b = A^{ib}\eta_{ia} = A_a^{\ b}$$

and if $a > b$ then

$$\sum_{i<j} A^{ij} (\mathsf{l}_{ij})_a{}^b = -\sum_{i>j} A^{ij} \delta_i{}^b \eta_{ja} = -A^{bj}\eta_{ja} = -A^b{}_a = A_a{}^b.$$

For $a = b$ both sides of the Eq. (1) vanish. It also follows immediately that

$$\sum_{i<j} A^{ij} \mathsf{l}_{ij} = 0 \Longrightarrow A^{ij} = 0 \text{ for all } i < j$$

showing that the l_{ij} are linearly independent and are a basis for the Lie algebra of $SO(p,q)$, The commutators are

$$\begin{aligned}
[\mathsf{l}_{ij}, \mathsf{l}_{kl}]_a{}^b &= (\mathsf{l}_{ij})_a{}^c (\mathsf{l}_{kl})_c{}^b - (\mathsf{l}_{kl})_a{}^c (\mathsf{l}_{ij})_c{}^b \\
&= (\eta_{ia}\delta_j{}^c - \delta_i{}^c \eta_{ja})(\eta_{kc}\delta_l{}^b - \delta_k{}^b \eta_{lc}) \\
&\quad - (\eta_{ka}\delta_l{}^c - \delta_k{}^c \eta_{la})(\eta_{ic}\delta_j{}^b - \delta_i{}^b \eta_{jc}) \\
&= \eta_{ia}\eta kj\delta_l{}^b - \eta_{ia}\eta lj\delta_k{}^b - \eta_{ik}\eta ja\delta_l{}^b + \eta_{il}\eta ja\delta_k{}^b \\
&\quad - \eta_{ka}\eta il\delta_j{}^b + \eta_{ka}\eta jl\delta_i{}^b + \eta_{ki}\eta la\delta_j{}^b - \eta_{kj}\eta la\delta_i{}^b \\
&= (\eta_{ia}\delta_l{}^b - \delta_i{}^b \eta la)\eta_{jk} + (\eta_{ja}\delta_k{}^b - \delta_j{}^b \eta ka)\eta_{il} \\
&\quad - (\eta_{ia}\delta_k{}^b - \delta_i{}^b \eta ka)\eta_{jl} - (\eta_{ja}\delta_l{}^b - \delta_j{}^b \eta la)\eta_{ik} \\
&= (\mathsf{l}_{il})_a{}^b \eta_{jk} + (\mathsf{l}_{jk})_a{}^b \eta_{il} - (\mathsf{l}_{ik})_a{}^b \eta_{jl} - (\mathsf{l}_{jl})_a{}^b \eta_{ik},
\end{aligned}$$

giving the desired result.