MATHEMATICS FOR QUANTUM CHEMISTRY

EXERCISES

Version dated: September 9, 2024

SIMEN KVAAL
HYLLERAAS CENTRE
UNIVERSITY OF OSLO
simen.kvaal@kjemi.uio.no

European Summerschool in Quantum Chemistry 2024

Contents

Introduction

This a set of exercises for the mathematics lectures.

Since the mathematics exercises are during the *first* day, the main part of the exercises should be related to material covered early on, ideally the first day. On the other hand, the mathematics lectures are not really intended to *teach* mathematics. Therefore, some students will find the exercises too simple, and some may find them too hard, depending on their background. To remedy the fact that some find may find them too easy, I have added a couple of “nuggets”, things I find interesting that are not so commonly known in the quantum chemistry community – at least to my knowledge. Hopefully, every student will find a challenge or two, and also have some fun.

Simen Kvaal, 2024

Exercise set 1: Introduction

1 Sets etc.

a) RECOMMENDED Write the days of the week as a set

b) RECOMMENDED Write the five first natural numbers using set notation as indicated in the lecture notes

c) An *ordered pair* (x, y) is a two-element "set" where the order matters. Two ordered pairs (x, y) and (u, v) are equal if and only if $x = u$ and $y = v$. From this it follows that $(x, y) = (y, x)$ if and only if $x = y$. A definition of an ordered pair as a *set* is

$$(x, y) = \{\{x\}, \{x, y\}\}. \tag{1.1}$$

Show that $(x, y) = (y, x)$ if and only if $x = y$ using the set theoretic definition.

d) A nonnegative rational number p/q, $p, q \in \mathbb{N}$, $q > 0$, can be written as an ordered pair (p, q). Write down the rational numbers $1/4$, $2/3$, 0, $1/3$ as ordered pairs using set theory, both for the pair and for each integer. (Strictly speaking we the number itself is an *equivalence class* of such pairs, i.e., $pn/qn = p/q$, so we must identify (p, q) and (np, nq). Ignore equivalence classes in this exercise.)

e) The *cartesian product* of two sets A and B is

$$A \times B = \{(a, b) \mid a \in A,\ \ b \in B\}. \tag{1.2}$$

Write down the cartesian product of $\{\heartsuit, \diamondsuit, \spadesuit, \clubsuit\}$ and $\{1, 2, 3\}$, using the set theoretic definition for the ordered pair. You can skip spelling out the set theoretic definition of the natural numbers.

f) Similar to the previous exercise, write down $\{1, 2, 3\} \times \{2, 3, 4\}$.

g) A *function* $f : A \to B$ from one set A to another set B is a rule that assigns to every $a \in A$ precisely one $b \in B$. In terms of set theory, a function $f : A \to B$ is a subset of $A \times B$, such that

- For all $a \in A$ there exists $b \in B$ such that $(a, b) \in f$.
- For all $a \in A$ and $b, b' \in B$, if $(a, b) \in f$ and $(a, b') \in f$, then $b = b'$.

Write down the function $f : \{1, 2, 3\} \to \{1, 2, 3\}$, $f(1) = 2$, $f(2) = 3$, $f(3) = 1$, using the set theoretic definition of ordered pairs and the cartesian product.

De Morgan's laws are useful when discussing subsets A, B of a larger set X. Recall that the complement of A relative to X is $A^\complement = X \setminus A$. De Morgan's laws state that:

- $(A \cup B)^\complement = A^\complement \cap B^\complement$.

- $(A \cap B)^{\mathsf{C}} = A^{\mathsf{C}} \cup B^{\mathsf{C}}$.

h) RECOMMENDED Draw a picture, representing X as the whole sheet, A and B as overlapping shapes, e.g., circles. Label A, B, $A \cap B$, and $A \cup B$, making another drawing if necessary. Convince yourself that De Morgan's laws are correct.

i) Prove De Morgan's laws mathematically. Note that the complement operation acts like negation of truth, i.e., $a \in A^{\mathsf{C}}$ if and only if $a \in X$ and $a \notin A$.

2 Cardinality of numbers

Recall that the *cardinality* $|S|$ of a set S is the number of elements in S. For finite sets, this is intuitive, but what about infinite sets?

One says that A and B have the same cardinality if there exists a bijection $f : A \to B$, i.e., f is one-to-one and onto. (Intuitively, you can draw lines between individual elements of A and B, only one line to/from each element. Precisely: Onto means, to every $b \in B$ there exists a $a \in A$ such that $f(a) = b$. One-to-one: $f(a) = f(a')$ implies $a = a'$.) One can think of the bijection as labelling each $b \in B$ by exactly one element $a \in A$, and that all labels from A are used up.

Let $\mathbb{N}_n = \{1, 2, \cdots, n\}$. This set has cardinality n.

The set $\mathbb{N}$ has an infinite number of elements. By definition, $|\mathbb{N}| = \aleph_0$ ("aleph-nought").

a) RECOMMENDED Show that $|\mathbb{Z}| = \aleph_0$ by explicitly constructing a bijection $f : \mathbb{N} \to \mathbb{Z}$.

b) Show that $|\mathbb{N} \times \mathbb{N}| = \aleph_0$. Hint: Draw a picture of $\mathbb{N} \times \mathbb{N}$, and try to draw a line through all the points in this set. How does this show the existence of a bijection between $\mathbb{N}$ and $\mathbb{N} \times \mathbb{N}$?

c) Show that $|\mathbb{Q}| = \aleph_0$.

d) Show that the interval $I = (-1, 1)$ has the same cardinality as $\mathbb{R}$. You must construct a function $f : (-1, 1) \to \mathbb{R}$ that is a bijection.

e) In this exercise, we show that $|\mathbb{R}| > \aleph_0$. The cardinality would be $\aleph_0$ if we could write a list of the real numbers. It suffices to find a list of the numbers in $I = [0, 1[$, i.e., infinite decimal expansions $0.d_1d_2d_3\cdots$, with $d_i \in \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$.

   We will look for a contradiction. Consider a list $x_1, x_2, x_3, \ldots$ of all real numbers in $[0, 1[$. This is essentially an infinite matrix of digits. Can you find a real number $x \in [0, 1]$ which is *not* in this list, i.e., $x \neq x_j$ for all j? Hint: consider the diagonal of the matrix.

   The cardinality of $\mathbb{R}$, "the continuum", is written $2^{\aleph_0}$.

f) Show that $|[0, 1[\times[0, 1[| = |[0, 1[|$, by constructing a one-to-one map. Hint: Use decimal expansions. Conclude that also $|\mathbb{R}^2| = |\mathbb{C}| = |\mathbb{R}|$.

3 Complex numbers

a) RECOMMENDED Let $z = x + iy \in \mathbb{C}$, and compute the real and imaginary parts of z^2, z^3, and z^4, as functions of x and y.

b) Compute a closed-form expression for the real and imaginary parts of z^n.

c) Compute closed-form expression for z^{-1}, exhibiting the real and imaginary parts as functions of x and y. (Assuming $z \neq 0$). Try to compute closed-form expressions of z^{-2} and z^n as well.

d) RECOMMENDED A complex numbers $z = x + \mathrm{i}y$ can be written on polar form,

$$x = r\cos(\theta), \quad y = r\sin(\theta). \tag{1.3}$$

Write down the rule for $z^n = R\cos\phi + \mathrm{i}R\sin\phi$ using angle θ and modulus r.

e) Consider the polynomial equation $f(z) = z^3 + 1 = 0$. Find all the roots in the complex plane using polar coordinates. Sketch the roots in a coordinate system.

4 Visualizing complex sets

Visualize the following subsets of $\mathbb{C}$. Decide which sets are open and which are closed, or neiter.

a) RECOMMENDED $\{z \mid |z - 1| = 1\}$

b) $\{z \mid |\mathrm{i}z - 1| < 1\}$

c) $\{z \mid |z - \mathrm{i}| + |z + \mathrm{i}| < 4\}$

d) $\{z^2 \mid |z - \mathrm{i}| = 1\}$

e) $\{z \mid \mathrm{Re}(e^{i\pi/2}z) > 0\}$

f) $\{z \mid \mathrm{Im}(z^2) > 0\}$

5 Dual numbers

FOR THE CURIOUS We now consider the number system called *the dual numbers*, often written $\mathbb{R}(\varepsilon)$. The dual numbers are the algebraical extension of $\mathbb{R}$ together with a *nonzero* ε such that $\varepsilon^2 = 0$. That is,

$$\mathbb{R}(\varepsilon) = \{x + \varepsilon y \mid x, y \in \mathbb{R}\}. \tag{1.4}$$

Thus, the dual numbers are similar to the complex numbers in their construction. Indeed, sometimes one writes $\mathbb{C} = \mathbb{R}(\mathrm{i})$. The dual numbers have applications in automatic differentiation of numerical codes.

g) Compute the multiplication law in terms of real and "dual" parts, i.e., $z_1z_2 = z = x + \varepsilon y$, and find x and y. Compute the addition law for $z_1 + z_2$.

h) Show that $\mathbb{R}(\varepsilon)$ is not a field.

i) Compute the law for all positive and negative powers of $z = x + \varepsilon y$.

j) Consider the special cases $z = x + \varepsilon$, and evaluate the positive and negative powers. What are the "dual" parts compared to the real part?

k) Let $p(x)$ be a polynomial over the reals, and extend them to polynomials over the duals, i.e., consider $x \in \mathbb{R}(\varepsilon)$. Show that

$$p(x + \varepsilon) = p(x) + \varepsilon p'(x).$$

l) Let $q(x)$ be a polynomial in negative powers of x. Show that $q(x + \varepsilon) = q(x) + \varepsilon q'(x)$.

m) Let $p(x)$ and $q(x)$ be polynomials, and consider the *rational function* $f(x) = p(x)/q(x)$. Extend the rational function to $\mathbb{R}(\varepsilon)$, and prove that $f(x + \varepsilon) = f(x) + \varepsilon f'(x)$ also in this case.

Exercise set 2: Some conceptual discussions

1 General vector spaces

RECOMMENDED

You may be very familiar with Euclidean space $\mathbb{R}^n$ as a vector space, and quite used to the notion that we have an inner product and a corresponding distance measure on this space. However, linear spaces are much more general. A general vector space does not have an inner product or a norm, just a "linear structure": Consider for example a general vector space over $\mathbb{R}$: An abstract set V with operations such that one can *add vectors*,

$$z = x + y \in V,$$

and *multiply vectors by constants* $c \in \mathbb{R}$,

$$z = cx \in V$$

where $c \in \mathbb{R}$. The full set of axioms is as follows:

Definition : Vector space

A *vector space over the field* $\mathbb{F}$ is a set V together with a binary *vector addition* $+ : V \times V \to V$ and *scalar multiplication* $\cdot : \mathbb{F} \times V \to V$ such that, for all x, y, $z \in V$ and all $\alpha, \beta \in \mathbb{F}$, the following axioms are true:

a) There exists a $0 \in V$ such that $0 + x = x$ for all $x \in V$ — ***identity element for addition***

b) $x + (y + z) = (x + y) + z$ — ***associativity for addition***

c) $x + y = y + x$ — ***commutativity for addition***

d) There exists x' such that $x + x' = 0$ — ***inverse element for addition***

e) $(\alpha\beta) \cdot x = \alpha \cdot (\beta \cdot x)$ — ***compatibility of scalar and field multiplications***

f) $1 \cdot x = x$ — ***identity for scalar multiplication***

g) $(\alpha + \beta) \cdot x = \alpha \cdot x + \beta \cdot x$ — ***distributivity of scalar multiplication***

h) $\alpha \cdot (x + y) = \alpha \cdot x + \alpha \cdot y$ — ***distributivity of scalar multiplication***

Example : Somerset Apple Cake

Ingredients:

- 170 g butter
- 170 g light brown sugar
- 3 eggs
- 1 tablespoon liquid honey
- 1.5 teaspoons cinnamon
- 0.5 teaspoon cloves
- 0.5 teaspoon nutmeg
- 1 teaspoon baking powder
- 240 g wheat flour
- 700 g apples, diced
- 100 ml apple cider, apple juice, or milk
- 100 g light sultana raisins (optional)

Instructions:

- Cream the room-temperature butter with the sugar until light and fluffy.
- Beat in the eggs and honey.
- Sift the flour, baking powder, and spices. Add them to the mixture and stir until the batter is smooth and lump-free.
- Stir in the milk or apple cider (and the raisins, if you choose to use them; they can be soaked in cider for a couple of hours beforehand, if desired).
- Finally, fold the apple pieces into the batter.
- Pour the batter into a greased round tin with parchment paper at the bottom, or use a deep, round ovenproof dish (24 cm in diameter).
- Bake the cake in the middle of the oven at 160°C for 1.5 hours (check with a cake tester to ensure it's fully baked).
- Serve warm or cold with a dollop of whipped cream, clotted cream, vanilla ice cream, or whatever you like.

a) Consider viewing the set L of cake recipe ingredient lists as a vector space. Thus, the components of L are various ingredients and their amounts. What is the dimension of this space? What are the challenges encountered?

b) How would you handle recipe lists with different units, such as pints and liters?

c) Consider the vector space L as a vector space of functions on the form $f : S \to \mathbb{R}$, where S is a set. Which set would S be?

d) Define addition of two recipe ingredient lists. Does the sum define a new cake recipe? What is missing?

e) Consider the scaling of a recipe list to produce a new ingredient list. What are the units of the scaling factor?

f) All elements in the the constructed vector space are not valid recipe lists. Which recipe lists must be discarded?

g) Can you think of a way to make the set R of recipes (with instructions) into a vector space? Which challenges are encountered?

2 Metric spaces

RECOMMENDED

A *metric* encodes the intuition behind measuring *distances* in a set M. It is a distance function $d(x, y)$, where $x, y \in M$.

The definition of a metric space is as follows:

Definition : Metric

Let M be a set. A function $f : S \times S \to \mathbb{R}$ is a *metric* if it satisfies the following axioms:

a) $d(x, y) = d(y, x)$ ***symmetry***

b) $d(x, y) \geq 0$, and $d(x, y) = 0$ if and only if $x = y$ ***positivity and nondegeneracy***

c) $d(x, y) \leq d(x, z) + d(z, y)$ ***triangle inequality***

The pair (M, d) is a *metric space*. If M is a vector space, we say that (M, d) is a *metric vector space*.

a) Consider two persons X and Y located at x and y in a set M. They both measure the distance to the other person. What would happen if axiom a) is violated? Is this reasonable?

b) What would happen if axiom b) is violated? Is this reasonable?

c) A third person Z located at z enters the picture. Consider axiom c). What happens if this axiom is violated? Is this reasonable?

d) Let $M = \mathbb{R}^2$, and let d be Euclidean distance, "as the crow flies". But there are other distance measures as well. Consider for exmaple the "Manhattan distance", given by

$$d(\mathbf{x}, \mathbf{y}) = |x_1 - y_1| + |x_2 - y_2|.$$

This has a standard interpretation in terms of cities with straight streets and taxis driving on these streets. Discuss.

e) Consider the same city, the same streets, the same taxis, but consider the fact that the driver is lazy, and wants to drive as long as they can before turning the car. Find the metric and discuss. Try to verify the axioms.

Exercise set 3: Linear algebra

The source for some of these exercises (indicated) is http://linear.ups.edu/version3/pdf/fcla-draft-solutions.pdf.

1 Basic vectors

a) We work in the space $V = \mathbb{R}^2$, plane vectors. We define two vectors,

$$\mathbf{b}_1 = [1, 2]^T, \quad \mathbf{b}_2 = [-1, 1]^T. \tag{3.1}$$

The notation $...^T$ means that we take the transpose, such that we have *column vectors*. Show by elementary means that these two vectors are linearly independent. Why do the vectors form a basis for V? Make a 2D drawing or plot of the two individual vectors as arrows from the origin.

b) RECOMMENDED Let $\mathbf{v} = [2, 2]^T$. Using gaussian elimination, compute the coefficients x_1 and x_2 such that

$$\mathbf{v} = x_1\mathbf{b}_1 + x_2\mathbf{b}_2.$$

Illustrate this decomposition in the graph from the previous point, using the parallelogram rule: The sum of two vectors is obained by placing the origin of one at the end of the second, or vice versa.

c) RECOMMENDED A person walks from home. First, they walk to the subway station. This is 1 km north and 500 m east as the crow flies. The subway takes the person to the city centre, which 3 km east and 2 km south of the subway station, as the crow flies. The person then walks 200 m south and 400 m west to their office.

What are the coordinates of the office, relative to the person's home?

What is the distance to home, as the crow flies?

d) A boat is sailing on the ocean, at a constant speed 20 km/h relative to the water surface, in the northeast direction. However, the water surface is moving too, at a constant velocity. After two hours, the boat is located 20 km north and 10 km east of the starting point. What is the speed and direction of ocean drift?

e) (Robert Beezer) A three-digit number has two properties. The tens-digit and the ones-digit add up to 5. If the number is written with the digits in the reverse order, and then subtracted from the original number, the result is 792. Use a system of equations to find all of the three-digit numbers with these properties.

2 More vectors and matrices

These exercises are to get practice with computing matrix-matrix products, and to get an intuitive feel for them.

a) RECOMMENDED Let

$$A = \begin{bmatrix} 1 & 3 \\ 0 & -2 \end{bmatrix}, \quad B = \begin{bmatrix} 7 & -1 & 0 \\ 8 & 1 & -1 \end{bmatrix}, \quad C = \begin{bmatrix} 0 & -1 \\ 5 & 7 \\ 0 & 7 \end{bmatrix} \tag{3.2}$$

Which matrix products between pairs of these matrices are defined? Compute the products.

b) RECOMMENDED Let A be the $n \times n$ matrix defined by $A_{i,i+1} = i^2$ for $1 \leq i < n$, and zero otherwise. Compute the action of A and A^T on the standard basis for $\mathbb{R}^n$. Compute the action of A and A^T on an arbitrary vector $\mathbf{x} \in \mathbb{R}^n$.

c) Let $A \in M(n, m, \mathbb{F})$, $B \in M(n, o, \mathbb{F})$. Write B in terms of column vectors, $B = [\mathbf{b}_1, \cdots, \mathbf{b}_o]$. Prove that

$$AB = [A\mathbf{b}_1, \cdots, A\mathbf{b}_o], \tag{3.3}$$

i.e., A acts on each individual column of B. Similarly, show that, if we write A in terms of its *row vectors*,

$$A = \begin{bmatrix} \mathbf{a}_1^T \\ \mathbf{a}_2^T \\ \cdots \\ \mathbf{a}_n^T \end{bmatrix}, \tag{3.4}$$

then

$$AB = \begin{bmatrix} \mathbf{a}_1^T B \\ \mathbf{a}_2^T B \\ \cdots \\ \mathbf{a}_n^T B \end{bmatrix}. \tag{3.5}$$

d) Show that the columns of $C = AB$ is a linear combination of the columns of A with coefficients determined by B,

$$\mathbf{c}_i = \sum_j \mathbf{a}_j B_{ji} \tag{3.6}$$

Show a similar result for the rows of C.

3 The complex numbers as a real vector space

FOR THE CURIOUS

The geometrical interpretation of the $\mathbb{C}$ as the plane $\mathbb{R}^2$ indicates that $\mathbb{C}$ can be viewed as a two-dimensional real vector space. Indeed, addition and multiplication with *real* scalars are compatible with the axioms for Euclidean space $\mathbb{R}^2$.

a) Show that $\mathbb{C}$ regarded can be regarded as the Euclidean plane: Check axioms for Euclidean space and check that vector addition and multiplication with real scalars are the same in the two spaces. Check also that the Euclidean norm is the modulus of the complex number. What are the complex numbers that correspond to the standard basis in $\mathbb{R}^2$? Conclude that $\mathbb{C}$ can be regarded as a *real* Euclidean vector space of dimension 2.

b) Show that multiplication with a *complex* number z is a linear operator on $\mathbb{R}^2$, and find its matrix in the standard basis. What is the matrix of multiplication with i?

c) Show that multiplication with a complex number of modulus 1 is a rotation in $\mathbb{R}^2$.

d) Show that the map $z \mapsto \bar{z}$ is a linear transformation in $\mathbb{R}^2$. What kind of linear transformation is this? Is it a linear transformation in $\mathbb{C}$?

4 Elementary row operations and Gaussian elimination

a) Compute the overlap matrix S of the basis in Eq. (3.1). Use gaussian elimination to compute the inverse of the matrix S.

b) Write down the dual basis of $B = \{\mathbf{b}_1, \mathbf{b}_2\}$.

c) RECOMMENDED (Robert Beezer) Consider the following system of equations:

$$2x_1 - 3x_2 + x_3 + 7x_4 = 14 \tag{3.7}$$
$$2x_1 + 8x_2 - 4x_3 + 5x_4 = -1 \tag{3.8}$$
$$x_1 + 3x_2 - 3x_3 = 4 \tag{3.9}$$
$$-5x_1 + 2x_2 + 3x_3 + 4x_4 = -19 \tag{3.10}$$

Use gaussian elimination to find all possible solutions. Write down the set of all solution using set notation.

d) (Robert Beezer) Find all possible solutions of the linear system

$$3x_1 + 4x_2 - x_3 + 2x_4 = 6 \tag{3.11}$$
$$x_1 - 2x_2 + 3x_3 + x_4 = 2 \tag{3.12}$$
$$10x_2 - 10x_3 - x_4 = 1 \tag{3.13}$$

Write down the solution set using set notation.

e) (Robert Beezer) Find all possible solutions of the linear system

$$2x_1 + 4x_2 + 5x_3 + 7x_4 = -26 \tag{3.14}$$
$$x_1 + 2x_2 + x_3 - x_4 = -4 \tag{3.15}$$
$$-2x_1 - 4x_2 + x_3 + 11x_4 = -10 \tag{3.16}$$

Write down the solution set using set notation.

f) Let $\mathbf{b}_i$, $i = 1, 2, \cdots, n$ be a basis for $\mathbb{R}^n$. Let

$$B = [\mathbf{b}_1, \mathbf{b}_2, \cdots, \mathbf{b}_n]$$

be the *basis matrix*, whose columns are precisely the $\mathbf{b}_i$. Explain that if $\mathbf{y} = x_1\mathbf{b}_1 + \cdots + x_n\mathbf{b}_n$, then

$$B\mathbf{x} = \mathbf{y}.$$

Solve for $\mathbf{x}$ using B in symbols.

g) Let n linearly independent vectors $B = \{\mathbf{b}_1, \cdots, \mathbf{b}_n\}$ of $\mathbb{F}^n$ be given. Write down a formula for the dual basis in terms of the overlap matrix.

5 Bra-ket notation

a) Let a basis $B = \{b_1, b_2, \cdots, b_n\}$ be a basis for a finite-dimensional Hilbert space, and let $\{|i\rangle\}$ be a the standard basis for $\mathbb{C}^n$. Consider the operator defined by

$$\hat{B} = \sum_{i=1}^{n} |b_i\rangle \langle i|, \tag{3.17}$$

What is the domain of $\hat{B}$, and the range of $\hat{B}$? In other words, which spaces does $\hat{B}$ map from and to?

b) Let $|x\rangle \in \mathbb{C}^n$, and consider

$$|\psi\rangle = \hat{B}\,|x\rangle\,.$$

Can you interpret this expression in terms of bases and expansion coefficients?

c) Show that

$$\hat{S} = \sum_{ij} |i\rangle\,\langle b_i|b_j\rangle\,\langle j|\,. \tag{3.18}$$

This is the overlap matrix expressed as an operator. Find an expression for the basis coefficients x_i in terms of the basis alone.

6 Eigenvalue decomposition

FOR THE CURIOUS

Of great importance to quantum chemistry is the time-independent Schrödinger equation: Given the Hamiltonian operator $\hat{H}$, which is usually Hermitian, find a nonzero $|\psi\rangle$, called an eigenvector, and a real number E, called an eigenvalue, such that

$$\hat{H}\,|\psi\rangle = E\,|\psi\rangle\,. \tag{3.19}$$

Choosing a particular orthonormal basis $\{|\phi_i\rangle\}$ for Hilbert space, a matrix eigenvalue problem arises,

$$A\mathbf{u} = E\mathbf{u}, \quad u_i = \langle\phi_i|\psi\rangle\,, \quad A_{ij} = \langle\phi_i|\hat{H}|\phi_j\rangle\,. \tag{3.20}$$

We prove the spectral theorem for normal operators over finite dimensional Hilbert spaces. The spectral theorem states that one can actually find a *basis* of eigenvectors for normal operators. The normal operators include the Hermitian operators.

In this section, we let $M(n) = M(n, n, \mathbb{C}) = \mathbb{C}^{n\times n}$ be the set of complex square matrices of dimension n.

Two matrices A and B are said to be *similar* if there is an invertible matrix U such that $A = UBU^{-1}$. If U is unitary, then A and B are *unitarily similar*. (U is unitary if and only if $U^H = U^{-1}$.)

a) What does it mean that $A \in M(n)$ is normal?

b) Prove that a diagonal matrix $D \in M(n)$ is normal. (A matrix A is diagonal if and only if $A_{ij} = 0$ whenever $i \neq j$.)

c) Prove that U is unitary if and only if its columns are orthonormal.

d) Prove that if A is a normal matrix, if B is unitarily equivalent to A, then B is normal.

e) Prove that if A is unitarily equivalent to a diagonal matrix, then A is normal.

The next part requires *Schur's Lemma*:

Lemma 1: Schur's Lemma

Let $A \in M(n)$. Then A is unitarily equivalent to an upper triangular matrix. (R is upper triangular if $R_{ij} = 0$ whenever $i > j$.)

The proof of Shur's Lemma can be found in, e.g., Fraleigh and Beauregard.

f) Prove that every normal matrix A is unitarily equivalent to a normal upper-triangular matrix B. (Use Schur's Lemma and exercise d).)

g) Prove that a normal upper triangular matrix B must be diagonal. Hint: Let $C = B^H B = BB^H$, and compute C_{11} to show that $B_{1j} = 0$ if $j > 1$. Continue to C_{22}, and all the way up to C_{nn}.

We have now proved:

Theorem 1: Spectral theorem for normal matrices

Suppose $A \in M(n)$ is normal, i.e., $AA^H = A^H A$. Then there is a unitary matrix $U \in M(b)$ and a diagnal matrix $D \in M(n)$ such that

$$A = UDU^H. \qquad \textit{spectral decomposition} \tag{3.21}$$

h) Let $\mathbf{u}_i = U_{:,i}$ be the ith column for U. Explain why these vectors form a basis for $\mathbb{C}^n$.

i) Show that

$$A = \sum_{i=1}^{n} d_i \mathbf{u}_i \mathbf{u}_i^H, \quad d_i = D_{ii}. \qquad \textit{spectral decomposition} \tag{3.22}$$

Write this formula also using bra-ket notation.

j) What does A do with the vector $\mathbf{u}_k$ for some k? What does A do to a linear combination of the basis vectors?

A special kind of normal matrix is the Hermitian matrix:

$$A^H = H. \tag{3.23}$$

k) Show that a Hermitian matrix is unitarily equivalent to a diagonal matrix with *real* numbers on the diagonal, i.e., $A = UDU^H$ with $D_{ii} \in \mathbb{R}$.

7 Diagonalization of a matrix

A matrix A has an eigenvalue E if and only if there is a nonzero solution to the linear system

$$(A - EI)\mathbf{x} = 0.$$

Thus, the matrix $A - EI$ must be singular, i.e., not invertible. The eigenvector $\mathbf{x}$ is then a vector in the null space of this matrix.

A sufficient and necessary criterion for invertibility of a matrix B is that the *determinant* is nonzero. The definition of the determinant can be stated as follows:

Definition 1: Determinant

Let $A \in M(n)$. The determinant, written $\det(A) \in \mathbb{C}$ is defined by

$$\det(A) = \sum_{p \in S_n} (-1)^{|p|} A_{1,p(1)} A_{2,p(2)} \cdots A_{n,p(n)}. \tag{3.24}$$

Here, S_N is the set of *permutations* of the numbers $\{1, 2, \cdots, n\}$. A permutation is by definition any bijective map.

Each permutation $p \in S_n$ is a rearrangement is thus completely specified by the new order of the numbers $\{1, 2, \cdots, n\}$, i.e., $(p(1), p(2), \cdots, p(n))$. For example, $(1, 2, 3, 4, 5, 6) \in S_6$ is the *identity* *permutation* that leaves any number alone. The permutation $(3, 2, 1, 4, 6, 5)$ makes 1 and 3 switch place, and 5 and 6.

Any permutation can be written as a product (composition) of *transpositions*, i.e., switching of exactly two elements. Thus, $(3, 1, 2) = (1, 2)(2, 3)$. Here, the two-element notation indicates which elements are swithced. It is a fact that all permutations can be decomposed in *either* an even number of permutations, or an odd number of permutations, denoted $|p|$. The *sign* of a permutation p is $(-1)^{|p|}$, and is unique for a permutation.

There are $n!$ unique permutations of an n-element set.

a) Write out the set of permutations S_1, S_2 and S_3.

b) Compute the sign of the permutations of the previous exercise.

c) Compute explicit expressions for the determinants of matrices of dimension 1, 2, and 3.

d) (Extra.) In Python, the module `itertools` contains an iterator class `permutations` that allow efficient looping over permutations. Can you write a function that computes the sign of a permutation? Use the internet for efficiency.

e) Compute explicit expressions for the determinants of matrices of dimension up to 3. By equating the determinant of $A - EI$ with zero, a polynomial equation in E is obtained. Find the solutions in the case $n = 2$.

f) Consider the matrix
$$A = \begin{bmatrix} -1 & z \\ z & 1 \end{bmatrix}.$$
For which $z \in \mathbb{C}$ is A Hermitian? Normal? Compute the eigenvalues as function of z. Also find the eigenvectors.

g) For which values of z does there exist a basis of eigenvectors?

8 Gram--Schmidt orthogonalization

FOR THE CURIOUS Let V be a vector space over $\mathbb{F}$ with $\dim(V) = n < +\infty$. Recall that every finite-dimensional vector space has a basis. In the lecture notes, it is claimed that every finite dimensional Hilbert space has a basis of orthonormal vectors! Recall, that in an orthonormal basis, computations are usually much simpler than in nonorthogonal bases.

In this exercise, we show that any set of linearly independent vectors can be orthonormalized. We work in the space $\mathbb{F}^n$ for simplicity. That is, let $B = [\mathbf{b}_1, \cdots, \mathbf{b}_m] \subset \mathbb{F}^n$ be a set of $m \leq n$ linearly independent vectors, spanning a subspace $V = \mathrm{sp}(B)$. We denote the vector set by its basis matrix B. Equivalently, B is an arbitrary matrix in $M(n, m, \mathbb{F})$ of full rank m, and the vectors $\mathbf{b}_i$ are the columns of B.

We will find a set of orthonormal vectors $U = [\mathbf{u}_1, \cdots, \mathbf{u}_m]$, again denoted by a matrix in $M(n, m, \mathbb{F})$, such that $\mathrm{sp}(B) = \mathrm{sp}(U)$. That is, the two sets of vectors are bases for the *same* subspace, but one of them is an orthonormal basis.

We first define the notion of orthogonal projection onto the span of a single vector: Let $\mathbf{y} \in \mathbb{F}^n$ be nonzero, $Y = \mathrm{sp}(\mathbf{y})$, and let $\mathbf{x} \in \mathbb{F}^n$.

a) Show that the vector $\mathbf{x}_Y$ given by
$$\mathbf{x}_Y = \mathbf{y}\frac{\langle \mathbf{y}, \mathbf{x} \rangle}{\langle \mathbf{y}, \mathbf{y} \rangle} \tag{3.25}$$

is the *unique* vector in Y, such that

$$\forall \mathbf{y}' \in Y, \quad \|\mathbf{x} - \mathbf{x}_Y\| \le \|\mathbf{x} - \mathbf{y}'\|. \tag{3.26}$$

Hint: Find a function $f : \mathbb{R} \to \mathbb{R}$ which you can study, by using the definition of $Y = \mathrm{sp}(\mathbf{y})$.

b) Show that

$$\mathbf{x} = \mathbf{x}_Y + \mathbf{x}_{Y^\perp}, \tag{3.27}$$

where

$$\langle \mathbf{x}_{Y^\perp} | \mathbf{x}_Y \rangle = 0. \tag{3.28}$$

Illustrate with a picture of the situation in $\mathbb{R}^2$.

The vector $\mathbf{x}_Y$ is called the *orthogonal projection of* $\mathbf{x}$ *onto* Y *or* $\mathbf{y}$.

c) Explain that the orthogonal projection is given by acting with an operator,

$$\mathbf{x}_Y = P_{\mathbf{y}}\mathbf{x}, \quad P_{\mathbf{y}} = \mathbf{y} \langle \mathbf{y}, \mathbf{y} \rangle^{-1} \mathbf{y}^H. \tag{3.29}$$

The operator $P_{\mathbf{y}}$ is called the *orthogonal projection operator onto* Y *or* $\mathbf{b}$.

d) Show that $P^2 = P$ and that $P^\dagger = P$.

The Gram–Schmidt orthogonalization proceeds recursively:

1. $\mathbf{u}_1 = \mathbf{b}_1 \langle \mathbf{b}_1, \mathbf{b}_1 \rangle^{-1/2}$ (normalization)
2. $\mathbf{u}_2' = \mathbf{b}_2 - P_{\mathbf{u}_1}\mathbf{b}_2$ (removal of projection), and $\mathbf{u}_2 = \mathbf{u}_2' \langle \mathbf{u}_2', \mathbf{u}_2' \rangle^{-1/2}$ (normalization)
3. $\mathbf{u}_3' = \mathbf{b}_3 - P_{\mathbf{u}_1}\mathbf{b}_3 - P_{\mathbf{u}_2}\mathbf{b}_3$ (removal of projection), and $\mathbf{u}_3 = \mathbf{u}_3' \langle \mathbf{u}_3', \mathbf{u}_3' \rangle^{-1/2}$ (normalization)
4. …
5. $\mathbf{u}_m' = \mathbf{b}_m - \sum_{i=1}^m P_{\mathbf{u}_i}\mathbf{b}_m$ (removal of projection), and $\mathbf{u}_m = \mathbf{u}_m' \langle \mathbf{u}_m', \mathbf{u}_m' \rangle^{-1/2}$ (normalization)

We see that each $\mathbf{u}_i$ is given by projecting away from $\mathbf{b}_i$ the components of all the *previously generated* $\mathbf{u}_j$, $j < i$.

a) Show that all the generated vectors $\mathbf{u}_i$ are orthonormal.

b) Show that a set of orthonormal vectors are linearly independent.

c) Show that the span of the first k original vectors is the same as the span of the first k generated vectors. Show that

$$B = UR, \tag{3.30}$$

where $R \in M(m, m, \mathbb{F})$ is *upper triangular*, i.e., $R_{ij} = 0$ whenever $i > j$. This is called the QR decomposition of the operator B. (The "Q" is just our U.)

d) We have shown the existence of the QR decomposition when B had full rank. Can you show the existence of the decomposition when B does *not* have full rank? It always exists. Hint: What happens when B does not have full rank?

We now turn to a computational exercise:

e) Write a Python function to compute the QR decomposition of a matrix with full rank. In SciPy, the function `scipy.linalg.qr` is an advanced implementation with pivoting. To compare with your implementation, use the call

```
Q,R = qr(A,mode='economic',pivoting=False)
```

With `pivoting=False`, the function computes essentially the same as our presented algorithm. (Turning on pivoting stabilizes the algorithm, rearranging the vectors before orthogonalizing.)

f) Let $n \geq m > 0$ be integers, and define the matrix $A \in \mathbb{R}^{n\times m}$ by

$$A_{ij} = \left(\frac{i-1}{n-1}\right)^{j-1}. \tag{3.31}$$

This the matrix of the m first monomials x^{j-1} evaluated at an equidistant grid of n points in the interval $[0, 1]$. Test your QR implementation on A for various n, m, say $m = 10$ and $m = 200$. In particular, compute the difference between your factorization and A, and compare also with the NumPy implementation.

9 The singular value decomposition as a compression tool

FOR THE CURIOUS

The SVD is a very powerful linear algebra tool. In quntum chemistry it is used to compress electron repulsion integrals, for example.

In this small exercise, we demonstrate its power applied to *image compression*. For this exercise, a good grayscale image is handy. You can find a nice picture of a parrot in Figure 3.1.

Recall first the *Hilbert–Schmidt* inner product on the set $M(n, m, \mathbb{F})$ of matrices:

$$\langle X, Y\rangle_{\mathrm{HS}} = \mathrm{Tr}(X^H Y) \tag{3.32}$$

The *trace* $\mathrm{Tr}(A)$ is the sum of the diagonal elements of A.

Recall also the SVD: For $A \in M(n, m, \mathbb{F})$, and $k = \min(n, m)$, there are unitary matrices $U \in M(n)$ and $V \in M(m)$, such that $U^H U = I_n$ and $V^H V = I_m$, and numbers$\sigma_1 \geq \sigma_2 \geq \cdots \sigma k$, such that

$$A = U\Sigma V^H = \sum_{i=1}^{k} \mathbf{u}_i \sigma_i \mathbf{v}_i^H, \tag{3.33}$$

where $\Sigma \in M(n, m, \mathbb{R})$ is such that the upper $k \times k$ block is diagonal, with the σ_is along the diagonal, and the rest is zero. Note that U and V may have columns that do not enter the expansion on the right-hand side. Thus, the "full SVD" on the left hand side can be reduced to the "economical SVD",

$$A = U\Sigma V = U_{:,:k}\Sigma_{:k,:k}V^H_{:,:k}, \tag{3.34}$$

where the index ": k" is short for the selection of the k first indices, and ":" is everything.

a) How is the Hilbert–Schmidt inner product related to the Euclidean inner product and the standard Euclidean vector space?

b) Show that $\mathrm{Tr}(ABC) = \mathrm{Tr}(CAB)$ (cyclic invariance).

c) Let $\mathbf{u}_i$ and $\mathbf{v}_i$ be the left and right singular vectors of a matrix A. Show that the matrices $X^{ij} = \mathbf{u}_i \mathbf{v}_j^T$ form an orthonormal set in the Hilbert–Schmidt inner product.

d) Conclude that the SVD is an expansion of a matrix in an orthonormal basis.

Figure 3.1: A parrot, https://www.dropbox.com/s/8r4svs6uuhgwawt/parrot.png

It is a fact, that the *truncated* SVD, i.e.,

$$A^{(m)} = \sum_{i=1}^{m} \mathbf{u}_i \sigma_i \mathbf{v}_i^H, \tag{3.35}$$

is the best rank m approximation of A in the sense of the Hilbert–Schmidt norm, i.e.,

$$A^{(m)} = \operatorname*{argmin}_B \|B - A\|_{\mathrm{HS}}, \tag{3.36}$$

subject to the condition that rank of B is m.

e) Write a Python program/notebook that reads an image file and converts it to a matrix $A \in M(n, m, \mathbb{R})$. The module `imageio` can be useful.

f) Continue writing the program, so it computes the truncated and full SVD of A. You can use, say, `scipy.linalg.svd` to find the full SVD.

g) Plot the singular values. Discuss.

h) Show that the error in the Hilbert–Schmidt norm is

$$\|A^{(m)} - A\|_{\mathrm{HS}} = \sqrt{\sum_{i=m+1}^{k} \sigma_i^2}. \tag{3.37}$$

i) Make visualizations of the truncated SVD of the image A, with various m. Compute the relative error $\|A^{(m)} - A\|_{\mathrm{HS}} / \|A\|_{\mathrm{HS}}$ for each m you consider. Discuss the image quality.

10 Space of polynomials

In this exercise, we study a vector space that is not $\mathbb{F}^n$. We start out without an inner product, but supply this later.

Recall that a polynomial of degree n with coefficients in $\mathbb{F}$ is a function $p : \mathbb{R} \to \mathbb{C}$ given by

$$p(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n, \tag{3.38}$$

for some $n \geq 0$, with $a_i \in \mathbb{C}$, and $a_n \neq 0$.

a) RECOMMENDED Let P_n be the set of all polynomials of degree less than or equal to n. Show that P_n is a vector space of dimension $n+1$ over $\mathbb{C}$.

b) RECOMMENDED Let $D : P_n \to P_n$ be the differentiation operator, i.e.,

$$(Dp)(x) = p'(x). \tag{3.39}$$

Show that D is indeed a linear operator.

c) Does there exist an inverse of D?

d) Let $B = \{1, x, x^2, \cdots, x^n\}$ be the basis of *monomials*. Using bra-ket notation, we write the basis operator as

$$\hat{B} = \sum_{i=1}^{n} |x^{i-1}\rangle \langle i|, \tag{3.40}$$

where $|i\rangle$ is a standard basis vector for $\mathbb{R}^n$. Compute the matrix of D relative to this basis. (Note carefully that we do not have an inner product defined.) That is, find a matrix T such that

$$D\hat{B} = \hat{B}T. \tag{3.41}$$

We now restrict our attention to the interval $x \in [-1, 1]$. That is,

$$V_n = \{p : [-1, 1] \to \mathbb{C} \mid p \text{ a polynomial of deg} \leq n\}. \tag{3.42}$$

The above subexercises did not depend on the domain of the polynomial.

We define an inner product on V,

$$\langle p|q\rangle := \int_{-1}^{1} \overline{p(x)}q(x)\,\mathrm{d}x. \tag{3.43}$$

We will now define the *Legendre polynomials*.

e) Show that $\langle\cdot|\cdot\rangle$ is indeed an inner product by checking the axioms.

f) Compute the overlap matrix S of the monomial basis. Is the basis orthonormal?

g) Formulate the Gram–Schmidt procedure using the bra-ket notation. Explain how the procedure generates an upper triangular matrix R such that

$$\hat{B} = \hat{U}R, \tag{3.44}$$

where $\hat{U}$ is the basis matrix of the orthonormal set of vectors.

h) Use the Gram–Schmidt decomposition by hand to compute the first three *normalized Legendre polynomials* $|\hat{P}_i\rangle$, $i = 0, 1, 2$, defined using the Gram–Schmidt orthogonalization of $\{1, x, x^2\}$.

The Legendre polynomials are defined as the set of orthogonal polynomials $P_i : [-1,-1] \to \mathbb{R}$ obtained by Gram–Schmidt orthogonalization of the monomials, but normalized according to $P_i(x) = 1$. Thus, they are not ortho*normal*. The endpoint-normalization is equivalent to

$$\int_{-1}^{1} P_i(x)P_j(x)\,\mathrm{d}x = \frac{2}{2i+1}\delta_{ij}. \tag{3.45}$$

It is also a fact that the Legendre polynomials satisfy a *recurrence relation*: With $P_0(x) = 1$ and $P_{-1} \equiv 0$, we have

$$(i+1)P_{i+1}(x) = (2i+1)xP_i(x) + iP_{i-1}(x). \tag{3.46}$$

From this, one can show:

$$P'_{i+1}(x) = \frac{2P_i(x)}{\|P_i\|^2} + \frac{2P_{i-2}(x)}{\|P_{i-2}\|^2} + \frac{2P_{i-4}(x)}{\|P_{i-4}\|^2} + \ldots \tag{3.47}$$

i) Compute the matrix of the differential operator D in the *normalized* Legendre polynomial basis, both by using matrix multiplication involving X, R, and R^{-1}, and also by using Eq. (3.47). Compare the two explicitly using pen-and-paper calculations.

j) Repeat the previous exercise with the operator D^2.

One of the most useful uses of orthogonal polynomials is that they define very efficient *quadrature rules*. Consider approximating the integral of some $f(x)$ by a linear combination of point samples,

$$\int_{-1}^{1} f(x)\,\mathrm{d}x \approx \sum_{i=1}^{n} f(x_i)w_i, \tag{3.48}$$

where the w_i are called *weights* and the x_i are called *nodes*. Then, it is a fact that there is a unique choice of weights and nodes such that the quadrature rule is *exact* for any polynomial of degree $2n-1$. This is quite remarkable. Even more remarkable perhaps, is that the nodes are the zeroes of P_n. The weights are also related to the Legendre polynomials,

$$w_i = \frac{2}{(1-x_i)^2[P'_i(x_i)]^2} \tag{3.49}$$

The Wikipedia page on Gauss–Legendre quadrature contains more details.

The so-called Golub–Welsch algorithm is a simple algorithm based on diagonalization of a tridiagonal Hermitian matrix for computing the nodes and weights of Gaussian quadratures, see the Wikipedia page on Gaussian quadrature. The interested student should check it out!

Exercise set 4: Vector calculus

1 Visualization of functions

Some paths and their curves:

a) RECOMMENDED Let $f : \mathbb{R} \to \mathbb{R}^2$ be a path defined by

$$f(t) = [r(t)\cos(4t), r(t)\sin(4t)], \quad r(t) = \exp(-t)].$$

Sketch the curve traced by the path for $0 \leq t \leq 2\pi$. You can check your sketch against, say, a Python plot.

b) Let $f :]0, +\infty[\subset \mathbb{R} \to \mathbb{R}^2$ be a path defined by

$$f(t) = [\ln(t), t].$$

Sketch the curve traced by the path. You can check your sketch against, say, a Python plot.

c) Let $f : [0, \pi] \to [x(t), y(t)]$, with $x(t) = 16\sin(t)^3$ and $y(t) = 13\cos(t) - 5\cos(2t) - 2\cos(3t) - \cos(4t)$. Sketch the curve. Here, it is probably easiest to just use Python or a similar tool.

Level curves:

d) RECOMMENDED Let $f(x, y) = x^2 + y^2$. Sketch the level curves $\{(x, y) \mid f(x, y) = n\}$, for $n = 1$ and $n = 2$, $n = 3$, and $n = 4$.

e) Let $f(x, y) = x + y + 2$. Sketch the level curves where $f(x, y) = 0$, 2 and 4. Can you sketch the graph of f?

f) Let $f(x, y) = x^2 - y^2$. The graph is called the *hyperbolic paraboloid*, or a saddle. Sketch the graph. Sketch the level curves $\{(x, y) \mid f(x, y) = n\}$, for $n = 0$ and $n = -1$, $n = 1$. Make sure that you get all "pieces".

g) Write a Python program for sketching the level curves in f), but use more closely spaced constants. There are tools in Matplotlib that are handy.

h) Adapt your program to draw level curves of $(x, y) \mapsto (x^2 + 3y^2)e^{1-x^2-y^2}$.

Sections and level surfaces:

Consider the hydrogen atom eigenfunctions. These are defined in terms of the nodal quantum number $n \in \mathbb{N}$, and the orbital angular momentum quantum numbers $\ell \in \mathbb{N}$ and $m \in \mathbb{Z}$, $|m| \leq \ell$. The general expression for the orbitals is

$$\psi(r, \theta, \phi) = R(r)Y_{\ell m}(\theta, \phi), \tag{4.1}$$

where $Y_{\ell m}$ is a spherical harmonic, and where $R(r)$ is the radial solution of the Schrödinger equation, given by

$$R(r) = N_{\ell m}\rho^{\ell} L_{n+\ell}^{2\ell+1}(\rho)e^{-\rho/2}, \quad \rho = 2r/n. \tag{4.2}$$

Here, $N_{\ell m}$ is a (more or less) irrelevant normalization constant. Useful conversion functions are

$$r(x, y, z) = \sqrt{x^2 + y^2 + z^2}, \tag{4.3}$$

$$x = r\sin\theta\cos\phi, \quad y = r\sin\theta\sin\phi, \quad z = r\cos\theta. \tag{4.4}$$

You can look up the spherical harmonics $Y_{\ell m}$ on, for example, Wikipedia.

In particular, the $1s$ and $2p_z$ functions are given by

$$\psi_{1s}(x, y, z) = \frac{1}{\sqrt{\pi}} e^{-r(x,y,z)}, \tag{4.5}$$

$$\psi_{2p_z}(x, y, z) = \frac{1}{4\sqrt{2\pi}} z e^{-r(x,y,z)/2}, \tag{4.6}$$

i) Write a Python program to draw the level *surfaces* of hydrogen orbitals, $\{(x, y, z) \mid \psi(x, y, z) = c\}$. Choose interesting values of c. It can be interesting to color code surfaces of opposite positive and negative c.

j) Write a program to visualize the *sections* of the graph of the orbitals. Use for example the xz plane for varying fixed values of y_{const}, and plot $(x, y) \mapsto f(x, y_{\text{const}}, z)$.

2 Open and closed sets in Euclidean space

a) Show that any ε-ball is open.

b) Show that $A \cup \partial A$ is closed.

c) Show that if $A \subset \mathbb{R}^n$ is closed, then if a sequence $\mathbf{x}_i \in A$ converges to some $\mathbf{x} \in \mathbb{R}^n$, then $\mathbf{x} \in A$. [This is an alternative definition of A being closed]

d) Show that the closure of A, the smallest closed set $\bar{A}$ that contains A, is equal to $A \cup \partial A$.

e) Show that

$$A = \{(x, 0) \mid x \in [-1, 1]\} \subset \mathbb{R}^2\} \tag{4.7}$$

is closed.

f) Show that $\mathbb{Q} \subset \mathbb{R}$ neither open nor closed. Show that the boundary of $\mathbb{Q}$ is $\mathbb{R}$. Show that the interior of $\mathbb{Q}$ is empty. What is the closure of $\mathbb{Q}$?

g) Let

$$A = \{(x, y) \in \mathbb{R}^2 \mid 0 \le x < 1, \quad 0 \le y < 1\}. \tag{4.8}$$

Compute the interior of A, the boundary of A, the closure of A.

3 Continuity

a) RECOMMENDED Let $f : \mathbb{R} \to \mathbb{R}$ be given by

$$f(x) = \begin{cases} 1 & x > 0 \\ -1 & x \leq 0 \end{cases} \tag{4.9}$$

For which $x \in \mathbb{R}$ does the limit

$$\lim_{h \to 0} f(x+h) \tag{4.10}$$

not exist? Why?

b) Show that any polynomial $p : \mathbb{R} \to \mathbb{R}$ is continuous, by using the theorem on properties of continuous functions.

c) (Marsden and Tromba) Consider the function

$$f(x,y) = \frac{\sin(x^2+y^2)}{x^2+y^2}. \tag{4.11}$$

This function is defined on $\mathbb{R}^2 \setminus \{(0,0)\}$. Determine whether $f(x,y)$ approaches some value as $(x,y) \to (0,0)$. What is this value?

d) RECOMMENDED (Marsden and Tromba) Does $\lim_{(x,y)\to(0,0)} x^2/(x^2+x^2)$ exist? If you want, plotting the function can help.

e) (Marsden and Tromba) Prove that $\lim_{(x,y)\to(0,0)} 2x^2y/(x^2+y^2) = 0$ using an ε-δ argument.

4 Differentiability

a) Compute the partial derivatives of

$$f(x,y) = \frac{xy}{(x^2+y^2)^{1/2}}. \tag{4.12}$$

5 Taylor polynomials

a) Consider the *matrix-valued function*

$$f(t) = e^{-tA} B e^{tA}, \tag{4.13}$$

where A and B are square matrices. The matrix exponential is defined in terms of the Taylor series,

$$\exp(X) = \mathbb{1} + X + \frac{1}{2!}X^2 + \frac{1}{3!}X^3 + \cdots, \tag{4.14}$$

which is convergent. It is a fact that if $XY - YX = [X,Y] = 0$ (commuting matrices) then $\exp(X+Y) = \exp(X)\exp(Y)$, by the same proof as for the exponential function over a field $\mathbb{F}$.

b) Explain why we may view f as a vector-valued function

c) Show that $t \mapsto e^{tA}$ is differentiable.

d) Show that f is differentiable. Compute its derivatives to arbitrary order

6 Newton--Rhapson method

In the next exercise, we derive the Newton–Rhapson method for root finding.

let $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}^n$ be a function of class C^1. We seek $\mathbf{x} \in \Omega$ such that $f(\mathbf{x}) = 0$. Assume we have an initial guess $\mathbf{x}_0 \in \Omega$.

e) Explain why the following polynomial is a good approximation to $f(\mathbf{x})$ near $\mathbf{x}_0$:

$$p(\mathbf{x}) = f(\mathbf{x}_0) + J(\mathbf{x} - \mathbf{x}_0), \quad J = Df(\mathbf{x}_0). \tag{4.15}$$

f) Newton's method now finds a new guess $\mathbf{x}_1$ by finding the root of the Taylor polynomial, i.e., $p(\mathbf{x}_1) = 0$. Write down a formula for $\mathbf{x}_1$, assuming that the matrix J has an inverse.

g) Suppose the true solution is $\mathbf{x}_* \in \Omega$, and assume that the error in $\mathbf{x}_0$ is sufficiently small. Find an estimate for the error of $\mathbf{x}_1$. Assuming that the error can be converted to significant digits n, what is the number of significant digits in $\mathbf{x}_1$? Hint: Look at the remainder term.

h) Suppose successive iterations are performed. How many digits do you have after k iterations?

i) Make a Python implementation of Newton's method and test it on the following function:

$$f(x, y) = (e^x y^3 - 1, y^2 - \sin(x) - 1). \tag{4.16}$$

It can be an idea to try and visualize the function. Two roots are:

$$\{(4.38277027, 0.23202192), \ (0, 1)\} \tag{4.17}$$

Try different initial guesses $\mathbf{x}_0$, close to the roots and further away. How many iterations do you need to achieve machine precision for your guesses? Can you find more roots?

Exercise set 5: Complex analysis

1 Basic calculations

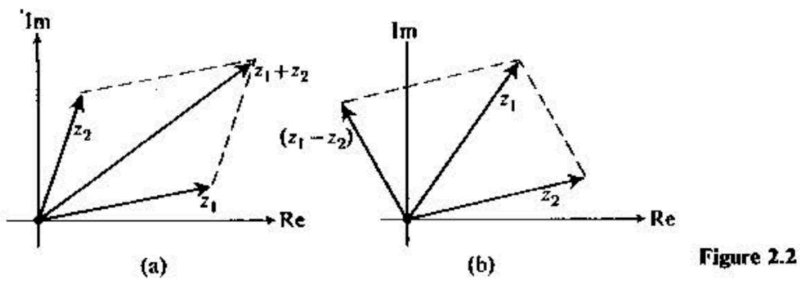

Figure 5.1: The parallelogram rule. (From Butkov.)

a) RECOMMENDED Let $x = x + iy \in \mathbb{C}$. Illustrate the addition $w = z + \bar{z}$ in a coordinate system using the parallelogram rule.

b) Recall our earlier exercise, where $\mathbb{C}$ was regarded as the real vector space $\mathbb{R}^2$. Let $z_i = x_i + iy_i$, and denote the corresponding vectors in $\mathbb{R}^2$ by $\mathbf{v}_i$. Show that

$$\langle \mathbf{v}_1, \mathbf{v}_2 \rangle = \mathrm{Re}(\overline{z_1} z_2) = \mathrm{Re}(z_1 \overline{z}_2). \tag{5.1}$$

c) Another product in $\mathbb{R}^2$ (and indeed $\mathbb{R}^3$) is the *cross product*:

$$\mathbf{v}_1 \times \mathbf{v}_2 = \begin{vmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{vmatrix}, \tag{5.2}$$

the matrix determinant. Show that

$$\mathbf{v}_1 \times \mathbf{v}_2 = \mathrm{Im}(\overline{z_1} z_2) = -\mathrm{Im}(z_1 \overline{z}_2). \tag{5.3}$$

d) Show that the function $z \mapsto iz$ is a counterclockwise rotation by $\pi/2$. Find the function that performs a clockwise rotation by the same angle.

2 Complex functions

In this section, we give some relatively easy exercises to illustrate the important concepts of complex analytic functions.

2.1 The complex exponential function

In this exercise, we derive the complex exponential function. The starting point is the real exponential function, which is the unique function $f(x) = \exp(x)$ that satisfies

$$f(x_1 + x_2) = f(x_1)f(x_2), \quad f(0) = 1. \tag{5.4}$$

We desire to generalize $\exp(x)$ to the complex plane, so that $\exp(x + 0\mathrm{i}) = \exp(x) \in \mathbb{R}$.

To avoid confusion, we call this unknown function $f : \mathbb{C} \to \mathbb{C}$.

a) Show that $f(x + \mathrm{i}y) = \exp(x)f(\mathrm{i}y)$. (Thus, we need to determine $f(\mathrm{i}y)$.)

b) We set $f(\mathrm{i}y) = A(y) + \mathrm{i}B(y)$. Show that

$$A(y) = B'(y), \quad B(y) = -A'(y), \tag{5.5}$$

and hence that $A''(y) = -A(y)$. Hint: Cauchy–Riemann.

c) The general solution to the ODE $A'' = -A$ is

$$A(y) = \alpha \cos(t) + \beta \sin(y), \tag{5.6}$$

where α and β are constants to be determined. Show that $\alpha = 1$ and $\beta = 0$ are the only constants compatible with $f(z)$ being a generalization of the real exponential function.

d) Conclude that

$$f(x + \mathrm{i}y) = e^x(\cos y + \mathrm{i} \sin y) \tag{5.7}$$

e) Show that $(\exp(z))' = \exp(z)$, using Cauchy–Riemann, and that $\exp(z)$ is everywhere analytic.

f) Show that the equation

$$\exp(z) = w \tag{5.8}$$

has infinitely many solutions for any $w \neq 0$.

g) Show that $\exp(z) \neq 0$.

2.2 Taylor series

a) Consider the Taylor series

$$\sum_{n=0}^{\infty} z^n = 1 + z + z^2 + \ldots \longrightarrow \frac{1}{1 - z}. \tag{5.9}$$

Show that the Nth partial sum is

$$\sum_{n=1}^{N} \frac{1 - z^{N+1}}{1 - z}. \tag{5.10}$$

Here, the polynomial long division algorithm is useful.

b) Find an upper bound for $|z|$ such that the partial sum converges to $1/(1 - z)$.

c) The Taylor series for the exponential function is

$$\sum_{n=0}^{\infty} \frac{1}{n!} z^n. \tag{5.11}$$

Using the definition of the complex exponential, find the Taylor series for $\cos(x)$ and $\sin(x)$, for $x \in \mathbb{R}$.

d) Make plots of the partial sums of the sine and cosine Taylor series to verify your result.

e) Find the Taylor series for the function $f(z) = (e^z - 1)/z$

f) Find the Taylor series for the function $f(z) = \sin(z)/z$ for $z \neq 0$ and $f(0) = 1$.

2.3 Singularities

a) RECOMMENDED Consider the function $f(z) = \frac{1}{z(z-1)}$. How many poles does f have, and where are they? What are the order of the poles? Can you write down the Laurent series of f around $z = 0$?

b) Does the function $f(z) = 1/\sin(z)$ have a pole? Where? What order?

c) Find the singularities of $f(z) = e^{-1/(z-1)^2}$.

2.4 Line integrals

Let

$$f(z) = \frac{1}{z^3 - (2+\mathrm{i})z^2 + (1+2\mathrm{i})z - \mathrm{i}} \tag{5.12}$$

a) Let $\Gamma_\varepsilon(w)$ be the simple closed curve defined by the counter-clockwise border of the ε-ball $B_\varepsilon(w)$. Write down a parameterization (a path) for this curve.

b) Identify the poles $P = \{w_1, w_2, \cdots\}$ of $f(z)$ and their order. Hint: A zero of the denominator is $z = \mathrm{i}$.

c) What is the value of

$$\oint_{\Gamma_\varepsilon(0)} f(z)\,\mathrm{d}z \quad ? \tag{5.13}$$

fpr $\varepsilon = 1/2$?

d) Find the value for the line integrals,

$$\oint_{\Gamma_\varepsilon(w_i)} f(z)\,\mathrm{d}z, \tag{5.14}$$

for $\varepsilon = 1/2$. Hint: Don't attempt the integral directly, but instead compute consider the Laurent series. It can be useful to use the Taylor expansion

$$\frac{1}{a-z} = \frac{1}{a}\frac{1}{1-z/a} = \frac{1}{a}\sum_{n=0}^{\infty}(z/a)^n. \tag{5.15}$$

3 Complex step method

FOR THE CURIOUS

Let $f(x)$ be a real-valued function, assumed to be analytic near some x, i.e., it agrees with its Taylor series in the vicinity of x.

a) Explain why there exists a complex analytic function near $x + 0\mathrm{i}$ that agrees with f on the real axis.

A classical method for computing a numerical derivative of a C^1 function is the finite difference approximation:

$$f'(x) \approx \delta_h f(x) := \frac{f(x+h) - f(x-h)}{2h}. \tag{5.16}$$

The error is $O(h^2)$. In the rest of the exercise, let f be given by

$$f(x) = \exp(x), \tag{5.17}$$

and we wish to differentiate around $x = 1$.

b) Write a small program that uses standard double precision floating point arithmetic to compute the finite difference derivative for step lengths from the interval $h \in [10^{-9}, 10^{-1}]$.

c) Make a plot of the absolute value of the error in the approximation as function of h. Use log scales.

The *complex step method* utilizes complex analyticity to avoid the cancellation errors seen in the finite difference scheme. It relies on the function in question being implemented/implementable using complex arithmetic.

d) Show that

$$f(x + \mathrm{i}h) = f(x) + \mathrm{i}hf'(x) - \frac{1}{2}h^2 f''(x) - \frac{\mathrm{i}h^3}{6} f'''(x) + O(h^4). \tag{5.18}$$

e) Solve for $f'(x)$, and show that

$$f'(x) = \operatorname{Im} \frac{f(x + \mathrm{i}h)}{h} + O(h^2). \tag{5.19}$$

This is the complex step method.

f) In the plot from above, add a plot of the error of the complex step method. Compare.