Mathematical Techniques III

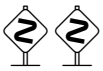
José Figueroa-O'Farrill
⟨j.m.figueroa@qmw.ac.uk⟩

Version of December 5, 2004

Preface

This the first draft of the Lecture Notes for Mathematical Techniques III (PHY 317), a course offered in the Physics Department of Queen Mary and Westfield College (University of London). These notes are loosely based on pre-existing notes by Professor John Charap. The notes contain all that is said in Lecture and sometimes more. The extra bits are typeset in smaller font and are adorned with one or two "dangerous bend" signs as in the next paragraphs.

Most paragraphs like this fill gaps in the main presentation (e.g., proofs, mathematical remarks,...). They contain material which, although necessary for the logical coherence of the presentation, may be skipped at a first reading or ignored by the less mathematically inclined student who is not interested in proofs,.... They are not an essential part of the course, although I believe they are an essential part of the topic.



Most paragraphs like this contain material which is generally more advanced than the rest of the lectures, but which I personally find interesting and have found useful at one time or other. They are not an essential part of the course, but I have included them in the hope that some of you might find them interesting enough to make the detour.

Some remarks about notation. Terms which are being defined for the first time appear in **bold sans-serif** type. Although the notation will be introduced as we go, here is a summary of the main notational conventions:

- $\mathbb{R}$ and $\mathbb{C}$ stand for the sets of real and complex numbers, respectively;
- vector spaces, subspaces,... are denoted by so-called "blackboard bold" uppercase Latin letters: $\mathbb{V}$, $\mathbb{W}$,...;
- abstract vectors are denoted by bold lowercase Latin letters: $\boldsymbol{v}$, $\boldsymbol{w}$,...;
- linear maps are denoted by uppercase Latin letters A, B,..., except for the identity map which is denoted $\mathbb{1}$.
- column vectors are denoted by sans-serif lowercase Latin letters: v, w,...;

- matrices are denoted by sans-serif uppercase Latin letters: A, B,.... The identity matrix will be denoted I.

The notes are not yet complete: in particular many of the asides are still to be completed, and the introductions have to be rewritten in light of what they are meant to introduce: they were written in advance in most cases. Many diagrams are missing, and many more examples and applications need to be added. The next stage in the development of the notes will consist in some changes in the visual layout, to break the monotony of the present style, and to make the exercises and the problems an integral part of the notes. The solutions, of course, will be available separately.

Contents

Chapter 1

Linear Algebra

In this part of the course we will review some basic linear algebra. The topics covered include: real and complex vector spaces and linear maps, bases, matrices, inner products, eigenvalues and eigenvectors. We start from the familiar setting in two dimensions and introduce the necessary formalism to be able to work with vectors in an arbitrary number of dimensions. We end the chapter with a physical application: the study of normal modes of an oscillatory system.

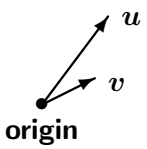
1.1 Vector spaces

Physics requires both *scalar* quantities like mass, temperature, charge which are uniquely specified by its magnitude in some units, e.g., $300°K$, 7 Kg,... and also *vectorial* quantities like velocity, force, angular momentum, which are specified both by a magnitude and a direction.

In the first part of the course we will study the general features shared by these vectorial quantities. As this is a course in mathematical techniques, we must abstract what these quantities have in common (the 'mathematical' part) while at the same time keeping a pragmatic perspective throughout (the 'techniques' part). This is not a mathematics course, but nevertheless a certain amount of formalism is needed. Some of you may not have seen formal definitions before, so we will start by motivating the notion of a vector space. For definiteness we will consider displacements in two dimensions; that is, in the plane.

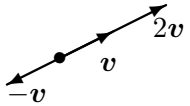
1.1.1 Displacements in the plane

Every displacement in the plane has an initial or starting point and a final point. We will only consider displacements which have a common starting point: the **origin**.



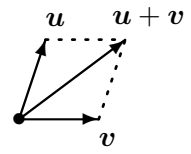
Any point in the plane is then understood as the final point of a displacement from the origin. We will depict such displacements by an arrow starting at the origin and ending at the final point. We will denote such displacements by boldfaced letters, like $\boldsymbol{u}$, $\boldsymbol{v}$. In lecture it is hard to write in boldface, so we use the notation $\vec{u}$, $\vec{v}$ which is not just easier to write but has the added benefit of being mnemonic, since the arrow reminds us that it is a displacement. We will say that displacements like $\boldsymbol{u}$, $\boldsymbol{v}$ are **vectors**.

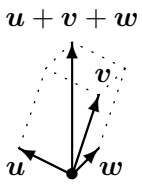
What can one do with vectors?



For example, vectors can be multiplied by real numbers (the **scalars**). If $\lambda > 0$ is a positive real number and $\boldsymbol{v}$ is a vector, then $\lambda\,\boldsymbol{v}$ is a vector pointing in the same direction as $\boldsymbol{v}$ but λ times as long as $\boldsymbol{v}$, e.g., $2\boldsymbol{v}$ is twice as long as $\boldsymbol{v}$ but points in the same direction. In the same manner, $-\lambda\boldsymbol{v}$ is a vector pointing in the direction opposite to $\boldsymbol{v}$ but λ times as long as $\boldsymbol{v}$. We call this operation **scalar multiplication**. This operation satisfies two properties which are plain to see from the pictures. The first says that if $\boldsymbol{v}$ is any vector and λ and μ are real numbers, then $\lambda\,(\mu\,\boldsymbol{v}) = (\lambda\mu)\,\boldsymbol{v}$. The second property is totally obvious from the picture: $1\,\boldsymbol{v} = \boldsymbol{v}$.

You should also be familiar from the study of, say, forces, with the fact that vectors can be added.





Indeed, if $\boldsymbol{u}$ and $\boldsymbol{v}$ are vectors, then their sum $\boldsymbol{u} + \boldsymbol{v}$ is the diagonal from the origin to the opposite vertex in the parallelogram defined by $\boldsymbol{u}$ and $\boldsymbol{v}$, as in the picture. This operation is called **vector addition** or simply **addition**. It follows from the picture that $\boldsymbol{u} + \boldsymbol{v} = \boldsymbol{v} + \boldsymbol{u}$, so that we get the same result regardless of the order in which we add the vectors. One says that *vector addition is commutative*.

Vector addition is also *associative*. This means that, as can be seen in the picture, when adding three vectors $\boldsymbol{u}$, $\boldsymbol{v}$, and $\boldsymbol{w}$ it does not matter whether we first add $\boldsymbol{u}$ and $\boldsymbol{v}$ and add $\boldsymbol{w}$ to the result: $(\boldsymbol{u} + \boldsymbol{v}) + \boldsymbol{w}$ or whether we first add $\boldsymbol{v}$ and $\boldsymbol{w}$ and add the result to $\boldsymbol{u}$: $\boldsymbol{u} + (\boldsymbol{v} + \boldsymbol{w})$.

Another easy property of vector addition is the existence of a vector $\mathbf{0}$ such that when added to any vector $\boldsymbol{v}$ gives back $\boldsymbol{v}$ again; that is,

$$\mathbf{0} + \boldsymbol{v} = \boldsymbol{v} \text{ for all vectors } \boldsymbol{v}.$$

Clearly the zero vector $\mathbf{0}$ corresponds to the trivial displacement which starts and ends at the origin, or in other words, to no displacement at all.

Similarly, given any vector $\boldsymbol{v}$ there is a vector $-\boldsymbol{v}$ which obeys $\boldsymbol{v}+(-\boldsymbol{v}) = \mathbf{0}$. We will often employ the notation $\boldsymbol{u} - \boldsymbol{v}$ to denote $\boldsymbol{u} + (-\boldsymbol{v})$.

Finally, notice that scalar multiplication and addition are compatible: scalar multiplication and addition can be performed in any order:

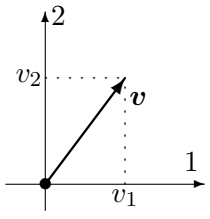
$$\lambda\,(\boldsymbol{u} + \boldsymbol{v}) = \lambda\,\boldsymbol{u} + \lambda\,\boldsymbol{v} \qquad \text{and} \qquad (\lambda + \mu)\,\boldsymbol{v} = \lambda\,\boldsymbol{v} + \mu\,\boldsymbol{v}\ .$$

The former identity says that *scalar multiplication is distributive over vector addition*. Notice that, in particular, it follows that $0\,\boldsymbol{v} = \mathbf{0}$ for all $\boldsymbol{v}$.

1.1.2 Displacements in the plane (revisited)

There is no conceptual reason why one should not consider displacements in space, i.e., in three dimensions, as opposed to the plane. The pictures get a little harder to draw, but in principle it can still be done with better draughtsmanship than mine. In physics, though, one needs to work with vectors in more than three dimensions—in fact, as in Quantum Mechanics, one often needs to work with vectors in an infinite number of dimensions. Pictures like the ones above then become of no use, and one needs to develop a notation we can calculate with.

Let us consider again the displacements in the plane, but this time with a more algebraic notation.



The first thing we do is to draw two cartesian axes centred at the origin: axis 1 and axis 2. Then every displacement $\boldsymbol{v}$ from the origin can be written as an ordered pair (v_1, v_2) of real numbers, corresponding to the components of the displacement $\boldsymbol{v}$ along the cartesian axes, as in the figure.

Let us define the set

$$\mathbb{R}^2 = \{(v_1, v_2) \mid v_i \in \mathbb{R} \text{ for } i = 1, 2\}$$

of ordered pairs of real numbers.

The above notation may need some explaining. The notation '$v_i \in \mathbb{R}$' is simply shorthand for the phrase 'v_i is a real number;' whereas the notation

'$\{(v_1, v_2) \mid v_i \in \mathbb{R} \text{ for } i = 1, 2\}$' is shorthand for the phrase 'the set consisting of pairs (v_1, v_2) such that both v_1 and v_2 are real numbers.'

The set $\mathbb{R}^2$ is in one-to-one correspondence with the set of displacements, for clearly every displacement gives rise to one such pair and every such pair gives rise to a displacement. We can therefore try to guess how to define the operations of vector addition and scalar multiplication in $\mathbb{R}^2$ in such a way that they correspond to the way they are defined for displacements.

From the pictures defining addition and scalar multiplication, one sees that if $\lambda \in \mathbb{R}$ is a real number, then

$$\lambda\,(v_1, v_2) = (\lambda\, v_1, \lambda\, v_2)\ , \qquad \text{(scalar multiplication)}$$

and also

$$(u_1, u_2) + (v_1, v_2) = (u_1 + v_1, u_2 + v_2)\ . \qquad \text{(addition)}$$

The zero vector corresponds with no displacement at all, hence it is given by the pair corresponding to the origin $(0, 0)$. It follows from the addition rule that

$$(0, 0) + (v_1, v_2) = (v_1, v_2)\ .$$

Similarly, $-(v_1, v_2) = (-v_1, -v_2)$. In fact it is not hard to show (do it!) that addition and scalar multiplication obey the same properties as they did for displacements.

The good thing about this notation is that there is no reason why we should restrict ourselves to *pairs*. Indeed, why not consider the set

$$\mathbb{R}^N = \{(v_1, v_2, \cdots, v_N) \mid v_i \in \mathbb{R} \text{ for } i = 1, 2, \ldots, N\}\ ,$$

of ordered *N-tuples* of real numbers? We can define addition and scalar multiplication in the same way as above:

(addition)

$$(u_1, u_2, \ldots, u_N) + (v_1, v_2, \ldots, v_N) \\ = (u_1 + v_1, u_2 + v_2, \ldots, u_N + v_N)\ ,$$

(multiplication by scalars)

$$\lambda\,(v_1, v_2, \ldots, v_N) = (\lambda\, v_1, \lambda\, v_2, \ldots, \lambda\, v_N) \quad \text{for } \lambda \in \mathbb{R}.$$

In the homework you are asked to prove that these operations on $\mathbb{R}^N$ obey the same properties that displacements do: commutativity, associativity, distributivity,... These properties can be formalised in the concept of an abstract vector space.

1.1.3 Abstract vector spaces

We are finally ready to formalise the observations made above into the definition of an abstract vector space. We say that this is an abstract vector space, because it does not refer to any concrete example.

A **real vector space** consists of the following data:

- Two sets:
  - the set of vectors, which we shall denote $\mathbb{V}$, and whose elements we will write as $\boldsymbol{u}$, $\boldsymbol{v}$, $\boldsymbol{w}$, ..., and
  - the set of scalars, which for a real vector space is simply the set $\mathbb{R}$ of real numbers. We will use lowercase Greek letters from the middle of the alphabet: λ, μ, ... to represent real numbers.
- Two operations:
  - Scalar multiplication, which takes a scalar λ and a vector $\boldsymbol{v}$ and produces another vector $\lambda\,\boldsymbol{v}$. One often abbreviates this as

$$\begin{aligned} \text{scalar multiplication} : \mathbb{R} \times \mathbb{V} &\to \mathbb{V} \\ (\lambda, \boldsymbol{v}) &\mapsto \lambda\,\boldsymbol{v} \ . \end{aligned}$$

  - Vector addition, which takes two vectors $\boldsymbol{u}$ and $\boldsymbol{v}$ and produces a third vector denoted $\boldsymbol{u} + \boldsymbol{v}$. Again one can abbreviate this as

$$\begin{aligned} \text{vector addition} : \mathbb{V} \times \mathbb{V} &\to \mathbb{V} \\ (\boldsymbol{u}, \boldsymbol{v}) &\mapsto \boldsymbol{u} + \boldsymbol{v} \ . \end{aligned}$$

- Eight properties (or *axioms*):

  V1 (*associativity*) $(\boldsymbol{u} + \boldsymbol{v}) + \boldsymbol{w} = \boldsymbol{u} + (\boldsymbol{v} + \boldsymbol{w})$ for all $\boldsymbol{u}$, $\boldsymbol{v}$ and $\boldsymbol{w}$;

  V2 (*commutativity*) $\boldsymbol{u} + \boldsymbol{v} = \boldsymbol{v} + \boldsymbol{u}$ for all $\boldsymbol{u}$ and $\boldsymbol{v}$;

  V3 There exists a zero vector $\boldsymbol{0}$ which obeys $\boldsymbol{0} + \boldsymbol{v} = \boldsymbol{v}$ for all $\boldsymbol{v}$;

  V4 For any given $\boldsymbol{v}$, there exists a vector $-\boldsymbol{v}$ such that $\boldsymbol{v} + (-\boldsymbol{v}) = \boldsymbol{0}$;

  V5 $\lambda\,(\mu\,\boldsymbol{v}) = (\lambda\,\mu)\,\boldsymbol{v}$ for all $\boldsymbol{v}$, λ and μ;

  V6 $1\,\boldsymbol{v} = \boldsymbol{v}$ for all $\boldsymbol{v}$;

  V7 $(\lambda + \mu)\,\boldsymbol{v} = \lambda\,\boldsymbol{v} + \mu\,\boldsymbol{v}$ for all λ and μ and $\boldsymbol{v}$;

  V8 (*distributivity*) $\lambda\,(\boldsymbol{u} + \boldsymbol{v}) = \lambda\,\boldsymbol{u} + \lambda\,\boldsymbol{v}$ for all λ, $\boldsymbol{u}$ and $\boldsymbol{v}$.

This formidable looking definition might at first seem to be something you had rather forget about. Actually you will see that after using it in practice it will become if not intuitive at least more sensible. Formal definitions like this one above are meant to capture the essence of what is being defined. Every vector space is an instance of an abstract vector space, and it will inherit all the properties of an abstract vector space. In other words, we can be sure that any result that we obtain for an abstract vector space will also hold for any concrete example.

A typical use of the definition is recognising vector spaces. To go about this one has to identify the sets of vectors and scalars, and the operations of scalar multiplication and vector addition and then check that all eight axioms are satisfied. In the homework I ask you to do this for two very different looking spaces: $\mathbb{R}^N$ which we have already met, and the set consisting of real-valued functions on the interval $[-1, 1]$. In the course of these lectures we will see many others.

You may wonder whether all eight axioms are necessary. For example, you may question the necessity of V4, given V3. Consider the following subset of $\mathbb{R}^2$:

$$\{(v_1, v_2) \mid v_i \in \mathbb{R} \text{ and } v_2 \geq 0\} \subset \mathbb{R}^2$$

consisting of pairs of real numbers where the second real number in the pair is non-negative. In terms of displacements, it corresponds to the upper half-plane. You can check that the first two axioms V1 and V2 are satisfied, and that the zero vector $(0, 0)$ belongs to this subset. However $-(v_1, v_2) = (-v_1, -v_2)$ whence if v_2 is non-negative, $-v_2$ cannot be non-negative unless $v_2 = 0$. Therefore V4 is not satisfied. In fact, neither are V5, V7 and V8 unless we restrict the scalars to be non-negative real numbers. A more challenging exercise is to determine whether V6 is really necessary.

The zero vector $\mathbf{0}$ of axiom V3 is unique. To see this notice that if there were another $\mathbf{0}'$ which also satisfies V3, then

$$\begin{aligned} \mathbf{0}' &= \mathbf{0} + \mathbf{0}' && \text{(by V3 for } \mathbf{0}) \\ &= \mathbf{0}' + \mathbf{0} && \text{(by V2)} \\ &= \mathbf{0} \ . && \text{(by V3 for } \mathbf{0}') \end{aligned}$$

Similarly the vector $-\boldsymbol{v}$ in V4 is also unique. In fact, suppose that there are two vectors $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$ which satisfy: $\boldsymbol{v} + \boldsymbol{u}_1 = \mathbf{0}$ and $\boldsymbol{v} + \boldsymbol{u}_2 = \mathbf{0}$. Then they are equal:

$$\begin{aligned} \boldsymbol{u}_1 &= \mathbf{0} + \boldsymbol{u}_1 && \text{(by V3)} \\ &= (\boldsymbol{v} + \boldsymbol{u}_2) + \boldsymbol{u}_1 && \text{(by hypothesis)} \\ &= \boldsymbol{v} + (\boldsymbol{u}_2 + \boldsymbol{u}_1) && \text{(by V1)} \\ &= \boldsymbol{v} + (\boldsymbol{u}_1 + \boldsymbol{u}_2) && \text{(by V2)} \\ &= (\boldsymbol{v} + \boldsymbol{u}_1) + \boldsymbol{u}_2 && \text{(by V1)} \\ &= \mathbf{0} + \boldsymbol{u}_2 && \text{(by hypothesis)} \\ &= \boldsymbol{u}_2 \ . && \text{(by V3)} \end{aligned}$$

A final word on notation: although we have defined a real vector space as two sets, vectors $\mathbb{V}$ and real scalars $\mathbb{R}$, and two operations satisfying some axioms, one often simply says that '$\mathbb{V}$ is a real vector space' leaving the other bits in the definition implicit. Similarly in what follows, and unless otherwise stated, we will implicitly assume that the scalars are real, so that whenever we say '$\mathbb{V}$ is a vector space' we shall mean that $\mathbb{V}$ is a real vector space.

1.1.4 Vector subspaces

A related notion to a vector space is that of a vector subspace. Suppose that $\mathbb{V}$ is a vector space and let $\mathbb{W} \subset \mathbb{V}$ be a subset. This means that $\mathbb{W}$ consists of some (but not necessarily all) of the vectors in $\mathbb{V}$. Since $\mathbb{V}$ *is* a vector space, we know that we can add vectors in $\mathbb{W}$ and multiply them by scalars, but does that make $\mathbb{W}$ into a vector space in its own right? As we saw above with the example of the upper half-plane, not every subset $\mathbb{W}$ will itself be a vector space. For this to be the case we have to make sure that the following two axioms are satisfied:

S1 If $\boldsymbol{v}$ and $\boldsymbol{w}$ are vectors in $\mathbb{W}$, then so is $\boldsymbol{v} + \boldsymbol{w}$; and

S2 For any scalar $\lambda \in \mathbb{R}$, if $\boldsymbol{w}$ is any vector in $\mathbb{W}$, then so is $\lambda\,\boldsymbol{w}$.

If these two properties are satisfied we say that $\mathbb{W}$ is a **vector subspace** of $\mathbb{V}$. One also often sees the phrases '$\mathbb{W}$ is a subspace of $\mathbb{V}$' and '$\mathbb{W}$ is a linear subspace of $\mathbb{V}$.'

Let us make sure we understand what these two properties mean. For $\boldsymbol{v}$ and $\boldsymbol{w}$ in $\mathbb{W}$, $\boldsymbol{v} + \boldsymbol{w}$ belongs to $\mathbb{V}$ because $\mathbb{V}$ is a vector space. The question is whether $\boldsymbol{v} + \boldsymbol{w}$ belongs to $\mathbb{W}$, and S1 says that it does. Similarly, if $\boldsymbol{w} \in \mathbb{W}$ is a vector in $\mathbb{W}$ and $\lambda \in \mathbb{R}$ is any scalar, then $\lambda\,\boldsymbol{w}$ belongs to $\mathbb{V}$ because $\mathbb{V}$ is a vector space. The question is whether $\lambda\,\boldsymbol{w}$ also belongs to $\mathbb{W}$, and S2 says that it does.

You may ask whether we should not also require that the zero vector $\mathbf{0}$ also belongs to $\mathbb{W}$. In fact this is guaranteed by S2, because for any $\boldsymbol{w} \in \mathbb{W}$, $\mathbf{0} = 0\,\boldsymbol{w}$ (why?) which belongs to $\mathbb{W}$ by S2. From this point of view, it is S2 that fails in the example of the upper half-plane, since scalar multiplication by a negative scalar $\lambda < 0$ takes vectors in the upper half-plane to vectors in the lower half-plane.

Let us see a couple of examples. Consider the set $\mathbb{R}^3$ of ordered triples of real numbers:

$$\mathbb{R}^3 = \{(v_1, v_2, v_3) \mid v_i \in \mathbb{R} \text{ for } i = 1, 2, 3\} \ ,$$

and consider the following subsets

- $\mathbb{W}_1 = \{(v_1, v_2, 0) \mid v_i \in \mathbb{R} \text{ for } i = 1, 2\} \subset \mathbb{R}^3$,
- $\mathbb{W}_2 = \{(v_1, v_2, v_3) \mid v_i \in \mathbb{R} \text{ for } i = 1, 2, 3 \text{ and } v_3 \geq 0\} \subset \mathbb{R}^3$, and
- $\mathbb{W}_3 = \{(v_1, v_2, 1) \mid v_i \in \mathbb{R} \text{ for } i = 1, 2\} \subset \mathbb{R}^3$.

I will leave it to you as an exercise to show that $\mathbb{W}_1$ obeys both S1 and S2 whence it is a vector subspace of $\mathbb{R}^3$, whereas $\mathbb{W}_2$ does not obey S2, and $\mathbb{W}_3$ does not obey either one. Can you think of a subset of $\mathbb{R}^3$ which obeys S2 but not S1?

1.1.5 Linear independence

In this section we will introduce the concepts of linear independence and basis for a vector space; but before doing so we must introduce some preliminary notation.

Let $\mathbb{V}$ be a vector space, $\boldsymbol{v}_1$, $\boldsymbol{v}_2$, ..., $\boldsymbol{v}_N$ nonzero vectors in $\mathbb{V}$, and λ_1, λ_2, ..., λ_N scalars, i.e., real numbers. Then the vector in $\mathbb{V}$ given by

$$\sum_{i=1}^{N} \lambda_i \, \boldsymbol{v}_i := \lambda_1 \, \boldsymbol{v}_1 + \lambda_2 \, \boldsymbol{v}_2 + \cdots + \lambda_N \, \boldsymbol{v}_N \ ,$$

is called a **linear combination** of the $\{\boldsymbol{v}_i\}$. The set $\mathbb{W}$ of *all* possible linear combinations of the $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_N\}$ is actually a vector subspace of $\mathbb{V}$, called the **linear span** of the $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_N\}$ or the vector subspace **spanned** by the $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_N\}$.

Recall that in order to show that a subset of a vector space is a vector subspace it is necessary and sufficient to show that it is closed under vector addition and under scalar multiplication. Let us check this for the subset $\mathbb{W}$ of all linear combinations of the $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_N\}$. Let $\boldsymbol{w}_1 = \sum_{i=1}^{N} \alpha_i \, \boldsymbol{v}_i$ and $\boldsymbol{w}_2 = \sum_{i=1}^{N} \beta_i \, \boldsymbol{v}_i$ be any two elements of $\mathbb{W}$. Then

$$\begin{aligned} \boldsymbol{w}_1 + \boldsymbol{w}_2 &= \sum_{i=1}^{N} \alpha_i \, \boldsymbol{v}_i + \sum_{i=1}^{N} \beta_i \, \boldsymbol{v}_i \\ &= \sum_{i=1}^{N} (\alpha_i \, \boldsymbol{v}_i + \beta_i \, \boldsymbol{v}_i) && \text{(by V2)} \\ &= \sum_{i=1}^{N} (\alpha_i + \beta_i) \ \boldsymbol{v}_i \ , && \text{(by V7)} \end{aligned}$$

which is clearly in $\mathbb{W}$, being again a linear combination of the $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_N\}$. Also, if λ

is any real number and $\boldsymbol{w} = \sum_{i=1}^{N} \alpha_i \, \boldsymbol{v}_i$ is any vector in $\mathbb{W}$,

$$\begin{aligned} \lambda \, \boldsymbol{w} &= \lambda \sum_{i=1}^{N} \alpha_i \, \boldsymbol{v}_i \\ &= \sum_{i=1}^{N} \lambda \, (\alpha_i \, \boldsymbol{v}_i) && \text{(by V8)} \\ &= \sum_{i=1}^{N} (\lambda \, \alpha_i) \, \boldsymbol{v}_i \ , && \text{(by V5)} \end{aligned}$$

which is again in $\mathbb{W}$.

A set $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_N\}$ of nonzero vectors is said to be **linearly independent** if the equation

$$\sum_{i=1}^{N} \lambda_i \, \boldsymbol{v}_i = \boldsymbol{0}$$

has only the trivial solution $\lambda_i = 0$ for all $i = 1, 2, \ldots, N$. Otherwise the $\{\boldsymbol{v}_i\}$ are said to be **linearly dependent**.

It is easy to see that if a set $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_N\}$ of nonzero vectors is linearly dependent, then one of the vectors, say, $\boldsymbol{v}_i$, can be written as a linear combination of the remaining $N-1$ vectors. Indeed, suppose that $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_N\}$ is linearly dependent. This means that the equation

$$\sum_{i=1}^{N} \lambda_i \, \boldsymbol{v}_i = \boldsymbol{0} \tag{1.1}$$

must have a nontrivial solution where at least one of the $\{\lambda_i\}$ is different from zero. Suppose, for definiteness, that it is λ_1. Because $\lambda_1 \neq 0$, we can divide equation (1.1) by λ_1 to obtain:

$$\boldsymbol{v}_1 + \sum_{i=2}^{N} \frac{\lambda_i}{\lambda_1} \, \boldsymbol{v}_i = \boldsymbol{0} \ ,$$

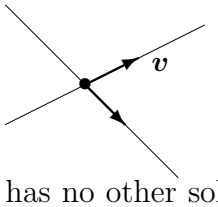
whence

$$\boldsymbol{v}_1 = -\frac{\lambda_2}{\lambda_1} \, \boldsymbol{v}_2 - \frac{\lambda_3}{\lambda_1} \, \boldsymbol{v}_3 - \cdots - \frac{\lambda_N}{\lambda_1} \, \boldsymbol{v}_N \ .$$

In other words, $\boldsymbol{v}_1$ is a linear combination of the $\{\boldsymbol{v}_2, \ldots, \boldsymbol{v}_N\}$. In general and in the same way, if $\lambda_i \neq 0$ then $\boldsymbol{v}_i$ is a linear combination of $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_{i-1}, \boldsymbol{v}_{i+1}, \ldots, \boldsymbol{v}_N\}$.

Let us try to understand these definitions by working through some examples.

We start, as usual, with displacements in the plane. Every nonzero displacement defines a line through the origin. We say that two displacements are **collinear** if they define the same line. In other words, $\boldsymbol{u}$ and $\boldsymbol{v}$ are collinear if and only if $\boldsymbol{u} = \lambda\,\boldsymbol{v}$ for some $\lambda \in \mathbb{R}$. Clearly, any two displacements in the plane are linearly independent provided they are not collinear, as in the figure.



Now consider $\mathbb{R}^2$ and let (u_1, u_2) and (v_1, v_2) be two nonzero vectors. When will they be linearly independent? From the definition, this will happen provided that the equation

$$\lambda_1\,(u_1, u_2) + \lambda_2\,(v_1, v_2) = (0, 0)$$

has no other solutions but $\lambda_1 = \lambda_2 = 0$. This is a system of linear homogeneous equations for the $\{\lambda_i\}$:

$$\begin{aligned} u_1\,\lambda_1 + v_1\,\lambda_2 &= 0 \\ u_2\,\lambda_1 + v_2\,\lambda_2 &= 0\ . \end{aligned}$$

What must happen for this system to have a nontrivial solution? It will turn out that the answer is that $u_1 v_2 = u_2 v_1$. We can see this as follows. Multiply the top equation by u_2 and the bottom equation by u_1 and subtract to get

$$(u_1 v_2 - u_2 v_1)\,\lambda_2 = 0\ ,$$

whence either $u_1 v_2 = u_2 v_1$ or $\lambda_2 = 0$. Now multiply the top equation by v_2 and the bottom equation by v_1 and subtract to get

$$(u_1 v_2 - u_2 v_1)\,\lambda_1 = 0\ ,$$

whence either $u_1 v_2 = u_2 v_1$ or $\lambda_1 = 0$. Since a nontrivial solution must have at least one of λ_1 or λ_2 nonzero, we are forced to have $u_1 v_2 = u_2 v_1$.

1.1.6 Bases

Let $\mathbb{V}$ be a vector space. A set $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots\}$ of nonzero vectors is said to be a **basis** for $\mathbb{V}$ if the following two axioms are satisfied:

- B1 The vectors $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots\}$ are linearly independent; and
- B2 The linear span of the $\{be_1, \boldsymbol{e}_2, \ldots\}$ is all of $\mathbb{V}$; in other words, any $\boldsymbol{v}$ in $\mathbb{V}$ can be written as a linear combination of the $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots\}$.

The vectors $\boldsymbol{e}_i$ in a basis are known as the **basis elements**.

There are two basic facts about bases which we mention without proof. First of all, *every vector space has a basis*, and in fact, unless it is the trivial vector space consisting only of $\mathbf{0}$, it has infinitely many bases. However not every vector space has a *finite* basis; that is, a basis with a finite number of elements. If a vector space *does* possess a finite basis $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_N\}$ then it is said to be **finite-dimensional**. Otherwise it is said to be **infinite-dimensional**. We will deal mostly with finite-dimensional vector spaces in this part of the course, although we will have the chance of meeting some infinite-dimensional vector spaces later on.

The second basic fact is that if $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_N\}$ and $\{\boldsymbol{f}_1, \boldsymbol{f}_2, \ldots, \boldsymbol{f}_M\}$ are two bases for a vector space $\mathbb{V}$, then $M = N$. In other words, every basis has the same number of elements, which is therefore an intrinsic property of the vector space in question. This number is called the **dimension** of the vector space. One says that $\mathbb{V}$ has dimension N or that it is N-dimensional. In symbols, one writes this as $\dim \mathbb{V} = N$.

From what we have said before, any two displacements which are non-collinear provide a basis for the displacements on the plane. Therefore this vector space is two-dimensional.

Similarly, any (v_1, v_2) in $\mathbb{R}^2$ can be written as a linear combination of $\{(1,0),(0,1)\}$:

$$(v_1, v_2) = v_1\,(1,0) + v_2\,(0,1)\ .$$

Therefore since $\{(1,0),(0,1)\}$ are linearly independent, they form a basis for $\mathbb{R}^2$. This shows that $\mathbb{R}^2$ is also two-dimensional.

More generally for $\mathbb{R}^N$, the set given by the N vectors

$$\{(1,0,\ldots,0),(0,1,\ldots,0),\ldots,(0,0,\ldots,1)\}$$

is a basis for $\mathbb{R}^N$, called the **canonical basis**. This shows that $\mathbb{R}^N$ has dimension N.

Let $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_p\}$ be a set of p linearly independent vectors in a vector space $\mathbb{V}$ of dimension $N \geq p$. Then they are a basis for the vector subspace $\mathbb{W}$ of $\mathbb{V}$ which they span. If $p = N$ they span the full space $\mathbb{V}$, whence they are a basis for $\mathbb{V}$. It is another basic fact that any set of linearly independent vectors can be completed to a basis.

One final remark: the property B2 satisfied by a basis guarantees that any vector $\boldsymbol{v}$ can be written as a linear combination of the basis elements, but does not say whether this can be done in more than one way. In fact, the linear combination turns out to be unique.

Let us prove this. For simplicity, let us work with a finite-dimensional vector space $\mathbb{V}$ with a basis $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_N\}$. Suppose that a vector $\boldsymbol{v} \in \mathbb{V}$ can be written as a linear

combination of the $\{\boldsymbol{e}_i\}$ in two ways:

$$\boldsymbol{v} = \sum_{i=1}^{N} v_i \boldsymbol{e}_i \quad \text{and} \quad \boldsymbol{v} = \sum_{i=1}^{N} v'_i \boldsymbol{e}_i \ .$$

We will show that $v_i = v'_i$ for all i. To see this consider

$$\begin{aligned} \mathbf{0} &= \boldsymbol{v} - \boldsymbol{v} \\ &= \sum_{i=1}^{N} v_i \boldsymbol{e}_i - \sum_{i=1}^{N} v'_i \boldsymbol{e}_i \\ &= \sum_{i=1}^{N} \ v_i - v'_i \ \ \boldsymbol{e}_i \ . \end{aligned}$$

But because of B1, the $\{\boldsymbol{e}_i\}$ are linearly independent, and by definition this means that the last of the above equations admits only the trivial solution $v_i - v'_i = 0$ for all i. The numbers $\{v_i\}$ are called the **components** of $\boldsymbol{v}$ relative to the basis $\{\boldsymbol{e}_i\}$.

Bases can be extremely useful in calculations with vector spaces. A clever choice of basis can help tremendously towards the solution of a problem, just like a bad choice of basis can make the problem seem very complicated. We will see more of them later, but first we need to introduce the second main concept of linear algebra, that of a linear map.

1.2 Linear maps

In the previous section we have learned about vector spaces by studying objects (subspaces, bases,...) living in a fixed vector space. In this section we will look at objects which relate different vector spaces. These objects are called linear maps.

1.2.1 Linear maps

Let $\mathbb{V}$ and $\mathbb{W}$ be two vector spaces, and consider a map $A : \mathbb{V} \to \mathbb{W}$ assigning to each vector $\boldsymbol{v}$ in $\mathbb{V}$ a unique vector $A(\boldsymbol{v})$ in $\mathbb{W}$. We say that A is a **linear map** (or a **homomorphism**) if it satisfies the following two properties:

L1 For all $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$ in $\mathbb{V}$, $A(\boldsymbol{v}_1 + \boldsymbol{v}_2) = A(\boldsymbol{v}_1) + A(\boldsymbol{v}_2)$; and

L2 For all $\boldsymbol{v}$ in $\mathbb{V}$ and $\lambda \in \mathbb{R}$, $A(\lambda\, \boldsymbol{v}) = \lambda\, A(\boldsymbol{v})$.

In other words, a linear map is compatible with the operations of vector addition and scalar multiplication which define the vector space; that is, it does not matter whether we apply the map A before or after performing these operations: we will get the same result. One says that 'linear maps respect addition and scalar multiplication.'

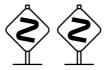
Any linear map $A : \mathbb{V} \to \mathbb{W}$ sends the zero vector in $\mathbb{V}$ to the zero vector in $\mathbb{W}$. Let us see this. (We will use the notation $\mathbf{0}$ *both* for the zero vector in $\mathbb{V}$ and for the zero vector in $\mathbb{W}$ as it should be clear from the context which one we mean.) Let $\boldsymbol{v}$ be any vector in $\mathbb{V}$ and let us apply A to $\mathbf{0} + \boldsymbol{v}$:

$$A(\mathbf{0} + \boldsymbol{v}) = A(\mathbf{0}) + A(\boldsymbol{v}) \ ; \qquad \text{(by L1)}$$

but because $\mathbf{0} + \boldsymbol{v} = \boldsymbol{v}$,

$$A(\boldsymbol{v}) = A(\mathbf{0}) + A(\boldsymbol{v}) \ ,$$

which says that $A(\mathbf{0}) = \mathbf{0}$, since the zero vector is unique.



Any linear map $A : \mathbb{V} \to \mathbb{W}$ gives rise to a vector subspace of $\mathbb{V}$, known as the **kernel** of A, and written $\ker A$. It is defined as the subspace of $\mathbb{V}$ consisting of those vectors in $\mathbb{V}$ which get mapped to the zero vector of $\mathbb{W}$. In other words,

$$\ker A := \{\boldsymbol{v} \in \mathbb{V} \mid A(\boldsymbol{v}) = \mathbf{0} \in \mathbb{W}\} \ .$$

To check that $\ker A \subset \mathbb{W}$ is really a vector subspace, we have to make sure that axioms S1 and S2 are satisfied. Suppose that $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$ belong to $\ker A$. Let us show that so does their sum $\boldsymbol{v}_1 + \boldsymbol{v}_2$:

$$\begin{aligned} A(\boldsymbol{v}_1 + \boldsymbol{v}_2) &= A(\boldsymbol{v}_1) + A(\boldsymbol{v}_2) && \text{(by L1)} \\ &= \mathbf{0} + \mathbf{0} && \text{(because } A(\boldsymbol{v}_i) = \mathbf{0}\text{)} \\ &= \mathbf{0} \ , && \text{(by V3 for } \mathbb{W}\text{)} \\ \therefore \quad \boldsymbol{v}_1 + \boldsymbol{v}_2 &\in \ker A \ . \end{aligned}$$

This shows that S1 is satisfied. Similarly, if $\boldsymbol{v} \in \ker A$ and $\lambda \in \mathbb{R}$ is any scalar, then

$$\begin{aligned} A(\lambda\, \boldsymbol{v}) &= \lambda\, A(\boldsymbol{v}) && \text{(by L2)} \\ &= \lambda\, \mathbf{0} && \text{(because } A(\boldsymbol{v}) = \mathbf{0}\text{)} \\ &= \mathbf{0} \ , && \text{(follows from V7 for } \mathbb{W}\text{)} \\ \therefore \quad \lambda\, \boldsymbol{v} &\in \ker A \ ; \end{aligned}$$

whence S2 is also satisfied. Notice that we used both properties L1 and L2 of a linear map.

There is also a vector subspace, this time of $\mathbb{W}$, associated with $A : \mathbb{V} \to \mathbb{W}$. It is called the **image** of A, and written $\operatorname{im} A$. It consists of those vectors in $\mathbb{W}$ which can be written as $A(\boldsymbol{v})$ for some $\boldsymbol{v} \in \mathbb{V}$. In other words,

$$\operatorname{im} A := \{\boldsymbol{w} \in \mathbb{W} \mid \boldsymbol{w} = A(\boldsymbol{v}) \text{ for some } \boldsymbol{v} \in \mathbb{V}\}.$$

To check that $\operatorname{im} A \subset \mathbb{W}$ is a vector subspace we must check that S1 and S2 are satisfied. Let us do this. Suppose that $\boldsymbol{w}_1$ and $\boldsymbol{w}_2$ belong to the image of A. This means that there are vectors $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$ in $\mathbb{V}$ which obey $A(\boldsymbol{v}_i) = \boldsymbol{w}_i$ for $i = 1, 2$. Therefore,

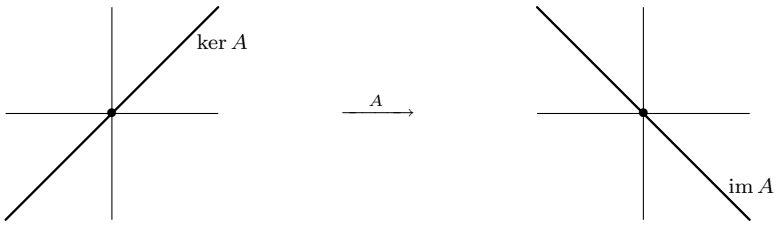
$$\begin{aligned} A(\boldsymbol{v}_1 + \boldsymbol{v}_2) &= A(\boldsymbol{v}_1) + A(\boldsymbol{v}_2) && \text{(by L1)} \\ &= \boldsymbol{w}_1 + \boldsymbol{w}_2 \ , \end{aligned}$$

whence $\boldsymbol{w}_1 + \boldsymbol{w}_2$ belong to the image of A. Similarly, if $\boldsymbol{w} = A(\boldsymbol{v})$ belongs to the image of A and $\lambda \in \mathbb{R}$ is any scalar,

$$\begin{aligned} A(\lambda\, \boldsymbol{v}) &= \lambda\, A(\boldsymbol{v}) && \text{(by L2)} \\ &= \lambda\, \boldsymbol{w} \ , \end{aligned}$$

whence $\lambda \boldsymbol{w}$ also belongs to the image of A.

As an example, consider the linear transformation $A : \mathbb{R}^2 \to \mathbb{R}^2$ defined by $(x, y) \mapsto (x - y, y - x)$. Its kernel and image are pictured below:



A linear map $A : \mathbb{V} \to \mathbb{W}$ is said to be **one-to-one** (or **injective** or a **monomorphism**) if $\ker A = \mathbf{0}$. The reason for the name is the following. Suppose that $A(\boldsymbol{v}_1) = A(\boldsymbol{v}_2)$. Then because of linearity, $A(\boldsymbol{v}_1 - \boldsymbol{v}_2) = \mathbf{0}$, whence $\boldsymbol{v}_1 - \boldsymbol{v}_2$ belongs to the kernel. Since the kernel is zero, we have that $\boldsymbol{v}_1 = \boldsymbol{v}_2$.

Similarly a linear map $A : \mathbb{V} \to \mathbb{W}$ is said to be **onto** (or **surjective** or an **epimorphism**) if $\operatorname{im} A = \mathbb{W}$, so that every vector of $\mathbb{W}$ is the image under A of some vector in $\mathbb{V}$. If this vector is unique, so that A is also one-to-one, we say that A is an **isomorphism**. If $A : \mathbb{V} \to \mathbb{W}$ is an isomorphism, one says that $\mathbb{V}$ is **isomorphic** to $\mathbb{W}$, and we write this as $\mathbb{V} \cong \mathbb{W}$. As we will see below, 'being isomorphic to' is an equivalence relation.

Notice that if $\mathbb{V}$ is an N-dimensional real vector space, any choice of basis $\{\boldsymbol{e}_i\}$ induces an isomorphism $A : \mathbb{V} \to \mathbb{R}^N$, defined by sending the vector $\boldsymbol{v} = \sum_{i=1}^N v_i \, \boldsymbol{e}_i$ to the ordered N-tuple made out from its components $(v_1, v_2, \ldots, v_N)$ relative to the basis. Therefore we see that all N-dimensional vector spaces are isomorphic to $\mathbb{R}^N$, and hence to each other.

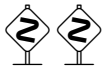
An important property of linear maps is that once we know how they act on a basis, we know how they act on any vector in the vector space. Indeed, suppose that $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_N\}$ is a basis for an N-dimensional vector space $\mathbb{V}$. Any vector $\boldsymbol{v} \in \mathbb{V}$ can be written uniquely as a linear combination of the basis elements:

$$\boldsymbol{v} = \sum_{i=1}^{N} v_i \, \boldsymbol{e}_i \ .$$

Let $A : \mathbb{V} \to \mathbb{W}$ be a linear map. Then

$$\begin{aligned} A(\boldsymbol{v}) &= A\left(\sum_{i=1}^{N} v_i \, \boldsymbol{e}_i\right) \\ &= \sum_{i=1}^{N} A\left(v_i \, \boldsymbol{e}_i\right) && \text{(by L1)} \\ &= \sum_{i=1}^{N} v_i \, A\left(\boldsymbol{e}_i\right) \ . && \text{(by L2)} \end{aligned}$$

Therefore if we know $A(\boldsymbol{e}_i)$ for $i = 1, 2, \ldots, N$ we know A on *any* vector.



The dual space.

1.2.2 Composition of linear maps

Linear maps can be composed to produce new linear maps. Let $A : \mathbb{V} \to \mathbb{W}$ and $B : \mathbb{U} \to \mathbb{V}$ be linear maps connecting three vectors spaces $\mathbb{U}$, $\mathbb{V}$ and $\mathbb{W}$. We can define a third map $C : \mathbb{U} \to \mathbb{W}$ by composing the two maps:

$$\mathbb{U} \xrightarrow{B} \mathbb{V} \xrightarrow{A} \mathbb{W} \ .$$

In other words, if $\boldsymbol{u} \in \mathbb{U}$ is any vector, then the action of C on it is defined by first applying B to get $B(\boldsymbol{u})$ and then applying A to the result to obtain $A(B(\boldsymbol{u}))$. The resulting map is written $A \circ B$, so that one has the composition rule:

$$(A \circ B)(\boldsymbol{u}) := A\,(B(\boldsymbol{u})) \ . \tag{1.2}$$

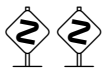
This new map is linear because B and A are, as we now show. It respects addition:

$$\begin{aligned}
(A \circ B)(\boldsymbol{u}_1 + \boldsymbol{u}_2) &= A\,(B(\boldsymbol{u}_1 + \boldsymbol{u}_2)) \\
&= A\,(B(\boldsymbol{u}_1) + B(\boldsymbol{u}_2)) && \text{(by L1 for } B\text{)} \\
&= A\,(B(\boldsymbol{u}_1)) + A\,(B(\boldsymbol{u}_2)) && \text{(by L1 for } A\text{)} \\
&= (A \circ B)(\boldsymbol{u}_1) + (A \circ B)(\boldsymbol{u}_2) \ ;
\end{aligned}$$

and it also respects scalar multiplication:

$$\begin{aligned}
(A \circ B)(\lambda\,\boldsymbol{u}) &= A\,(B(\lambda\,\boldsymbol{u})) \\
&= A\,(\lambda\,B(\boldsymbol{u})) && \text{(by L2 for } B\text{)} \\
&= \lambda\,A\,(B(\boldsymbol{u})) && \text{(by L2 for } A\text{)} \\
&= \lambda\,(A \circ B)(\boldsymbol{u}) \ .
\end{aligned}$$

Thus $A \circ B$ is a linear map, known as the **composition** of A and B. One usually reads $A \circ B$ as 'B composed with A' (notice the order!) or 'A pre-composed with B.'

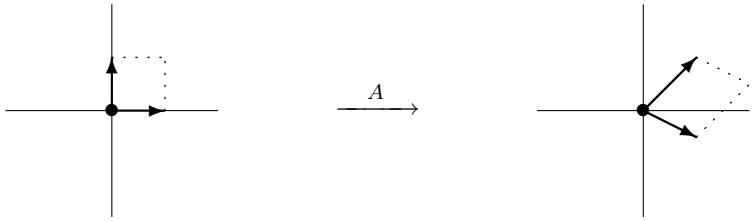


Notice that if A and B are isomorphisms, then so is $A \circ B$. In other words, composition of isomorphisms is an isomorphism. This means that if $\mathbb{U} \cong \mathbb{V}$ and $\mathbb{V} \cong \mathbb{W}$, then $\mathbb{U} \cong \mathbb{W}$, so that the property of being isomorphic is transitive. This property is also symmetric: if $A : \mathbb{V} \to \mathbb{W}$ is an isomorphism, $A^{-1} : \mathbb{W} \to \mathbb{V}$ is too, so that $\mathbb{V} \cong \mathbb{W}$ implies $\mathbb{W} \cong \mathbb{V}$. Moreover it is also reflexive, the identity map $\mathbb{1} : \mathbb{V} \to \mathbb{V}$ provides an isomorphism $\mathbb{V} \cong \mathbb{V}$. Hence the property of being isomorphic is an equivalence relation.

1.2.3 Linear transformations

An important special case of linear maps are those which map a vector space to itself: $A : \mathbb{V} \to \mathbb{V}$. These linear maps are called **linear transformations**

(or **endomorphisms**). Linear transformations are very easy to visualise in two dimensions:



A linear transformation sends the origin to the origin, straight lines to straight lines, and parallelograms to parallelograms.

Composition of two linear transformation is another linear transformation. In other words, we can think of composition of linear transformations as some sort of multiplication. This multiplication obeys a property reminiscent of the associativity V1 of vector addition. Namely, given three linear transformations A, B and C, then

$$(A \circ B) \circ C = A \circ (B \circ C) \ . \tag{1.3}$$

To see this simply apply both sides of the equation to $\boldsymbol{v} \in \mathbb{V}$ and use equation (1.2) to obtain in both cases simply $A(B(C(\boldsymbol{v})))$. By analogy, we say that composition of linear transformations is **associative**. Unlike vector addition, composition is *not* commutative; that is, in general, $A \circ B \neq B \circ A$.

Let $\mathbb{1} : \mathbb{V} \to \mathbb{V}$ denote the **identity transformation**, defined by $\mathbb{1}(\boldsymbol{v}) = \boldsymbol{v}$ for all $\boldsymbol{v} \in \mathbb{V}$. Clearly,

$$\mathbb{1} \circ A = A \circ \mathbb{1} = A \ , \tag{1.4}$$

for any linear transformations A. In other words, $\mathbb{1}$ is an **identity** for the composition of linear transformations. Given a linear transformation $A : \mathbb{V} \to \mathbb{V}$, it may happen that there is a linear transformation $B : \mathbb{V} \to \mathbb{V}$ such that

$$B \circ A = A \circ B = \mathbb{1} \ . \tag{1.5}$$

If this is the case, we say that A is **invertible**, and we call B its **inverse**. We then write $B = A^{-1}$.

The composition of two invertible linear transformations is again invertible. Indeed one has

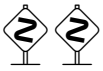
$$(A \circ B)^{-1} = B^{-1} \circ A^{-1} \ .$$

To show this we compute

$$\begin{aligned}
B^{-1} \circ A^{-1} \ \circ (A \circ B) &= B^{-1} \circ \ A^{-1} \circ (A \circ B) && \text{(by equation (1.3))} \\
&= B^{-1} \circ \ \ A^{-1} \circ A \ \circ B && \text{(by equation (1.3))} \\
&= B^{-1} \circ (\mathbb{1} \circ B) && \text{(by equation (1.5))} \\
&= B^{-1} \circ B && \text{(by equation (1.4))} \\
&= \mathbb{1} \ , && \text{(by equation (1.5))}
\end{aligned}$$

and similarly

$$\begin{aligned}
(A \circ B) \circ (B^{-1} \circ A^{-1}) &= A \circ \ B \circ \ B^{-1} \circ A^{-1} && \text{(by equation (1.3))} \\
&= A \circ \quad B \circ B^{-1} \ \circ A^{-1}) && \text{(by equation (1.3))} \\
&= A \circ \ \mathbb{1} \circ A^{-1}) && \text{(by equation (1.5))} \\
&= A \circ A^{-1} && \text{(by equation (1.4))} \\
&= \mathbb{1} \ . && \text{(by equation (1.5))}
\end{aligned}$$



This shows that the invertible transformations of a vector space $\mathbb{V}$ form a **group**, called the **general linear group** of $\mathbb{V}$ and written $\mathrm{GL}(\mathbb{V})$.

A group is a set G whose elements are called **group elements**, together with an operation called **group multiplication** and written simply as

$$\begin{aligned}
\text{group multiplication} : G \times G &\to G \\
(x, y) &\mapsto xy
\end{aligned}$$

satisfying the following three axioms:

G1 group multiplication is **associative**:

$$(xy)z = x(yz) \quad \text{for all group elements } x, y \text{ and } z.$$

G2 there exists an **identity element** $e \in G$ such that

$$ex = xe = x \quad \text{for all group elements } x.$$

G3 every group element x has an **inverse**, denoted x^{-1} and obeying

$$x^{-1}x = xx^{-1} = e \ .$$

If in addition the group obeys a fourth axiom

G4 group multiplication is **commutative**:

$$xy = yx \quad \text{for all group elements } x \text{ and } y,$$

then we say that the group is **commutative** or **abelian**, in honour of the Norwegian mathematician Niels Henrik Abel (1802-1829).

When the group is abelian, the group multiplication is usually written as a group addition: $x + y$ instead of xy. Notice that axioms V1—V4 for a vector space say that, under vector addition, a vector space is an abelian group.

Groups are extremely important objects in both mathematics and physics. It is an 'algebraic' concept, yet its uses transcend algebra; for example, it was using the theory of groups that quarks were originally postulated in particle physics. The fact that we now think of quarks as elementary particles and not simply as mathematical construct is proof of how far group theory has become a part of our description of nature at its most fundamental.

1.2.4 The vector space of linear maps

Now we point out that linear maps themselves also form a vector space! In order to do this, we have to produce the two operations: vector addition and scalar multiplication, and show that they satisfy the eight axioms V1—V8. Let A and B be linear maps $\mathbb{V} \to \mathbb{W}$, let $\lambda \in \mathbb{R}$ be a scalar, and let $\boldsymbol{v} \in \mathbb{V}$ be any vector. Then we define the two operations by

(addition)

$$(A + B)(\boldsymbol{v}) = A(\boldsymbol{v}) + B(\boldsymbol{v}) \ , \tag{1.6}$$

(scalar multiplication)

$$(\lambda\, A)(\boldsymbol{v}) = \lambda\, A(\boldsymbol{v}) \ . \tag{1.7}$$

Having defined these two operations we must check that the axioms are satisfied. We leave this as an exercise, except to note that the zero vector is the transformation which sends every $\boldsymbol{v} \in \mathbb{V}$ to $\mathbf{0} \in \mathbb{W}$. The rest of the axioms follow from the fact that $\mathbb{W}$ is a vector space.

This is a general mathematical fact: the space of functions $f : X \to Y$ always inherits whatever algebraic structures Y possesses simply by defining the operations pointwise in X.

Let $\mathcal{L}(\mathbb{V}, \mathbb{W})$ denote the vector space of linear maps $\mathbb{V} \to \mathbb{W}$. What is its dimension? We will see in the next section when we talk about matrices that its dimension is given by the product of the dimensions of $\mathbb{V}$ and $\mathbb{W}$:

$$\dim \mathcal{L}(\mathbb{V}, \mathbb{W}) = \dim \mathbb{V} \, \dim \mathbb{W} \ . \tag{1.8}$$

In particular the space $\mathcal{L}(\mathbb{V}, \mathbb{V})$ of linear transformations of $\mathbb{V}$ has dimension $(\dim \mathbb{V})^2$. We will call this space $\mathcal{L}(\mathbb{V})$ from now on.

Because $\mathcal{L}(\mathbb{V})$ is a vector space, its elements can be added and as we saw above, composition allows us to multiply them too. It turns out that these two operations are compatible:

$$A \circ (B + C) = (A \circ B) + (A \circ C) \tag{1.9}$$

$$(A + B) \circ C = (A \circ C) + (B \circ C) \ . \tag{1.10}$$

Let us prove the left and right distributivity properties. Let A, B, and C be linear transformations of a vector space $\mathbb{V}$ and let $\boldsymbol{v} \in \mathbb{V}$ be an arbitrary vector. Then

$$\begin{aligned}
(A \circ (B + C))\,(\boldsymbol{v}) &= A\,((B + C)(\boldsymbol{v})) && \text{(by equation (1.2))} \\
&= A\,(B(\boldsymbol{v}) + C(\boldsymbol{v})) && \text{(by equation (1.6))} \\
&= A(B(\boldsymbol{v})) + A(C(\boldsymbol{v})) && \text{(because } A \text{ is linear)} \\
&= (A \circ B)(\boldsymbol{v}) + (A \circ C)(\boldsymbol{v}) \ , && \text{(by equation (1.2))}
\end{aligned}$$

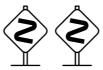
which proves (1.9). Similarly

$$\begin{aligned}((A+B)\circ C)\,(\boldsymbol{v}) &= (A+B)(C(\boldsymbol{v})) && \text{(by equation (1.2))}\\ &= (A(C(\boldsymbol{v})) + B(C(\boldsymbol{v})) && \text{(by equation (1.6))}\\ &= (A\circ C)(\boldsymbol{v}) + (B\circ C)(\boldsymbol{v})\ , && \text{(by equation (1.2))}\end{aligned}$$

which proves (1.10).

Composition of linear transformations is also compatible with scalar multiplication:

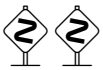
$$(\lambda\, A)\circ B = A\circ(\lambda\, B) = \lambda\,(A\circ B)\ . \tag{1.11}$$



In fact, we can summarise the properties (1.9), (1.10) and (1.11) in a very simple way using concepts we have already introduced. Given a linear transformation A of $\mathbb{V}$ we will define two operations on $\mathcal{L}(\mathbb{V})$, left and right multiplication by A, as follows:

$$\begin{aligned} L_A : \mathcal{L}(\mathbb{V}) &\to \mathcal{L}(\mathbb{V}) \qquad\qquad \text{and} \qquad R_A : \mathcal{L}(\mathbb{V}) \to \mathcal{L}(\mathbb{V})\\ B &\mapsto A\circ B \qquad\qquad\qquad\qquad\quad B \mapsto B\circ A\ . \end{aligned}$$

Then equations (1.9), (1.10) and (1.11) simply say that L_A and R_A are linear transformations of $\mathcal{L}(\mathbb{V})$!



The vector space $\mathcal{L}(\mathbb{V})$ of linear transformations of $\mathbb{V}$ together with the operation of composition, the identity $\mathbb{1}$, the distributive properties (1.9) and (1.10), and the condition (1.11) is an **associative algebra with identity**.

An **algebra** is a vector space $\mathbb{A}$ together with a multiplication

$$\begin{aligned}\text{multiplication} : \mathbb{A}\times\mathbb{A} &\to \mathbb{A}\\ (A,B) &\mapsto A\,B\ ,\end{aligned}$$

obeying the following axioms, where $A, B, C \in \mathbb{A}$ and $\lambda \in \mathbb{R}$:

- A1 (left distributivity) $A\,(B+C) = A\,B + A\,C$;
- A2 (right distributivity) $(A+B)\,C = A\,C + B\,C$;
- A3 $A\,(\lambda\,B) = (\lambda\,A)\,B = \lambda\,(A\,B)$.

If in addition $\mathbb{A}$ obeys the axiom

- A4 (identity) There exists $\mathbb{1} \in \mathbb{A}$ such that $\mathbb{1}\,A = A\,\mathbb{1} = A$;

then it is an **algebra with identity**. If instead $\mathbb{A}$ obeys the axiom

- A5 (associativity) $A\,(B\,C) = (A\,B)\,C$;

it is an **associative algebra**. Finally if it obeys all five axioms, it is an **associative algebra with identity**.

It is a general fact that the invertible elements of an associative algebra with identity form a group.

1.2.5 Matrices

Matrices are intimately linked to linear maps. Let $A : \mathbb{V} \to \mathbb{W}$ be a linear map between two *finite-dimensional* vector spaces. Let $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_N\}$ be a basis for $\mathbb{V}$ and let $\{\boldsymbol{f}_1, \boldsymbol{f}_2, \ldots, \boldsymbol{f}_M\}$ be a basis for $\mathbb{W}$. Let us write each $A(\boldsymbol{e}_i)$ as a linear combination of the basis elements $\{\boldsymbol{f}_j\}$:

$$A(\boldsymbol{e}_i) = \sum_{j=1}^{M} A_{ji}\, \boldsymbol{f}_j \ , \tag{1.12}$$

where have introduced a real number A_{ji} for each $i = 1, 2, \ldots, M$ and $j = 1, 2, \ldots, M$, a total of $N\,M$ real numbers. Now let $\boldsymbol{v}$ be a vector in $\mathbb{V}$ and consider its image $\boldsymbol{w} = A(\boldsymbol{v})$ under A. We can expand both $\boldsymbol{v}$ and $\boldsymbol{w}$ as linear combinations of the respective bases:

$$\boldsymbol{v} = \sum_{i=1}^{N} v_i\, \boldsymbol{e}_i \quad \text{and} \quad \boldsymbol{w} = \sum_{j=1}^{M} w_j\, \boldsymbol{f}_j \ . \tag{1.13}$$

Let us now express the w_j in terms of the v_i:

$$\begin{aligned}
\boldsymbol{w} &= A(\boldsymbol{v}) \\
&= A\left(\sum_{i=1}^{N} v_i\, \boldsymbol{e}_i\right) && \text{(by the first equation in (1.13))} \\
&= \sum_{i=1}^{N} A\left(v_i\, \boldsymbol{e}_i\right) && \text{(by L1)} \\
&= \sum_{i=1}^{N} v_i\, A\left(\boldsymbol{e}_i\right) && \text{(by L2)} \\
&= \sum_{i=1}^{N} v_i \sum_{j=1}^{M} A_{ji}\, \boldsymbol{f}_j && \text{(by equation (1.12))} \\
&= \sum_{j=1}^{M} \left(\sum_{i=1}^{N} A_{ji}\, v_i\right)\, \boldsymbol{f}_j \ , && \text{(rearranging the sums)}
\end{aligned}$$

whence comparing with the second equation in (1.13) we obtain the desired result:

$$w_j = \sum_{i=1}^{N} A_{ji}\, v_i \ . \tag{1.14}$$

To visualise this equation, let us arrange the components $\{v_i\}$ and $\{w_j\}$ of $\boldsymbol{v}$ and $\boldsymbol{w}$ as 'column vectors' v and w, and the real numbers A_{ji} as an $M \times N$ matrix A. Then equation (1.14) can be written as

$$\mathsf{w} = \mathsf{A}\,\mathsf{v} \ ,$$

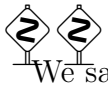
or explicitly as

$$\begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_M \end{pmatrix} = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1N} \\ A_{21} & A_{22} & \cdots & A_{2N} \\ \vdots & \vdots & & \vdots \\ A_{M1} & A_{M2} & \cdots & A_{MN} \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_N \end{pmatrix} \ .$$

Therefore the matrix

$$\mathsf{A} = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1N} \\ A_{21} & A_{22} & \cdots & A_{2N} \\ \vdots & \vdots & & \vdots \\ A_{M1} & A_{M2} & \cdots & A_{MN} \end{pmatrix}$$

represents the linear map $A : \mathbb{V} \to \mathbb{W}$ relative to the bases $\{\boldsymbol{e}_i\}$ and $\{\boldsymbol{f}_j\}$ of $\mathbb{V}$ and $\mathbb{W}$. It is important to stress that the linear map A is more fundamental than the matrix A. If we choose different basis, the matrix for the linear map will change (we will see this in detail below), but the map itself does not. However if we fix bases for $\mathbb{V}$ and $\mathbb{W}$, then there is a one-to-one correspondence between linear maps $\mathbb{V} \to \mathbb{W}$ and $M \times N$ matrices.



The commuting square: linear maps to matrices...

We saw in Section 1.2.4 that the space $\mathcal{L}(\mathbb{V}, \mathbb{W})$ of linear maps $\mathbb{V} \to \mathbb{W}$ is a vector space in its own right. How are the operations of vector addition and scalar multiplication defined for the matrices? It turns out that they are defined entry-wise as for real numbers. Let us see this. The matrix corresponding to the sum of two linear maps A and A' is given by

$$(A + A')(\boldsymbol{e}_i) = \sum_{j=1}^{M} (A + A')_{ji}\, \boldsymbol{f}_j \ .$$

On the other hand, from equation (1.6) we have that

$$\begin{aligned} (A + A')(\boldsymbol{e}_i) &= A(\boldsymbol{e}_i) + A'(\boldsymbol{e}_i) \\ &= \sum_{j=1}^{M} A_{ji}\, \boldsymbol{f}_j + \sum_{j=1}^{M} A'_{ji}\, \boldsymbol{f}_j \\ &= \sum_{j=1}^{M} \left(A_{ji} + A'_{ji}\right)\, \boldsymbol{f}_j \ . \end{aligned}$$

Therefore we see that the matrix of the sum is the sum of the matrices:

$$(A + A')_{ji} = A_{ji} + A'_{ji} \ ;$$

or in other words, the sum of two matrices is performed entry-by-entry:

$$\begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1N} \\ A_{21} & A_{22} & \cdots & A_{2N} \\ \vdots & \vdots & & \vdots \\ A_{M1} & A_{M2} & \cdots & A_{MN} \end{pmatrix} + \begin{pmatrix} A'_{11} & A'_{12} & \cdots & A'_{1N} \\ A'_{21} & A'_{22} & \cdots & A'_{2N} \\ \vdots & \vdots & & \vdots \\ A'_{M1} & A'_{M2} & \cdots & A'_{MN} \end{pmatrix}$$
$$= \begin{pmatrix} A_{11} + A'_{11} & A_{12} + A'_{12} & \cdots & A_{1N} + A'_{1N} \\ A_{21} + A'_{21} & A_{22} + A'_{22} & \cdots & A_{2N} + A'_{2N} \\ \vdots & \vdots & & \vdots \\ A_{M1} + A'_{M1} & A_{M2} + A'_{M2} & \cdots & A_{MN} + A'_{MN} \end{pmatrix} .$$

Similarly, scalar multiplication is also performed entry-by-entry. If $\lambda \in \mathbb{R}$ is a scalar and A is a linear map, then on the one hand we have

$$(\lambda\, A)(\boldsymbol{e}_i) = \sum_{j=1}^{M} (\lambda\, A)_{ji}\, \boldsymbol{f}_j \ ,$$

but from equation (1.7) we have that

$$\begin{aligned} (\lambda\, A)(\boldsymbol{e}_i) &= \lambda\, A(\boldsymbol{e}_i) \\ &= \lambda \sum_{j=1}^{M} A_{ji}\, \boldsymbol{f}_j \\ &= \sum_{j=1}^{M} \lambda\, A_{ji}\, \boldsymbol{f}_j \ , \end{aligned}$$

so that the matrix of $\lambda\, A$ is obtained from the matrix of A by multiplying each entry by λ:

$$(\lambda\, A)_{ji} = \lambda\, A_{ji} \ ;$$

explicitly,

$$\lambda \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1N} \\ A_{21} & A_{22} & \cdots & A_{2N} \\ \vdots & \vdots & & \vdots \\ A_{M1} & A_{M2} & \cdots & A_{MN} \end{pmatrix} = \begin{pmatrix} \lambda\, A_{11} & \lambda\, A_{12} & \cdots & \lambda\, A_{1N} \\ \lambda\, A_{21} & \lambda\, A_{22} & \cdots & \lambda\, A_{2N} \\ \vdots & \vdots & & \vdots \\ \lambda\, A_{M1} & \lambda\, A_{M2} & \cdots & \lambda\, A_{MN} \end{pmatrix} .$$

The vector space of $M \times N$ matrices has a 'canonical' basis given by the matrices E_{ji} all of whose entries are zero except for the entry sitting in the intersection of the jth column and the ith row, which is 1. They are clearly linearly independent and if A is any matrix with entries A_{ji} then

$$\mathsf{A} = \sum_{i=1}^{N} \sum_{j=1}^{M} A_{ji} \mathsf{E}_{ji} \ ,$$

so that their span is the space of all $M \times N$ matrices. Therefore they form a basis for this space. The matrices E_{ji} are known as **elementary matrices**. Clearly there are $M\,N$ such matrices, whence the dimension of the space of $M \times N$ matrices, and hence of $\mathfrak{L}(\mathbb{V}, \mathbb{W})$, is $M\,N$ as claimed in equation (1.8).

Now consider a third vector space $\mathbb{U}$ of dimension P and with basis $\{\boldsymbol{g}_1, \boldsymbol{g}_2, \ldots, \boldsymbol{g}_P\}$. Then a linear map $B : \mathbb{U} \to \mathbb{V}$ will be represented by an $N \times P$ matrix

$$\mathsf{B} = \begin{pmatrix} B_{11} & B_{12} & \cdots & B_{1P} \\ B_{21} & B_{22} & \cdots & B_{2P} \\ \vdots & \vdots & & \vdots \\ B_{N1} & B_{N2} & \cdots & B_{NP} \end{pmatrix}$$

relative to the chosen bases for $\mathbb{U}$ and $\mathbb{V}$; that is,

$$B(\boldsymbol{g}_k) = \sum_{i=1}^{N} B_{ik}\, \boldsymbol{e}_i \ . \tag{1.15}$$

The composition $A \circ B : \mathbb{U} \to \mathbb{W}$ will now be represented by an $M \times P$ matrix whose entries C_{jk} are given by

$$(A \circ B)(\boldsymbol{g}_k) = \sum_{j=1}^{M} C_{jk}\, \boldsymbol{f}_j \ . \tag{1.16}$$

The matrix of $A \circ B$ can be expressed in terms of the matrices A and B. To

see this, let us compute

$$\begin{aligned}
(A \circ B)(\boldsymbol{g}_k) &= A(B(\boldsymbol{g}_k)) && \text{(by equation (1.2))} \\
&= A\left(\sum_{i=1}^{N} B_{ik}\, \boldsymbol{e}_i\right) && \text{(by equation (1.15))} \\
&= \sum_{i=1}^{N} B_{ik}\, A(\boldsymbol{e}_i) && \text{(since } A \text{ is linear)} \\
&= \sum_{i=1}^{N} B_{ik} \sum_{j=1}^{M} A_{ji}\, \boldsymbol{f}_j && \text{(by equation (1.12))} \\
&= \sum_{j=1}^{M} \left(\sum_{i=1}^{N} A_{ji}\, B_{ik}\right) \boldsymbol{f}_j\ . && \text{(rearranging sums)}
\end{aligned}$$

Therefore comparing with equation (1.16) we see that

$$C_{jk} = \sum_{i=1}^{N} A_{ji}\, B_{ik}\ , \tag{1.17}$$

which is nothing else but matrix multiplication:

$$\begin{pmatrix} C_{11} & C_{12} & \cdots & C_{1P} \\ C_{21} & C_{22} & \cdots & C_{2P} \\ \vdots & \vdots & & \vdots \\ C_{M1} & C_{M2} & \cdots & C_{MP} \end{pmatrix} = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1N} \\ A_{21} & A_{22} & \cdots & A_{2N} \\ \vdots & \vdots & & \vdots \\ A_{M1} & A_{M2} & \cdots & A_{MN} \end{pmatrix} \begin{pmatrix} B_{11} & B_{12} & \cdots & B_{1P} \\ B_{21} & B_{22} & \cdots & B_{2P} \\ \vdots & \vdots & & \vdots \\ B_{N1} & B_{N2} & \cdots & B_{NP} \end{pmatrix}\ .$$

In other words,

the matrix of $A \circ B$ is the matrix product $\mathsf{A}\,\mathsf{B}$.

Let us consider now linear transformations $\mathcal{L}(\mathbb{V})$ of an N-dimensional vector space $\mathbb{V}$ with basis $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_N\}$. Matrices representing linear transformations $\mathbb{V} \to \mathbb{V}$ are now a square $N \times N$ matrices. We can add them and multiply them as we do real numbers, except that multiplication is not

commutative: for two matrices A and B one has that, in general, $\mathsf{AB} \neq \mathsf{BA}$. Let A be an $N \times N$ matrix. If there exists another matrix B which obeys

$$\mathsf{A}\,\mathsf{B} = \mathsf{B}\,\mathsf{A} = \mathsf{I}$$

where I is the identity matrix, then we say that A is **invertible**. Its **inverse** B is written A^{-1}. A matrix which is not invertible is called **singular**. Clearly a matrix is singular if and only if its determinant is zero.

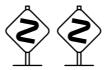
A useful fact is that a matrix is invertible if and only if its determinant is different from zero. This allows us to show that the product of invertible elements is again invertible. To see this notice that the determinant of a product is the product of the determinants:

$$\boxed{\det(\mathsf{A}\,\mathsf{B}) = \det\mathsf{A}\,\det\mathsf{B}\ ,} \tag{1.18}$$

and that this is not zero because neither are $\det\mathsf{A}$ nor $\det\mathsf{B}$. In fact, the inverse of a product $\mathsf{A}\,\mathsf{B}$ is given by

$$\boxed{(\mathsf{A}\,\mathsf{B})^{-1} = \mathsf{B}^{-1}\,\mathsf{A}^{-1}\ .} \tag{1.19}$$

(Notice the order!)



Matrices, just like $\mathfrak{L}(\mathbb{V})$, form an associative algebra with identity. The algebra of $N \times N$ real matrices is denoted $\mathrm{Mat}_N(\mathbb{R})$. The invertible elements form a group, which is denoted $\mathrm{GL}_N(\mathbb{R})$, the general linear group of $\mathbb{R}^N$.

1.2.6 Change of basis

We mentioned above that a linear map is more fundamental than the matrix representing it relative to a chosen basis, for the matrix changes when we change the basis but the linear map remains unchanged. In this Section we will explore how the matrix of a linear map changes as we change the basis. We will restrict ourselves to linear transformations, but the results here extend straightforwardly to linear maps between different vector spaces.

Let $\mathbb{V}$ be an N-dimensional vector space with basis $\{\boldsymbol{e}_i\}$, and let $A : \mathbb{V} \to \mathbb{V}$ be a linear transformation with matrix A relative to this basis. Let $\{\boldsymbol{e}'_i\}$ be another basis. We want to know what the matrix A' representing A relative this new basis is. By definition, the matrix A' has entries A'_{ji} given by

$$A(\boldsymbol{e}'_i) = \sum_{j=1}^{N} A'_{ji}\,\boldsymbol{e}'_j\ . \tag{1.20}$$

Because $\{\boldsymbol{e}_i\}$ is a basis, we can express each element $\boldsymbol{e}'_i$ of the primed basis in terms of them:

$$\boldsymbol{e}'_i = \sum_{j=1}^{N} S_{ji}\, \boldsymbol{e}_j \ , \tag{1.21}$$

for some N^2 numbers S_{ji}. We have written this equation in such a way that it looks as if S_{ji} are the entries of a matrix. This is with good reason. Let $S : \mathbb{V} \to \mathbb{V}$ be the linear transformation defined by $S(\boldsymbol{e}_i) = \boldsymbol{e}'_i$ for $i = 1, 2, \ldots, N$. Then using the explicit expression for $\boldsymbol{e}'_i$ we see that

$$S(\boldsymbol{e}_i) = \sum_{j=1}^{N} S_{ji}\, \boldsymbol{e}_j \ ,$$

so that S_{ji} are indeed the entries of a matrix S relative to the basis $\{\boldsymbol{e}_i\}$. We can compute both sides of equation (1.20) separately and compare. The right-hand side gives

$$\begin{aligned}
A(\boldsymbol{e}'_i) &= A\left(\sum_{j=1}^{N} S_{ji}\, \boldsymbol{e}_j\right) && \text{(by equation (1.21))} \\
&= \sum_{j=1}^{N} S_{ji}\, A(\boldsymbol{e}_j) && \text{(since } A \text{ is linear)} \\
&= \sum_{j=1}^{N} S_{ji} \sum_{k=1}^{N} A_{kj}\, \boldsymbol{e}_k && \text{(by equation (1.12))} \\
&= \sum_{k=1}^{N} \sum_{j=1}^{N} A_{kj} S_{ji}\, \boldsymbol{e}_k \ . && \text{(rearranging sums)}
\end{aligned}$$

On the other hand, the left-hand side gives

$$\begin{aligned}
\sum_{j=1}^{N} A'_{ji}\, \boldsymbol{e}'_j &= \sum_{j=1}^{N} A'_{ji} \sum_{k=1}^{N} S_{kj}\, \boldsymbol{e}_k && \text{(by equation (1.21))} \\
&= \sum_{k=1}^{N} \sum_{j=1}^{N} S_{kj} A'_{ji}\, \boldsymbol{e}_k \ . && \text{(rearranging sums)}
\end{aligned}$$

Comparing the two sides, we see that

$$\sum_{j=1}^{N} A_{kj} S_{ji} = \sum_{j=1}^{N} S_{kj} A'_{ji} \ ,$$

or in terms of matrices,

$$\mathsf{A}\,\mathsf{S} = \mathsf{S}\,\mathsf{A}' \ . \tag{1.22}$$

Now, S is invertible. To see this use the fact that because $\{\boldsymbol{e}'_i\}$ is also a basis, we can write each $\boldsymbol{e}_i$ in terms of the $\{\boldsymbol{e}'_i\}$:

$$\boldsymbol{e}_i = \sum_{j=1}^{N} T_{ji}\,\boldsymbol{e}'_j \ . \tag{1.23}$$

By the same argument as above, the N^2 numbers T_{ji} are the entries of a matrix which, relative to the primed basis, represents the linear transformation $T : \mathbb{V} \to \mathbb{V}$ defined by $T(\boldsymbol{e}'_i) = \boldsymbol{e}_i$. The linear transformations S and T are mutual inverses:

$$S(T(\boldsymbol{e}'_i)) = S(\boldsymbol{e}_i) = \boldsymbol{e}'_i \quad \text{and} \quad T(S(\boldsymbol{e}_i)) = T(\boldsymbol{e}'_i) = \boldsymbol{e}_i \ ,$$

so that $T \circ S = S \circ T = \mathbb{1}$; or in other words, $T = S^{-1}$.

Therefore, we can multiply both sides of equation (1.22) by S^{-1} on the left to obtain

$$\boxed{\mathsf{A}' = \mathsf{S}^{-1}\,\mathsf{A}\,\mathsf{S} \ .} \tag{1.24}$$

The operation above taking A to A' is called **conjugation** by S. One says that the matrices A and A' are **conjugate**. (This is not be confused with the notion of *complex conjugation*.)

Change of basis for linear maps.

1.2.7 Matrix invariants

Certain properties of square matrices do not change when we change the basis; one says that they are **invariants** of the matrix or, more precisely, of the linear map that the matrix represents.

For example, the determinant is one such invariant. This can be seen by computing the determinant to both sides of equation (1.24) and using equation (1.18), to obtain that $\det \mathsf{A}' = \det \mathsf{A}$. This implies that also the property of being invertible is invariant.

Another invariant is the **trace** of a matrix, defined as the sum of the diagonal elements, and written $\operatorname{tr} \mathsf{A}$. Explicitly, if A is given by

$$\mathsf{A} = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1N} \\ A_{21} & A_{22} & \cdots & A_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ A_{N1} & A_{N2} & \cdots & A_{NN} \end{pmatrix}$$

then its trace $\operatorname{tr} \mathsf{A}$ is given by

$$\operatorname{tr} \mathsf{A} = \sum_{i=1}^{N} A_{ii} = A_{11} + A_{22} + \cdots + A_{NN} \ . \tag{1.25}$$

A matrix whose trace vanishes is said to be **traceless**.

The fact that the trace is indeed an invariant, will follow from some fundamental properties of the trace, which we discuss now. The trace satisfies the following property:

$$\boxed{\operatorname{tr} (\mathsf{A}\,\mathsf{B}) = \operatorname{tr} (\mathsf{B}\,\mathsf{A}) \ .} \tag{1.26}$$

Let us prove this. Let $A, B : \mathbb{V} \to \mathbb{V}$ be linear maps with matrices A and B relative to some fixed basis. The matrix product $\mathsf{A}\,\mathsf{B}$ is the matrix of the composition $A \circ B$. Computing the trace of the product, using equations (1.17) and (1.25), we find

$$\begin{aligned}
\operatorname{tr} (\mathsf{A}\,\mathsf{B}) &= \sum_{i=1}^{N} \sum_{j=1}^{N} A_{ij}\, B_{ji} \\
&= \sum_{j=1}^{N} \sum_{i=1}^{N} B_{ji}\, A_{ij} && \text{(rearranging the sums)} \\
&= \sum_{i=1}^{N} \sum_{j=1}^{N} B_{ij}\, A_{ji} && \text{(relabelling the sums)} \\
&= \operatorname{tr} (\mathsf{B}\,\mathsf{A}) \ .
\end{aligned}$$

The fact which allows us to relabel the summation indices is known as the **Shakespeare Theorem**: "a dummy index by any other name..." The modern version of this theorem is due to Gertrude Stein: "a dummy index is a dummy index is a dummy index."

It follows from equation (1.26) that

$$\boxed{\operatorname{tr} (\mathsf{A}\,\mathsf{B}\,\mathsf{C}) = \operatorname{tr} (\mathsf{C}\,\mathsf{A}\,\mathsf{B}) = \operatorname{tr} (\mathsf{B}\,\mathsf{C}\,\mathsf{A}) \ ,} \tag{1.27}$$

which is often called the **cyclic property** of the trace. Using this property and computing the trace to both sides of equation (1.24), we see that $\operatorname{tr} \mathsf{A}' = \operatorname{tr} \mathsf{A}$, as claimed. Notice that the trace of the identity $N \times N$ matrix I is $\operatorname{tr} \mathsf{I} = N$.

Because the trace is an invariant, it actually defines a function on the vector space of linear maps $\mathcal{L}(\mathbb{V})$. The trace of a linear map $A : \mathbb{V} \to \mathbb{V}$ is defined as the trace of any matrix of A relative to some basis. Invariance says that it does not depend on which basis we choose to compute it with respect to. As a function $\operatorname{tr} : \mathcal{L}(\mathbb{V}) \to \mathbb{R}$, the trace is actually linear. It is an easy exercise to prove that

$$\operatorname{tr}(\mathsf{A} + \mathsf{B}) = \operatorname{tr} \mathsf{A} + \operatorname{tr} \mathsf{B} \qquad \text{and} \qquad \operatorname{tr}(\lambda\, \mathsf{A}) = \lambda \operatorname{tr} \mathsf{A} \ .$$

There are other properties of a matrix which are *not* invariant under arbitrary change of basis; but are nevertheless important. For example, given a matrix A let its **transpose**, denoted A^t, be the matrix whose (i,j) entry equals the (j,i) entry of A. Explicitly,

$$\mathsf{A} = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1N} \\ A_{21} & A_{22} & \cdots & A_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ A_{N1} & A_{N2} & \cdots & A_{NN} \end{pmatrix} \Rightarrow \mathsf{A}^t = \begin{pmatrix} A_{11} & A_{21} & \cdots & A_{N1} \\ A_{12} & A_{22} & \cdots & A_{N2} \\ \vdots & \vdots & \ddots & \vdots \\ A_{1N} & A_{2N} & \cdots & A_{NN} \end{pmatrix} .$$

In other words, A^t is obtained from A by reflecting the matrix on the main diagonal, and because reflection is an involution, it follows that

$$\boxed{\left(\mathsf{A}^t\right)^t = \mathsf{A} \ .} \tag{1.28}$$

It follows from the expression for A^t that the diagonal entries are not changed, and hence that

$$\boxed{\operatorname{tr} \mathsf{A}^t = \operatorname{tr} \mathsf{A} \ .} \tag{1.29}$$

It is also easy to see that

$$\boxed{(\mathsf{A}\,\mathsf{B})^t = \mathsf{B}^t\,\mathsf{A}^t} \tag{1.30}$$

and also that

$$\boxed{(\mathsf{A}+\mathsf{B})^t = \mathsf{A}^t + \mathsf{B}^t} \qquad \text{and} \qquad \boxed{(\lambda\mathsf{A})^t = \lambda\mathsf{A}^t \ .} \tag{1.31}$$

From the former equation it follows that

$$\boxed{\left(\mathsf{A}^{-1}\right)^t = \left(\mathsf{A}^t\right)^{-1} \ .} \tag{1.32}$$

A less obvious identity is

$$\boxed{\det \mathsf{A}^t = \det \mathsf{A} \ ,} \tag{1.33}$$

which follows from the fact that the row expansion of the determinant of A^t is precisely the column expansion of the determinant of A. A matrix is said to be **symmetric** if $\mathsf{A}^t = \mathsf{A}$. It is said to be **antisymmetric** or **skew-symmetric** if $\mathsf{A}^t = -\mathsf{A}$. Notice that an antisymmetric matrix is traceless, since

$$\operatorname{tr} \mathsf{A} = \operatorname{tr} \mathsf{A}^t = \operatorname{tr}(-\mathsf{A}) = -\operatorname{tr} \mathsf{A} \ .$$

The converse is of course false: a traceless matrix need not be antisymmetric.

Generic matrices are neither symmetric nor antisymmetric, yet any matrix is the sum of a symmetric matrix and an antisymmetric matrix. Indeed, adding and subtracting $\frac{1}{2}\mathsf{A}^t$ in a clever way, we see that

$$\mathsf{A} = \tfrac{1}{2}\left(\mathsf{A} + \mathsf{A}^t\right) + \tfrac{1}{2}\left(\mathsf{A} - \mathsf{A}^t\right) \ .$$

But now, using equations (1.28) and (1.31), we see that $\frac{1}{2}(\mathsf{A}+\mathsf{A}^t)$ is symmetric and $\frac{1}{2}(\mathsf{A} - \mathsf{A}^t)$ antisymmetric.

A matrix O is said to be **orthogonal** if its transpose is its inverse:

$$\mathsf{O}^t\,\mathsf{O} = \mathsf{O}\,\mathsf{O}^t = \mathsf{I} \ .$$

The property of being symmetric or antisymmetric is not invariant under arbitrary changes of basis, but it will be preserved under certain types of changes of basis, e.g., under orthogonal changes of basis.

1.3 Inner products

Vectors in physics are usually defined as objects which have both a magnitude and a direction. In that sense, they do not quite correspond to the mathematical notion of a vector as we have been discussing above. In our definition of an abstract vector space as in the discussion which followed, there is no mention of how to compute the magnitude of a vector. In this section we will remedy this situation. Geometrically the magnitude of a vector is simply its length. If we think of vectors as displacement, the magnitude is the distance away from the origin. In order to define distance we will need to introduce an *inner product* or *scalar product*, as it is often known.

1.3.1 Norms and inner products

Let us start by considering displacements in the plane. The length $\|\boldsymbol{v}\|$ of the displacement $\boldsymbol{v} = (v_1, v_2)$ is given by the Pythagorean theorem: $\|\boldsymbol{v}\|^2 = v_1^2 + v_2^2$. This length obeys the following properties which are easily verified. First of all it is a non-negative quantity $\|\boldsymbol{v}\|^2 \geq 0$, vanishing precisely for the zero displacement $\mathbf{0} = (0, 0)$. If we rescale $\boldsymbol{v}$ by a real number λ: $\lambda\,\boldsymbol{v} = (\lambda\, v_1, \lambda\, v_2)$, its length rescales by the absolute value of λ: $\|\lambda\,\boldsymbol{v}\| = |\lambda|\,\|\boldsymbol{v}\|$. Finally, the length obeys the so-called triangle inequality: $\|\boldsymbol{v} + \boldsymbol{w}\| \leq \|\boldsymbol{v}\| + \|\boldsymbol{w}\|$. This is obvious pictorially, since the shortest distance between two points in the plane is the straight line which joins them. In any case we will prove it later in much more generality.

Now consider $\mathbb{R}^N$. We can define a notion of length by generalising slightly what was done above: if $(v_1, v_2, \ldots, v_N) \in \mathbb{R}^N$, then define its length by

$$\|(v_1, v_2, \ldots, v_N)\| = \sqrt{v_1^2 + v_2^2 + \cdots + v_N^2} \ .$$

It again satisfies the same three properties described above.

We can formalise this into the notion of a norm in a vector space. By a **norm** in a real vector space $\mathbb{V}$ we mean a function $\|\cdot\| : \mathbb{V} \to \mathbb{R}$ assigning a real number to every vector in $\mathbb{V}$ in such a way that the following three properties are satisfied for every vector $\boldsymbol{v}$ and $\boldsymbol{w}$ and every scalar λ:

N1 $\|\boldsymbol{v}\| \geq 0$, and $\|\boldsymbol{v}\| = 0$ if and only if $\boldsymbol{v} = \boldsymbol{0}$;

N2 $\|\lambda\, \boldsymbol{v}\| = |\lambda|\, \|\boldsymbol{v}\|$; and

N3 (triangle inequality) $\|\boldsymbol{v} + \boldsymbol{w}\| \leq \|\boldsymbol{v}\| + \|\boldsymbol{w}\|$.

The study of normed vector spaces is an important branch of modern mathematics (cf., one of the 1998 Fields Medals). In physics, however, it is fair to say that the more important notion is that of an inner product. If a norm allows us to calculate lengths, an inner product will allow us to also calculate angles.

Consider again the case of displacements in two dimensions, or equivalent $\mathbb{R}^2$. Let us define now a function which assigns a real number to two displacements $\boldsymbol{v} = (v_1, v_2)$ and $\boldsymbol{w} = (w_1, w_2)$:

$$\langle \boldsymbol{v}, \boldsymbol{w} \rangle := v_1\, w_1 + v_2\, w_2 \ .$$

This is usually called the **dot product** and is written $\boldsymbol{v} \cdot \boldsymbol{w}$. We will not use this notation.

Clearly, $\langle \boldsymbol{v}, \boldsymbol{v} \rangle = \|\boldsymbol{v}\|^2$, so that this construction also incorporates a norm. If we write the displacements using polar coordinates: $\boldsymbol{v} = \|\boldsymbol{v}\|\, (\cos\theta_1, \sin\theta_1)$ and similarly for $\boldsymbol{w} = \|\boldsymbol{w}\|\, (\cos\theta_2, \sin\theta_2)$, then we can compute:

$$\langle \boldsymbol{v}, \boldsymbol{w} \rangle = \|\boldsymbol{v}\|\, \|\boldsymbol{w}\|\, \cos(\theta_1 - \theta_2) \ . \tag{1.34}$$

In other words, $\langle \cdot, \cdot \rangle$ is essentially the angle between the two displacements. More generally we can consider $\mathbb{R}^N$ and define its dot product as follows. If $\boldsymbol{v} = (v_1, v_2, \ldots, v_N)$ and $\boldsymbol{w} = (w_1, w_2, \ldots, w_N)$, then

$$\langle \boldsymbol{v}, \boldsymbol{w} \rangle := \sum_{i=1}^{N} v_i\, w_i = v_1\, w_1 + v_2\, w_2 + \cdots + v_N\, w_N \ .$$

The dot product satisfies the following properties. First of all it is symmetric: $\langle \boldsymbol{v}, \boldsymbol{w} \rangle = \langle \boldsymbol{w}, \boldsymbol{v} \rangle$. It is also linear in the right-hand slot: $\langle \boldsymbol{v}, \boldsymbol{w} + \boldsymbol{w} \rangle = \langle \boldsymbol{v}, \boldsymbol{w} \rangle + \langle \boldsymbol{v}, \boldsymbol{w} \rangle$ and $\langle \boldsymbol{v}, \lambda\, \boldsymbol{w} \rangle = \lambda \langle \boldsymbol{v}, \boldsymbol{w} \rangle$; and using the symmetry also in the left-hand slot. It is also important that the function $\|\boldsymbol{v}\| := \sqrt{\langle \boldsymbol{v}, \boldsymbol{v} \rangle}$ is a norm. The only non-obvious thing is to prove the triangle inequality for the norm, but we will do this below in all generality. The vector space $\mathbb{R}^N$ with the dot product defined above is called N-**dimensional Euclidean space**, and is denoted $\mathbb{E}^N$. As a vector space, of course, $\mathbb{E}^N = \mathbb{R}^N$, but $\mathbb{E}^N$ serves to remind us that we are talking about the vector space with the dot product. Notice that in terms of column vectors:

$$\mathsf{v} = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_N \end{pmatrix} \qquad \text{and} \qquad \mathsf{w} = \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_N \end{pmatrix} \ ,$$

the dot product is given by

$$\langle \mathsf{v}, \mathsf{w} \rangle = \mathsf{v}^t\, \mathsf{w} = \begin{pmatrix} v_1 & v_2 & \cdots & v_N \end{pmatrix} \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_N \end{pmatrix} = \sum_{i=1}^{N} v_i\, w_i \ .$$

More generally, we define an **inner product** (or **scalar product**) on a real vector space $\mathbb{V}$ to be a function $\langle \cdot, \cdot \rangle : \mathbb{V} \times \mathbb{V} \to \mathbb{R}$ taking pairs of vectors to real numbers and obeying the following axioms:

IP1 $\langle \boldsymbol{v}, \boldsymbol{w} \rangle = \langle \boldsymbol{w}, \boldsymbol{v} \rangle$;

IP2 $\langle \boldsymbol{u}, \lambda\, \boldsymbol{v} + \mu\, \boldsymbol{w} \rangle = \lambda \langle \boldsymbol{u}, \boldsymbol{v} \rangle + \mu \langle \boldsymbol{u}, \boldsymbol{w} \rangle$; and

IP3 $\|\boldsymbol{v}\|^2 = \langle \boldsymbol{v}, \boldsymbol{v} \rangle > 0$ for all $\boldsymbol{v} \neq \boldsymbol{0}$.

Notice that IP1 and IP2 together imply that $\langle \lambda\, \boldsymbol{u} + \mu\, \boldsymbol{v}, \boldsymbol{w} \rangle = \lambda \langle \boldsymbol{u}, \boldsymbol{w} \rangle + \mu \langle \boldsymbol{v}, \boldsymbol{w} \rangle$.

Let $\{\boldsymbol{e}_i\}$ be a basis for $\mathbb{V}$. Because of IP1 and IP2, it is enough to know what the inner product of any two basis elements is to be know what it is on any two vectors. Indeed, let $\boldsymbol{v} = \sum_{i=1}^{N} v_i\, \boldsymbol{e}_i$ and $\boldsymbol{w} = \sum_{i=1}^{N} w_i\, \boldsymbol{e}_i$ be any two vectors. Then their inner product is given by

$$\begin{aligned} \langle \boldsymbol{v}, \boldsymbol{w} \rangle &= \langle \sum_{i=1}^{N} v_i\, \boldsymbol{e}_i, \sum_{j=1}^{N} v_j\, \boldsymbol{e}_j \rangle \\ &= \sum_{i,j=1}^{N} v_i\, w_j\, \langle \boldsymbol{e}_i, \boldsymbol{e}_j \rangle \ . \qquad \text{(using IP1,2)} \end{aligned}$$

In other words, all we need to know in order to compute this are the real numbers $G_{ij} := \langle \boldsymbol{e}_i, \boldsymbol{e}_j \rangle$. These can be thought of as the entries of a matrix G. If we think of $\boldsymbol{v}$ as a column vector v in $\mathbb{R}^N$ whose entries are the components of $\boldsymbol{v}$ relative to the basis $\{\boldsymbol{e}_i\}$, and the same for $\boldsymbol{w}$, we can compute their inner product using matrix multiplication:

$$\langle \boldsymbol{v}, \boldsymbol{w} \rangle = \mathsf{v}^t \, \mathsf{G} \, \mathsf{w} \ .$$

The matrix G is not arbitrary. First of all from IP1 it follows that it is symmetric:

$$G_{ij} = \langle \boldsymbol{e}_i, \boldsymbol{e}_j \rangle = \langle \boldsymbol{e}_j, \boldsymbol{e}_i \rangle = G_{ji} \ .$$

Furthermore IP3 imposes a strong condition known as **positive-definiteness**. We will see at the end of this section what this means explicitly. Let us simply mention that IP3 implies that the only vector which is orthogonal to all vectors is the zero vector. This condition is weaker than IP3. It is often desirable to relax IP3 in terms of this condition. Such inner products are called **non-degenerate**. Non-degeneracy means that the matrix G is invertible, so that its determinant is non-zero.

Here comes a point which confuses many people, so *pay attention*! Both inner products and linear transformations are represented by matrices relative to a basis, but they are very different objects. In particular, they transform different under a change of basis and this means that even if the matrices for a linear transformation and an inner product agree numerically in a given basis, they will generically *not* agree with respect to a different basis. Let us see this in detail. Let $\{\boldsymbol{e}'_i\}$ be a new basis, with $\boldsymbol{e}'_i = S(\boldsymbol{e}_i)$ for some linear transformation S. Relative to $\{\boldsymbol{e}_i\}$ the linear transformation S is represented by a matrix S with entries S_{ji} given by equation (1.21). Let G' denote the matrix describing the inner product in the new basis: its entries G'_{ij} are given by

$$\begin{aligned}
G'_{ij} &= \langle \boldsymbol{e}'_i, \boldsymbol{e}'_j \rangle && \text{(by definition)} \\
&= \langle \sum_{k=1}^{N} S_{ki} \, \boldsymbol{e}_k, \sum_{l=1}^{N} S_{lj} \, \boldsymbol{e}_l \rangle && \text{(by equation (1.21))} \\
&= \sum_{k,l=1}^{N} S_{ki} \, S_{lj} \, \langle \boldsymbol{e}_k, \boldsymbol{e}_l \rangle && \text{(using IP1,2)} \\
&= \sum_{k,l=1}^{N} S_{ki} \, G_{kl} \, S_{lj} \ .
\end{aligned}$$

In other words,

$$\boxed{\mathsf{G}' = \mathsf{S}^t \, \mathsf{G} \, \mathsf{S} \ ,} \tag{1.35}$$

to be contrasted with the analogous formula (1.24) for the behaviour of the matrix of a linear transformation under a change of basis.

Notice, however, that under an orthogonal change of basis, so that $\mathsf{S}^{-1} = \mathsf{S}^t$, then both inner products and linear maps transform the same way.

1.3.2 The Cauchy–Schwartz and triangle inequalities

In this section we prove that $\|\boldsymbol{v}\| = \sqrt{\langle \boldsymbol{v}, \boldsymbol{v} \rangle}$ is indeed a norm. Because axioms N1 and N2 are obvious from the axioms of the inner product, all we really need to prove is the triangle inequality. This inequality will follow trivially from another inequality called the Cauchy–Schwartz inequality, and which is itself quite useful. Consider equation (1.34). Because the cosine function obeys $|\cos\theta| \leq 1$, we can deduce an inequality from equation (1.34). Namely that for any two displacements $\boldsymbol{v}$ and $\boldsymbol{w}$ in the plane,

$$|\langle \boldsymbol{v}, \boldsymbol{w} \rangle| \leq \|\boldsymbol{v}\| \, \|\boldsymbol{w}\| \ ,$$

with equality if and only if the angle between the two displacements is zero; in other words, if the displacements are collinear. The above inequality is called the two-dimensional Cauchy–Schwartz inequality. This inequality actually holds in any vector space with an inner product (even if it is infinite-dimensional).

Let $\boldsymbol{v}$ and $\boldsymbol{w}$ be any two vectors in a vector space $\mathbb{V}$ with an inner product $\langle \cdot, \cdot \rangle$. Let λ be a real number and let us consider the following inequality:

$$\begin{aligned}
0 &\leq \|\boldsymbol{v} - \lambda \, \boldsymbol{w}\|^2 \\
&= \langle \boldsymbol{v} - \lambda \, \boldsymbol{w}, \boldsymbol{v} - \lambda \, \boldsymbol{w} \rangle && \text{(by definition)} \\
&= \|\boldsymbol{v}\|^2 + \|\lambda \, \boldsymbol{w}\|^2 - 2\langle \boldsymbol{v}, \lambda \, \boldsymbol{w} \rangle && \text{(expanding and using IP1,2)} \\
&= \|\boldsymbol{v}\|^2 + \lambda^2 \|\boldsymbol{w}\|^2 - 2\lambda \, \langle \boldsymbol{v}, \boldsymbol{w} \rangle \ . && \text{(using IP2)}
\end{aligned}$$

Now we want to make a clever choice of λ which allows us to partially cancel the last two terms against each other. This way we can hope to get an inequality involving only two terms. The clever choice of λ turns out to be $\lambda = \langle \boldsymbol{v}, \boldsymbol{w} \rangle / \|\boldsymbol{w}\|^2$. Inserting this into the above equation and rearranging the terms a little, we obtain the following inequality

$$\|\boldsymbol{v}\|^2 \geq \frac{\langle \boldsymbol{v}, \boldsymbol{w} \rangle^2}{\|\boldsymbol{w}\|^2} \ .$$

Taking the (positive) square root and rearranging we arrive at the **Cauchy–Schwartz inequality**:

$$\boxed{|\langle \boldsymbol{v}, \boldsymbol{w}\rangle| \leq \|\boldsymbol{v}\|\,\|\boldsymbol{w}\|\ .} \tag{1.36}$$

The triangle inequality now follows easily. Let us expand $\|\boldsymbol{v}+\boldsymbol{w}\|^2$ as follows:

$$\begin{aligned}
\|\boldsymbol{v}+\boldsymbol{w}\|^2 &= \langle \boldsymbol{v}+\boldsymbol{w}, \boldsymbol{v}+\boldsymbol{w}\rangle \\
&= \|\boldsymbol{v}\|^2 + \|\boldsymbol{w}\|^2 + 2\langle \boldsymbol{v}, \boldsymbol{w}\rangle && \text{(using IP1,2)} \\
&\leq \|\boldsymbol{v}\|^2 + \|\boldsymbol{w}\|^2 + 2|\langle \boldsymbol{v}, \boldsymbol{w}\rangle| && \text{(since } x \leq |x|\text{)} \\
&\leq \|\boldsymbol{v}\|^2 + \|\boldsymbol{w}\|^2 + 2\|\boldsymbol{v}\|\,\|\boldsymbol{w}\| && \text{(using Cauchy–Schwartz)} \\
&= \left(\|\boldsymbol{v}\| + \|\boldsymbol{w}\|\right)^2\ .
\end{aligned}$$

Taking the (positive) square root we arrive at the **triangle inequality**:

$$\boxed{\|\boldsymbol{v}+\boldsymbol{w}\| \leq \|\boldsymbol{v}\| + \|\boldsymbol{w}\|\ .} \tag{1.37}$$

1.3.3 Orthonormal bases and Gram–Schmidt

Throughout this section we will let $\mathbb{V}$ be an N-dimensional real vector space with an inner product $\langle \cdot, \cdot \rangle$.

We say that two vectors $\boldsymbol{v}$ and $\boldsymbol{w}$ are **orthogonal** (written $\boldsymbol{v} \perp \boldsymbol{w}$) if their inner product vanishes: $\langle \boldsymbol{v}, \boldsymbol{w}\rangle = 0$. Any nonzero vector can be normalised to have unit norm simply dividing by its norm: $\boldsymbol{v}/\|\boldsymbol{v}\|$ has unit norm. A basis $\{\boldsymbol{e}_i\}$ is said to be **orthonormal** if

$$\langle \boldsymbol{e}_i, \boldsymbol{e}_j\rangle = \delta_{ij} := \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise.} \end{cases} \tag{1.38}$$

In other words, the basis elements in an orthonormal basis are mutually orthogonal and are normalised to unit norm. Notice that the matrix representing the inner product relative to an orthonormal basis is the identity matrix.

The components of a vector $\boldsymbol{v}$ relative to an orthonormal basis $\{\boldsymbol{e}_i\}$ are very easy to compute. Let $\boldsymbol{v} = \sum_{i=1}^{N} v_i\, \boldsymbol{e}_i$, and take its inner product with

$\boldsymbol{e}_j$:

$$\begin{aligned}\langle \boldsymbol{e}_j, \boldsymbol{v}\rangle &= \langle \boldsymbol{e}_j, \sum_{i=1}^{N} v_i\, \boldsymbol{e}_i\rangle \\ &= \sum_{i=1}^{N} v_i\, \langle \boldsymbol{e}_j, \boldsymbol{e}_i\rangle && \text{(using IP2)} \\ &= v_j\ . && \text{(using equation (1.38))}\end{aligned}$$

This shows that orthonormal vectors are automatically linearly independent. Indeed, suppose that $\{\boldsymbol{e}_i\}$ are orthonormal vectors. Then suppose that a linear combination is the zero vector:

$$\sum_i \lambda_i\, \boldsymbol{e}_i = \boldsymbol{0}\ .$$

Taking the inner product of both sides of this equality with $\boldsymbol{e}_j$ we find, on the left-hand side λ_j and on the right-hand side 0, hence $\lambda_j = 0$ and thus the $\{\boldsymbol{e}_i\}$ are linearly independent.

We now discuss an algorithmic procedure by which *any* basis can be modified to yield an orthonormal basis. Let $\{\boldsymbol{f}_i\}$ be any basis whatsoever for $\mathbb{V}$. We will define iteratively a new basis $\{\boldsymbol{e}_i\}$ which will be orthonormal. The procedure starts as follows. We define

$$\boldsymbol{e}_1 = \frac{\boldsymbol{f}_1}{\|\boldsymbol{f}_1\|}\ ,$$

which has unit norm by construction. We now define $\boldsymbol{e}_2$ starting from $\boldsymbol{f}_2$ but making it orthogonal to $\boldsymbol{e}_1$ and normalising it to unit norm. A moment's thought reveals that the correct definition is

$$\boldsymbol{e}_2 = \frac{\boldsymbol{f}_2 - \langle \boldsymbol{f}_2, \boldsymbol{e}_1\rangle\, \boldsymbol{e}_1}{\|\boldsymbol{f}_2 - \langle \boldsymbol{f}_2, \boldsymbol{e}_1\rangle\, \boldsymbol{e}_1\|}\ .$$

It has unit norm by construction, and it is clearly orthogonal to $\boldsymbol{e}_1$ because

$$\langle \boldsymbol{f}_2 - \langle \boldsymbol{f}_2, \boldsymbol{e}_1\rangle, \boldsymbol{e}_1\rangle = \langle \boldsymbol{f}_2, \boldsymbol{e}_1\rangle - \langle \boldsymbol{f}_2, \boldsymbol{e}_1\rangle\, \|\boldsymbol{e}_1\|^2 = 0\ .$$

We can continue in this fashion and at each step define $\boldsymbol{e}_i$ as $\boldsymbol{f}_i + \cdots$ divided by its norm, where the omitted terms are a linear combination of the $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_{i-1}\}$ defined in such a way that the $\boldsymbol{e}_i$ is orthogonal to them. For a finite-dimensional vector space, this procedure stops in a finite time

and we are left with an orthonormal basis $\{\boldsymbol{e}_i\}$. The general formulae for the $\boldsymbol{e}_i$ is

$$\boxed{\boldsymbol{e}_i = \frac{\boldsymbol{f}_i - \sum_{j=1}^{i-1} \langle \boldsymbol{f}_i, \boldsymbol{e}_j \rangle \, \boldsymbol{e}_j}{\| \boldsymbol{f}_i - \sum_{j=1}^{i-1} \langle \boldsymbol{f}_i, \boldsymbol{e}_j \rangle \, \boldsymbol{e}_j \|} \ .} \tag{1.39}$$

Notice that this formula is *recursive*: it defines $\boldsymbol{e}_i$ in terms of $\boldsymbol{f}_i$ and the $\{\boldsymbol{e}_{j<i}\}$.

Studying this formula we see that each $\boldsymbol{e}_i$ is a linear combination

$$\boldsymbol{e}_i = \sum_{j=1}^{i} S_{ji} \boldsymbol{f}_j \ , \tag{1.40}$$

where S_{ii} is positive, since it is given by $S_{ii} = 1/\| \boldsymbol{f}_i + \cdots \|$. Now let S be the linear transformation defined by $S(\boldsymbol{f}_i) = \boldsymbol{e}_i$. Relative to the original basis $\{\boldsymbol{f}_i\}$, S has a matrix S with entries S_{ji} defined by

$$\boldsymbol{e}_i = \sum_{j=1}^{N} S_{ji} \, \boldsymbol{f}_j \ .$$

Comparing with equation (1.40) we see that $S_{ji} = 0$ for $j > i$, so that all the entries of S below the main diagonal are zero. We say that S is **upper triangular**. The condition $S_{ii} > 0$ says that the diagonal entries are positive.

We can turn equation (1.39) around and notice that $\boldsymbol{f}_i$ is in turn given as a linear combination of $\{\boldsymbol{e}_{j \le i}\}$. The linear transformation T defined by $\boldsymbol{f}_i = T(\boldsymbol{e}_i)$, which is the inverse of S, has a matrix T relative to the $\{\boldsymbol{e}_i\}$ basis which is also upper triangular with positive entries on the main diagonal. Now the matrix G with entries $G_{ij} = \langle \boldsymbol{f}_i, \boldsymbol{f}_j \rangle$ representing the inner product on the $\{\boldsymbol{f}_i\}$ basis, is now given by

$$\mathsf{G} = \mathsf{T}^t \, \mathsf{T} \ .$$

In other words, since the $\{\boldsymbol{f}_i\}$ were an arbitrary basis, G is an arbitrary matrix representing an inner product. We have learned then that this matrix can always be written as a "square" $\mathsf{T}^t \, \mathsf{T}$, where T is an upper triangular matrix with positive entries in the main diagonal.

1.3.4 The adjoint of a linear transformation

Throughout this section we will let $\mathbb{V}$ be an N-dimensional real vector space with an inner product $\langle \cdot, \cdot \rangle$.

Let $A : \mathbb{V} \to \mathbb{V}$ be a linear transformation. A linear transformation is uniquely defined by its *matrix elements* $\langle A(\boldsymbol{v}), \boldsymbol{w} \rangle$. Indeed, if A' is another linear transformation with $\langle A'(\boldsymbol{v}), \boldsymbol{w} \rangle = \langle A(\boldsymbol{v}), \boldsymbol{w} \rangle$ for all $\boldsymbol{v}$ and $\boldsymbol{w}$, then we claim that $A = A'$. To see this notice that

$$\begin{aligned} 0 &= \langle A'(\boldsymbol{v}), \boldsymbol{w} \rangle - \langle A(\boldsymbol{v}), \boldsymbol{w} \rangle \\ &= \langle A'(\boldsymbol{v}) - A(\boldsymbol{v}), \boldsymbol{w} \rangle \ . \qquad \text{(using IP1,2)} \end{aligned}$$

Since this is true for all $\boldsymbol{w}$, it says that the vector $A'(\boldsymbol{v}) - A(\boldsymbol{v})$ is orthogonal to all vectors, and in particular to itself. Therefore it has zero norm and by IP3 it is the zero vector. In other words, $A'(\boldsymbol{v}) = A(\boldsymbol{v})$ for all $\boldsymbol{v}$, which means that $A = A'$.

Given a linear transformation $A : \mathbb{V} \to \mathbb{V}$ we define its **adjoint** relative to the inner product, as the linear transformation $A^\dagger : \mathbb{V} \to \mathbb{V}$ with matrix elements

$$\boxed{\langle A^\dagger(\boldsymbol{v}), \boldsymbol{w}\rangle = \langle \boldsymbol{v}, A(\boldsymbol{w})\rangle \ .} \tag{1.41}$$

The adjoint operation obeys several properties. First of all, taking adjoint is an involution:

$$\boxed{A^{\dagger\dagger} = A \ .} \tag{1.42}$$

Moreover it is a linear operation

$$\boxed{(\lambda\, A + \mu\, B)^\dagger = \lambda\, A^\dagger + \mu\, B^\dagger \ ,} \tag{1.43}$$

which reverses the order of a composition:

$$\boxed{(A \circ B)^\dagger = B^\dagger \circ A^\dagger \ .} \tag{1.44}$$

These properties are easily proven. The method of proof consists in showing that both sides of each equation have the same matrix elements. For example, the matrix elements of the double adjoint $A^{\dagger\dagger}$ are given by

$$\begin{aligned}
\langle A^{\dagger\dagger}(\boldsymbol{v}), \boldsymbol{w}\rangle &= \langle \boldsymbol{v}, A^\dagger(\boldsymbol{w})\rangle && \text{(by equation (1.41))}\\
&= \langle A^\dagger(\boldsymbol{w}), \boldsymbol{v}\rangle && \text{(by IP1)}\\
&= \langle \boldsymbol{w}, A(\boldsymbol{v})\rangle && \text{(by equation (1.41))}\\
&= \langle A(\boldsymbol{v}), \boldsymbol{w}\rangle \ ; && \text{(by IP1)}
\end{aligned}$$

whence they agree with the matrix elements of A.

Similarly, the matrix elements of $(\lambda\, A + \mu\, B)^\dagger$ are given by

$$\begin{aligned}
\langle (\lambda\, A + \mu\, B)^\dagger(\boldsymbol{v}), \boldsymbol{w}\rangle &= \langle \boldsymbol{v}, (\lambda\, A + \mu\, B)(\boldsymbol{w})\rangle && \text{(by equation (1.41))}\\
&= \lambda\, \langle \boldsymbol{v}, A(\boldsymbol{w})\rangle + \mu\, \langle \boldsymbol{v}, B(\boldsymbol{w})\rangle && \text{(using IP2)}\\
&= \lambda\, \langle A^\dagger(\boldsymbol{v}), \boldsymbol{w}\rangle + \mu\, \langle B^\dagger(\boldsymbol{v}), \boldsymbol{w}\rangle && \text{(by equation (1.41))}\\
&= \langle (\lambda\, A^\dagger + \mu\, B^\dagger)(\boldsymbol{v}), \boldsymbol{w}\rangle \ , && \text{(using IP1,2)}
\end{aligned}$$

which agree with the matrix elements of $\lambda\, A^\dagger + \mu\, B^\dagger$.

Finally, the matrix elements of $(A \circ B)^\dagger$ are given by

$$\begin{aligned}
\langle (A \circ B)^\dagger(\boldsymbol{v}), \boldsymbol{w}\rangle &= \langle \boldsymbol{v}, (A \circ B)(\boldsymbol{w})\rangle && \text{(by equation (1.41))}\\
&= \langle \boldsymbol{v}, A(B(\boldsymbol{w}))\rangle && \text{(by equation (1.2))}\\
&= \langle A^\dagger(\boldsymbol{v}), B(\boldsymbol{w})\rangle && \text{(by equation (1.41))}\\
&= \langle B^\dagger(A^\dagger(\boldsymbol{v})), \boldsymbol{w}\rangle && \text{(by equation (1.41))}\\
&= \langle (B^\dagger \circ A^\dagger)(\boldsymbol{v}), \boldsymbol{w}\rangle \ , && \text{(by equation (1.2))}
\end{aligned}$$

which agree with the matrix elements of $B^\dagger \circ A^\dagger$.

A linear transformation is said to be **symmetric** if $A^\dagger = A$. It is said to be **orthogonal** if $A^\dagger \circ A = A \circ A^\dagger = \mathbb{1}$. In particular, orthogonal transformations preserve inner products:

$$\begin{aligned}
\langle A(\boldsymbol{v}), A(\boldsymbol{w})\rangle &= \langle \boldsymbol{v}, A^\dagger(A(\boldsymbol{w}))\rangle && \text{(by equation (1.41))}\\
&= \langle \boldsymbol{v}, (A^\dagger \circ A)(\boldsymbol{w})\rangle && \text{(by equation (1.2))}\\
&= \langle \boldsymbol{v}, \boldsymbol{w}\rangle \ . && \text{(since } A \text{ is orthogonal)}
\end{aligned}$$

Notice that in the above we only used the condition $A^\dagger \circ A = \mathbb{1}$ but not $A \circ A^\dagger = \mathbb{1}$. In a finite-dimensional vector space one implies the other, but in infinite dimensional vector spaces it may happen that a linear transformation which preserves the inner product obeys $A^\dagger \circ A = \mathbb{1}$ but does *not* obey $A \circ A^\dagger = \mathbb{1}$. (Maybe an example?)

To justify these names, notice that *relative to an orthonormal basis* the matrix of a symmetric transformation is symmetric and the matrix of an orthogonal transformation is orthogonal, as defined in Section 1.2.7. This follows because the matrix of the adjoint of a linear transformation is the transpose of the matrix of the linear transformation.

Let us prove this. Let $\{\boldsymbol{e}_i\}$ be an orthonormal basis and let $A : \mathbb{V} \to \mathbb{V}$ be a linear transformation. The matrix A of A relative to this basis has entries A_{ij} defined by

$$A(\boldsymbol{e}_i) = \sum_{j=1}^{N} A_{ji}\, \boldsymbol{e}_j \ .$$

The entries A_{ij} are also given by matrix elements:

$$\begin{aligned}
\langle A(\boldsymbol{e}_i), \boldsymbol{e}_j\rangle &= \langle \sum_{k=1}^{N} A_{ki}\, \boldsymbol{e}_k, \boldsymbol{e}_j\rangle\\
&= \sum_{k=1}^{N} A_{ki}\, \langle \boldsymbol{e}_k, \boldsymbol{e}_j\rangle && \text{(using IP1,2)}\\
&= A_{ji} \ . && \text{(using equation (1.38))}
\end{aligned}$$

In other words, relative to an orthonormal basis, we have the following useful formula:

$$\boxed{A_{ij} = \langle \boldsymbol{e}_i, A(\boldsymbol{e}_j)\rangle \ .} \tag{1.45}$$

From this it follows that the matrix of the adjoint $\mathsf{A}^\dagger$ relative to this basis is given by A^t. Indeed,

$$\begin{aligned}
A^\dagger_{ij} &= \langle A^\dagger(\boldsymbol{e}_j), \boldsymbol{e}_i\rangle\\
&= \langle \boldsymbol{e}_j, A(\boldsymbol{e}_i)\rangle && \text{(using equation (1.41))}\\
&= \langle A(\boldsymbol{e}_i), \boldsymbol{e}_j\rangle && \text{(using IP1)}\\
&= A_{ji} \ .
\end{aligned}$$

Therefore if $A^\dagger = A$, then $\mathsf{A} = \mathsf{A}^t$, and the matrix is symmetric. Similarly, if $A \circ A^\dagger = A^\dagger \circ A = \mathbb{1}$, then $\mathsf{A}^t\,\mathsf{A} = \mathsf{A}\,\mathsf{A}^t = \mathsf{I}$, and the matrix is orthogonal.

Notice that equations (1.42), (1.43) and (1.44) for the linear transformations are now seen to be consequences of equations (1.28), (1.31) and (1.30) applied to their matrices relative to an orthonormal basis.

1.3.5 Complex vector spaces

Much of what we have been saying about vector spaces remains true if we substitute the scalars and instead of real numbers consider complex numbers. Only the notion of an inner product will have to be changed in order for it to become useful. Inner products on complex vector spaces will be the subject of the next section; in this one, we want to emphasise those aspects of vectors spaces which remain unchanged when we extend the scalars from the real to the complex numbers.

As you know, complex numbers themselves can be understood as a real vector space of dimension two; that is, as $\mathbb{R}^2$. If $z = x + i\,y$ is a complex number with x, y real and $i = \sqrt{-1}$, then we can think of z as the pair $(x, y) \in \mathbb{R}^2$. Addition of complex numbers corresponds to vector addition in $\mathbb{R}^2$. Indeed, if $z = x + i\,y$ and $w = u + i\,v$ then $z + w = (x + u) + i\,(y + v)$, which is precisely what we expect from the vector addition $(x, y) + (u, v) = (x + u, y + v)$. Similarly, multiplication by a *real* number λ corresponds to scalar multiplication in $\mathbb{R}^2$. Indeed, $\lambda\,z = (\lambda\,x) + i\,(\lambda\,y)$, which is in agreement with $\lambda\,(x, y) = (\lambda\,x, \lambda\,y)$. However the complex numbers have more structure than that of a mere vector space. Unlike vectors in a general vector space, complex numbers can be multiplied: if $z = x + i\,y$ and $w = u + i\,v$, then $zw = (xu - yv) + i\,(xv + yu)$. Multiplication is commutative: $wz = zw$.

In a sense, complex numbers are more like matrices than vectors. Indeed, consider the 2×2 matrices of the form

$$\begin{matrix} a & -b \\ b & a \end{matrix} .$$

If we take the matrix product

$$\begin{matrix} x & -y & u & -v \\ y & x & v & u \end{matrix} = \begin{matrix} xu - yv & -(xv + yu) \\ xv + yu & xu - yv \end{matrix} ,$$

we see that we recover the multiplication of complex numbers. Notice that the complex number i is represented by the matrix

$$\mathsf{J} = \begin{matrix} 0 & -1 \\ 1 & 0 \end{matrix} ,$$

which obeys $\mathsf{J}^2 = -\mathsf{I}$.

A real matrix J obeying $\mathsf{J}^2 = -\mathsf{I}$ is called a **complex structure**.

We now briefly review some basic facts about complex numbers. Although you should be familiar with the following concepts, I will briefly review them here just to set the notation. As we have seen complex number can be added and multiplied. So far that is as with the real numbers, but in addition there is a notion of **complex conjugation**: $z = x + i\,y \mapsto z^* = x - i\,y$. Clearly conjugation is an involution: $(z^*)^* = z$. It also obeys $(zw)^* = z^*w^*$. A complex number z is said to be **real** if it is invariant under conjugation: $z^* = z$. Similarly a complex number is said to be **imaginary** if $z^* = -z$. Given $z = x + i\,y$, z is real if and only if $y = 0$, whereas z is imaginary if and only if $x = 0$. If $z = x + i\,y$, x is said to be the **real part** of z, written $x = \mathrm{Re}\,z$, and y is said to be the **imaginary part** of z, written $y = \mathrm{Im}\,z$. Notice that the imaginary part of a complex number is a real number, not an imaginary number! Given a complex number z, the product zz^* is real: $(zz^*)^* = zz^*$. It is written $|z|^2$ and it is called the **modulus** of z. If $z = x+i\,y$, then $|z|^2 = x^2 + y^2$, which coincides with the squared norm $\|(x, y)\|^2$ of the corresponding vector in the plane. Notice that the modulus is multiplicative: $|zw| = |z||w|$ and invariant under conjugation: $|z^*| = |z|$.

After this flash review of complex numbers, it is possible to define the notion of a complex vector space. There is really very little to do. Everything that was said in Sections 1.1 and 1.2 still holds provided we replace *real* with *complex* everywhere. An abstract complex vector space satisfies the same axioms, except that the scalars are now complex numbers as opposed to real numbers. Vector subspaces work the same way. Bases and linear independence also work in the same way, linear combinations being now complex linear combinations. The canonical example of a complex vector space is $\mathbb{C}^N$, the set of ordered N-tuples of complex numbers: $(z_1, z_2, \ldots, z_N)$, with the operations defined slot-wise as for $\mathbb{R}^N$. The canonical basis $\{(1, 0, \ldots, 0), (0, 1, \ldots, 0), \ldots, (0, 0, \ldots, 1)\}$ still spans $\mathbb{C}^N$, but where we now take complex linear combinations. As a result $\mathbb{C}^N$ has (complex) dimension N. If we only allowed ourselves to take real linear combinations, then in order to span $\mathbb{C}^N$ we would need in addition the N vectors $\{(i, 0, \ldots, 0), (0, i, \ldots, 0), \ldots, (0, 0, \ldots, i)\}$, showing that as a *real* vector space, $\mathbb{C}^N$ is $2N$-dimensional.

Linear maps and linear transformations are now complex linear and matrices and column vectors now have complex entries instead of real entries. Matrix invariants like the trace and the determinant are now complex numbers instead of real numbers. There is one more operation we can do with complex matrices, and that is to take complex conjugation. If A is a complex $N \times M$ matrix, then A^* is the $N \times M$ matrix whose entries are simply the

complex conjugates of the entries in A. Clearly, for square matrices,

$$\det(\mathsf{A}^*) = (\det \mathsf{A})^* \qquad \text{and} \qquad \operatorname{tr}(\mathsf{A}^*) = (\operatorname{tr} \mathsf{A})^* \ .$$

The only significant difference between real and complex vector spaces is when we introduce inner products, which we do now.

1.3.6 Hermitian inner products

We motivated the introduction of inner products as a way to measure, in particular, lengths of vectors. The need to compute lengths was motivated in turn by the fact that the vectorial quantities used in physics have a magnitude as well as a direction. Magnitudes, like anything else that one ever measures experimentally, are positive (or at least non-negative) real numbers. However if were to simply extend the dot product from $\mathbb{R}^N$ to $\mathbb{C}^N$, we would immediately notice that for $\boldsymbol{z} = (z_1, z_2, \ldots, z_N) \in \mathbb{C}^N$, the dot product with itself

$$\boldsymbol{z} \cdot \boldsymbol{z} = \sum_{i=1}^{N} z_i z_i \ ,$$

gives a complex number, not a real number. Hence we cannot understand this as a length. One way to generate a positive real number is to define the following inner product on $\mathbb{C}^N$:

$$\langle \boldsymbol{z}, \boldsymbol{w} \rangle = \sum_{i=1}^{N} z_i^* w_i \ ,$$

where $\boldsymbol{z} = (z_1, z_2, \ldots, z_N)$ and $\boldsymbol{w} = (w_1, w_2, \ldots, w_N)$. It is then easy to see that now

$$\langle \boldsymbol{z}, \boldsymbol{z} \rangle = \sum_{i=1}^{N} z_i^* z_i = \sum_{i=1}^{N} |z_i|^2 \ ,$$

so that this is a non-negative real number, so that it can be interpreted as a norm. The above inner product obeys the following property, in contrast with the dot product in $\mathbb{R}^N$: it is not symmetric, so rather than IP1 it obeys $\langle \boldsymbol{z}, \boldsymbol{w} \rangle = \langle \boldsymbol{w}, \boldsymbol{z} \rangle^*$.

This suggests the following definition. A complex valued function $\langle \cdot, \cdot \rangle : \mathbb{V} \times \mathbb{V} \to \mathbb{C}$ taking pairs of vectors to complex numbers is called a **hermitian inner product** if the following axioms are satisfied:

HIP1 $\langle \boldsymbol{z}, \boldsymbol{w} \rangle = \langle \boldsymbol{w}, \boldsymbol{z} \rangle^*$;

HIP2 $\langle \boldsymbol{x}, \lambda\, \boldsymbol{z} + \mu\, \boldsymbol{w} \rangle = \lambda \langle \boldsymbol{x}, \boldsymbol{z} \rangle + \mu \langle \boldsymbol{x}, \boldsymbol{w} \rangle$; and

HIP3 $\|\boldsymbol{z}\|^2 = \langle \boldsymbol{z}, \boldsymbol{z} \rangle > 0$ for all $\boldsymbol{z} \neq \boldsymbol{0}$,

where here λ and μ are complex scalars.

Except for the fact that $\langle \cdot, \cdot \rangle$ is a complex function, the only obvious difference is HIP1. Using HIP1 and HIP2 we see that

$$\begin{aligned} \langle \lambda\, \boldsymbol{z} + \mu\, \boldsymbol{w}, \boldsymbol{x} \rangle &= \langle \boldsymbol{x}, \lambda\, \boldsymbol{z} + \mu\, \boldsymbol{w} \rangle^* && \text{(by HIP1)} \\ &= (\lambda \langle \boldsymbol{x}, \boldsymbol{z} \rangle + \mu \langle \boldsymbol{x}, \boldsymbol{w} \rangle)^* && \text{(by HIP2)} \\ &= \lambda^* \langle \boldsymbol{x}, \boldsymbol{z} \rangle^* + \mu^* \langle \boldsymbol{x}, \boldsymbol{w} \rangle^* \\ &= \lambda^* \langle \boldsymbol{z}, \boldsymbol{x} \rangle + \mu \langle \boldsymbol{w}, \boldsymbol{x} \rangle\ , && \text{(using HIP1)} \end{aligned}$$

so that $\langle \cdot, \cdot \rangle$ is complex linear in the second slot but only conjugate linear in the first. One says that hermitian inner products are **sesquilinear**, which means 'one and a half' linear.

Just as in the real case, the inner product of any two vectors is determined by the matrix of inner products relative to any basis. Let $\{\boldsymbol{e}_i\}$ be a basis for $\mathbb{V}$. Let $\boldsymbol{v} = \sum_{i=1}^N v_i\, \boldsymbol{e}_i$ and $\boldsymbol{w} = \sum_{i=1}^N w_i\, \boldsymbol{e}_i$ be any two vectors. Then their inner product is given by

$$\begin{aligned} \langle \boldsymbol{v}, \boldsymbol{w} \rangle &= \langle \sum_{i=1}^N v_i\, \boldsymbol{e}_i, \sum_{j=1}^N v_j\, \boldsymbol{e}_j \rangle \\ &= \sum_{i,j=1}^N v_i^*\, w_j\, \langle \boldsymbol{e}_i, \boldsymbol{e}_j \rangle\ . && \text{(using HIP1,2)} \end{aligned}$$

In other words, all we need to know in order to compute this are the complex numbers $H_{ij} := \langle \boldsymbol{e}_i, \boldsymbol{e}_j \rangle$, which can be thought of as the entries of a matrix H. If we think of $\boldsymbol{v}$ as a column vector v in $\mathbb{C}^N$ whose entries are the components of $\boldsymbol{v}$ relative to the basis $\{\boldsymbol{e}_i\}$, and the same for $\boldsymbol{w}$, we can compute their inner product using matrix multiplication:

$$\langle \boldsymbol{v}, \boldsymbol{w} \rangle = (\mathsf{v}^*)^t\, \mathsf{H}\, \mathsf{w}\ .$$

We saw in the real case that the analogues matrix there was symmetric and positive-definite, reflecting the similar properties of the inner product. In the complex case, we expect that H should still be positive-definite but that instead of symmetry it should obey a property based on HIP1. Indeed, it follows from HIP1 that

$$H_{ij} = \langle \boldsymbol{e}_i, \boldsymbol{e}_j \rangle = \langle \boldsymbol{e}_j, \boldsymbol{e}_i \rangle^* = H_{ji}^*\ .$$

This means that the matrix H is equal to its conjugate transpose:

$$\mathsf{H} = (\mathsf{H}^*)^t\ . \tag{1.46}$$

Such matrices are called **hermitian**. Property HIP3 means that H is positive-definite, so that in particular it is non-degenerate.

Let us see how H transforms under a change of basis. Let $\{\boldsymbol{e}'_i\}$ be a new basis, with $\boldsymbol{e}'_i = S(\boldsymbol{e}_i)$ for some complex linear transformation S. Relative to $\{\boldsymbol{e}_i\}$ the linear transformation S is represented by a matrix S with entries S_{ji} given by equation (1.21). Let H′ denote the matrix describing the inner product in the new basis: its entries H'_{ij} are given by

$$\begin{aligned} H'_{ij} &= \langle \boldsymbol{e}'_i, \boldsymbol{e}'_j \rangle && \text{(by definition)} \\ &= \langle \sum_{k=1}^{N} S_{ki}\, \boldsymbol{e}_k, \sum_{l=1}^{N} S_{lj}\, \boldsymbol{e}_l \rangle && \text{(by equation (1.21))} \\ &= \sum_{k,l=1}^{N} S^*_{ki}\, S_{lj}\, \langle \boldsymbol{e}_k, \boldsymbol{e}_l \rangle && \text{(using HIP1,2)} \\ &= \sum_{k,l=1}^{N} S^*_{ki}\, H_{kl}\, S_{lj} \ . \end{aligned}$$

In other words,

$$\boxed{\mathsf{H}' = (\mathsf{S}^*)^t\ \mathsf{H}\,\mathsf{S}\ ,} \tag{1.47}$$

to be contrasted with the analogous formula (1.35).

The Cauchy–Schwartz and triangle inequalities are still valid for hermitian inner products. The proofs are essentially the same as for the real case. We will therefore be brief.

In order to prove the Cauchy–Schwarz inequality, we start the following inequality, which follows from HIP3,

$$\|\boldsymbol{v} - \lambda\, \boldsymbol{w}\|^2 \geq 0\ ,$$

and choose $\lambda \in \mathbb{C}$ appropriately. Expanding this out using HIP1 and HIP2 we can rewrite it as

$$\|\boldsymbol{v}\|^2 + |\lambda|^2\, \|\boldsymbol{w}\|^2 - \lambda\, \langle \boldsymbol{v}, \boldsymbol{w} \rangle - \lambda^*\, \langle \boldsymbol{w}, \boldsymbol{v} \rangle \geq 0\ .$$

Hence if we choose $\lambda = \langle \boldsymbol{w}, \boldsymbol{v} \rangle / \|\boldsymbol{w}\|^2$, we turn the inequality into

$$\|\boldsymbol{v}\|^2 - \frac{|\langle \boldsymbol{v}, \boldsymbol{w} \rangle|^2}{\|\boldsymbol{w}\|^2} \geq 0\ ,$$

which can be rewritten as

$$|\langle \boldsymbol{v}, \boldsymbol{w} \rangle|^2 \leq \|\boldsymbol{v}\|^2\, \|\boldsymbol{w}\|^2\ .$$

Taking square roots (all quantities are positive) we obtain the Cauchy–Schwarz inequality (1.36).

In order to prove the triangle inequality, we start with

$$\begin{aligned}
\|\boldsymbol{v}+\boldsymbol{w}\|^2 &= \langle \boldsymbol{v}+\boldsymbol{w}, \boldsymbol{v}+\boldsymbol{w}\rangle \\
&= \|\boldsymbol{v}\|^2 + \|\boldsymbol{w}\|^2 + 2\,\mathrm{Re}\langle \boldsymbol{v}, \boldsymbol{w}\rangle \\
&\le \|\boldsymbol{v}\|^2 + \|\boldsymbol{w}\|^2 + 2|\langle \boldsymbol{v}, \boldsymbol{w}\rangle| && \text{(since } \mathrm{Re}\, z \le |z| \ \forall z \in \mathbb{C}\text{)} \\
&\le \|\boldsymbol{v}\|^2 + \|\boldsymbol{w}\|^2 + 2\|\boldsymbol{v}\|\,\|\boldsymbol{w}\| && \text{(by Cauchy–Schwarz)} \\
&= \left(\|\boldsymbol{v}\| + \|\boldsymbol{w}\|\right)^2 \ ;
\end{aligned}$$

whence taking square roots we obtain the triangle inequality (1.37).

The complex analogue of an orthonormal basis is a unitary basis. Explicitly, a basis $\{\boldsymbol{e}_i\}$ is said to be **unitary** if

$$\langle \boldsymbol{e}_i, \boldsymbol{e}_j\rangle = \delta_{ij} := \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise.} \end{cases} \tag{1.48}$$

The components of a vector $\boldsymbol{v}$ relative to a unitary basis $\{\boldsymbol{e}_i\}$ can be computed by taking inner products, just as in the real case. Let $\boldsymbol{v} = \sum_{i=1}^N v_i\,\boldsymbol{e}_i$, and take its inner product with $\boldsymbol{e}_j$:

$$\begin{aligned}
\langle \boldsymbol{e}_j, \boldsymbol{v}\rangle &= \langle \boldsymbol{e}_j, \sum_{i=1}^N v_i\,\boldsymbol{e}_i\rangle \\
&= \sum_{i=1}^N v_i\,\langle \boldsymbol{e}_j, \boldsymbol{e}_i\rangle && \text{(using HIP2)} \\
&= v_j\ . && \text{(using equation (1.48))}
\end{aligned}$$

This shows that unitary vectors are automatically linearly independent.

One still has the Gram–Schmidt procedure for hermitian inner products. It works essentially in the same way as in the real case, so we will not spend much time on this. Consider a basis $\{\boldsymbol{f}_i\}$ for $\mathbb{V}$. Define the following vectors:

$$\boldsymbol{e}_i = \frac{\boldsymbol{f}_i - \sum_{j=1}^{i-1}\langle \boldsymbol{e}_j, \boldsymbol{f}_i\rangle\,\boldsymbol{e}_j}{\|\boldsymbol{f}_i - \sum_{j=1}^{i-1}\langle \boldsymbol{e}_j, \boldsymbol{f}_i\rangle\,\boldsymbol{e}_j\|}\ .$$

It is easily checked that they are a unitary basis. First of all each $\boldsymbol{e}_i$ is clearly normalised, because it is defined as a vector divided by its norm; and moreover if $i > j$, then $\boldsymbol{e}_i$ is clearly orthogonal to $\boldsymbol{e}_j$.

Finally, we discuss the adjoint of a complex linear map relative to a hermitian inner product. Let $A : \mathbb{V} \to \mathbb{V}$ be a complex linear map. We

define its adjoint $A^\dagger$ by equation (1.41), where now $\langle \cdot, \cdot \rangle$ is a hermitian inner product. The properties (1.42) and (1.44) still hold, and are proven in exactly the same way.

Only property (1.43) changes, reflecting the sesquilinear nature of the inner product. Indeed notice that

$$\begin{aligned}
\langle (\lambda\, A + \mu\, B)^\dagger\, \boldsymbol{v}, \boldsymbol{w} \rangle &= \langle \boldsymbol{v}, (\lambda\, A + \mu\, B)\, \boldsymbol{w} \rangle && \text{(by (1.41))} \\
&= \lambda\, \langle \boldsymbol{v}, A\, \boldsymbol{w} \rangle + \mu\, \langle \boldsymbol{v}, B\, \boldsymbol{w} \rangle && \text{(by HIP2)} \\
&= \lambda\, \langle A^\dagger\, \boldsymbol{v}, \boldsymbol{w} \rangle + \mu\, \langle B^\dagger\, \boldsymbol{v}, \boldsymbol{w} \rangle && \text{(by (1.41))} \\
&= \left(\lambda^*\, \langle \boldsymbol{w}, A^\dagger\, \boldsymbol{v} \rangle + \mu^*\, \langle \boldsymbol{w}, B^\dagger\, \boldsymbol{v} \rangle \right)^* && \text{(by HIP1)} \\
&= \langle \boldsymbol{w}, \left(\lambda^*\, A^\dagger + \mu^*\, B^\dagger \right)\, \boldsymbol{v} \rangle^* && \text{(by HIP2)} \\
&= \langle \left(\lambda^*\, A^\dagger + \mu^*\, B^\dagger \right)\, \boldsymbol{v}, \boldsymbol{w} \rangle\ ; && \text{(by HIP1)}
\end{aligned}$$

whence

$$\boxed{(\lambda\, A + \mu\, B)^\dagger = \lambda^*\, A^\dagger + \mu^*\, B^\dagger\ .} \tag{1.49}$$

A complex linear transformation A is said to be **hermitian** if $A^\dagger = A$, and it is said to be **anti-hermitian** (also **skew-hermitian**) if $A^\dagger = -A$. As in the real case, the nomenclature can be justified by noticing that the matrix of a hermitian transformation relative to a unitary basis is hermitian, as defined in equation (1.46). The proof is similar to the proof of the analogous statement in the real case. Indeed,

$$\begin{aligned}
A^\dagger_{ij} &= \langle A^\dagger(\boldsymbol{e}_j), \boldsymbol{e}_i \rangle && \text{(by equation (1.45))} \\
&= \langle \boldsymbol{e}_j, A(\boldsymbol{e}_i) \rangle && \text{(using equation (1.41))} \\
&= \langle A(\boldsymbol{e}_i), \boldsymbol{e}_j \rangle^* && \text{(using HIP1)} \\
&= A^*_{ji}\ . && \text{(by equation (1.45))}
\end{aligned}$$

Therefore if $A^\dagger = A$, then $\mathsf{A} = (\mathsf{A}^*)^t$, and the matrix is hermitian. Notice that if A is a hermitian matrix, then $i\,\mathsf{A}$ is antihermitian, hence unlike the real case, the distinction between hermitian and anti-hermitian is trivial.

Let us say that a linear transformation U is **unitary** if $U^\dagger \circ U = U \circ U^\dagger = \mathbb{1}$. In this case, its matrix U relative to a unitary basis obeys $(\mathsf{U}^*)^t\, \mathsf{U} = \mathsf{U}\, (\mathsf{U}^*)^t = \mathsf{I}$. This means that the conjugate transpose is the inverse,

$$\mathsf{U}^{-1} = (\mathsf{U}^*)^t\ . \tag{1.50}$$

Not surprisingly, such matrices are called **unitary**. Finally let us notice that

just as in the real case, a unitary transformation preserves the inner product:

$$\begin{aligned}\langle U(\boldsymbol{v}), U(\boldsymbol{w})\rangle &= \langle \boldsymbol{v}, U^\dagger(U(\boldsymbol{w}))\rangle && \text{(by equation (1.41))}\\ &= \langle \boldsymbol{v}, (U^\dagger \circ U)(\boldsymbol{w})\rangle && \text{(by equation (1.2))}\\ &= \langle \boldsymbol{v}, \boldsymbol{w}\rangle\ . && \text{(since } U \text{ is unitary)}\end{aligned}$$

1.4 The eigenvalue problem and applications

In this section we study perhaps the most important aspect of linear algebra from a physical perspective: the so-called eigenvalue problem. We mentioned when we introduced the notion of a basis that a good choice of basis can often simplify the solution of a problem involving linear transformations. Given a linear transformation, it is hard to imagine a better choice of basis than one in which the matrix is diagonal. However not all linear transformations admit such a basis. Understanding which transformations admit such basis is an important part of linear algebra; but one whose full solution requires more machinery than the one we will have available in this course. We will content ourselves with showing that certain types of linear transformation of use in physics do admit a diagonal basis. We will finish this section with two applications of these results: one to mathematics (quadratic forms) and one to physics (normal modes).

1.4.1 Eigenvectors and eigenvalues

Throughout this section $\mathbb{V}$ shall be an N-dimensional complex vector space.

Let $A : \mathbb{V} \to \mathbb{V}$ be a complex linear transformation. Let $\boldsymbol{v} \in \mathbb{V}$ be a *nonzero* vector which obeys

$$A\,\boldsymbol{v} = \lambda\,\boldsymbol{v} \qquad \text{for some } \lambda \in \mathbb{C}. \tag{1.51}$$

We say that $\boldsymbol{v}$ is an **eigenvector** of A with **eigenvalue** λ. Let $\{\boldsymbol{e}_i\}$ be a basis for $\mathbb{V}$. Let v be the column vector whose entries are the components v_i of $\boldsymbol{v}$ relative to this basis: $\boldsymbol{v} = \sum_i v_i \boldsymbol{e}_i$; and let A be the matrix representing A relative to this basis. Then equation (1.51) becomes

$$\mathsf{A}\,\mathsf{v} = \lambda\,\mathsf{v}\ . \tag{1.52}$$

Rewriting this as

$$(\mathsf{A} - \lambda\,\mathsf{I})\ \mathsf{v} = \mathsf{0}\ ,$$

we see that the matrix $\mathsf{A} - \lambda\,\mathsf{I}$ annihilates a nonzero vector, whence it must have zero determinant:

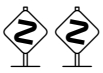
$$\det(\mathsf{A} - \lambda\,\mathsf{I}) = 0\ . \tag{1.53}$$

Let λ be an eigenvalue of A. The set of eigenvectors of A with eigenvalue λ, together with the zero vector, form a vector subspace $\mathbb{V}_\lambda$ of $\mathbb{V}$, known as the **eigenspace** of A with eigenvalue λ.

It is easy to prove this: all one needs to show is that $\mathbb{V}_\lambda$ is closed under vector addition and scalar multiplication. Indeed, let $\boldsymbol{v}$ and $\boldsymbol{w}$ be eigenvectors of A with eigenvalue λ and let α, β be scalars. Then

$$\begin{aligned} A(\alpha\,\boldsymbol{v}+\beta\,\boldsymbol{w}) &= \alpha\,A(\boldsymbol{v})+\beta\,A(\boldsymbol{w}) && \text{(by L1,2)}\\ &= \alpha\,\lambda\,\boldsymbol{v}+\beta\,\lambda\,\boldsymbol{w} && \text{(by equation (1.51))}\\ &= \lambda\,(\alpha\,\boldsymbol{v}+\beta\,\boldsymbol{w})\ , \end{aligned}$$

whence $\alpha\,\boldsymbol{v}+\beta\,\boldsymbol{w}$ is also an eigenvector of A with eigenvalue λ.



That $\mathbb{V}_\lambda$ is a subspace also follows trivially from the fact that it is the kernel of the linear transformation $A-\lambda\,\mathbb{1}$.

The dimension of the eigenspace $\mathbb{V}_\lambda$ is called the **multiplicity** of the eigenvalue λ. One says that an eigenvalue λ is **non-degenerate** if $\mathbb{V}_\lambda$ is one-dimensional and **degenerate** otherwise.

A linear transformation $A:\mathbb{V}\to\mathbb{V}$ is **diagonalisable** if there exists a basis $\{\boldsymbol{e}_i\}$ for $\mathbb{V}$ made up of eigenvectors of A. In this basis, the matrix A representing A is a diagonal matrix:

$$\mathsf{A}=\begin{pmatrix}\lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_N\end{pmatrix}\ ,$$

where not all of the λ_i need be distinct. In this basis we can compute the trace and the determinant very easily. We see that

$$\begin{aligned}\mathrm{tr}(\mathsf{A}) &= \lambda_1+\lambda_2+\cdots+\lambda_N=\sum_{i=1}^{N}\lambda_i\\ \det(\mathsf{A}) &= \lambda_1\lambda_2\cdots\lambda_N=\prod_{i=1}^{N}\lambda_i\ .\end{aligned}$$

Therefore the trace is the sum of the eigenvalues and the determinant is their product. This is independent of the basis, since both the trace and the determinant are invariants.

This has a very interesting consequence. Consider the identity:

$$\prod_{i=1}^{N}\exp(\lambda_i)=\exp\ \Big(\sum_{i=1}^{N}\lambda_i\Big)\ .$$

We can interpret this identity as an identity involving the diagonal matrix A:

$$\det\left(\exp(\mathsf{A})\right) = \exp\left(\mathrm{tr}(\mathsf{A})\right) \ ,$$

where the exponential of a matrix is defined via its Taylor series expansion:

$$\exp(\mathsf{A}) = \mathsf{I} + \mathsf{A} + \tfrac{1}{2}\mathsf{A}^2 + \tfrac{1}{3!}\mathsf{A}^3 + \cdots = \sum_{n=1}^{\infty} \tfrac{1}{n!}\,\mathsf{A}^n \ ,$$

so that for a diagonal matrix, it is simply the exponential of its diagonal entries. Now notice that under a change of basis given by $\mathsf{A} \mapsto \mathsf{A}'$, where A' is given by equation (1.24),

$$\begin{aligned}
\exp(\mathsf{A}') &= \sum_{n=1}^{\infty} \tfrac{1}{n!}\,(\mathsf{A}')^n \\
&= \sum_{n=1}^{\infty} \tfrac{1}{n!}\,(\mathsf{S}^{-1}\,\mathsf{A}\,\mathsf{S})^n && \text{(by equation (1.24))} \\
&= \sum_{n=1}^{\infty} \tfrac{1}{n!}\,\mathsf{S}^{-1}\,\mathsf{A}^n\,\mathsf{S} \\
&= \mathsf{S}^{-1}\,\exp(\mathsf{A})\,\mathsf{S} \ ;
\end{aligned}$$

whence because the trace and determinant are invariants

$$\det\ \exp(\mathsf{A}')\ = \exp\ \mathrm{tr}(\mathsf{A}')\ \ .$$

Hence this equation is still true for diagonalisable matrices. In fact, it follows from the fact (see next section) that diagonalisable matrices are dense in the space of matrices, that this identity is true for arbitrary matrices:

$$\boxed{\det\left(\exp(\mathsf{A})\right) = \exp\left(\mathrm{tr}(\mathsf{A})\right)\ .} \tag{1.54}$$

This is an extremely useful formula, particularly in quantum field theory and statistical mechanics, where it is usually applied to define the determinant of infinite-dimensional matrices.

1.4.2 Diagonalisability

Throughout this section $\mathbb{V}$ is an N-dimensional complex vector space.

It turns out that not every linear transformation is diagonalisable, but many of the interesting ones in physics will be. In this section, which lies somewhat outside the main scope of this course, we will state the condition for a linear transformation to be diagonalisable.

Fix a basis for $\mathbb{V}$ and let A be the matrix representing A relative to this basis. Let us define the following polynomial

$$\boxed{\chi_{\mathsf{A}}(t) = \det\left(\mathsf{A} - t\,\mathsf{I}\right)\ ,} \tag{1.55}$$

known as the **characteristic polynomial** of the matrix A. Under a change of basis, the matrix A changes to the matrix A' given by equation (1.24). The characteristic polynomial of the transformed matrix A' is given by

$$\begin{aligned}
\chi_{\mathsf{A}'}(t) &= \det\left(\mathsf{A}' - t\,\mathsf{I}\right) \\
&= \det\left(\mathsf{S}^{-1}\,\mathsf{A}\,\mathsf{S} - t\mathsf{I}\right) && \text{(by equation (1.24))} \\
&= \det\left(\mathsf{S}^{-1}\left(\mathsf{A} - t\mathsf{I}\right)\mathsf{S}\right) && \text{(since } \mathsf{S}^{-1}\,\mathsf{I}\,\mathsf{S} = \mathsf{I}\text{)} \\
&= \det\left(\mathsf{S}^{-1}\right)\,\det\left(\mathsf{A} - t\,\mathsf{I}\right)\,\det\left(\mathsf{S}\right) && \text{(by equation (1.18))} \\
&= \frac{1}{\det\left(\mathsf{S}\right)}\,\chi_{\mathsf{A}}(t)\,\det\left(\mathsf{S}\right) \\
&= \chi_{\mathsf{A}}(t)\ .
\end{aligned}$$

In other words, the characteristic polynomial is a matrix invariant and hence is a property of the linear transformation A. We will therefore define the **characteristic polynomial** $\chi_A(t)$ of a linear transformation $A : \mathbb{V} \to \mathbb{V}$ as the polynomial $\chi_{\mathsf{A}}(t)$ of the matrix which represents it relative to any basis. By the above calculation it does not depend on the basis.

The characteristic polynomial is a polynomial of order N where N is the complex dimension of $\mathbb{V}$. Its highest order term is of the form $(-1)^N t^N$ and its zeroth order term is the determinant of A, as can be seen by evaluating $\chi_A(t)$ at $t = 0$. In other words,

$$\chi_A(t) = \det(A) + \cdots + (-1)^N t^N\ .$$

Equation (1.53) implies that every eigenvalue λ of A is a root of its characteristic polynomial: $\chi_A(\lambda) = 0$. Conversely it is possible to prove that every root of the characteristic polynomial is an eigenvalue of A; although the multiplicities need not correspond: the multiplicity of the eigenvalue is never larger than that of the root.

This gives a method to compute the eigenvalues and eigenvectors of a linear transformation A. We simply choose a basis and find the matrix A representing A. We compute its characteristic polynomial and find its roots. For each root λ we solve the system of linear homogeneous equations:

$$\left(\mathsf{A} - \lambda\,\mathsf{I}\right)\,\mathsf{v} = \mathsf{0}\ .$$

This approach rests on the following general fact, known as the *Fundamental Theorem of Algebra*: every complex polynomial has a root. In fact, any complex polynomial of order N has N roots counted with multiplicity. In particular, the characteristic polynomial factorises into a product of monomials:

$$P(t) = (\lambda_1 - t)^{m_1}(\lambda_2 - t)^{m_2}\cdots(\lambda_k - t)^{m_k}\ ,$$

where all the λ_i are distinct and where $m_i \geq 1$ are positive integers. Clearly each λ_i is a root and m_i is its multiplicity. Each λ_i is also an eigenvalue of A, but m_i is *not* necessarily the multiplicity of the eigenvalue λ_i. Consider the matrix

$$\mathsf{A} = \begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix} ,$$

where $a \neq 0$ is any complex number. Its characteristic polynomial is given by

$$\chi_{\mathsf{A}}(t) = \det(\mathsf{A} - t\,\mathsf{I}) = \begin{Vmatrix} 1-t & a \\ 0 & 1-t \end{Vmatrix} = (1-t)^2 = 1 - 2t + t^2 .$$

Hence the only root of this polynomial is 1 with multiplicity 2. The number 1 is also an eigenvalue of A. For example, an eigenvector v is given by

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

However, the multiplicity of the eigenvalue 1 is only 1. Indeed, if it were 2, this would mean that there are two linearly independent eigenvectors with eigenvalue 1. These eigenvectors would then form a basis, relative to which A would be the identity matrix. But if $\mathsf{A} = \mathsf{I}$ relative to some basis, $\mathsf{A}' = \mathsf{I}$ relative to any other basis, since the identity matrix is invariant under change of basis. This violates the explicit expression for A above.

A result known as the **Cayley–Hamilton Theorem** states that any matrix A satisfies the following polynomial equation:

$$\chi_{\mathsf{A}}(\mathsf{A}) = \mathsf{0} ,$$

where $\mathsf{0}$ means the matrix all of whose entries are zero, and where a scalar a is replaced by the scalar matrix $a\,\mathsf{I}$. For example, consider the matrix A above:

$$\begin{aligned} \chi_{\mathsf{A}}(\mathsf{A}) &= \mathsf{I} - 2\mathsf{A} + \mathsf{A}^2 \\ &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - 2 \begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix}^2 \\ &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \begin{pmatrix} 2 & 2a \\ 0 & 2 \end{pmatrix} + \begin{pmatrix} 1 & 2a \\ 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} . \end{aligned}$$

The Cayley–Hamilton theorem shows that any $N \times N$ matrix obeys an N-th order polynomial equation. However in some cases an $N \times N$ matrix

A obeys a polynomial equation of smaller order. The polynomial $\mu_{\mathsf{A}}(t)$ of smallest order such that

$$\mu_{\mathsf{A}}(\mathsf{A}) = \mathsf{0} \ ,$$

is called the **minimal polynomial** of the matrix A. One can show that the minimal polynomial divides the characteristic polynomial. In fact, if the characteristic polynomial has the factorisation

$$\chi_{\mathsf{A}}(t) = (\lambda_1 - t)^{m_1}(\lambda_2 - t)^{m_2} \cdots (\lambda_k - t)^{m_k} \ ,$$

the minimal polynomial has the factorisation

$$\mu_{\mathsf{A}}(t) = (\lambda_1 - t)^{n_1}(\lambda_2 - t)^{n_2} \cdots (\lambda_k - t)^{n_k} \ ,$$

where $1 \leq n_i \leq m_i$. The main result in this topic is that a matrix A is diagonalisable if and only if all $n_i = 1$. For the non-diagonalisable matrix above, we see that its characteristic polynomial equals its minimal polynomial, since $\mathsf{A} \neq \mathsf{I}$.

In particular this shows that if all eigenvalues of a linear transformation are non-degenerate, then the linear transformation is diagonalisable. Given any matrix, one need only perturb it infinitesimally to lift any degeneracy its eigenvalues might have. This then implies that the diagonalisable matrices are dense in the space of matrices; that is, infinitesimally close to any non-diagonalisable matrix there is one which is diagonalisable. This is key to proving many identities involving matrices. If an identity of the form $f(\mathsf{A}) = \mathsf{0}$ holds for diagonalisable matrices then it holds for any matrix *provided* that f is a *continuous* function.

Computing the minimal polynomial of a linear transformation is not an easy task, hence it is in practice not very easy to decide whether or not a given linear transformation is diagonalisable. Luckily large classes of linear transformations can be shown to be diagonalisable, as we will now discuss.

1.4.3 Spectral theorem for hermitian transformations

Throughout this section $\mathbb{V}$ is an N-dimensional complex vector space with a hermitian inner product $\langle \cdot, \cdot \rangle$.

Let $A : \mathbb{V} \to \mathbb{V}$ be a hermitian linear transformation: $A^\dagger = A$. We will show that it is diagonalisable. As a corollary we will see that unitary transformations $U : \mathbb{V} \to \mathbb{V}$ such that $U^\dagger \circ U = U \circ U^\dagger = \mathbb{1}$ are also diagonalisable. These results are known as the spectral theorems for hermitian and unitary transformations.

We will first need to show two key results about the eigenvalues and eigenvectors of a hermitian transformation. First we will show that the eigenvalues

of a hermitian transformation are real. Let $\boldsymbol{v}$ be an eigenvector of A with eigenvalue λ. Then on the one hand,

$$\begin{aligned} \langle A(\boldsymbol{v}), \boldsymbol{v} \rangle &= \langle \lambda\, \boldsymbol{v}, \boldsymbol{v} \rangle \\ &= \lambda^* \langle \boldsymbol{v}, \boldsymbol{v} \rangle \ ; && \text{(by sesquilinearity)} \end{aligned}$$

whereas on the other hand,

$$\begin{aligned} \langle A(\boldsymbol{v}), \boldsymbol{v} \rangle &= \langle \boldsymbol{v}, A^\dagger(\boldsymbol{v}) \rangle && \text{(by equation (1.41))} \\ &= \langle \boldsymbol{v}, A(\boldsymbol{v}) \rangle && \text{(since } A \text{ is hermitian)} \\ &= \langle \boldsymbol{v}, \lambda\, \boldsymbol{v} \rangle \\ &= \lambda \langle \boldsymbol{v}, \boldsymbol{v} \rangle \ . && \text{(by HIP2)} \end{aligned}$$

Hence,

$$(\lambda - \lambda^*) \|\boldsymbol{v}\|^2 = 0 \ .$$

Since $\boldsymbol{v} \neq \boldsymbol{0}$, HIP3 implies that $\|\boldsymbol{v}\|^2 \neq 0$, whence $\lambda = \lambda^*$.

The second result is that eigenvectors corresponding to different eigenvalues are orthogonal. Let $\boldsymbol{v}$ and $\boldsymbol{w}$ be eigenvectors with distinct eigenvalues λ and μ, respectively. Then on the one hand,

$$\begin{aligned} \langle A(\boldsymbol{v}), \boldsymbol{w} \rangle &= \langle \lambda\, \boldsymbol{v}, \boldsymbol{w} \rangle \\ &= \lambda \langle \boldsymbol{v}, \boldsymbol{w} \rangle \ . && \text{(since } \lambda \text{ is real)} \end{aligned}$$

On the other hand,

$$\begin{aligned} \langle A(\boldsymbol{v}), \boldsymbol{w} \rangle &= \langle \boldsymbol{v}, A^\dagger(\boldsymbol{w}) \rangle && \text{(by equation (1.41))} \\ &= \langle \boldsymbol{v}, A(\boldsymbol{w}) \rangle && \text{(since } A \text{ is hermitian)} \\ &= \langle \boldsymbol{v}, \mu\, \boldsymbol{w} \rangle \\ &= \mu \langle \boldsymbol{v}, \boldsymbol{w} \rangle \ . && \text{(by HIP2)} \end{aligned}$$

Hence,

$$(\lambda - \mu) \langle \boldsymbol{v}, \boldsymbol{w} \rangle \ ,$$

whence if $\lambda \neq \mu$, $\boldsymbol{v} \perp \boldsymbol{w}$.

Now we need a basic fact: every hermitian transformation has at least one eigenvalue.

This can be shown using variational calculus. Consider the expression

$$f(\boldsymbol{v}) \equiv \langle \boldsymbol{v}, A(\boldsymbol{v}) \rangle \ .$$

We claim that $f(\boldsymbol{v})$ is a real number:

$$\begin{aligned} f(\boldsymbol{v})^* &= \langle \boldsymbol{v}, A(\boldsymbol{v})\rangle^* \\ &= \langle A(\boldsymbol{v}), \boldsymbol{v}\rangle && \text{(by HIP1)} \\ &= \langle \boldsymbol{v}, A^\dagger(\boldsymbol{v})\rangle && \text{(by equation (1.41))} \\ &= \langle \boldsymbol{v}, A(\boldsymbol{v})\rangle && \text{(since } A \text{ is hermitian)} \\ &= f(v)\ . \end{aligned}$$

Therefore f defines a continuous quadratic function from $\mathbb{V}$ to the real numbers. We would like to extremise this function. Clearly,

$$f(\alpha\,\boldsymbol{v}) = |\alpha|^2 f(\boldsymbol{v})\ ,$$

and this means that by rescaling $\boldsymbol{v}$ we can make $f(\boldsymbol{v})$ be as large or as small as we want. This is not the type of extremisation that we are interested: we want to see in which direction is $f(\boldsymbol{v})$ extremal. One way to do this is to restrict ourselves to vectors such that $\|\boldsymbol{v}\|^2 = 1$. This can be imposed using a Lagrange multiplier λ. Extremising $f(\boldsymbol{v})$ subject to the constraint $\|\boldsymbol{v}\|^2 = 1$, can be done by extremising the expression

$$I(\boldsymbol{v}, \lambda) = f(\boldsymbol{v}) - \lambda\,(\|\boldsymbol{v}\|^2 - 1)\ .$$

The variation of I yields the following expression:

$$\delta I = 2\,\langle \delta\boldsymbol{v}, (\mathsf{A} - \lambda\,\mathsf{I})\,\boldsymbol{v}\rangle - \delta\lambda\,(\|\boldsymbol{v}\|^2 - 1)\ .$$

Therefore the variational equations are $\|\boldsymbol{v}\|^2 = 1$ and

$$\mathsf{A}\,\boldsymbol{v} = \lambda\,\boldsymbol{v}\ ,$$

where we have used the non-degeneracy of the inner product and the fact that we want $\delta I = 0$ for all $\delta\lambda$ and $\delta\boldsymbol{v}$. Therefore this says that the extrema of I are the pairs $(\boldsymbol{v}, \lambda)$ where $\boldsymbol{v}$ is a normalised eigenvalue of A with eigenvalue λ. The function $I(\boldsymbol{v}, \lambda)$ takes the value $I(\boldsymbol{v}, \lambda) = \lambda$ at such a pair; whence the maxima and minima correspond to the largest and smallest eigenvalues. It remains to argue that the variational problem has solution. This follows from the compactness of the space of normalised vectors, which is the unit sphere in $\mathbb{V}$. The function $f(\boldsymbol{v})$ is continuous on the unit sphere and hence attains its maxima and minima in it.

We are now ready to prove the spectral theorem. We will first assume that the eigenvalues are non-degenerate, for ease of exposition and then we will relax this hypothesis and prove the general result.

Let $\boldsymbol{v}_1$ be a normalised eigenvector of A with eigenvalue λ_1. It exists from the above discussion and it is the only such eigenvector, up to scalar multiplication, by the non-degeneracy hypothesis. The eigenvalue is real as we saw above. Choose vectors $\{\boldsymbol{e}_2, \boldsymbol{e}_3, \ldots\}$ such that $\{\boldsymbol{v}_1, \boldsymbol{e}_2, \ldots\}$ is a basis for $\mathbb{V}$ and apply the Gram–Schmidt procedure if necessary so that it is a unitary basis. Let us look at the matrix A of A in such a basis. Because $\boldsymbol{e}_1$ is an eigenvector, one has

$$\langle A(\boldsymbol{v}_1), \boldsymbol{e}_j\rangle = \langle \lambda_1\,\boldsymbol{v}_1, \boldsymbol{e}_j\rangle = \lambda_1\,\langle \boldsymbol{v}_1, \boldsymbol{e}_j\rangle = 0\ ,$$

and similarly

$$\langle A(\boldsymbol{e}_j), \boldsymbol{v}_1 \rangle = \langle \boldsymbol{e}_j, A(\boldsymbol{v}_1) \rangle = \langle \boldsymbol{e}_j, \lambda_1 \, \boldsymbol{v}_1 \rangle = \lambda_1 \, \langle \boldsymbol{e}_j, \boldsymbol{v}_1 \rangle = 0 \ .$$

Moreover

$$\langle \boldsymbol{v}_1, A(\boldsymbol{v}_1) \rangle = \langle \boldsymbol{v}_1, \lambda_1 \, \boldsymbol{v}_1 \rangle = \lambda_1 \, \|\boldsymbol{e}_1\|^2 = \lambda_1 \ .$$

This means that the matrix takes the form

$$\begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & A_{22} & \cdots & A_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & A_{N2} & \cdots & A_{NN} \end{pmatrix} \ . \tag{1.56}$$

The submatrix

$$\begin{pmatrix} A_{22} & \cdots & A_{2N} \\ \vdots & \ddots & \vdots \\ A_{N2} & \cdots & A_{NN} \end{pmatrix} \ ,$$

is still hermitian, since for $i, j = 2, \ldots, N$,

$$A_{ij} = \langle \boldsymbol{e}_i, A(\boldsymbol{e}_j) \rangle = \langle A(\boldsymbol{e}_j), \boldsymbol{e}_i \rangle^* = \langle \boldsymbol{e}_j, A(\boldsymbol{e}_i) \rangle^* = A^*_{ji} \ .$$

Now we can apply the procedure again to this $(N-1) \times (N-1)$ matrix: we find a normalised eigenvector $\boldsymbol{v}_2$, which by assumption corresponds to a non-degenerate eigenvalue λ_2. Starting with this eigenvector we build a unitary basis $\{\boldsymbol{v}_2, \boldsymbol{e}'_3, \ldots\}$ for the $(N-1)$-dimensional subspace spanned by the $\{\boldsymbol{e}_2, \boldsymbol{e}_3, \ldots\}$. The submatrix $\mathsf{A}^{(N-1)}$ then takes the form analogous to the one in equation (1.56), leaving an $(N-2) \times (N-2)$ submatrix which is again still hermitian. We can apply the same procedure to this smaller matrix, and so on until we are left with a 1×1 hermitian matrix, i.e., a real number: λ_N. The basis $\{\boldsymbol{v}_i\}$ formed by the eigenvectors is clearly unitary, since each $\boldsymbol{v}_i$ is normalised by definition and is orthogonal to the preceding $\{\boldsymbol{v}_{j<i}\}$ by the way they were constructed. The matrix of A relative to this basis is then

$$\mathsf{A} = \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_N \end{pmatrix} \ ,$$

with real eigenvalues λ_i.

The case with degenerate eigenvalues works along similar lines. We start with an eigenvalue λ_1 and consider the eigenspace $\mathbb{V}_{\lambda_1}$. It may be that the

dimension m_1 of $\mathbb{V}_{\lambda_1}$ is larger than 1. In any case, every vector in $\mathbb{V}_{\lambda_1}$ is an eigenvector of A. Use Gram–Schmidt to find a unitary basis $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_{m_1}\}$ for $\mathbb{V}_{\lambda_1}$. Complete this basis to a unitary basis $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_{m_1}, \boldsymbol{e}_{m_1+1}, \ldots, \boldsymbol{e}_N\}$ for $\mathbb{V}$, which can be done using Gram–Schmidt if necessary again. The matrix A representing A in this basis is given by

$$\mathsf{A} = \begin{pmatrix} \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_1 \end{pmatrix} & \\ & \begin{pmatrix} A_{m_1+1,m_1+1} & \cdots & A_{m_1+1,N} \\ \vdots & \ddots & \vdots \\ A_{N,m_1+1} & \cdots & A_{NN} \end{pmatrix} \end{pmatrix} ,$$

where the off-diagonal blocks have vanishing entries because

$$\langle \boldsymbol{e}_i, A(\boldsymbol{v}_j) \rangle = \lambda_1 \langle \boldsymbol{e}_i, \boldsymbol{v}_j \rangle = 0 \ .$$

The submatrix

$$\begin{pmatrix} A_{m_1+1,m_1+1} & \cdots & A_{m_1+1,N} \\ \vdots & \ddots & \vdots \\ A_{N,m_1+1} & \cdots & A_{NN} \end{pmatrix}$$

is again hermitian, so we can apply the procedure again to it, until we are left with a basis $\{\boldsymbol{v}_i\}$ of eigenvectors of A, so that the matrix is diagonal.

In summary, suppose that we start with a hermitian matrix A, thought of as the matrix of a hermitian linear transformation A relative to a unitary basis. Then the above iterative procedure produces a unitary basis relative to which the matrix for A is diagonal. Because the initial and final basis are unitary, the change of basis transformation U is unitary. In other words, given a hermitian matrix A there is a unitary matrix U such that

$$\mathsf{A}' = \mathsf{U}^{-1} \, \mathsf{A} \, \mathsf{U} = (\mathsf{U}^*)^t \, \mathsf{A} \, \mathsf{U}$$

is diagonal. In other words,

Every hermitian matrix can be diagonalised by a unitary transformation.

In fact the unitary matrix U above can be written down explicitly in terms of the normalised eigenvectors of A. Let $\{\mathsf{v}_i\}$ be a set of normalised eigenvectors which are mutually orthogonal. This is guaranteed if they have

different eigenvalues, and in the case of degenerate eigenvalues by Gram–Schmidt. Consider the matrix

$$\mathsf{U} = \begin{pmatrix} \uparrow & \uparrow & \cdots & \uparrow \\ \mathsf{v}_1 & \mathsf{v}_2 & \cdots & \mathsf{v}_N \\ \downarrow & \downarrow & \cdots & \downarrow \end{pmatrix} .$$

We claim first of all that U is unitary. Indeed,

$$(\mathsf{U}^*)^t = \begin{pmatrix} \leftarrow & (\mathsf{v}_1^*)^t & \rightarrow \\ \leftarrow & (\mathsf{v}_2^*)^t & \rightarrow \\ \vdots & \vdots & \vdots \\ \leftarrow & (\mathsf{v}_N^*)^t & \rightarrow \end{pmatrix} .$$

Hence

$$\begin{aligned} (\mathsf{U}^*)^t\,\mathsf{U} &= \begin{pmatrix} \leftarrow & (\mathsf{v}_1^*)^t & \rightarrow \\ \leftarrow & (\mathsf{v}_2^*)^t & \rightarrow \\ \vdots & \vdots & \vdots \\ \leftarrow & (\mathsf{v}_N^*)^t & \rightarrow \end{pmatrix} \begin{pmatrix} \uparrow & \uparrow & \cdots & \uparrow \\ \mathsf{v}_1 & \mathsf{v}_2 & \cdots & \mathsf{v}_N \\ \downarrow & \downarrow & \cdots & \downarrow \end{pmatrix} \\ &= \begin{pmatrix} & (\mathsf{v}_i^*)^t \mathsf{v}_j & \\ & & \end{pmatrix} \\ &= \mathsf{I} , \end{aligned}$$

since the $\{\mathsf{v}_i\}$ form a unitary basis. Moreover, paying attention to the way matrix multiplication is defined and using that the $\{\mathsf{v}_i\}$ are eigenvectors of A, we find

$$\begin{aligned} \mathsf{A}\,\mathsf{U} &= \begin{pmatrix} \uparrow & \uparrow & \cdots & \uparrow \\ \lambda_1\,\mathsf{v}_1 & \lambda_2\,\mathsf{v}_2 & \cdots & \lambda_N\,\mathsf{v}_N \\ \downarrow & \downarrow & \cdots & \downarrow \end{pmatrix} \\ &= \begin{pmatrix} \uparrow & \uparrow & \cdots & \uparrow \\ \mathsf{v}_1 & \mathsf{v}_2 & \cdots & \mathsf{v}_N \\ \downarrow & \downarrow & \cdots & \downarrow \end{pmatrix} \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_N \end{pmatrix} \\ &= \mathsf{U}\,\mathsf{A}' . \end{aligned}$$

In other words, $\mathsf{A}' = \mathsf{U}^{-1}\,\mathsf{A}\,\mathsf{U}$ just as above.

There is a real version of this result: if A is real and symmetric, then it is hermitian. We can diagonalise it with a unitary transformation, which is

also real, whence it is orthogonal. This then yields the spectral theorem for symmetric real matrices, which says that any real symmetric matrix can be diagonalised by an orthogonal transformation.

We can now understand the positive-definiteness condition on the matrix representing a inner product on a vector space. We saw in Section 1.3.1 that the matrix G of an inner product in a real vector space is symmetric. Hence it can be diagonalised by an orthogonal transformation. From equation (1.35), it follows that there is a basis relative to which the matrix of the inner product is diagonal. Let $\{\boldsymbol{e}_i\}$ be such a basis and let $\langle \boldsymbol{e}_i, \boldsymbol{e}_j \rangle = \lambda_i \, \delta_{ij}$. If $\boldsymbol{v} = \sum_i v_i \boldsymbol{e}_i$ is any vector, then

$$\|\boldsymbol{v}\|^2 = \langle \boldsymbol{v}, \boldsymbol{v} \rangle = \sum_i \lambda_i v_i^2 \ .$$

Axiom IP3 says that this quantity has to be positive for all nonzero vectors $\boldsymbol{v}$, which clearly implies that $\lambda_i > 0$ for all i. Therefore a symmetric matrix is positive definite if and only if all its eigenvalues are positive. A similar statement also holds for hermitian inner products, whose proof is left as an exercise.

It is not just hermitian matrices that can be diagonalised by unitary transformations. Let us say that a linear transformation N is **normal** if it commutes with its adjoint

$$N^\dagger \circ N = N \circ N^\dagger \ . \tag{1.57}$$

Then it can be proven that a N is diagonalisable by a unitary transformation. As an example consider the 3×3 matrix

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

considered in the Exercises. We saw that it was diagonalisable by a unitary transformation, yet it is clearly not hermitian. Nevertheless it is easy to check that it is normal. Indeed,

$$(\mathsf{P}^*)^t = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \ ,$$

so that

$$\mathsf{P}\,(\mathsf{P}^*)^t = (\mathsf{P}^*)^t\,\mathsf{P} = \mathsf{I} \ ;$$

in other words, it is unitary.

It follows from the spectral theorem for hermitian transformations that unitary transformations can also be diagonalised by unitary transformations. This is known as the **Cayley transformation**, which is discussed in detail in the Problems. It follows from the Cayley transformation that the eigenvalues of a unitary matrix take values in the unit circle in the complex plane. This can also be seen directly as follows. Let U be a unitary transformation and

let $\boldsymbol{v}$ be an eigenvector with eigenvalue λ. Then consider $\|U(\boldsymbol{v})\|^2$. Because U is unitary,

$$\|U(\boldsymbol{v})\|^2 = \|\boldsymbol{v}\|^2 \ ,$$

but because $\boldsymbol{v}$ is an eigenvector,

$$\|U(\boldsymbol{v})\|^2 = \|\lambda\, \boldsymbol{v}\|^2 = |\lambda|^2\, \|\boldsymbol{v}\|^2 \ ,$$

whence $|\lambda|^2 = 1$.

Spectral theorems are extremely powerful in many areas of physics and mathematics, and in the next sections we will discuss two such applications. However the real power of the spectral theorem manifests itself in quantum mechanics, although the version of the theorem used there is the one for self-adjoint operators in an infinite-dimensional Hilbert space, which we will not have the opportunity to discuss in this course.

1.4.4 Application: quadratic forms

In this section we discuss a mathematical application of the spectral theorem for real symmetric transformations.

Let us start with the simplest case of a two-dimensional quadratic form. By a **quadratic form** on two variables (x_1, x_2) we mean a quadratic polynomial of the form

$$Q(x_1, x_2) = a x_1^2 + 2b x_1 x_2 + c x_2^2 \ , \tag{1.58}$$

for some real constants a, b, c. By a **quadric** we mean the solutions (x_1, x_2) of an equation of the form

$$Q(x_1, x_2) = d \ ,$$

where d is some real number and Q is a quadratic form. For example, we can take

$$Q_1(x_1, x_2) = x_1^2 + x_2^2 \ ,$$

in which case the quadrics $Q_1(x_1, x_2) = d$ for $d > 0$ describe a circle of radius $\sqrt{d}$ in the plane coordinatised by (x_1, x_2). To investigate the type of quadric that a quadratic form gives rise to, it is convenient to diagonalise it: that it, change to coordinates (y_1, y_2) for which the mixed term $y_1\, y_2$ in the quadratic form is not present. To tie this to the spectral theorem, it is convenient to rewrite this in terms of matrices. In terms of the column vector $\mathsf{x} = (x_1, x_2)^t$, the general two-dimensional quadratic form in equation (1.58) can be written as

$$Q(x_1, x_2) = \mathsf{x}^t\, \mathsf{Q}\, \mathsf{x} \ ,$$

where Q is the matrix

$$\mathsf{Q} = \begin{pmatrix} a & b \\ b & c \end{pmatrix} \ .$$

Because Q is symmetric, it can be diagonalised by an orthogonal transformation which is built out of the normalised eigenvectors as was explained in the previous section. Hence there is an orthogonal matrix O such that $\mathsf{Q} = \mathsf{O}\,\mathsf{D}\,\mathsf{O}^t$, where D is a diagonal matrix with entries λ_i, for $i = 1, 2$. That means that in terms of the new coordinates

$$\mathsf{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \mathsf{O}^t\,\mathsf{x} \ ,$$

the quadratic form is diagonal

$$Q(y_1, y_2) = \mathsf{y}^t\,\mathsf{D}\,\mathsf{y} = \lambda_1\, y_1^2 + \lambda_2\, y_2^2 \ .$$

We can further rescale the coordinates $\{y_i\}$: $z_i = \mu_i y_i$, where μ_i is real. This means that relative to the new coordinates z_i, the quadratic form takes the form

$$Q(z_1, z_2) = \varepsilon_1 z_1^2 + \varepsilon_2\, y_2^2 \ ,$$

where ε_i are $0, \pm 1$.

We can distinguish three types of quadrics, depending on the relative signs of the eigenvalues:

1. ($\varepsilon_1\varepsilon_2 = 1$) In this case the eigenvalues have the same sign and the quadric is an ellipse.

2. ($\varepsilon_1\varepsilon_2 = -1$) In this case the eigenvalues have different sign and the quadric is a hyperbola.

3. ($\varepsilon_1\varepsilon_2 = 0$) In this case one of the eigenvalues is zero, and the quadric consists of a pair of lines.

The general case is not much more complicated. Let $\mathbb{V}$ be a real vector space of dimension N with an inner product. By a **quadratic form** we mean a symmetric bilinear form $Q : \mathbb{V} \times \mathbb{V} \to \mathbb{R}$. In other words, Q satisfies axioms IP1 and IP2 of an inner product, but IP3 need *not* be satisfied. Associated to every quadratic form there is a linear transformation in $\mathbb{V}$, which we also denote Q, defined as follows

$$\langle \boldsymbol{v}, Q(\boldsymbol{w}) \rangle = Q(\boldsymbol{v}, \boldsymbol{w}) \ .$$

Symmetry of the bilinear form implies that the linear transformation Q is also symmetric:

$$\langle \boldsymbol{v}, Q(\boldsymbol{w})\rangle = Q(\boldsymbol{v},\boldsymbol{w}) = Q(\boldsymbol{w},\boldsymbol{v}) = \langle \boldsymbol{w}, Q(\boldsymbol{v})\rangle = \langle Q(\boldsymbol{v}), \boldsymbol{w}\rangle \ .$$

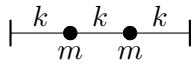
Hence it can be diagonalised by an orthogonal transformation. Relative to an orthonormal basis $\{\boldsymbol{e}_i\}$ for $\mathbb{V}$, Q is represented by a symmetric matrix Q. Let O be an orthogonal matrix which diagonalises Q; that is, $\mathsf{Q} = \mathsf{O}\,\mathsf{D}\,\mathsf{O}^t$, with D diagonal.

We can further change basis to an orthogonal basis whose elements are however no longer normalised, in such a way that the resulting matrix D' is still diagonal with all its entries either ± 1 or 0. Let (n_+, n_-, n_0) denote, respectively, the number of positive, negative and zero diagonal entries of D'. There is a result, known as **Sylvester's Law of Inertia**, which says that the numbers (n_+, n_-, n_0) are an invariant of the quadratic form, so that they can be computed from the matrix of the quadratic form relative to any basis. A quadratic form is said to be **non-degenerate** if $n_0 = 0$. It is said to be **positive-definite** if $n_- = n_0 = 0$, and **negative-definite** if $n_+ = n_0 = 0$. Clearly a quadratic form is an inner product when it is positive-definite. A non-degenerate quadratic form, which is not necessarily positive- or negative-definite, defines a **generalised inner product** on $\mathbb{V}$. There are two integers which characterise a non-degenerate quadratic form: the dimension N of the vector space, and the **signature** $n_+ - n_-$. Notice that if the signature is bounded above by the dimension: the bound being saturated when the quadratic form is positive-definite. There are plenty of interesting non-degenerate quadratic forms which are not positive-definite. For example, Minkowski spacetime in the theory of special relativity possesses a quadratic form with dimension 4 and signature 2.

1.4.5 Application: normal modes

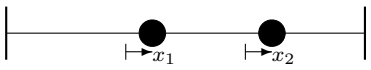
This section discusses the powerful method of normal modes to decouple interacting mechanical systems near equilibrium. It is perhaps not too exaggerated to suggest that theoretical physicists spend a large part of their lives studying the problem of normal modes in one way or another.

We start with a simple example.



Consider an idealised one-dimensional mechanical system consisting of two point masses each of mass m connected by springs to each other and to two fixed ends. We will neglect gravity, friction and the mass of the springs. The springs obey Hooke's law with spring constant k. We assume that the system is at equilibrium when the springs are relaxed, and we want to study the system around

equilibrium; that is, we wish to study small displacements of the masses. We let x_i for $i = 1, 2$ denote the displacements from equilibrium for each of the two point masses, as shown below.



Then the potential energy due to the springs is the sum of the potential energies of each of the springs:

$$\begin{aligned} V &= \tfrac{1}{2}k\,x_1^2 + \tfrac{1}{2}k\,(x_2 - x_1)^2 + \tfrac{1}{2}k\,x_2^2 \\ &= k\,\left(x_1^2 + x_2^2 - x_1 x_2\right)\ . \end{aligned}$$

The kinetic energy is given by

$$T = \tfrac{1}{2}m\dot{x}_1^2 + \tfrac{1}{2}m\dot{x}_2^2\ .$$

The equations of motion are then, for $i = 1, 2$,

$$\frac{d}{dt}\,\frac{\partial T}{\partial \dot{x}_i} = -\frac{\partial V}{\partial x_i}\ .$$

Explicitly, we have the following coupled system of second order ordinary differential equations:

$$\begin{aligned} m\ddot{x}_1 &= -2kx_1 + kx_2 \\ m\ddot{x}_2 &= -2kx_2 + kx_1\ . \end{aligned}$$

Let us write this in matrix form. We introduce a column vector $\mathsf{x}^t = (x_1, x_2)$. Then the above system of equations becomes

$$\ddot{\mathsf{x}} = -\omega^2\,\mathsf{K}\,\mathsf{x}\ , \tag{1.59}$$

where K is the matrix

$$\mathsf{K} = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}\ ,$$

and where we have introduced the notation

$$\omega \equiv \sqrt{\frac{k}{m}}\ .$$

Notice that K is symmetric, hence it can be diagonalised by an orthogonal transformation. Let us find its eigenvalues and its eigenvectors. The characteristic polynomial of K is given by

$$\chi_{\mathsf{K}}(\lambda) = \begin{Vmatrix} 2-\lambda & -1 \\ -1 & 2-\lambda \end{Vmatrix} = (2-\lambda)^2 - 1 = (\lambda - 1)(\lambda - 3) \ ,$$

from which it follows that it has as roots $\lambda = 1, 3$. The normalised eigenvectors corresponding to these eigenvalues are

$$\mathsf{v}_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \ , \quad \text{and} \quad \mathsf{v}_3 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix} \ ,$$

respectively. We build the following matrix O out of the normalised eigenvectors

$$\mathsf{O} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \ .$$

One can check that O is orthogonal: $\mathsf{O}^t = \mathsf{O}^{-1}$. One can also check that

$$\mathsf{K} = \mathsf{O}\,\mathsf{D}\,\mathsf{O}^t \ ,$$

where D is the diagonal matrix

$$\mathsf{D} = \begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix} \ .$$

Inserting this expression into equation (1.59), we see that

$$\ddot{\mathsf{x}} = -\omega^2 \mathsf{O}\,\mathsf{D}\,\mathsf{O}^t\,\mathsf{x} \ .$$

In terms of the new variables

$$\mathsf{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \mathsf{O}^t\,\mathsf{x} \ ,$$

the equation of motion (1.59) becomes

$$\ddot{\mathsf{y}} = -\omega^2\,\mathsf{D}\,\mathsf{y} \ . \tag{1.60}$$

Because the matrix D is diagonal, the equations of motion for the new variables y_i are now decoupled:

$$\ddot{y}_1 = -\omega^2 y_1 \qquad \text{and} \qquad \ddot{y}_2 = -3\omega^2 y_2 \ .$$

One can now easily solve these equations,

$$y_1(t) = A_1 \cos(\omega_1 t + \varphi_1)$$
$$y_2(t) = A_2 \cos(\omega_2 t + \varphi_2) \ ,$$

where $\omega_1 = \omega$, $\omega_2 = \sqrt{3}\,\omega$ and A_i and φ_i are constants to be determined from the initial conditions. The physical variables in the original problem are the displacements x_i of each of the point masses. They can be found in terms of the new decoupled variables y_i simply by inverting the change of variables (1.60). Explicitly,

$$x_1(t) = \frac{A_1}{\sqrt{2}} \cos(\omega_1 t + \varphi_1) + \frac{A_2}{\sqrt{2}} \cos(\omega_2 t + \varphi_2)$$
$$x_2(t) = \frac{A_1}{\sqrt{2}} \cos(\omega_1 t + \varphi_1) - \frac{A_2}{\sqrt{2}} \cos(\omega_2 t + \varphi_2) \ .$$

Variables like the y_i which decouple the equations of motion are called the **normal modes** of the mechanical system. Their virtue is that they reduce an interacting (i.e., coupled) mechanical system around equilibrium to a set of independent free oscillators. Each of these free oscillators are mathematical constructs: the normal modes do not generally correspond to the motion of any of the masses in the original system, but they nevertheless possess a certain "physicality" and it is fruitful to work with them as if they were physical. The original physical variables can then be understood as linear combinations of the normal modes as we saw above. The frequencies ω_i of the normal modes are known as the **characteristic frequencies** of the mechanical system. In particle physics, for example, the elementary particles are the normal modes and their masses are the characteristic frequencies.

To illustrate the simplification in the dynamics which results from considering the normal modes, in Figure 1.1 we have sketched the motion of the two masses in the problem and of the two normal modes, with time running horizontally to the right.

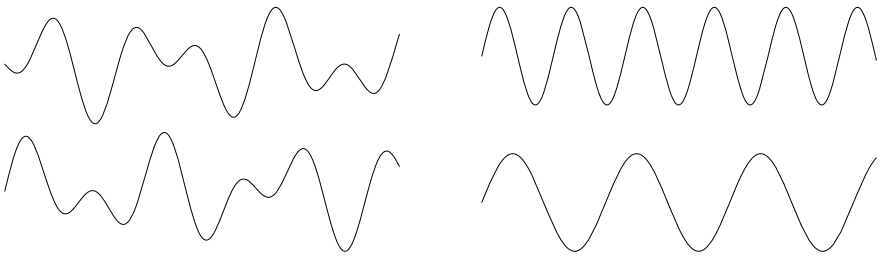
Notice also that although the motion of each of the normal modes is periodic, the system as a whole is not. This is due to the fact that the characteristic frequencies are not rational multiples of each other.

Let us see this. Suppose that we have to oscillators with frequencies ω_1 and ω_2. That means that the oscillators are periodic with periods $T_1 = 2\pi/\omega_1$ and $T_2 = 2\pi/\omega_2$. The combined system will be periodic provided that $N_1 T_1 = N_2 T_2$ for some integers N_i. But this means that

$$\frac{\omega_1}{\omega_2} = \frac{N_1}{N_2} \ ,$$

which is a rational number. In the problem treated above, the ratio

$$\frac{\omega_1}{\omega_2} = \frac{1}{\sqrt{3}} = \frac{\sqrt{3}}{3}$$



(a) Point masses (b) Normal modes

Figure 1.1: Dynamics of point masses and normal modes.

is irrational. Therefore the motion is aperiodic.

If we were to plot the trajectory of the system in the plane, with the trajectory of one of the point masses along the x-axis and the trajectory of the other point mass along the y-axis, we see that the orbit never repeats, and that we end up filling up the available configuration space. In Figure 1.2 we have plotted the cumulative trajectory of the system after letting it run for T units of time, for different values of T. As you can see, as T grows the system has visited more and more points in the available configuration space. Asymptotically, as $T \to \infty$, the system will have visited the whole available space.

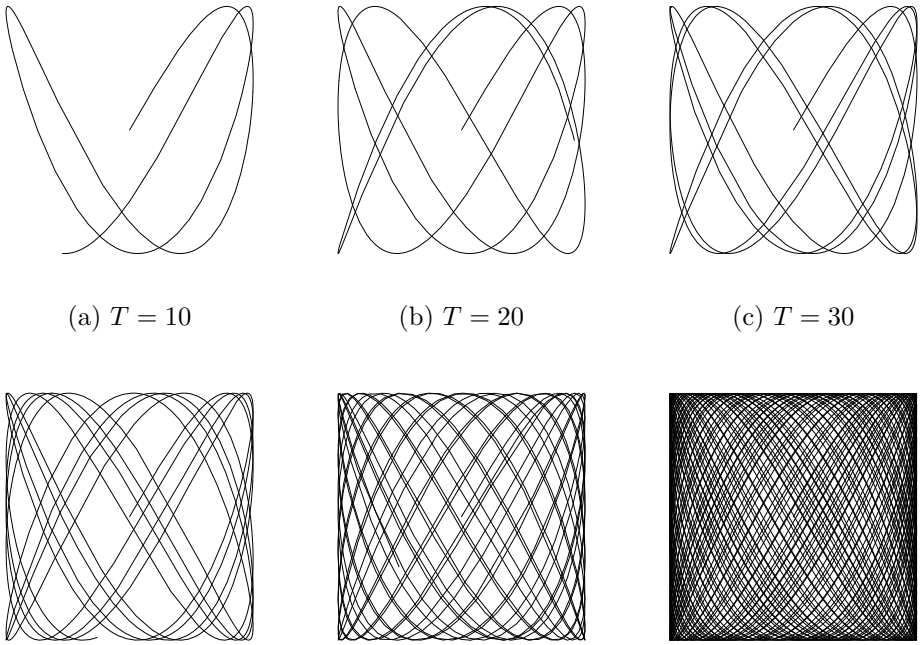
1.4.6 Application: near equilibrium dynamics

In this section we will consider a more general mechanical system near equilibrium. Throughout the section $\mathbb{V}$ will be a real finite-dimensional vector space with an inner product.

Consider a mechanical system whose configuration space is $\mathbb{V}$. For example, it could be a system of n point particles in d dimensions, and then $\mathbb{V}$ would be an (nd)-dimensional vector space. In the previous section we discussed the case of a one-dimensional system consisting of two point particles, so that $\mathbb{V}$ was two-dimensional. In the Problems we looked at systems with three-dimensional $\mathbb{V}$. In this section we are letting $\mathbb{V}$ be arbitrary but finite-dimensional.

The potential energy is given by a function $V : \mathbb{V} \to \mathbb{R}$. The configurations of mechanical equilibrium are those for which the gradient of the potential vanishes. Hence let us consider one such equilibrium configuration $\boldsymbol{q}_0 \in \mathbb{V}$:

$$\nabla V|_{\boldsymbol{q}_0} = \mathbf{0} \ .$$



(a) $T = 10$ (b) $T = 20$ (c) $T = 30$

(d) $T = 50$ (e) $T = 100$ (f) $T = 300$

Figure 1.2: Trajectory of the mechanical system at different times.

Because the potential energy is only defined up to an additive constant, we are free to choose it such that $V(\boldsymbol{q}_0) = 0$. We can therefore expand the potential function V about $\boldsymbol{q}_0$ and the first contribution will be quadratic:

$$\begin{aligned} V(\boldsymbol{q}) &= V(\boldsymbol{q}_0) + \langle \nabla V|_{\boldsymbol{q}_0}, \boldsymbol{q} - \boldsymbol{q}_0 \rangle + \tfrac{1}{2}\langle \boldsymbol{q} - \boldsymbol{q}_0, H(\boldsymbol{q} - \boldsymbol{q}_0)\rangle + \cdots \\ &= \tfrac{1}{2}\langle \boldsymbol{q} - \boldsymbol{q}_0, H(\boldsymbol{q} - \boldsymbol{q}_0)\rangle \ , \end{aligned}$$

where $H : \mathbb{V} \to \mathbb{V}$ is a symmetric linear transformation known as the **Hessian** of V at $\boldsymbol{q}_0$. Explicitly, if we choose an orthonormal basis $\{\boldsymbol{e}_i\}$ for $\mathbb{V}$, then let $\boldsymbol{q} = \sum_i q_i \boldsymbol{e}_i$ define some coordinates q_i for the configuration space. Then relative to this basis the Hessian of V has matrix elements

$$H_{ij} = \langle \boldsymbol{e}_i, H(\boldsymbol{e}_j)\rangle = \left.\frac{\partial^2 V}{\partial q_i \partial q_j}\right|_{\boldsymbol{q}_0} \ ,$$

which shows manifestly that it is symmetric: $H_{ij} = H_{ji}$. Let us define $\boldsymbol{x} = \boldsymbol{q} - \boldsymbol{q}_0$ to be the displacements about equilibrium. These will be our dynamical variables. The potential energy in the quadratic approximation is given by

$$V = \tfrac{1}{2}\langle \boldsymbol{x}, H(\boldsymbol{x})\rangle \ .$$

We will make the assumption that the kinetic energy is quadratic in the velocities $\dot{\boldsymbol{x}}$:

$$T = \tfrac{1}{2}\langle \dot{\boldsymbol{x}}, M(\dot{\boldsymbol{x}})\rangle \ ,$$

where the **mass matrix** M is assumed to be symmetric and positive-definite; that is, all its eigenvalues are positive.

We will now analyse the dynamics of small displacements from equilibrium following the following prescription:

1. we will standardise the kinetic energy by diagonalising and normalising the mass matrix; and

2. we will then diagonalise the potential energy and solve for the normal modes and characteristic frequencies of this system.

Both steps make use of the spectral theorem for symmetric transformations. To do the first step notice that relative to an orthonormal basis $\{\boldsymbol{e}_i\}$ for $\mathbb{V}$, $\boldsymbol{x} = \sum_i x_i \boldsymbol{e}_i$ and we can form a column vector

$$\mathsf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{pmatrix}$$

out of the components of $\boldsymbol{x}$. Relative to this basis, the mass matrix M and the Hessian H have matrices M and H, respectively. By assumption both are symmetric, and M is in addition positive-definite. The kinetic and potential energies become

$$T = \tfrac{1}{2}\dot{\mathsf{x}}^t\,\mathsf{M}\,\dot{\mathsf{x}} \qquad \text{and} \qquad V = \tfrac{1}{2}\mathsf{x}^t\,\mathsf{H}\,\mathsf{x} \ .$$

Because M is symmetric, there is an orthogonal matrix O_1 such that $\mathsf{M}' = \mathsf{O}_1^t\,\mathsf{M}\,\mathsf{O}_1$ is diagonal with positive entries. Let D_1 be the diagonal matrix whose entries are the (positive) square roots of the diagonal entries of M'. In other words, $\mathsf{M}' = \mathsf{D}_1^2$. We can therefore write

$$\mathsf{M} = \mathsf{O}_1\,\mathsf{D}_1^2\,\mathsf{O}_1^t = (\mathsf{O}_1\,\mathsf{D}_1)\,(\mathsf{O}_1\,\mathsf{D}_1)^t \ ,$$

where we have used that $\mathsf{D}_1^t = \mathsf{D}_1$ since it is diagonal. Introduce then the following variables

$$\mathsf{y} = (\mathsf{O}_1\,\mathsf{D}_1)^t\,\mathsf{x} = \mathsf{D}_1\,\mathsf{O}_1^t\,\mathsf{x} \ .$$

We can invert this change of variables as follows:

$$\mathsf{x} = \mathsf{O}_1\,\mathsf{D}_1^{-1}\,\mathsf{y} \ ,$$

where we have used that O_1 is orthogonal, so that $\mathsf{O}_1^t = \mathsf{O}_1^{-1}$. This change of variables accomplishes the first step outlined above, since in terms of y, the kinetic energy becomes simply

$$T = \tfrac{1}{2}\dot{\mathsf{y}}^t\,\dot{\mathsf{y}} = \tfrac{1}{2}\|\dot{\mathsf{y}}\|^2 \ .$$

Similarly, the potential energy has become

$$V = \tfrac{1}{2}\mathsf{y}^t\,\mathsf{K}\,\mathsf{y} \ ,$$

where the matrix K is defined by

$$\mathsf{K} = \mathsf{D}_1^{-1}\,\mathsf{O}_1^t\,\mathsf{H}\,\mathsf{O}_1\,\mathsf{D}_1^{-1} \ ,$$

which is clearly symmetric since H and D_1 are. Therefore we can find a second orthogonal matrix O_2 such that $\mathsf{O}_2^t\,\mathsf{K}\,\mathsf{O}_2$ is diagonal; call this matrix D. Let us define a new set of variables

$$\mathsf{z} = \mathsf{O}_2^t\,\mathsf{y} \ ,$$

relative to which the kinetic energy remains simple

$$T = \tfrac{1}{2}\|\mathsf{O}_2\,\mathsf{z}\|^2 = \tfrac{1}{2}\|\mathsf{z}\|^2 \ ,$$

since orthogonal matrices preserve norms, and the potential energy diagonalises

$$V = \tfrac{1}{2}\mathsf{z}^t\,\mathsf{D}\,\mathsf{z} \ .$$

Because D is diagonal, the equations of motion of the z are decoupled:

$$\ddot{\mathsf{z}} = -\mathsf{D}\,\mathsf{z} \ ,$$

whence the z are the normal modes of the system. Let D have entries

$$\mathsf{D} = \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_N \end{pmatrix} \ ,$$

Then the equations of motion for the normal modes are

$$\ddot{z}_i = -\lambda_i\,z_i \ .$$

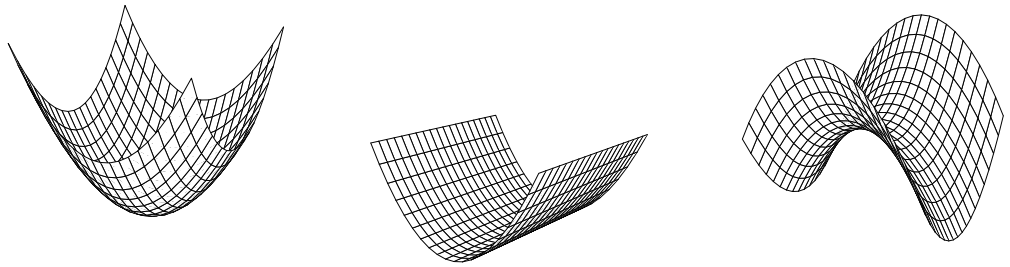
We can distinguish three types of solutions:

1. ($\lambda_i > 0$) The solution is oscillatory with characteristic frequency $\omega_i = \sqrt{\lambda_i}$:
$$z_i(t) = A_i \cos(\omega_i t + \varphi_i) \ .$$

2. ($\lambda_i = 0$) The solution is linear
$$z_i(t) = a_i + b_i t \ .$$
Such a normal mode is said to be a **zero mode**, since it has zero characteristic frequency.

3. ($\lambda_i < 0$) The solution is exponential
$$z_i(t) = A_i \exp\left(\sqrt{|\lambda_i|}t\right) + B_i \exp\left(-\sqrt{|\lambda_i|}t\right) \ .$$

If all eigenvalues λ_i are positive the equilibrium point is said to be **stable**, if they are all non-negative then it is **semi-stable**, whereas if there is a negative eigenvalue, then the equilibrium is **unstable**. The signs of the eigenvalues of the matrix D agree with the sign of the eigenvalues of the Hessian matrix of the potential at the equilibrium point. The different types of equilibria are illustrated in Figure 1.3, which shows the behaviour of the potential function around an equilibrium point in the simple case of a two-dimensional configuration space $\mathbb{V}$. The existence of zero modes is symptomatic of *flat directions* in the potential along which the system can evolve without spending any energy. This usually signals the existence of some continuous symmetry in the system. In the Figure we see that the semi-stable equilibrium point indeed has a flat direction along which the potential is constant. In other words, translation along the flat direction is a symmetry of the potential function.



(a) stable (b) semi-stable (c) unstable

Figure 1.3: Different types of equilibrium points.

Chapter 2

Complex Analysis

In this part of the course we will study some basic complex analysis. This is an extremely useful and beautiful part of mathematics and forms the basis of many techniques employed in many branches of mathematics and physics. We will extend the notions of derivatives and integrals, familiar from calculus, to the case of complex functions of a complex variable. In so doing we will come across analytic functions, which form the centerpiece of this part of the course. In fact, to a large extent complex analysis is the study of analytic functions. After a brief review of complex numbers as points in the complex plane, we will first discuss analyticity and give plenty of examples of analytic functions. We will then discuss complex integration, culminating with the generalised Cauchy Integral Formula, and some of its applications. We then go on to discuss the power series representations of analytic functions and the residue calculus, which will allow us to compute many real integrals and infinite sums very easily via complex integration.

2.1 Analytic functions

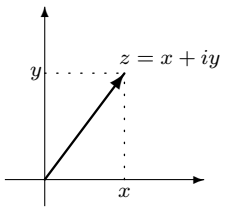
In this section we will study complex functions of a complex variable. We will see that differentiability of such a function is a non-trivial property, giving rise to the concept of an analytic function. We will then study many examples of analytic functions. In fact, the construction of analytic functions will form a basic *leitmotif* for this part of the course.

2.1.1 The complex plane

We already discussed complex numbers briefly in Section 1.3.5. The emphasis in that section was on the algebraic properties of complex numbers, and

although these properties are of course important here as well and will be used all the time, we are now also interested in more geometric properties of the complex numbers.

The set $\mathbb{C}$ of complex numbers is naturally identified with the plane $\mathbb{R}^2$. This is often called the **Argand plane**.

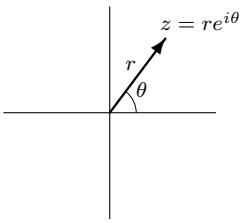


Given a complex number $z = x + i\,y$, its real and imaginary parts define an element (x, y) of $\mathbb{R}^2$, as shown in the figure. In fact this identification is one of real vector spaces, in the sense that adding complex numbers and multiplying them with real scalars mimic the similar operations one can do in $\mathbb{R}^2$. Indeed, if $\alpha \in \mathbb{R}$ is real, then to $\alpha\, z = (\alpha\, x) + i\,(\alpha\, y)$ there corresponds the pair $(\alpha\, x, \alpha\, y) = \alpha\,(x, y)$. Similarly, if $z_1 = x_1 + i\, y_1$ and $z_2 = x_2 + i\, y_2$ are complex numbers, then $z_1 + z_2 = (x_1 + x_2) + i\,(y_1 + y_2)$, whose associated pair is $(x_1 + x_2, y_1 + y_2) = (x_1, y_1) + (x_2, y_2)$. In fact, the identification is even one of euclidean spaces. Given a complex number $z = x + i\, y$, its modulus $|z|$, defined by $|z|^2 = zz^*$, is given by $\sqrt{x^2 + y^2}$ which is precisely the norm $\|(x, y)\|$ of the pair (x, y). Similarly, if $z_1 = x_1 + i\, y_1$ and $z_2 = x_2 + i\, y_2$, then $\mathrm{Re}(z_1^* z_2) = x_1 x_2 + y_1 y_2$ which is the dot product of the pairs (x_1, y_1) and (x_2, y_2). In particular, it follows from these remarks and the triangle inequality for the norm in $\mathbb{R}^2$, that complex numbers obey a version of the triangle inequality:

$$\boxed{|z_1 + z_2| \leq |z_1| + |z_2|\ .} \tag{2.1}$$

Polar form and the argument function

Points in the plane can also be represented using polar coordinates, and this representation in turn translates into a representation of the complex numbers.



Let (x, y) be a point in the plane. If we define $r = \sqrt{x^2 + y^2}$ and θ by $\theta = \arctan(y/x)$, then we can write $(x, y) = (r\,\cos\theta, r\,\sin\theta) = r\,(\cos\theta, \sin\theta)$. The complex number $z = x + i\, y$ can then be written as $z = r\,(\cos\theta + i\,\sin\theta)$. The real number r, as we have seen, is the modulus $|z|$ of z, and the complex number $\cos\theta + i\,\sin\theta$ has unit modulus. Comparing the Taylor series for the cosine and sine functions and the exponential functions we notice that $\cos\theta + i\sin\theta = e^{i\theta}$. The angle θ is called the **argument** of z and is written $\arg(z)$. Therefore we

have the following **polar form** for a complex number z:

$$\boxed{z = |z|\, e^{i \arg(z)} \ .} \tag{2.2}$$

Being an angle, the argument of a complex number is only defined up to the addition of integer multiples of 2π. In other words, it is a **multiple-valued function**. This ambiguity can be resolved by defining the **principal value** Arg of the arg function to take values in the interval $(-\pi, \pi]$; that is, for any complex number z, one has

$$\boxed{-\pi < \mathrm{Arg}(z) \leq \pi \ .} \tag{2.3}$$

Notice, however, that Arg is not a continuous function: it has a discontinuity along the negative real axis. Approaching a point on the negative real axis from the upper half-plane, the principal value of its argument approaches π, whereas if we approach it from the lower half-plane, the principal value of its argument approaches $-\pi$. Notice finally that whereas the modulus is a multiplicative function: $|zw| = |z||w|$, the argument is additive: $\arg(z_1\, z_2) = \arg(z_1) + \arg(z_2)$, provided that we understand the equation to hold up to integer multiples of 2π. Also notice that whereas the modulus is invariant under conjugation $|z^*| = |z|$, the argument changes sign $\arg(z^*) = -\arg(z)$, again up to integer multiples of 2π.

Some important subsets of the complex plane

We end this section with a brief discussion of some very important subsets of the complex plane. Let z_0 be any complex number, and consider all those complex numbers z which are a distance at most ε away from z_0. These points form a disk of radius ε centred at z_0. More precisely, let us define the **open ε-disk around** z_0 to be the subset $D_\varepsilon(z_0)$ of the complex plane defined by

$$D_\varepsilon(z_0) = \{z \in \mathbb{C} \mid |z - z_0| < \varepsilon\} \ . \tag{2.4}$$

Similarly one defines the **closed ε-disk** around z_0 to be the subset

$$\bar{D}_\varepsilon(z_0) = \{z \in \mathbb{C} \mid |z - z_0| \leq \varepsilon\} \ , \tag{2.5}$$

which consists of the open ε-disk and the circle $|z - z_0| = \varepsilon$ which forms its boundary. More generally a subset $U \subset \mathbb{C}$ of the complex plane is said to be **open** if given any $z \in U$, there exists some positive real number $\varepsilon > 0$ (which can depend on z) such that the open ε-disk around z also belongs to U. A set C is said to be **closed** if its complement $C^c = \{z \in \mathbb{C} \mid z \notin C\}$—that is, all

those points not in C—is open. One should keep in mind that generic subsets of the complex plane are neither closed nor open. By a **neighbourhood** of a point z_0 in the complex plane, we will mean any open set containing z_0. For example, any open ε-disk around z_0 is a neighbourhood of z_0.

Let us see that the open and closed ε-disks are indeed open and closed, respectively. Let $z \in D_\varepsilon(z_0)$. This means that $|z - z_0| = \delta < \varepsilon$. Consider the disk $D_{\varepsilon-\delta}(z)$. We claim that this disk is contained in $D_\varepsilon(z_0)$. Indeed, if $|w - z| < \varepsilon - \delta$ then,

$$\begin{aligned} |w - z_0| &= |(w - z) + (z - z_0)| && \text{(adding and subtracting } z\text{)} \\ &\leq |w - z| + |z - z_0| && \text{(by the triangle inequality (2.1))} \\ &< \varepsilon - \delta + \delta \\ &= \varepsilon \ . \end{aligned}$$

Therefore the disk $D_\varepsilon(z_0)$ is indeed open. Consider now the subset $\bar{D}_\varepsilon(z_0)$. Its complement is the subset of points z in the complex plane such that $|z - z_0| > \varepsilon$. We will show that it is an open set. Let z be such that $|z - z_0| = \eta > \varepsilon$. Then consider the open disk $D_{\eta-\varepsilon}(z)$, and let w be a point in it. Then

$$\begin{aligned} |z - z_0| &= |(z - w) + (w - z_0)| && \text{(adding and subtracting } w\text{)} \\ &\leq |z - w| + |w - z_0| \ . && \text{(by the triangle inequality (2.1))} \end{aligned}$$

We can rewrite this as

$$\begin{aligned} |w - z_0| &\geq |z - z_0| - |z - w| \\ &> \eta - (\eta - \varepsilon) && \text{(since } |z - w| = |w - z| < \eta - \varepsilon\text{)} \\ &= \varepsilon \ . \end{aligned}$$

Therefore the complement of $\bar{D}_\varepsilon(z_0)$ is open, whence $\bar{D}_\varepsilon(z_0)$ is closed.

We should remark that the closed disk $\bar{D}_\varepsilon(z_0)$ is not open, since any open disk around a point z at the boundary of $\bar{D}_\varepsilon(z_0)$—that is, for which $|z - z_0| = \varepsilon$—contains points which are not included in $D_\varepsilon(z_0)$.

Notice that it follows from this definition that every open set is made out of the union of (a possibly uncountable number of) open disks.

2.1.2 Complex-valued functions

In this section we will discuss complex-valued functions.

We start with a rather trivial case of a complex-valued function. Suppose that f is a complex-valued function of a real variable. That means that if x is a real number, $f(x)$ is a complex number, which can be decomposed into its real and imaginary parts: $f(x) = u(x) + i\, v(x)$, where u and v are real-valued functions of a real variable; that is, the objects you are familiar with from calculus. We say that f is **continuous** at x_0 if u and v are continuous at x_0.

Let us recall the definition of continuity. Let f be a real-valued function of a real variable. We say that f is continuous at x_0, if for every $\varepsilon > 0$, there is a $\delta > 0$ such that $|f(x) - f(x_0)| < \varepsilon$ whenever $|x - x_0| < \delta$. A function is said to be continuous if it is continuous at all points where it is defined.

Now consider a complex-valued function f of a complex variable z. We say that f is **continuous** at z_0 if given any $\varepsilon > 0$, there exists a $\delta > 0$ such that $|f(z) - f(z_0)| < \varepsilon$ whenever $|z - z_0| < \delta$. Heuristically, another way of saying that f is continuous at z_0 is that $f(z)$ tends to $f(z_0)$ as z approaches z_0. This is equivalent to the continuity of the real and imaginary parts of f thought of as real-valued functions on the complex plane. Explicitly, if we write $f = u + i\,v$ and $z = x + i\,y$, $u(x, y)$ and $v(x, y)$ are real-valued functions on the complex plane. Then the continuity of f at $z_0 = x_0 + i\,y_0$ is equivalent to the continuity of u and v at the point (x_0, y_0).

"Graphing" complex-valued functions

Complex-valued functions of a complex variable are harder to visualise than their real analogues. To visualise a real function $f : \mathbb{R} \to \mathbb{R}$, one simply graphs the function: its graph being the curve $y = f(x)$ in the (x, y)-plane. A complex-valued function of a complex variable $f : \mathbb{C} \to \mathbb{C}$ maps complex numbers to complex numbers, or equivalently points in the (x, y)-plane to points in the (u, v) plane. Hence its graph defines a surface $u = u(x, y)$ and $v = v(x, y)$ in the four-dimensional space with coordinates (x, y, u, v), which is not so easy to visualise. Instead one resorts to investigating what the function does to regions in the complex plane. Traditionally one considers two planes: the z-plane whose points have coordinates (x, y) corresponding to the real and imaginary parts of $z = x + i\,y$, and the w-plane whose points have coordinates (u, v) corresponding to $w = u + i\,v$. Any complex-valued function f of the complex variable z maps points in the z-plane to points in the w-plane via $w = f(z)$. A lot can be learned from a complex function by analysing the image in the w-plane of certain sets in the z-plane. We will have plenty of opportunities to use this throughout the course of these lectures.

With the picture of the z- and w-planes in mind, one can restate the continuity of a function very simply in terms of open sets. In fact, this was the historical reason why the notion of open sets was introduced in mathematics. As we saw, a complex-valued function f of a complex variable z defines a mapping from the complex z-plane to the complex w-plane. The function f is continuous at z_0 if for every neighbourhood U of $w_0 = f(z_0)$ in the w-plane, the set

$$f^{-1}(U) = \{z \mid f(z) \in U\}$$

is open in the z-plane. Checking that both definitions of continuity agree is left as an exercise.

2.1.3 Differentiability and analyticity

Let us now discuss differentiation of complex-valued functions. Again, if $f = u + i\,v$ is a complex-valued function of a *real* variable x, then the derivative

of f at the point x_0 is defined by

$$f'(x_0) = u'(x_0) + i\,v'(x_0)\ ,$$

where u' and v' are the derivatives of u and v respectively. In other words, we extend the operation of differentiation complex-linearly. There is nothing novel here.

Differentiability and the Cauchy–Riemann equations

The situation is drastically different when we consider a complex-valued function $f = u+i\,v$ of a complex variable $z = x+i\,y$. As is calculus, let us attempt to define its derivative by

$$\boxed{f'(z_0) \equiv \lim_{\Delta z\to 0} \frac{f(z_0+\Delta z) - f(z_0)}{\Delta z}\ .} \tag{2.6}$$

The first thing that we notice is that Δz, being a complex number, can approach zero in more than one way. If we write $\Delta z = \Delta x + i\,\Delta y$, then we can approach zero along the real axis $\Delta y = 0$ or along the imaginary axis $\Delta x = 0$, or indeed along any direction. For the derivative to exist, the answer should not depend on how Δz tends to 0. Let us see what this entails. Let us write $f = u + i\,v$ and $z_0 = x_0 + i\,y_0$, so that

$$\begin{aligned} f(z_0) &= u(x_0,y_0) + i\,v(x_0,y_0) \\ f(z_0+\Delta z) &= u(x_0+\Delta x, y_0+\Delta y) + i\,v(x_0+\Delta x, y_0+\Delta y)\ . \end{aligned}$$

Then

$$f'(z_0) = \lim_{\substack{\Delta x\to 0\\ \Delta y\to 0}} \frac{\Delta u(x_0,y_0) + i\,\Delta v(x_0,y_0)}{\Delta x + i\Delta y}\ ,$$

where

$$\begin{aligned} \Delta u(x_0,y_0) &= u(x_0+\Delta x, y_0+\Delta y) - u(x_0,y_0) \\ \Delta v(x_0,y_0) &= v(x_0+\Delta x, y_0+\Delta y) - v(x_0,y_0)\ . \end{aligned}$$

Let us first take the limit $\Delta z \to 0$ by first taking $\Delta y \to 0$ and then $\Delta x \to 0$; in other words, we let $\Delta z \to 0$ along the real axis. Then

$$\begin{aligned} f'(z_0) &= \lim_{\Delta x\to 0} \lim_{\Delta y\to 0} \frac{\Delta u(x_0,y_0) + i\,\Delta v(x_0,y_0)}{\Delta x + i\Delta y} \\ &= \lim_{\Delta x\to 0} \frac{\Delta u(x_0,y_0) + i\,\Delta v(x_0,y_0)}{\Delta x} \\ &= \left.\frac{\partial u}{\partial x}\right|_{(x_0,y_0)} + i\ \left.\frac{\partial v}{\partial x}\right|_{(x_0,y_0)}\ . \end{aligned}$$

Now let us take the limit $\Delta z \to 0$ by first taking $\Delta x \to 0$ and then $\Delta y \to 0$; in other words, we let $\Delta z \to 0$ along the imaginary axis. Then

$$\begin{aligned} f'(z_0) &= \lim_{\Delta y \to 0} \lim_{\Delta x \to 0} \frac{\Delta u(x_0, y_0) + i\,\Delta v(x_0, y_0)}{\Delta x + i \Delta y} \\ &= \lim_{\Delta y \to 0} \frac{\Delta u(x_0, y_0) + i\,\Delta v(x_0, y_0)}{i\,\Delta y} \\ &= -i \left.\frac{\partial u}{\partial y}\right|_{(x_0,y_0)} + \left.\frac{\partial v}{\partial y}\right|_{(x_0,y_0)} . \end{aligned}$$

These two expressions for $f'(z_0)$ agree if and only if the following equations are satisfied at (x_0, y_0):

$$\boxed{\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y} \qquad \text{and} \qquad \frac{\partial v}{\partial x} = -\frac{\partial u}{\partial y} \; .} \tag{2.7}$$

These equations are called the **Cauchy–Riemann equations**.

We say that the function f is **differentiable** at z_0 if $f'(z_0)$ is well-defined at z_0. For a differentiable function f we have just seen that

$$f'(z) = \frac{\partial u}{\partial x} + i \frac{\partial v}{\partial x} = \frac{\partial v}{\partial y} - i \frac{\partial u}{\partial y} \; .$$

We have just shown that a necessary condition for f to be differentiable at z_0 is that its real and imaginary parts obey the Cauchy–Riemann equations at (x_0, y_0). Conversely, it can be shown that this condition is also sufficient provided that the the partial derivatives of u and v are continuous.

We say that the function f is **analytic** in a neighbourhood U of z_0 if it is differentiable everywhere in U. We say that a function is **entire** if it is analytic in the whole complex plane. Often the terms **regular** and **holomorphic** are used as synonyms for analytic.

For example, the function $f(z) = z$ is entire. We can check this either by verifying the Cauchy–Riemann equations or else simply by noticing that

$$\begin{aligned} f'(z_0) &= \lim_{\Delta z \to 0} \frac{f(z_0 + \Delta z) - f(z_0)}{\Delta z} \\ &= \lim_{\Delta z \to 0} \frac{z_0 + \Delta z - z_0}{\Delta z} \\ &= \lim_{\Delta z \to 0} \frac{\Delta z}{\Delta z} \\ &= \lim_{\Delta z \to 0} 1 \\ &= 1 \; ; \end{aligned}$$

whence it is well-defined for all z_0.

On the other hand, the function $f(z) = z^*$ is not differentiable anywhere:

$$\begin{aligned} f'(z_0) &= \lim_{\Delta z \to 0} \frac{f(z_0 + \Delta z) - f(z_0)}{\Delta z} \\ &= \lim_{\Delta z \to 0} \frac{z_0^* + (\Delta z)^* - z_0^*}{\Delta z} \\ &= \lim_{\Delta z \to 0} \frac{(\Delta z)^*}{\Delta z} \ ; \end{aligned}$$

whence if we let Δz tend to zero along real values, we would find that $f'(z_0) = 1$, whereas if we would let Δz tend to zero along imaginary values we would find that $f'(z_0) = -1$. We could have reached the same conclusion via the Cauchy–Riemann equations with $u(x, y) = x$ and $v(x, y) = -y$, which violates the first of the Cauchy–Riemann equations.

It is important to realise that analyticity, unlike differentiability, is not a property of a function at a point, but on an open set of points. The reason for this is to able to eliminate from the class of interesting functions, functions which may be differentiable at a point but nowhere else. Whereas this is a rarity in calculus[1], it is a very common occurrence for complex-valued functions of a complex variables. For example, consider the function $f(z) = |z|^2$. This function has $u(x, y) = x^2 + y^2$ and $v(x, y) = 0$. Therefore the Cauchy–Riemann equations are only satisfied at the origin in the complex plane:

$$\frac{\partial u}{\partial x} = 2x = \frac{\partial v}{\partial y} = 0 \qquad \text{and} \qquad \frac{\partial v}{\partial x} = 0 = -\frac{\partial u}{\partial y} = -2y \ .$$

Relation with harmonic functions

Analytic functions are intimately related to harmonic functions. We say that a real-valued function $h(x, y)$ on the plane is **harmonic** if it obeys **Laplace's equation**:

$$\boxed{\frac{\partial^2 h}{\partial x^2} + \frac{\partial^2 h}{\partial y^2} = 0 \ .} \tag{2.8}$$

In fact, as we now show, the real and imaginary parts of an analytic function are harmonic. Let $f = u + i\,v$ be analytic in some open set of the complex

[1]Try to come up with a real-valued function of a real variable which is differentiable only at the origin, for example.

plane. Then,

$$\begin{aligned}\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2} &= \frac{\partial}{\partial x}\frac{\partial u}{\partial x}+\frac{\partial}{\partial y}\frac{\partial u}{\partial y}\\ &= \frac{\partial}{\partial x}\frac{\partial v}{\partial y}-\frac{\partial}{\partial y}\frac{\partial v}{\partial x} \qquad \text{(using Cauchy–Riemann)}\\ &= \frac{\partial^2 v}{\partial x\,\partial y}-\frac{\partial^2 v}{\partial y\,\partial x}\\ &= 0\ .\end{aligned}$$

A similar calculation shows that v is also harmonic. This result is important in applications because it shows that one can obtain solutions of a second order partial differential equation by solving a system of first order partial differential equations. It is particularly important in this case because we will be able to obtain solutions of the Cauchy–Riemann equations without really solving these equations.

Given a harmonic function u we say that another harmonic function v is its **harmonic conjugate** if the complex-valued function $f = u+i\,v$ is analytic. For example, consider the function $u(x,y) = xy-x+y$. It is clearly harmonic since

$$\frac{\partial u}{\partial x} = y-1 \qquad \text{and} \qquad \frac{\partial u}{\partial y} = x+1\ ,$$

whence

$$\frac{\partial^2 u}{\partial x^2}=\frac{\partial^2 u}{\partial y^2}=0\ .$$

By a harmonic conjugate we mean any function $v(x,y)$ which together with $u(x,y)$ satisfies the Cauchy–Riemann equations:

$$\frac{\partial v}{\partial x}=-\frac{\partial u}{\partial y}=-x-1 \qquad \text{and} \qquad \frac{\partial v}{\partial y}=\frac{\partial u}{\partial x}=y-1\ .$$

We can integrate the first of the above equations, to obtain

$$v(x,y) = -\tfrac{1}{2}x^2-x+\psi(y)\ ,$$

for ψ an arbitrary function of y which is to be determined from the second of the Cauchy–Riemann equations. Doing this one finds

$$\psi'(y) = y-1\ ,$$

which is solved by $\psi(y)=\frac{1}{2}y^2-y+c$, where c is any constant. Therefore, the function $f = u+i\,v$ becomes

$$f(x,y) = xy-x+y+i\,(-\tfrac{1}{2}x^2+\tfrac{1}{2}y^2-x-y+c)\ .$$

We can try to write this down in terms of z and z^* by making the substitutions $x = \frac{1}{2}(z + z^*)$ and $y = -i\,\frac{1}{2}(z - z^*)$. After a little bit of algebra, we find

$$f(z) = -iz^2 - (1+i)\,z + i\,c\ .$$

Notice that all the z^* dependence has dropped out. We will see below that this is a sign of analyticity.

2.1.4 Polynomials and rational functions

We now start to build up some examples of analytic functions. We have already seen that the function $f(z) = z$ is entire. In this section we will generalise this to show that so is any polynomial $P(z)$. We will also see that ratios of polynomials are also analytic everywhere but on a finite set of points in the complex plane where the denominator vanishes.

There are many ways to do this, but one illuminating way is to show that complex linear combinations of analytic functions are analytic and that products of analytic functions are analytic functions. Let $f(z)$ be an analytic function on some open subset $U \subset \mathbb{C}$, and let α be a complex number. Then it is easy to see that the function $\alpha\, f(z)$ is also analytic on U. Indeed, from the definition (2.6) of the derivative, we see that

$$(\alpha\, f)'(z_0) = \alpha\, f'(z_0)\ , \tag{2.9}$$

which exists whenever $f'(z_0)$ exists.

Let $f(z)$ and $g(z)$ be analytic functions on the same open subset $U \subset \mathbb{C}$. Then the functions $f(z) + g(z)$ and $f(z)g(z)$ are also analytic. Again from the definition (2.6) of the derivative,

$$(f+g)'(z_0) = f'(z_0) + g'(z_0) \tag{2.10}$$

$$(f\,g)'(z_0) = f'(z_0)\,g(z_0) + f(z_0)\,g'(z_0)\ , \tag{2.11}$$

which exist whenever $f'(z_0)$ and $g'(z_0)$ exist.

The only tricky bit in the above result is that we have to make sure that f and g are analytic in the same open set U. Normally it happens that f and g are analytic in different open sets, say, U_1 and U_2 respectively. Then the sum $f(z) + g(z)$ and product $f(z)\,g(z)$ are analytic in the intersection $U = U_1 \cap U_2$, which is also open. This is easy to see. Let us assume that U is not empty, otherwise the statement is trivially satisfied. Let $z \in U$. This means that $z \in U_1$ and $z \in U_2$. Because each U_i is open there are positive real numbers ε_i such that $D_{\varepsilon_i}(z)$ lies inside U_i. Let $\varepsilon = \min(\varepsilon_1, \varepsilon_2)$ be the smallest of the ε_i. Then $D_\varepsilon(z) \subseteq D_{\varepsilon_i}(z) \subset U_i$ for $i = 1, 2$. Therefore $D_\varepsilon(z) \subset U$, and U is open.

It is important to realise that only finite intersections of open sets will again be open in general. Consider, for example, the open disks $D_{1/n}(0)$ of radius $1/n$ about the origin, for $n = 1, 2, 3, \ldots$. Their intersection consists of the points z with $|z| < 1/n$ for all $n = 1, 2, 3, \ldots$. Clearly, if $z \neq 0$ then there will be some positive integer n for which

$|z| > 1/n$. Therefore the only point in the intersection of all the $D_{1/n}(0)$ is the origin itself. But this set is clearly not open, since it does not contain any open disk with nonzero radius. More generally, sets consisting of a finite number of points are never open; although they are closed.

Therefore we see that (finite) sums and products of analytic functions are analytic with the same domain of analyticity. In particular, sums and products of entire functions are again entire. As a result, from the fact that the function $f(z) = z$ is entire, we see that any polynomial $P(z) = \sum_{n=0}^{N} a_n z^n$ of finite degree N is also an entire function. Indeed, its derivative is given by

$$P'(z_0) = \sum_{n=1}^{N} n\, a_n\, z_0^{n-1} \ ,$$

as follows from the above formulae for the derivatives of sums and products.

We will see later on in the course that to some extent we will be able to describe all analytic functions (at least locally) in terms of polynomials, provided that we allow the polynomials to have arbitrarily high degree; in other words, in terms of power series.

There are two more constructions which start from analytic functions and yield an analytic function: quotients and composition. Let $f(z)$ and $g(z)$ be analytic functions on some open subset $U \subset \mathbb{C}$. Then the quotient $f(z)/g(z)$ is continuous away from the zeros of $g(z)$, which can be shown (see below) to be an open set. If $g(z_0) \neq 0$, then from the definition of the derivative (2.6), it follows that

$$\left(\frac{f}{g}\right)'(z_0) = \frac{f'(z_0)\, g(z_0) - f(z_0)\, g'(z_0)}{g(z_0)^2} \ .$$

To see that the subset of points z for which $g(z) \neq 0$ is open, we need only realise that this set is the inverse image $g^{-1}(\{0\}^c)$ under g of the complement of 0. The result then follows because the complement of 0 is open and g is continuous, so that $g^{-1}(\text{open})$ is open.

By a **rational function** we mean the ratio of two polynomials. Let $P(z)$ and $Q(z)$ be two polynomials. Then the rational function

$$R(z) = \frac{P(z)}{Q(z)}$$

is analytic away from the zeros of $Q(z)$.

We have been tacitly assuming that every (non-constant) polynomial $Q(z)$ has zeros. This result is known as the *Fundamental Theorem of Algebra* and although it is of course intuitive and in agreement with our daily experience with polynomials, its proof is surprisingly difficult. There are three standard proofs: one is purely algebraic, but it is long and arduous, one uses algebraic topology and the other uses complex analysis. We will in fact see this third proof later on in Section 2.2.6.

Finally let $g(z)$ be analytic in an open subset $U \subset \mathbb{C}$ and let $f(z)$ be analytic in some open subset containing $g(U)$, the image of U under g. Then the composition $f \circ g$ defined by $(f \circ g)(z) = f(g(z))$ is also analytic in U. In fact, its derivative can be computed using the chain rule,

$$(f \circ g)'(z_0) = f'(g(z_0))\, g'(z_0) \ . \tag{2.12}$$

You may wonder whether $g(U)$ is an open set, for U open and g analytic. This is indeed true: it is called the open mapping property of analytic functions. We may see this later on in the course.

It is clear that if f and g are rational functions so will be its composition $f \circ g$, so one only ever constructs new functions this way when one of the functions being composed is not rational. We will see plenty of examples of this as the lectures progress.

Another look at the Cauchy–Riemann equations

Finally let us mention an a different way to understand the Cauchy–Riemann equations, at least for the case of rational functions. Notice that the above polynomials and rational functions share the property that they do not depend on z^* but only on z. Suppose that one is given a rational function where the dependence on x and y has been made explicit. For example,

$$f(x,y) = \frac{x - 1 - i\,y}{(x-1)^2 + y^2} \ .$$

In order to see whether f is analytic one would have to apply the Cauchy–Riemann equations, which can get rather messy when the rational function is complicated. It turns out that it is not necessary to do this. Instead one can try to re-express the function in terms of z and z^* by the substitutions

$$x = \frac{z + z^*}{2} \qquad \text{and} \qquad y = \frac{z - z^*}{2i} \ .$$

Then, the rational function $f(x,y)$ is analytic if and only if the z^* dependence cancels. In the above example, one can see that this is indeed the case. Indeed, rewriting $f(x,y)$ in terms of z and z^* we see that

$$f(x,y) = \frac{z^* - 1}{z z^* - z - z^* + 1} = \frac{1}{z - 1} \ ,$$

whence the z^* dependence has dropped out. We therefore expect that the Cauchy–Riemann equations will be satisfied. Indeed, one has that

$$u(x,y) = \frac{x-1}{(x-1)^2 + y^2} \qquad \text{and} \qquad v(x,y) = \frac{-y}{(x-1)^2 + y^2} \ ,$$

and after some algebra,

$$\frac{\partial u}{\partial x} = \frac{-\left(x-1\right)^2 + y^2}{\left(\left(x-1\right)^2 + y^2\right)^2} = \frac{\partial v}{\partial y}$$
$$\frac{\partial u}{\partial y} = \frac{-2\left(x-1\right)\, y}{\left(\left(x-1\right)^2 + y^2\right)^2} = -\frac{\partial v}{\partial x} \ .$$

The reason this works is the following. Let us think formally of z and z^* as independent variables for the plane, like x and y. Then we have that

$$\frac{\partial f}{\partial z^*} = \frac{\partial f}{\partial (x - i\, y)} = \frac{\partial f}{\partial x} + i\, \frac{\partial f}{\partial y} \ .$$

Let us break up f into its real and imaginary parts: $f(x,y) = u(x,y) + i\, v(x,y)$. Then,

$$\begin{aligned} \frac{\partial f}{\partial z^*} &= \frac{\partial u}{\partial x} + i\, \frac{\partial v}{\partial x} + i\, \frac{\partial u}{\partial y} - \frac{\partial v}{\partial y} \\ &= \left(\frac{\partial u}{\partial x} - \frac{\partial v}{\partial y}\right) + i\, \left(\frac{\partial v}{\partial x} + \frac{\partial u}{\partial y}\right) \ . \end{aligned}$$

Therefore we see that the Cauchy–Riemann equations are equivalent to the condition

$$\boxed{\frac{\partial f}{\partial z^*} = 0 \ .} \tag{2.13}$$

2.1.5 The complex exponential and related functions

There are many other analytic functions besides the rational functions. Some of them are related to the exponential function.

Let $z = x + i\, y$ be a complex number and define the **complex exponential** $\exp(z)$ (also written e^z) to be the function

$$\exp(z) = \exp(x + i\, y) \equiv e^x \left(\cos y + i\, \sin y\right) \ .$$

We will first check that this function is entire. Decomposing it into real and imaginary parts, we see that

$$u(x,y) = e^x\, \cos y \qquad \text{and} \qquad v(x,y) = e^x\, \sin y \ .$$

It is easy to check that the Cauchy–Riemann equations (2.7) are satisfied everywhere on the complex plane:

$$\frac{\partial u}{\partial x} = e^x\, \cos y = \frac{\partial v}{\partial y} \qquad \text{and} \qquad \frac{\partial v}{\partial x} = e^x\, \sin y = -\frac{\partial u}{\partial y} \ .$$

Therefore the function is entire and its derivative is given by

$$\begin{aligned}\exp'(z) &= \frac{\partial u}{\partial x} + i\,\frac{\partial v}{\partial x} \\ &= e^x\,\cos y + i\,e^x\,\sin y \\ &= \exp(z)\ .\end{aligned}$$

The exponential function obeys the following addition property

$$\boxed{\exp(z_1 + z_2) = \exp(z_1)\,\exp(z_2)\ ,} \tag{2.14}$$

which has as a consequence the celebrated **De Moivre's Formula**:

$$(\cos\theta + i\,\sin\theta)^n = \cos(n\theta) + i\,\sin(n\theta)\ ,$$

obtained simply by noticing that $\exp(i\,n\theta) = \exp(i\,\theta)^n$.

The exponential is also a periodic function, with period $2\pi\,i$. In fact from the periodicity of trigonometric functions, we see that $\exp(2\pi\,i) = 1$ and hence, using the addition property (2.14), we find

$$\boxed{\exp(z + 2\pi\,i) = \exp(z)\ .} \tag{2.15}$$

This means that the exponential is not one-to-one, in sharp contrast with the real exponential function. It follows from the definition of the exponential function that

$$\exp(z_1) = \exp(z_2) \quad \text{if and only if} \quad z_1 = z_2 + 2\pi\,i\,k \quad \text{for some integer } k.$$

We can divide up the complex plane into horizontal strips of height 2π in such a way that in each strip the exponential function is one-to-one. To see this define the following subsets of the complex plane

$$\mathcal{S}_k \equiv \{x + i\,y \in \mathbb{C} \mid (2k-1)\pi < y \leq (2k+1)\pi\}\ ,$$

for $k = 0, \pm1, \pm2, \ldots$, as shown in Figure 2.1.

Then it follows that if z_1 and z_2 belong to the same set $\mathcal{S}_k$, then $\exp(z_1) = \exp(z_2)$ implies that $z_1 = z_2$. Each of the sets $\mathcal{S}_k$ is known as a **fundamental region** for the exponential function. The basic property satisfied by a fundamental region of a periodic function is that if one knows the behaviour of the function on the fundamental region, one can use the periodicity to find out the behaviour of the function everywhere, and that it is the smallest region with that property. The periodicity of the complex exponential will have as a consequence that the complex logarithm will not be single-valued.

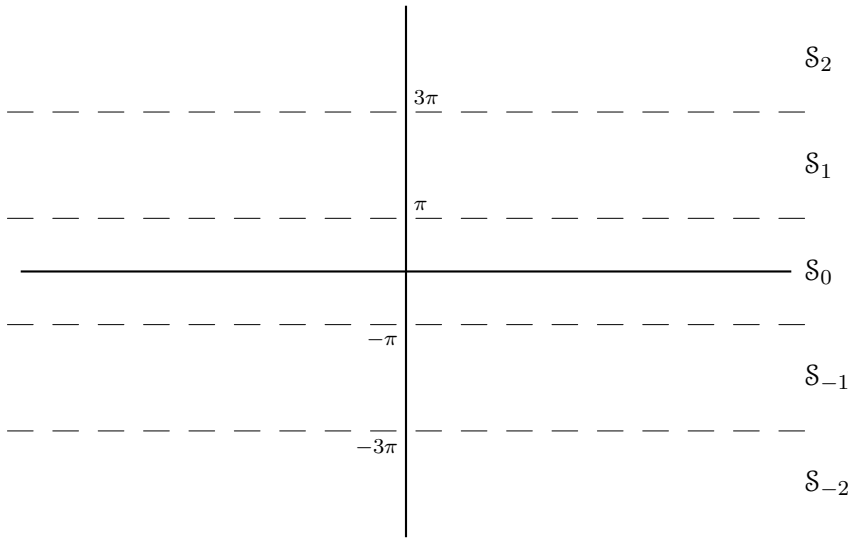


Figure 2.1: Fundamental regions of the complex exponential function.

Complex trigonometric functions

We can also define complex trigonometric functions starting from the complex exponential. Let $z = x + i\,y$ be a complex number. Then we define the complex sine and cosine functions as

$$\sin(z) \equiv \frac{e^{iz} - e^{-iz}}{2i} \qquad \text{and} \qquad \cos(z) \equiv \frac{e^{iz} + e^{-iz}}{2} \ .$$

Being linear combinations of the entire functions $\exp(\pm iz)$, they themselves are entire. Their derivatives are

$$\sin'(z) = \cos(z) \qquad \text{and} \qquad \cos'(z) = -\sin(z) \ .$$

The complex trigonometric functions obey many of the same properties of the real sine and cosine functions, with which they agree when z is real. For example,

$$\cos(z)^2 + \sin(z)^2 = 1 \ ,$$

and they are periodic with period 2π. However, there is one important difference between the real and complex trigonometric functions: whereas the real sine and cosine functions are bounded, their complex counterparts are not. To see this let us break them up into real and imaginary parts:

$$\begin{aligned} \sin(x + i\,y) &= \sin x\, \cosh y + i\, \cos x\, \sinh y \\ \cos(x + i\,y) &= \cos x\, \cosh y - i\, \sin x\, \sinh y \ . \end{aligned}$$

We see that the appearance of the hyperbolic functions means that the complex sine and cosine functions are not bounded.

Finally, let us define the complex hyperbolic functions. If $z = x + i\,y$, then let

$$\sinh(z) \equiv \frac{e^z - e^{-z}}{2} \qquad \text{and} \qquad \cosh(z) \equiv \frac{e^z + e^{-z}}{2} \ .$$

In contrast with the real hyperbolic functions, they are not independent from the trigonometric functions. Indeed, we see that

$$\sinh(iz) = i\sin(z) \qquad \text{and} \qquad \cosh(iz) = \cos(z) \ . \tag{2.16}$$

Notice that one can also define other complex trigonometric functions: $\tan(z)$, $\cot(z)$, $\sec(z)$ and $\csc(z)$ in the usual way, as well as their hyperbolic counterparts. These functions obey many other properties, but we will not review them here. Instead we urge you to play with these functions until you are familiar with them.

2.1.6 The complex logarithm

This section introduces the logarithm of a complex number. We will see that in contrast with the real logarithm function which is only defined for positive real numbers, the complex logarithm is defined for all nonzero complex numbers, but at a price: the function is not single-valued. This has to do with the periodicity (2.15) of the complex exponential or, equivalently, with the multiple-valuedness of the argument $\arg(z)$.

In this course we will use the notation 'log' for the natural logarithm, not for the logarithm base 10. Some people also use the notation 'ln' for the natural logarithm, in order to distinguish it from the logarithm base 10; but we will not be forced to do this since we will only be concerned with the natural logarithm.

By analogy with the real natural logarithm, we define the **complex logarithm** as an inverse to the complex exponential function. In other words, we say that a logarithm of a nonzero complex number z, is any complex number w such that $\exp(w) = z$. In other words, we define the function $\log(z)$ by

$$\boxed{w = \log(z) \quad \text{if} \quad \exp(w) = z \ .} \tag{2.17}$$

From the periodicity (2.15) of the exponential function it follows that if $w = \log(z)$ so is $w + 2\pi\,i\,k$ for any integer k. Therefore we see that $\log(z)$ is a multiple-valued function. We met a multiple-valued function before: the

argument function $\arg(z)$. Clearly if $\theta = \arg(z)$ then so is $\theta + 2\pi\, k$ for any integer k. This is no accident: the imaginary part of the $\log(z)$ function is $\arg(z)$. To see this, let us write z in polar form (2.2) $z = |z|\, \exp(i\, \arg(z))$ and $w = \log(z) = u + i\, v$. By the above definition and using the addition property (2.14), we have

$$\exp(u + i\, v) = e^u\, e^{i\, v} = |z|\, e^{i\, \arg(z)} \ ,$$

whence comparing polar forms we see that

$$e^u = |z| \qquad \text{and} \qquad e^{i\, v} = e^{i\, \arg(z)} \ .$$

Since u is a real number and $|z|$ is a positive real number, we can solve the first equation for u uniquely using the *real* logarithmic function, which in order to distinguish it from the complex function $\log(z)$ we will write as Log:

$$u = \mathrm{Log}\, |z| \ .$$

Similarly, we see that $v = \arg(z)$ solves the second equation. So does $v + 2\pi\, k$ for any integer k, but this is already taken into account by the multiple-valuedness of the $\arg(z)$ function. Therefore we can write

$$\boxed{\log(z) = \mathrm{Log}\, |z| + i\ \arg(z) \ ,} \tag{2.18}$$

where we see that it is a multiple-valued function as a result of the fact that so is $\arg(z)$. In terms of the principal value $\mathrm{Arg}(z)$ of the argument function, we can also write the $\log(z)$ as follows:

$$\boxed{\log(z) = \mathrm{Log}\, |z| + i\ \mathrm{Arg}(z) + 2\pi\, i\, k \qquad \text{for } k = 0, \pm 1, \pm 2, \ldots,} \tag{2.19}$$

which makes the multiple-valuedness manifest.

For example, whereas the real logarithm of 1 is simply 0, the complex logarithm is given by

$$\log(1) = \mathrm{Log}\, |1| + i\ \arg(1) = 0 + i\, 2\pi\, k \quad \text{for any integer } k.$$

As promised, we can now take the logarithm of negative real numbers. For example,

$$\log(-1) = \mathrm{Log}\, |-1| + i\ \arg(-1) = 0 + i\, \pi + i\, 2\pi\, k \quad \text{for any integer } k.$$

The complex logarithm obeys many of the algebraic identities that we expect from the real logarithm, only that we have to take into account its multiple-valuedness properly. Therefore an identity like

$$\boxed{\log(z_1\, z_2) = \log(z_1) + \log(z_2) \ ,} \tag{2.20}$$

for nonzero complex numbers z_1 and z_2, is still valid in the sense that having chosen a value (out of the infinitely many possible values) for $\log(z_1)$ and for $\log(z_2)$, then there is a value of $\log(z_1\, z_2)$ for which the above equation holds. Or said in a different way, the identity holds up to integer multiples of $2\pi\, i$ or, as it is often said, **modulo** $2\pi\, i$:

$$\log(z_1\, z_2) - \log(z_1) - \log(z_2) = 2\pi\, i\, k \quad \text{for some integer } k.$$

Similarly we have

$$\boxed{\log(z_1/z_2) = \log(z_1) - \log(z_2)\ ,} \tag{2.21}$$

in the same sense as before, for any two nonzero complex numbers z_1 and z_2.

Choosing a branch for the logarithm

We now turn to the discussion of the analyticity properties of the complex logarithm function. In order to discuss the analyticity of a function, we need to investigate its differentiability, and for this we need to be able to take its derivative as in equation (2.6). Suppose we were to try to compute the derivative of the function $\log(z)$ at some point z_0. Writing the derivative as the limit of a quotient,

$$\log'(z_0) = \lim_{\Delta z \to 0} \frac{\log(z_0 + \Delta z) - \log(z_0)}{\Delta z}\ ,$$

we encounter an immediate obstacle: since the function $\log(z)$ is multiple-valued we have to make sure that the two log functions in the numerator tend to the same value in the limit, otherwise the limit will not exist. In other words, we have to choose one of the infinitely many values for the log function in a consistent way. This way of restricting the values of a multiple-valued function to make it single-valued in some region (in the above example in some neighbourhood of z_0) is called choosing a branch of the function. For example, we define the **principal branch** Log of the logarithmic function to be

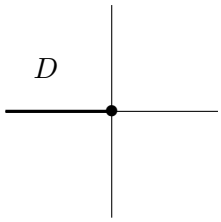
$$\mathrm{Log}(z) = \mathrm{Log}\,|z| + i\ \mathrm{Arg}(z)\ ,$$

where $\mathrm{Arg}(z)$ is the principal value of $\arg(z)$. Af first sight it might seem that this notation is inconsistent, since we are using Log both for the real logarithm and the principal branch of the complex logarithm. So let us make sure that this is not the case. If z is a positive real number, then $z = |z|$ and $\mathrm{Arg}(z) = 0$, whence $\mathrm{Log}(z) = \mathrm{Log}\,|z|$. Hence at least the notation is consistent. The function $\mathrm{Log}(z)$ is single-valued, but at a price: it is no

longer continuous in the whole complex plane, since $\mathrm{Arg}(z)$ is not continuous in the whole complex plane. As explained in Section 2.1.1, the principal branch $\mathrm{Arg}(z)$ of the argument function is discontinuous along the negative real axis. Indeed, let $z_\pm = -x \pm i\,\varepsilon$ where x and ε are positive numbers. In the limit $\varepsilon \to 0$, z_+ and z_- tend to the same point on the negative real axis from the upper and lower half-planes respectively. Hence whereas $\lim_{\varepsilon\to 0} z_\pm = -x$, the principal value of the logarithm obeys

$$\lim_{\varepsilon\to 0} \mathrm{Log}(z_\pm) = \mathrm{Log}(x) \pm i\,\pi \ ,$$

so that it is not a continuous function anywhere on the negative real axis, or at the origin, where the function itself is not well-defined. The non-positive real axis is known as a **branch cut** for this function and the origin is known as a **branch point**.



Let D denote all the points in the complex plane except for those which are real and non-positive; in other words, D is the complement of the non-positive real axis. It is easy to check that D is an open subset of the complex plane and by construction, $\mathrm{Log}(z)$ is single-valued and continuous for all points in D. We will now check that it is analytic there as well. For this we need to compute its derivative. So let $z_0 \in D$ be any point in D and consider $w_0 = \mathrm{Log}(z_0)$. Letting $\Delta z = z - z_0$, we can write the derivative of $w = \mathrm{Log}(z)$ at z_0 in the following form

$$\begin{aligned} \mathrm{Log}'(z_0) &= \lim_{z\to z_0} \frac{w - w_0}{z - z_0} \\ &= \lim_{z\to z_0} \frac{1}{\frac{z-z_0}{w-w_0}} \\ &= \lim_{w\to w_0} \frac{1}{\frac{z-z_0}{w-w_0}} \ , \end{aligned}$$

where to reach the second line we used the fact that $w = w_0$ implies $z = z_0$ (single-valuedness of the exponential function), and to reach the third line we used the continuity of $\mathrm{Log}(z)$ in D to deduce that $w \to w_0$ as $z \to z_0$. Now using that $z = e^w$ we see that what we have here is the reciprocal of the derivative of the exponential function, whence

$$\mathrm{Log}'(z_0) = \lim_{w\to w_0} \frac{1}{\frac{e^w - e^{w_0}}{w-w_0}} = \frac{1}{\exp'(w_0)} = \frac{1}{\exp(w_0)} = \frac{1}{z_0} \ .$$

Since this is well-defined everywhere but for $z_0 = 0$, which does not belong to D, we see that $\mathrm{Log}(z)$ is analytic in D.

Other branches

The choice of branch for the logarithm is basically that, a choice. It is certainly not the only one. We can make the logarithm function single-valued in other regions of the complex plane by choosing a different branch for the argument function.

For example, another popular choice is to consider the function $\mathrm{Arg}_0(z)$ which is the value of the argument function for which

$$0 \leq \mathrm{Arg}_0(z) < 2\pi \ .$$

This function, like $\mathrm{Arg}(z)$, is single-valued but discontinuous; however the discontinuity is now along the positive real axis, since approaching a positive real number from the upper half-plane we would conclude that its argument tends to 0 whereas approaching it from the lower half-plane the argument would tend to 2π. We can therefore define a branch $\mathrm{Log}_0(z)$ of the logarithm by

$$\mathrm{Log}_0(z) = \mathrm{Log}\,|z| + i\ \mathrm{Arg}_0(z) \ .$$

This branch then has a branch cut along the non-negative real axis, but it is continuous in its complement D_0 as shown in Figure 2.2. The same argument as before shows that $\mathrm{Log}_0(z)$ is analytic in D_0 with derivative given by

$$\mathrm{Log}_0'(z_0) = \frac{1}{z_0} \qquad \text{for all } z_0 \text{ in } D_0.$$

Figure 2.2: Two further branches of the logarithm.

There are of course many other branches. For example, let τ be any real number and define the branch $\mathrm{Arg}_\tau(z)$ of the argument function to take the values

$$\tau \leq \mathrm{Arg}_\tau(z) < \tau + 2\pi \ .$$

This gives rise to a branch $\mathrm{Log}_\tau(z)$ of the logarithm function defined by

$$\mathrm{Log}_\tau(z) = \mathrm{Log}\,|z| + i\ \mathrm{Arg}_\tau(z) \ ,$$

which has a branch cut emanating from the origin and consisting of all those points z with $\arg(z) = \tau$ modulo 2π. Again the same arguments show that $\mathrm{Log}_\tau(z)$ is analytic everywhere on the complement D_τ of the branch cut, as shown in Figure 2.2, and its derivative is given by

$$\mathrm{Log}'_\tau(z_0) = \frac{1}{z_0} \qquad \text{for all } z_0 \text{ in } D_\tau.$$

The choice of branch is immaterial for many properties of the logarithm, although it is important that a choice be made. Different applications may require choosing one branch over another. Provided one is consistent this should not cause any problems.

As an example suppose that we are faced with computing the derivative of the function $f(z) = \log(z^2 + 2iz + 2)$ at the point $z = i$. We need to choose a branch of the logarithm which is analytic in a region containing a neighbourhood of the point $i^2 + 2i\,i + 2 = -1$. The principal branch is not analytic there, so we have to choose another branch. Suppose that we choose $\mathrm{Log}_0(z)$. Then, by the chain rule

$$f'(i) = \left.\frac{2z + 2i}{z^2 + 2iz + 2}\right|_{z=i} = \frac{2\,i + 2\,i}{i^2 + 2i^2 + 2} = -4\,i\ .$$

Any other valid branch would of course give the same result.

2.1.7 Complex powers

With the logarithm function at our disposal, we are able to define complex powers of complex numbers. Let α be a complex number. The for all $z \neq 0$, we define the α-**th power** z^α of z by

$$\boxed{z^\alpha \equiv e^{\alpha\,\log(z)} = e^{\alpha\,\mathrm{Log}\,|z| + i\,\alpha\,\arg(z)}\ .} \tag{2.22}$$

The multiple-valuedness of the argument means that generically there will be an infinite number of values for z^α. We can rewrite the above expression a little to make this manifest:

$$z^\alpha = e^{\alpha\,\mathrm{Log}\,|z| + i\,\alpha\,\mathrm{Arg}(z) + i\,\alpha\,2\pi\,k} = e^{\alpha\,\mathrm{Log}(z)} e^{i\,\alpha\,2\pi\,k}\ ,$$

for $k = 0, \pm 1, \pm 2, \ldots$.

Depending on α we will have either one, finitely many or infinitely many values of $\exp(i\,2\pi\,\alpha\,k)$. Suppose that α is real. If $\alpha = n$ is an integer then

so is $\alpha\, k = nk$ and $\exp(i\, 2\pi\, \alpha\, k) = \exp(i\, 2\pi\, nk) = 1$. There is therefore only one value for z^n. This is as we expect, since in this case we have

$$z^n = \begin{cases} 1 & \text{for } n = 0, \\ \underbrace{z\, z\, \cdots\, z}_{n \text{ times}} & \text{for } n > 0, \\ \frac{1}{z^{-n}} & \text{for } n < 0. \end{cases}$$

If $\alpha = p/q$ is a rational number, where we have chosen the integers p and q to have no common factors (i.e., to be **coprime**), then $z^{p/q}$ will have a finite number of values. Indeed consider $\exp(i\, 2\pi\, kp/q)$ as k ranges over the integers. It is clear that this exponential takes the same values for k and for $k + q$:

$$e^{i\, 2\pi\, (k+q)p/q} = e^{i\, 2\pi\, (k(p/q)+p)} = e^{i\, 2\pi\, k(p/q)+i\, 2\pi\, p} = e^{i\, 2\pi\, kp/q}\ ,$$

where we have used the addition and periodicity properties (2.14) and (2.15) of the exponential function. Therefore $z^{p/q}$ will have at most q distinct values, corresponding to the above formula with, say, $k = 0, 1, 2, \ldots, q - 1$. In fact, it will have precisely q distinct values, as we will see below. Finally, if α is irrational, then z^α will possess an infinite number of values. To see this notice that if there are integers k and k' for which $e^{i\, \alpha\, 2\pi\, k} = e^{i\, \alpha\, 2\pi\, k'}$, then we must have that $e^{i\, \alpha\, 2\pi\, (k-k')} = 1$, which means that $\alpha\, (k - k')$ must be an integer. Since α is irrational, this can only be true if $k = k'$.

For example, let us compute $1^{1/q}$. According to the formula,

$$1^{1/q} = e^{\mathrm{Log}(1)/q}\, e^{i\, 2\pi\, (k/q)} = e^{i\, 2\pi\, (k/q)}\ ,$$

as k ranges over the integers. As discussed above only the q values $k = 0, 1, 2, \ldots, q - 1$ will be different. The values of $1^{1/q}$ are known as **q-th roots of unity**. They each have the property that their q-th power is equal to 1: $(1^{1/q})^q = 1$, as can be easily seen from the above expression. Let $\omega = \exp(i\, 2\pi/q)$ correspond to the $k = 1$ value of $1^{1/q}$. Then the q-th roots of unity are given by $1, \omega, \omega^2, \ldots, \omega^{q-1}$, and there are q of them. The q-th roots of unity lie in the unit circle $|z| = 1$ in the complex plane and define the vertices of a regular q-gon. For example, in Figure 2.3 we depict the q-th roots of unity for $q = 3, 5, 7, 11$.

Let z be a nonzero complex number and suppose that we are after its q-th roots. Writing z in polar form $z = |z|\, \exp(i\, \theta)$, we have

$$z^{1/q} = |z|^{1/q}\, e^{i\, \theta/q} \omega^k \qquad \text{for } k = 0, 1, 2, \ldots, q - 1.$$

In other words the q different values of $z^{1/q}$ are obtained from any one value by multiplying it by the q powers of the q-th roots of unity. If p is any integer,

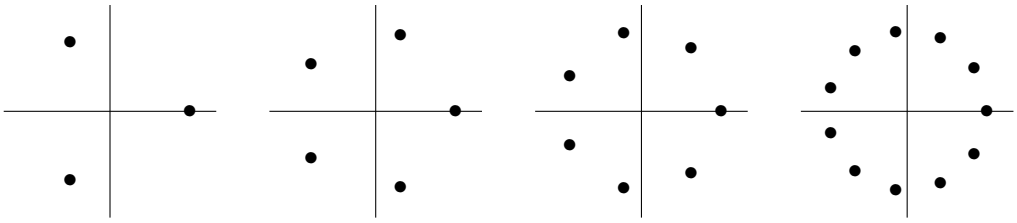


Figure 2.3: Some roots of unity.

we can then take the p-th power of the above formula:

$$z^{p/q} = |z|^{p/q}\, e^{i\,p\,\theta/q} \omega^{pk} \qquad \text{for } k = 0, 1, 2, \ldots, q-1.$$

If p and q are coprime, the ω^{pk} for $k = 0, 1, 2, \ldots, q-1$ are different. Indeed, suppose that $\omega^{pk} = \omega^{pk'}$, for k and k' between 0 and $q-1$. Then $\omega^{p(k-k')} = 1$, which means that $p(k-k')$ has to be a multiple of q. Because p and q are coprime, this can only happen when $k = k'$. Therefore we see that indeed a rational power p/q (with p and q coprime) of a complex number has precisely q values.

Let us now consider complex powers. If $\alpha = a + i\,b$ is not real (so that $b \neq 0$), then z^α will always have an infinite number of values. Indeed, notice that the last term in the following expression takes a different value for each integer k:

$$e^{i\,\alpha\,2\pi\,k} = e^{i\,(a+i\,b)\,2\pi\,k} = e^{i\,2\pi\,k\,a} e^{-2\pi\,k\,b}\ .$$

For examples, let us compute i^i. By definition,

$$i^i = e^{i\,\log(i)} = e^{i\,(\mathrm{Log}(i)+i\,2\pi\,k)} = e^{i\,(i\pi/2+i\,2\pi\,k)} = e^{-\pi/2}\, e^{-2\pi\,k}\ ,$$

for $k = 0. \pm 1, \pm 2, \ldots$, which interestingly enough is real.

Choosing a branch for the complex power

Every branch of the logarithm gives rise to a branch of z^α. In particular we define the **principal branch** of z^α to be $\exp(\alpha\ \mathrm{Log}(z))$. Since the exponential function is entire, the principal branch of z^α is analytic in the domain D where $\mathrm{Log}(z)$ is analytic. We can compute its derivative for any point z_0 in D using the chain rule (2.12):

$$\frac{d}{dz}\left(e^{\alpha\ \mathrm{Log}(z)}\right)\Big|_{z=z_0} = e^{\alpha\ \mathrm{Log}(z_0)}\ \frac{\alpha}{z_0}\ .$$

Given any nonzero z_0 in the complex plane, we can choose a branch of the logarithm so that the function z^α is analytic in a neighbourhood of z_0. We

can compute its derivative there and we see that the following equation holds

$$\frac{d}{dz}\,(z^\alpha)|_{z=z_0} = \alpha\, z_0^\alpha\, \frac{1}{z_0}\ ,$$

provided that we use the same branch of z^α on both sides of the equation.

One might be tempted to write the right-hand side of the above equation as $\alpha\, z_0^{\alpha-1}$, and indeed this is correct, since the complex powers satisfy many of the identities that we are familiar with from real powers. For example, one can easily show that for any complex numbers α and β

$$z^\alpha\, z^\beta = z^{\alpha+\beta}\ ,$$

provided that the same branch of the logarithm, and hence of the complex power, is chosen on both sides of the equation. Nevertheless, there is one identity that does *not* hold. Suppose that α is a complex number and let z_1 and z_2 be nonzero complex numbers. Then it is *not* true that $z_1^\alpha\, z_2^\alpha$ and $(z_1\, z_2)^\alpha$ agree, even if, as we always should, we choose the same branch of the complex power on both sides of the equation.

We end this section with the observation that the function z^z is analytic wherever the chosen branch of the logarithm function is defined. Indeed, $z^z = \exp(z\,\log(z))$ and its principal branch can is defined to be the function $\exp(z\,\mathrm{Log}(z))$, which as we now show is analytic in D. Taking the derivative we see that

$$\frac{d}{dz}\left(e^{z\,\mathrm{Log}(z)}\right)\Big|_{z=z_0} = e^{z_0\,\mathrm{Log}(z_0)}\,(\mathrm{Log}(z_0)+1)\ ,$$

which exists everywhere on D. Again a similar result holds for any other branch provided we are consistent and take the same branches of the logarithm in both sides of the following equation:

$$\frac{d}{dz}\,(z^z)|_{z=z_0} = z_0^{z_0}\,(\log(z_0)+1)\ .$$

2.2 Complex integration

Having discussed differentiation of complex-valued functions, it is time to now discuss integration. In real analysis differentiation and integration are roughly speaking inverse operations. We will see that something similar also happens in the complex domain; but in addition, and this is unique to complex analytic functions, differentiation and integration are also roughly equivalent operations, in the sense that we will be able to take derivatives by performing integrals.

2.2.1 Complex integrals

There is a sense in which the integral of a complex-valued function is a trivial extension of the standard integral one learns about in calculus. Suppose that f is a complex-valued function of a *real* variable t. We can decompose $f(t)$ into its real and imaginary parts $f(t) = u(t) + i\,v(t)$, where u and v are now real-valued functions of a real variable. We can therefore define the integral $\int_a^b f(t)\,dt$ of $f(t)$ on the interval $[a, b]$ as

$$\int_a^b f(t)\,dt = \int_a^b u(t)\,dt + i\,\int_a^b v(t)\,dt\ ,$$

provided that the functions u and v are integrable. We will not develop a formal theory of integrability in this course. You should nevertheless be aware of the fact that whereas not every function is integrable, a continuous function always is. Hence, for example, if f is a continuous function in the interval $[a, b]$ then the integral $\int_a^b f(t)\,dt$ will always exist, since u and v are continuous and hence integrable.

This integral satisfies many of the properties that real integrals obey. For instance, it is (complex) linear, so that if α and β are complex numbers and f and g are complex-valued functions of t, then

$$\int_a^b \left(\alpha\,f(t) + \beta\,g(t)\right)\,dt = \alpha\,\int_a^b f(t)\,dt + \beta\,\int_a^b g(t)\,dt\ .$$

It also satisfies a complex version of the **Fundamental Theorem of Calculus**. This theorem states that if $f(t)$ is continuous in $[a, b]$ and there exists a function $F(t)$ also defined on $[a, b]$ such that $\dot{F}(t) = f(t)$ for all $a \le t \le b$, where $\dot{F}(t) \equiv \frac{dF}{dt}$, then

$$\int_a^b f(t)\,dt = \int_a^b \frac{dF(t)}{dt}\,dt = F(b) - F(a)\ . \tag{2.23}$$

This follows from the similar theorem for real integrals, as we now show. Indeed, let us decompose both f and F into real and imaginary parts: $f(t) = u(t) + i\,v(t)$ and $F(t) = U(t) + i\,V(t)$. Then since F is an antiderivative $\dot{F}(t) = \dot{U}(t) + i\,\dot{V}(t) = f(t) = u(t) + i\,v(t)$, whence $\dot{U}(t) = u(t)$ and $\dot{V}(t) = v(t)$. Therefore, by definition

$$\begin{aligned}\int_a^b f(t)\,dt &= \int_a^b u(t)\,dt + i\,\int_a^b v(t)\,dt \\ &= U(b) - U(a) + i\,\left(V(b) - V(a)\right) \\ &= U(b) + i\,V(b) - \left(U(a) + i\,V(a)\right) \\ &= F(b) - F(a)\ ,\end{aligned}$$

where to reach the second line we used the real version of the fundamental theorem of calculus for the real and imaginary parts of the integral.

A final useful property of the complex integral is that

$$\left| \int_a^b f(t)\, dt \right| \leq \int_a^b |f(t)|\, dt \ . \tag{2.24}$$

This result makes sense intuitively because in integrating $f(t)$ one might encounter cancellations which do not occur while integrating the non-negative quantity $|f(t)|$.

This last property follows from the similar property of real integrals. Let us see this. Write the complex integral $\int_a^b f(t)\, dt$ in polar form:

$$\int_a^b f(t)\, dt = R\, e^{i\,\theta} \ ,$$

where

$$R = \int_a^b f(t)\, dt \quad .$$

On the other hand,

$$R = \int_a^b e^{-i\,\theta} f(t)\, dt \ .$$

Write $e^{-i\,\theta} f(t) = U(t) + i\, V(t)$ where $U(t)$ and $V(t)$ are real-valued functions. Then because R is real,

$$R = \int_a^b U(t)\, dt \ .$$

But now,

$$U(t) = \mathrm{Re}\ \ e^{-i\,\theta} f(t) \ \ \leq \ e^{-i\,\theta} f(t) \ = |f(t)| \ .$$

Therefore, from the properties of real integrals,

$$\int_a^b U(t)\, dt \leq \int_a^b |f(t)|\, dt \ ,$$

which proves the desired result.

2.2.2 Contour integrals

Much more interesting is the integration of complex-valued functions of a *complex* variable. We would like to be able to make sense out of something like

$$\int_{z_0}^{z_1} f(z)\, dz \ ,$$

where z_0 and z_1 are complex numbers. We are immediately faced with a difficulty. Unlike the case of an interval $[a, b]$ when it is fairly obvious how to go from a to b, here z_0 and z_1 are points in the complex plane and there are many ways to go from one point to the other. Therefore as it stands, the above integral is ambiguous. The way out of this ambiguity is to specify a path joining z_0 to z_1 and then integrate the function along the path. In order to do this we will have to introduce some notation.

The integral along a parametrised curve

Let z_0 and z_1 be two points in the complex plane. One has an intuitive notion of what one means by a curve joining z_0 and z_1. Physically, we can think of a point-particle moving in the complex plane, starting at some time t_0 at the point z_0 and ending at some later time t_1 at the point z_1. At any given instant in time $t_0 \leq t \leq t_1$, the particle is at the point $z(t)$ in the complex plane. Therefore we see that a curve joining z_0 and z_1 can be defined by a function $z(t)$ taking points t in the interval $[t_0, t_1]$ to points $z(t)$ in the complex plane in such a way that $z(t_0) = z_0$ and $z(t_1) = z_1$. Let us make this a little more precise. By a **(parametrised) curve** joining z_0 and z_1 we shall mean a continuous function $z : [t_0, t_1] \to \mathbb{C}$ such that $z(t_0) = z_0$ and $z(t_1) = z_1$. We can decompose z into its real and imaginary parts, and this is equivalent to two continuous real-valued functions $x(t)$ and $y(t)$ defined on the interval $[t_0, t_1]$ such that $x(t_0) = x_0$ and $x(t_1) = x_1$ and similarly for $y(t)$: $y(t_0) = y_0$ and $y(t_1) = y_1$, where $z_0 = x_0 + i\, y_0$ and $z_1 = x_1 + i\, y_1$. We say that the curve is **smooth** if its velocity $\dot{z}(t)$ is a continuous function $[t_0, t_1] \to \mathbb{C}$ which is never zero.

Let Γ be a smooth curve joining z_0 to z_1, and let $f(z)$ be a complex-valued function which is continuous on Γ. Then we define the **integral of f along Γ** by

$$\boxed{\int_\Gamma f(z)\, dz \equiv \int_{t_0}^{t_1} f(z(t))\, \dot{z}(t)\, dt\ .} \tag{2.25}$$

By hypothesis, the integrand, being a product of continuous functions, is itself continuous and hence the integral exists.

Let us compute some examples. Consider the function $f(z) = x^2 + i\, y^2$ integrated along the smooth curve parametrised by $z(t) = t + i\, t$ for $0 \leq t \leq 1$. As shown in Figure 2.4 this is the straight line segment joining the origin and the point $1+i$. Decomposing $z(t) = x(t) + i\, y(t)$ into real and imaginary parts, we see that $x(t) = y(t) = t$. Therefore $f(z(t)) = t^2 + i\, t^2$ and $\dot{z}(t) = 1 + i$. Putting it all together, using complex linearity of the integral and performing the elementary real integral, we find the following result

$$\int_\Gamma f(z)\, dz = \int_0^1 (t^2 + i\, t^2)(1+i)\, dt = \int_0^1 (1+i)^2\, t^2\, dt = 2i\ \frac{t^3}{3}\bigg|_0^1 = \frac{2i}{3}\ .$$

Consider now the function $f(z) = 1/z$ integrated along the smooth curve Γ parametrised by $z(t) = R \exp(i\, 2\pi\, t)$ for $0 \leq t \leq 1$, where $R \neq 0$. As shown in Figure 2.4, the resulting curve is the circle of radius R centred about the origin. Here $f(z(t)) = (1/R) \exp(-i\, 2\pi\, t)$ and $\dot{z}(t) = 2\pi\, i\, R \exp(i\, 2\pi\, t)$.

Figure 2.4: Two parametrised curves.

Putting it all together we obtain

$$\int_\Gamma f(z)\,dz = \int_0^1 \frac{2\pi\,i\,Re^{i\,2\pi\,t}}{Re^{i\,2\pi\,t}}\,dt = 2\pi\,i\,\int_0^1 dt = 2\pi\,i\ . \tag{2.26}$$

Notice that the result is independent of the radius. This is in sharp contrast with real integrals, which we are used to interpret physically in terms of area. In fact, the above integral behaves more like a charge than like an area.

Finally let us consider the function $f(z) \equiv 1$ along *any* smooth curve Γ parametrised by $z(t)$ for $0 \le t \le 1$. It may seem that we do not have enough information to compute the integral, but let us see how far we can get with the information given. The integral becomes

$$\int_\Gamma f(z)\,dz = \int_0^1 \dot{z}(t)\,dt\ .$$

Using the complex version of the fundamental theorem of calculus, we have

$$\int_0^1 \dot{z}(t)\,dt = z(1) - z(0)\ ,$$

independent of the actual curve used to join the two points! Notice that this integral is therefore *not* the length of the curve as one might think from the notation.

The length of a curve and a useful estimate

The length of the curve can be computed, but the integral is not related to the complex dz but the real $|dz|$. Indeed, if Γ is a curve parametrised by $z(t) = x(t) + i\,y(t)$ for $t \in [t_0, t_1]$, consider the real integral

$$\begin{aligned}\int_\Gamma |dz| &\equiv \int_{t_0}^{t_1} |\dot{z}(t)|\,dt \\ &= \int_{t_0}^{t_1} \sqrt{\dot{x}(t)^2 + \dot{y}(t)^2}\,dt\ ,\end{aligned}$$

which is the integral of the infinitesimal line element $\sqrt{dx^2 + dy^2}$ along the curve. Therefore, the integral is the (arc)length $\ell(\Gamma)$ of the curve:

$$\boxed{\int_\Gamma |dz| = \ell(\Gamma) \ .} \tag{2.27}$$

This immediately yields a useful estimate on integrals along curves, analogous to equation (2.24). Indeed, suppose that Γ is a curve parametrised by $z(t)$ for $t \in [t_0, t_1]$. Then,

$$\begin{aligned}\left|\int_\Gamma f(z)\,dz\right| &= \left|\int_{t_0}^{t_1} f(z(t))\,\dot{z}(t)\,dt\right| \\ &\leq \int_{t_0}^{t_1} |f(z(t))|\,|\dot{z}(t)|\,dt \qquad \text{(using (2.24))} \\ &\leq \max_{z\in\Gamma} |f(z)| \int_{t_0}^{t_1} |\dot{z}(t)|\,dt \ .\end{aligned}$$

But this last integral is simply the length $\ell(\Gamma)$ of the curve, whence we have

$$\boxed{\left|\int_\Gamma f(z)\,dz\right| \leq \int_\Gamma |f(z)|\,|dz| \leq \max_{z\in\Gamma} |f(z)|\,\ell(\Gamma) \ .} \tag{2.28}$$

Results of this type are the bread and butter of analysis and in this part of the course we will have ample opportunity to use this particular one.

Some further properties of the integrals along a curve

We have just seen that one of the above integrals does not depend on the actual path but just on the endpoints of the contour. We will devote the next two sections to studying conditions for complex integrals to be independent of the path; but before doing so, we discuss some general properties of the integrals $\int_\Gamma f(z)\,dz$.

The first important property is that the integral is complex linear. That is, if α and β are complex numbers and f and g are functions which are continuous on Γ, then

$$\int_\Gamma (\alpha\, f(z) + \beta\, g(z))\ dz = \alpha \int_\Gamma f(z)\,dz + \beta \int_\Gamma g(z)\,dz \ .$$

The proof is routine and we leave it as an exercise.

The first nontrivial property is that the integral $\int_\Gamma f(z)\,dz$ does not depend on the actual parametrisation of the curve Γ. In other words, it is a "physical" property of the curve itself, meaning the set of points $\Gamma \subset \mathbb{C}$ together with the direction along the curve, and not of the way in which we go about traversing them.

The only difficult thing in showing this is coming up with a mathematical statement to prove. Let $z(t)$ for $t_0 \le t \le t_1$ and $z'(t)$ for $t'_0 \le t \le t'_1$ be two smooth parametrisations of the same curve Γ. This means that $z(t_0) = z'(t'_0)$ and $z(t_1) = z'(t'_1)$. We will say that the parametrisations $z(t)$ and $z'(t)$ are **equivalent** if there exists a one-to-one differentiable function $\lambda : [t'_0, t'_1] \to [t_0, t_1]$ such that $z'(t) = z(\lambda(t))$. In particular, this means that $\lambda(t'_0) = t_0$ and $\lambda(t'_1) = t_1$. (It is possible to show that this is indeed an equivalence relation.)

The condition of **reparametrisation invariance** of $\int_\Gamma f(z)\,dz$ can then be stated as follows. Let z and z' be two equivalent parametrisations of a curve Γ. Then for any function $f(z)$ continuous on Γ, we have

$$\int_{t'_0}^{t'_1} f(z'(t))\,\dot{z}'(t)\,dt = \int_{t_0}^{t_1} f(z(t))\,\dot{z}(t)\,dt\ .$$

Let us prove this.

$$\begin{aligned}\int_{t'_0}^{t'_1} f(z'(t))\,\dot{z}'(t)\,dt &= \int_{t'_0}^{t'_1} f(z(\lambda(t)))\,\dot{z}(\lambda(t))\,dt \\ &= \int_{\lambda(t'_0)}^{\lambda(t'_1)} f(z(\lambda))\,\frac{dz(\lambda)}{d\lambda}\,d\lambda \\ &= \int_{t_0}^{t_1} f(z(\lambda))\,\frac{dz(\lambda)}{d\lambda}\,d\lambda\ ,\end{aligned}$$

which after changing the name of the variable of integration from λ to t (Shakespeare's Theorem!), is seen to agree with

$$\int_{t_0}^{t_1} f(z(t))\,\dot{z}(t)\,dt\ .$$

Because of reparametrisation invariance, we can always parametrise a curve in such a way that the initial time is $t = 0$ and the final time is $t = 1$. Indeed, let $z(t)$ for $t_0 \le t \le t_1$ be any smooth parametrisation of a curve Γ. Then define the parametrisation $z'(t) = z(t_0 + t(t_1 - t_0))$. Clearly, $z'(0) = z(t_0)$ and $z'(1) = z(t_1)$, and moreover $\dot{z}'(t) = (t_1 - t_0)\dot{z}(t_0 + t(t_1 - t_0))$ hence z' is also smooth.

Now let us notice that parametrised curves Γ have a natural notion of direction: this is the direction in which we traverse the curve. Choosing a parametrisation $z(t)$ for $0 \le t \le 1$, as we go from $z(0)$ to $z(1)$, we trace the points in the curve in a given order, which we depict by an arrowhead on the curve indicating the direction along which t increases, as in the curves in Figure 2.4. A curve with such a choice of direction is said to be **directed**.

Given any directed curve Γ, we let $-\Gamma$ denote the directed curve with the opposite direction; that is, with the arrow pointing in the opposite direction. The final interesting property of the integral $\int_\Gamma f(z)\,dz$ is that

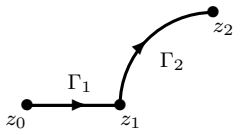
$$\boxed{\int_{-\Gamma} f(z)\,dz = -\int_{\Gamma} f(z)\,dz\ .} \tag{2.29}$$

To prove this it is enough to find two parametrisations for Γ and $-\Gamma$ and compute the integrals. By reparametrisation independence it does not matter which parametrisations we choose. If $z(t)$ for $0 \le t \le 1$ is a parametrisation for Γ, then $z'(t) = z(1-t)$ for $0 \le t \le 1$ is a parametrisation for $-\Gamma$. Indeed, $z'(0) = z(1)$ and $z'(1) = z(0)$ and they trace the same set of points. Let us compute:

$$\begin{aligned}
\int_{-\Gamma} f(z)\,dz &= \int_0^1 f(z'(t))\,\dot{z}'(t)\,dt \\
&= -\int_0^1 f(z(1-t))\,\dot{z}(1-t)\,dt \\
&= \int_1^0 f(z(t'))\,\dot{z}(t')\,dt' \\
&= -\int_0^1 f(z(t'))\,\dot{z}(t')\,dt' \\
&= -\int_\Gamma f(z)\,dz\ .
\end{aligned}$$

Piecewise smooth curves and contour integrals

Finally we have to generalise the integral $\int_\Gamma f(z)\,dz$ to curves which are not necessarily smooth, but which are made out of smooth curves. Curves can be composed: if Γ_1 is a curve joining z_0 to z_1 and Γ_2 is a curve joining z_1 to z_2, then we can make a curve Γ joining z_0 to z_2 by first going to the intermediate point z_1 via Γ_1 and then from there via Γ_2 to our destination z_2. The resulting curve Γ is still continuous, but it will generally fail to be smooth, since the velocity need not be continuous at the intermediate point z_1, as shown in the figure.



However such curve is **piecewise smooth**: which means that it is made out of smooth components by the composition procedure just outlined. In terms of parametrisations, if $z_1(t)$ and $z_2(t)$, for $0 \le t \le 1$, are smooth parametrisations for Γ_1 and Γ_2 respectively, then

$$z(t) = \begin{cases} z_1(2t) & \text{for } 0 \le t \le \frac{1}{2} \\ z_2(2t-1) & \text{for } \frac{1}{2} \le t \le 1 \end{cases}$$

is a parametrisation for Γ. Notice that it is well-defined and continuous at $t = \frac{1}{2}$ precisely because $z_1(1) = z_2(0)$; however it need not be smooth there

since $\dot{z}_1(1) \neq \dot{z}_2(0)$ necessarily. We can repeat this procedure and construct curves which are not smooth but which are made out of a finite number of smooth curves: one curve ending where the next starts. Such a piecewise smooth curve will be called a **contour** from now on. If a contour Γ is made out of composing a finite number of smooth curves $\{\Gamma_j\}$ we will say that each Γ_j is a **smooth component** of Γ.

Let Γ be a contour with n smooth components $\{\Gamma_j\}$ for $j = 1, 2, \ldots, n$. If $f(z)$ is a function continuous on Γ, then the **contour integral of f along** Γ is defined as

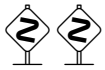
$$\int_\Gamma f(z)\,dz = \sum_{j=1}^{n} \int_{\Gamma_j} f(z)\,dz = \int_{\Gamma_1} f(z)\,dz + \int_{\Gamma_2} f(z)\,dz + \cdots + \int_{\Gamma_n} f(z)\,dz \ ,$$

with each of the $\int_{\Gamma_i} f(z)\,dz$ is defined by (2.25) relative to any smooth parametrisation.

2.2.3 Independence of path

In this section we will investigate conditions under which a contour integral only depends on the endpoints of the contour, and not not the contour itself. This is necessary preparatory material for Cauchy's integral theorem which will be discussed in the next section.

We will say that an open subset U of the complex plane is **connected**, if every pair of points in U can be joined by a contour. A connected open subset of the complex plane will be called a **domain**.



What we have called connected here is usually called **path-connected**. We can allow ourselves this abuse of notation because path-connectedness is easier to define and it can be shown that the two notions agree for subsets of the complex plane.

Fundamental Theorem of Calculus: contour integral version

First we start with a contour integral version of the fundamental theorem of calculus. Let D be a domain and let $f : D \to \mathbb{C}$ be a continuous complex-valued function defined on D. We say that f has an **antiderivative** in D if there exists some function $F : D \to \mathbb{C}$ such that

$$F'(z) = \frac{dF(z)}{dz} = f(z) \ .$$

Notice that F is therefore analytic in D. Now let Γ be any contour in D with endpoints z_0 and z_1. If f has an antiderivative F on D, the contour integral is given by

$$\int_\Gamma f(z)\,dz = F(z_1) - F(z_0) \ . \tag{2.30}$$

Let us first prove this for Γ a smooth curve, parametrised by $z(t)$ for $0 \leq t \leq 1$. Then

$$\int_\Gamma f(z)\,dz = \int_0^1 F'(z(t))\dot{z}(t)dt = \int_0^1 \frac{dF(z(t))}{dt}\,dt\ .$$

Using the complex version of the fundamental theorem of calculus (2.23), we see that

$$\int_\Gamma f(z)\,dz = F(z(1)) - F(z(0)) = F(z_1) - F(z_0)\ .$$

Now we consider the general case: Γ a contour with smooth components $\{\Gamma_j\}$ for $j = 1, 2, \ldots, n$. The curve Γ_1 starts in z_0 and ends in some intermediate point τ_1, Γ_2 starts in τ_1 and ends in a second intermediate point τ_2, and so so until Γ_n which starts in the intermediate point τ_{n-1} and ends in z_1. Then

$$\begin{aligned}\int_\Gamma f(z)dz &= \sum_{j=1}^n \int_{\Gamma_j} f(z)\,dz \\ &= \int_{\Gamma_1} f(z)\,dz + \int_{\Gamma_2} f(z)\,dz + \cdots + \int_{\Gamma_n} f(z)\,dz \\ &= F(\tau_1) - F(z_0) + F(\tau_2) - F(\tau_1) + \cdots + F(z_1) - F(\tau_{n-1}) \\ &= F(z_1) - F(z_0)\ ,\end{aligned}$$

where we have used the definition of the contour integral and the result proven above for each of the smooth components.

This result says that if a function f has an antiderivative, then its contour integrals do not depend on the precise path, but only on the endpoints. Path independence can also be rephrased in terms of closed contour integrals. We say that a contour is **closed** if its endpoints coincide. The contour integral along a closed contour Γ is sometimes denoted $\oint_\Gamma$ when we wish to emphasise that the contour is closed.

The path-independence lemma

As a corollary of the above result, we see that if Γ is a closed contour in some domain D and $f : D \to \mathbb{C}$ has an antiderivative in D, then

$$\oint_\Gamma f(z)\,dz = 0\ .$$

This is clear because if the endpoints coincide, so that $z_0 = z_1$, then $F(z_1) - F(z_0) = 0$.

In fact, let $f : D \to \mathbb{C}$ be a continuous function on some domain D. Then the following three statements are equivalent:

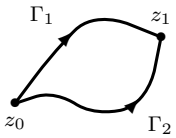
(a) f has an antiderivative F in D;

(b) The closed contour integral $\oint_\Gamma f(z)\,dz$ vanishes for all closed contours Γ in D; and

(c) The contour integrals $\int_\Gamma f(z)\,dz$ are independent of the path.

We shall call this result the **Path-independence Lemma**.

We have already proven that (a) implies (b) and (c). We will now show that in fact (b) and (c) are equivalent.



Let Γ_1 and Γ_2 be any two contours in D sharing the same initial and final endpoints: z_0 and z_1, say. Then consider the contour Γ obtained by composing Γ_1 with $-\Gamma_2$. This is a closed contour with initial and final endpoint z_0. Therefore, using (2.29) for the integral along $-\Gamma_2$,

$$\begin{aligned}\oint_\Gamma f(z)\,dz &= \int_{\Gamma_1} f(z)\,dz + \int_{-\Gamma_2} f(z)\,dz \\ &= \int_{\Gamma_1} f(z)\,dz - \int_{\Gamma_2} f(z)\,dz\ ,\end{aligned}$$

whence $\oint_\Gamma f(z)\,dz = 0$ if and only if $\int_{\Gamma_1} f(z)\,dz = \int_{\Gamma_2} f(z)\,dz$. This shows that (b) implies (c). Now we prove that, conversely, (c) implies (b). Let Γ be any closed contour with endpoints $z_1 = z_0$. By path-independence, we can evaluate the integral by taking the trivial contour which remains at z_0 for all $0 \le t \le 1$. This parametrisation is strictly speaking not smooth since $\dot{z}(t) = 0$ for all t, but the integrand $f(z(t))\dot{z}(t) = 0$ is certainly continuous, so that the integral exists and is clearly zero. Hence $\oint_\Gamma f(z)\,dz = 0$ for all closed contours Γ. Alternatively, we can pick any point τ in the contour not equal to $z_0 = z_1$. We can think of the contour as made out of two contours: Γ_1 from z_0 to τ and Γ_2 from τ to $z_1 = z_0$. We can therefore go from $z_0 = z_1$ to τ in two ways: one is along Γ_1 and the other one is along $-\Gamma_2$. Path-independence says that the result is the same:

$$\int_{\Gamma_1} f(z)\,dz = \int_{-\Gamma_2} f(z)\,dz = -\int_{\Gamma_2} f(z)\,dz\ ,$$

where we have used equation (2.29). Therefore,

$$0 = \int_{\Gamma_1} f(z)\,dz + \int_{\Gamma_2} f(z)\,dz = \int_{\Gamma} f(z)\,dz\ .$$

Finally we finish the proof of the path-independence lemma by showing that (c) implies (a); that is, if all contour integrals are path-independence, then the function f has an antiderivative. The property of path-independence suggests a way to define the antiderivative. Let us fix once and for all a point z_0 in the domain D. Let z be an arbitrary point in D. Because D is connected there will be a contour Γ joining z_0 and z. Define a function $F(z)$ by

$$F(z) \equiv \int_\Gamma f(\zeta)\, d\zeta \ ,$$

where we have changed notation in the integral (Shakespeare's Theorem again) not to confuse the variable of integration with the endpoint z of the contour. By path-independence this integral is independent of the contour and is therefore well-defined as a function of the endpoint z. We must now check that it is an antiderivative for f.

The derivative of $F(z)$ is computed by

$$F'(z) = \lim_{\Delta z \to 0} \frac{1}{\Delta z} \left[\int_{\Gamma'} f(\zeta)\, d\zeta - \int_\Gamma f(\zeta)\, d\zeta \right] \ ,$$

where Γ' is any contour from z_0 to $z+\Delta z$. Since we are interested in the limit of $\Delta z \to 0$, we can assume that Δz is so small that $z+\Delta z$ is contained in some open ε-disk about z which also belongs to D.[2] This means that the straight-line segment Γ'' from z to $z+\Delta z$ belongs to D. By path-independence we are free to choose the contour Γ', and we exercise this choice by taking Γ' to be the composition of Γ with this straight-line segment Γ''. Therefore,

$$\begin{aligned} \int_{\Gamma'} f(\zeta)\, d\zeta - \int_\Gamma f(\zeta)\, d\zeta &= \int_\Gamma f(\zeta)\, d\zeta + \int_{\Gamma''} f(\zeta)\, d\zeta - \int_\Gamma f(\zeta)\, d\zeta \\ &= \int_{\Gamma''} f(\zeta)\, d\zeta \ , \end{aligned}$$

whence

$$F'(z) = \lim_{\Delta z \to 0} \frac{1}{\Delta z} \int_{\Gamma''} f(\zeta)\, d\zeta \ .$$

We parametrise the contour Γ'' by $\zeta(t) = z + t\Delta z$ for $0 \le t \le 1$. Then we

[2] In more detail, since D is open we know that there exists some $\varepsilon > 0$ small enough so that $D_\varepsilon(z)$ belongs to D. We then simply take $|\Delta z| < \varepsilon$, which we can do since we are interested in the limit $\Delta z \to 0$.

have

$$\begin{aligned} F'(z) &= \lim_{\Delta z \to 0} \frac{1}{\Delta z} \int_0^1 f(z + t\Delta z)\, \dot{\zeta}(t)\, dt \\ &= \lim_{\Delta z \to 0} \frac{1}{\Delta z} \int_0^1 f(z + t\Delta z)\, \Delta z\, dt \\ &= \lim_{\Delta z \to 0} \int_0^1 f(z + t\Delta z)\, dt \ . \end{aligned}$$

One might be tempted now to simply sneak the limit inside the integral, use continuity of f and obtain

$$F'(z) \stackrel{?}{=} \int_0^1 \lim_{\Delta z \to 0} f(z + t\Delta z)\, dt = \int_0^1 f(z)\, dt = f(z) \ ,$$

which would finish the proof. However sneaking the limit inside the integral is not always allowed since integration itself is a limiting process and limits cannot always be interchanged.

A simple example showing that the order in which one takes limits matters is the following. Consider the following limit

$$\lim_{\substack{n \to \infty \\ m \to \infty}} \frac{m+n}{m} \ .$$

We can take this limit in two ways. On the one hand,

$$\lim_{n \to \infty} \lim_{m \to \infty} \frac{m}{m+n} = \lim_{n \to \infty} 1 = 1 \ ;$$

yet on the other

$$\lim_{m \to \infty} \lim_{n \to \infty} \frac{m}{m+n} = \lim_{m \to \infty} 0 = 0 \ .$$

Nevertheless, as we sketch below, in this case interchanging the limits turns out to be a correct procedure due to the continuity of the integrand.

We want to prove here that indeed

$$\lim_{\Delta z \to 0} \int_0^1 f(z + t\Delta z)\, dt = f(z) \ .$$

We do this by showing that in this limit, the quantity

$$\int_0^1 f(z + t\Delta z)\, dt \quad - f(z) = \int_0^1 \left[f(z + t\Delta z) - f(z) \right] dt$$

goes to zero. We will prove that its modulus goes to zero, which is clearly equivalent. By equation (2.24), we have

$$\int_0^1 \left[f(z + t\Delta z) - f(z) \right] dt \ \leq \int_0^1 \left| f(z + t\Delta z) - f(z) \right| dt \ .$$

By continuity of f we know that given any $\varepsilon > 0$ there exists a $\delta > 0$ such that

$$|f(z + t\Delta z) - f(z)| < \varepsilon \qquad \text{whenever} \qquad |\Delta z| < \delta \ .$$

Since we are taking the limit $\Delta z \to 0$, we can take $|\Delta z| < \delta$, whence

$$\lim_{\Delta z \to 0} \int_0^1 [f(z + t\Delta z) - f(z)] \, dt \ \leq \lim_{\Delta z \to 0} \int_0^1 |f(z + t\Delta z) - f(z)| \, dt < \int_0^1 \varepsilon \, dt = \varepsilon \ ,$$

for any $\varepsilon > 0$, where we have used equation (2.24) to arrive at the last inequality. Hence,

$$\lim_{\Delta z \to 0} \int_0^1 [f(z + t\Delta z) - f(z)] \, dt \ = 0 \ ,$$

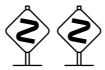
so that

$$\lim_{\Delta z \to 0} \int_0^1 [f(z + t\Delta z) - f(z)] \, dt = 0 \ .$$

2.2.4 Cauchy's Integral Theorem

We have now laid the groundwork to be able to discuss one of the key results in complex analysis. The path-independence lemma tells us that a continuous function $f : D \to \mathbb{C}$ in some domain D has an antiderivative if and only if all its closed contour integrals vanish. Unfortunately it is impractical to check this hypothesis explicitly, so one would like to be able to conclude the vanishing of the closed contour integrals some other way. Cauchy's integral theorem will tell us that, under some conditions, this is true if f is analytic. These conditions refer to the topology of the domain, so we have to first introduce a little bit of notation.

Let us say that a contour is **simple** if it has no self-intersections. We define a **loop** to be a closed simple contour. We start by mentioning the celebrated **Jordan curve lemma**, a version of which states that any loop in the complex plane separates the plane into two domains with the loop as common boundary: one of which is bounded and is called the **interior** and one of which is unbounded and is called the **exterior**.



This is a totally obvious statement and as most such statements extremely hard to prove, requiring techniques of algebraic topology.

We say that a domain D is **simply-connected** if the interior domain of every loop in D lies wholly in D. Hence for example, a disk is simply connected, while a punctured disk is not: any circle around the puncture contains the puncture in its interior, but this has been excised from the disk. Intuitively speaking, a domain is simply-connected if any loop in the domain can be continuously shrunk to a point without any point of the loop ever leaving the domain.

We are ready to state the **Cauchy Integral Theorem**: Let $D \subset \mathbb{C}$ be a *simply-connected* domain and let $f : D \to \mathbb{C}$ be an analytic function, then for any loop Γ, the contour integral vanishes:

$$\oint_\Gamma f(z)dz = 0 \ .$$

As an immediate corollary of this theorem and of the path-independence lemma, we see that an analytic function in a simply-connected domain has an antiderivative, which is itself analytic in D.

We will actually prove a slightly weaker version of the theorem which requires the stronger hypothesis that $f'(z)$ be continuous in D. Recall that analyticity only requires $f'(z)$ to exist. The proof uses a version of Green's theorem which is valid in the complex plane. This theorem states that if $\boldsymbol{V}(x,y) = P(x,y)\,dx + Q(x,y)\,dy$ is a continuously differentiable vector field in a simply-connected domain D in the complex plane, and if Γ is any positively oriented loop in D, then the line integral of $\boldsymbol{V}$ along Γ can be written as the area integral of the function $\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}$ on the interior $\mathrm{Int}(\Gamma)$ of Γ:

$$\oint_\Gamma \left(P(x,y)\,dx + Q(x,y)\,dy\right) = \iint\limits_{\mathrm{Int}(\Gamma)} \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) dx\,dy \ . \tag{2.31}$$

We will sketch a proof of this theorem below; but now let us use it to prove the Cauchy Integral Theorem. Let Γ be a loop in a simply-connected domain D in the complex plane, and let $f(z)$ be a function which is analytic in D. Computing the contour integral, we find

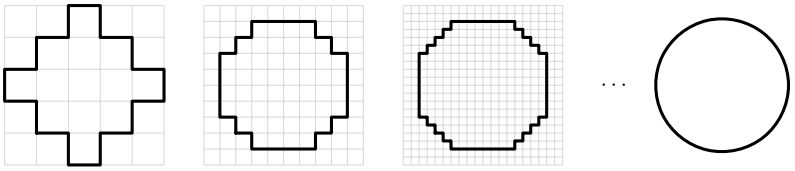
$$\begin{aligned}\oint_\Gamma f(z)\,dz &= \int_\Gamma \left(u(x,y) + i\,v(x,y)\right)\,(dx + i\,dy) \\ &= \int_\Gamma \left(u(x,y)\,dx - v(x,y)\,dy\right) + i\int_\Gamma \left(v(x,y)\,dx + u(x,y)\,dy\right) \ .\end{aligned}$$

By hypothesis, $f'(z)$ is continuous, which means that the vector fields $u\,dx - v\,dy$ and $v\,dx + u\,dy$ are continuously differentiable, whence we can use Green's Theorem (2.31) to deduce that

$$\oint_\Gamma f(z)\,dz = \iint\limits_{\mathrm{Int}(\Gamma)} \left(-\frac{\partial v}{\partial x} - \frac{\partial u}{\partial y}\right) dx\,dy + \iint\limits_{\mathrm{Int}(\Gamma)} \left(\frac{\partial u}{\partial x} - \frac{\partial v}{\partial y}\right) dx\,dy \ ,$$

which vanishes by the Cauchy–Riemann equations (2.7).

Here we will sketch a proof of Green's Theorem (2.31). The strategy will be the following. We will approximate the interior of the loop by tiny squares (*plaquettes*) in such a way that the loop itself is approximated by the straight line segments which make up the edges of the squares. As the size of the plaquettes decreases, the approximation becomes better and better. In the picture we have illustrated this by showing three approximations to the unit disk. For each we show the value of the length ℓ of the contour and of the area A of its interior.



$\ell = 9.6$
$A = 2.9952$

$\ell = 7.68$
$A = 2.9952$

$\ell = 7.68$
$A = 3.1104$

$\ell = 2\pi$
$A = \pi$

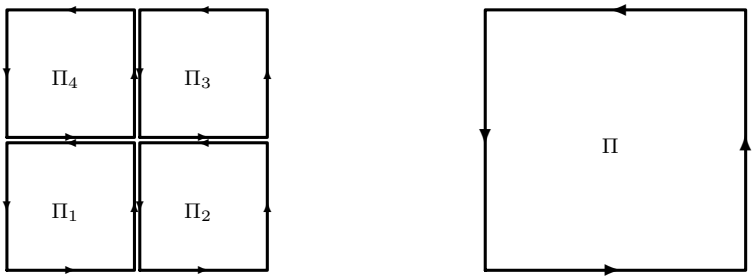
In fact, it is a simple matter of careful bookkeeping to prove that in the limit,

$$\iint_{\mathrm{Int}(\Gamma)} = \lim_{\mathrm{size}\to 0} \sum_{\text{plaquettes } \Pi} \iint_{\mathrm{Int}(\Pi)} .$$

Similarly for the contour integral,

$$\oint_{\Gamma} = \lim_{\mathrm{size}\to 0} \sum_{\text{plaquettes } \Pi} \oint_{\Pi} .$$

To see this notice that the contour integrals along internal edges common to two adjacent plaquettes cancel because of equation (2.29) and the fact that we integrated twice along them: once for each plaquette but in the opposite orientation, as shown in the picture below. Therefore we only receive contributions from the external edges. Since the region is simply-connected this means that boundary of the region covered by the plaquettes.

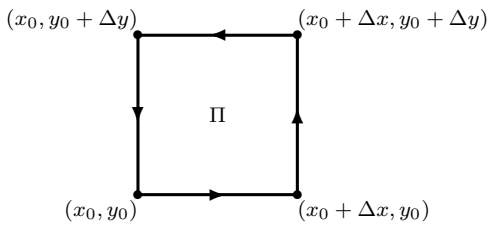


In the notation of the picture, then, one has

$$\oint_{\Pi_1} + \oint_{\Pi_2} + \oint_{\Pi_3} + \oint_{\Pi_4} = \oint_{\Pi} .$$

Therefore it is sufficient to prove formula (2.31) for the special case of one plaquette. To this effect we will choose our plaquette Π to have size $\Delta x \times \Delta y$ and whose lower left-hand

corner is at the point (x_0, y_0):



Performing the contour integral we have for $\boldsymbol{V}(x,y) = P(x,y)\,dx + Q(x,y)\,dy$,

$$\oint_\Pi \boldsymbol{V}(x,y) = \int_{(x_0,y_0)}^{(x_0+\Delta x,y_0)} \boldsymbol{V}(x,y) + \int_{(x_0+\Delta x,y_0)}^{(x_0+\Delta x,y_0+\Delta y)} \boldsymbol{V}(x,y) + \int_{(x_0+\Delta x,y_0+\Delta y)}^{(x_0,y_0+\Delta y)} \boldsymbol{V}(x,y) + \int_{(x_0,y_0+\Delta y)}^{(x_0,y_0)} \boldsymbol{V}(x,y)\ .$$

Along the first and third contour integrals the value of y is constant, whereas along the second and fourth integrals it is the value of x which is constant. Taking this into account, we can rewrite the integrals as follows

$$\oint_\Pi \boldsymbol{V}(x,y) = \int_{x_0}^{x_0+\Delta x} P(x,y_0)\,dx + \int_{y_0}^{y_0+\Delta y} Q(x_0+\Delta x,y)\,dy + \int_{x_0+\Delta x}^{x_0} P(x,y_0+\Delta y)\,dx + \int_{y_0+\Delta y}^{y_0} Q(x_0,y)\,dy\ .$$

Exchanging the limits of integration in the third and fourth integrals, and picking up a sign in each, we can rewrite the integrals as follows:

$$\oint_\Pi \boldsymbol{V}(x,y) = \int_{y_0}^{y_0+\Delta y} \left[Q(x_0+\Delta x,y) - Q(x_0,y)\right]\,dy - \int_{x_0}^{x_0+\Delta x} \left[P(x,y_0+\Delta y) - P(x,y_0)\right]\,dx\ .$$

But now we make use of the facts that

$$Q(x_0+\Delta x,y) - Q(x_0,y) = \int_{x_0}^{x_0+\Delta x} \frac{\partial Q(x,y)}{\partial x}\,dx$$

$$P(x,y_0+\Delta y) - P(x,y_0) = \int_{y_0}^{y_0+\Delta y} \frac{\partial P(x,y)}{\partial y}\,dy\ ;$$

whence the integrals become

$$\oint_\Pi \boldsymbol{V}(x,y) = \int_{y_0}^{y_0+\Delta y}\int_{x_0}^{x_0+\Delta x} \frac{\partial Q(x,y)}{\partial x}\,dx\,dy - \int_{x_0}^{x_0+\Delta x}\int_{y_0}^{y_0+\Delta y} \frac{\partial P(x,y)}{\partial y}\,dy\,dx$$

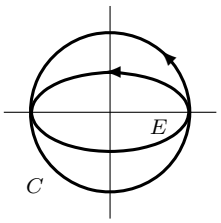
$$= \int_{x_0}^{x_0+\Delta x}\int_{y_0}^{y_0+\Delta y} \left(\frac{\partial Q(x,y)}{\partial x} - \frac{\partial P(x,y)}{\partial y}\right)\,dx\,dy$$

$$= \iint_{\mathrm{Int}(\Pi)} \left(\frac{\partial Q(x,y)}{\partial x} - \frac{\partial P(x,y)}{\partial y}\right)\,dx\,dy\ ,$$

which proves the formula for the plaquette Π.

Deforming the contour without changing the integral

The Cauchy Integral Theorem has a very important consequence for the computation of contour integrals. It basically says that contours can be moved about (or deformed) without changing the result of the integral, provided that in doing so we never cross a point where the integrand ceases to be analytic. Let us illustrate this with a few examples.



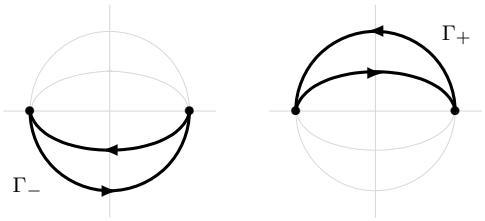
Let us compute the contour integral

$$\oint_E \frac{1}{z}\,dz \ ,$$

where E is the positively-oriented ellipse $x^2 + 4y^2 = 1$ depicted in the figure. Earlier we computed the same integral around a circular contour C of radius 1, centred at the origin, and we obtained

$$\oint_C \frac{1}{z}\,dz = 2\pi\, i \ .$$

We will argue, using the Cauchy Integral Theorem, that we get the same answer whether we integrate along E or along C. Consider the two domains in the interior of the circle C but in the exterior of the ellipse E. The integrand is analytic everywhere in the complex plane except for the origin, which lies outside these two regions. The Cauchy Integral Theorem says that the contour integral vanishes along either of the two contours which make up the boundary of these domains. Let us be more explicit and let us call these contours $\Gamma_\pm$ as in the figure below.



Then it is clear that

$$\oint_C \frac{1}{z}\,dz = \oint_{\Gamma_+} \frac{1}{z}\,dz + \oint_{\Gamma_-} \frac{1}{z}\,dz + \oint_E \frac{1}{z}\,dz \ .$$

Since the interior $\Gamma_\pm$ is simply-connected and the integrand $\frac{1}{z}$ is analytic in and on $\Gamma_\pm$, the Cauchy Integral Theorem says that

$$\oint_{\Gamma_\pm} \frac{1}{z}\,dz = 0 \ ,$$

whence

$$\oint_E \frac{1}{z}\,dz = \oint_C \frac{1}{z}\,dz = 2\pi\,i\ .$$

In other words, we could deform the contour from E to C without altering the result of the integral because in doing so we do not pass over any point where the integrand ceases to be analytic.

Let us illustrate this with another example, which generalises this one. Let Γ be any positively-oriented loop in the complex plane, let z_0 be any complex number which does *not* lie on Γ, and consider the following contour integral

$$\oint_\Gamma \frac{1}{z - z_0}\,dz\ .$$

We must distinguish two possibilities: z_0 is in the interior of Γ or in the exterior. In the latter case, the integral is zero because the integrand is analytic everywhere but at z_0, hence if z_0 lies outside Γ, Cauchy's Integral Theorem applies. On the other hand, if z_0 is in the interior of Γ we expect that we should obtain a nonzero answer—after all, if Γ were the circle $|z - z_0| = R > 0$, then the same calculation as in (2.26) yields a value of $2\pi\,i$ for the integral. In fact, as we will now show this is the answer we get for *any* positively-oriented loop containing z_0 in its interior.

In Figure 2.5 we have depicted the contour Γ and a circular contour C of radius r about the point z_0. We have also depicted two pairs of points (P_1, P_2) and (P_3, P_4): each pair having one point in each contour, as well as straight line segments joining the points in each pair.

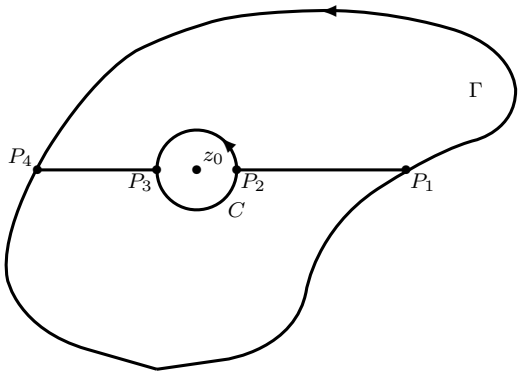


Figure 2.5: The contours Γ and C and some special points.

Now consider the following loop Γ_1 starting and ending at P_1, as illustrated in Figure 2.6. We start at P_1 and go to P_4 via the top half of Γ, call this, Γ_+; then we go to P_3 along the straight line segment joining them, call

it $-\gamma_{34}$; then to P_2 via the upper half of C in the negative sense, call it $-C_+$; and then back to P_1 via the straight line segment joining P_2 and P_1, call it $-\gamma_{12}$. The interior of this contour is simply-connected and does not contain the point z_0. Therefore Cauchy's Integral Theorem says that

$$\begin{aligned}\oint_{\Gamma_1} \frac{1}{z-z_0}\,dz &= \left(\int_{\Gamma_+} + \int_{-\gamma_{34}} + \int_{-C_+} + \int_{-\gamma_{12}}\right) \frac{1}{z-z_0}\,dz \\ &= \left(\int_{\Gamma_+} - \int_{\gamma_{34}} - \int_{C_+} - \int_{\gamma_{12}}\right) \frac{1}{z-z_0}\,dz \\ &= 0\ ,\end{aligned}$$

from where we deduce that

$$\int_{\Gamma_+} \frac{1}{z-z_0}\,dz = \left(\int_{\gamma_{34}} + \int_{C_+} + \int_{\gamma_{12}}\right) \frac{1}{z-z_0}\,dz\ .$$

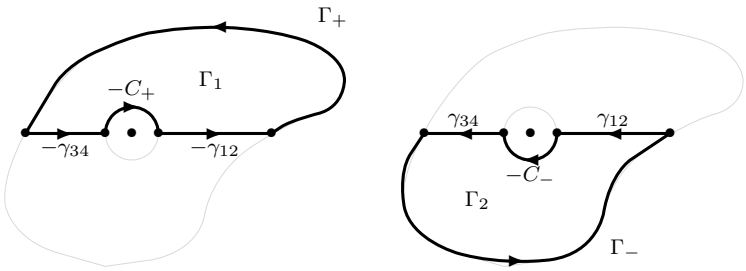


Figure 2.6: The contours Γ_1 and Γ_2.

Similarly consider the loop Γ_2 starting and ending at P_4. We start at P_4 and go to P_1 along the lower half of Γ, call it Γ_-; then we go to P_2 along γ_{12}; then to P_3 via the lower half of the circular contour in the negative sense $-C_-$; and then finally back to P_4 along γ_{34}. By the same argument as above, the interior of Γ_2 is simply-connected and z_0 lies in its exterior domain. Therefore by the Cauchy Integral Theorem,

$$\begin{aligned}\oint_{\Gamma_2} \frac{1}{z-z_0}\,dz &= \left(\int_{\Gamma_-} + \int_{\gamma_{34}} + \int_{-C_-} + \int_{-\gamma_{12}}\right) \frac{1}{z-z_0}\,dz \\ &= \left(\int_{\Gamma_-} + \int_{\gamma_{34}} - \int_{C_-} + \int_{-\gamma_{12}}\right) \frac{1}{z-z_0}\,dz \\ &= 0\ ,\end{aligned}$$

from where we deduce that

$$\int_{\Gamma_-} \frac{1}{z-z_0}\,dz = \left(-\int_{\gamma_{34}} + \int_{C_+} - \int_{\gamma_{12}}\right) \frac{1}{z-z_0}\,dz\ .$$

Putting the two results together, we find that

$$\begin{aligned}\int_\Gamma \frac{1}{z-z_0}\,dz &= \int_{\Gamma_+} \frac{1}{z-z_0}\,dz + \int_{\Gamma_-} \frac{1}{z-z_0}\,dz \\ &= \int_{C_+} \frac{1}{z-z_0}\,dz + \int_{C_-} \frac{1}{z-z_0}\,dz \\ &= \int_C \frac{1}{z-z_0}\,dz \\ &= 2\pi\, i\ .\end{aligned}$$

In summary, we find that if Γ is any positively-oriented loop in the complex plane and z_0 a point *not* in Γ, then

$$\int_\Gamma \frac{1}{z-z_0}\,dz = \begin{cases} 2\pi\, i & \text{for } z_0 \text{ in the interior of } \Gamma\text{; and} \\ 0 & \text{otherwise.}\end{cases} \tag{2.32}$$

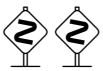
In the following section we will generalise this formula in a variety of ways.

2.2.5 Cauchy's Integral Formula

In this section we present several generalisations of the formula (2.32). Let $f(z)$ be analytic in a simply-connected domain D, and let Γ be a positively-oriented loop in D. Let z_0 be any point in the interior of Γ. Then the **Cauchy Integral Formula** reads

$$\boxed{f(z_0) = \frac{1}{2\pi\, i}\ \oint_\Gamma \frac{f(z)}{z-z_0}\,dz\ .} \tag{2.33}$$

This is a remarkable formula. It says that an analytic function in a simply-connected domain is determined by its behaviour on the boundary. In other words, if two analytic functions $f(z)$ and $g(z)$ agree on the boundary of a simply-connected domain they agree everywhere in the domain.



Cauchy's Integral Formula is a mathematical analogue of a notion that is very much in vogue in today's theoretical physics, namely 'holography'. You all know what the idea of an optical hologram is: it is a two-dimensional film which contains enough information to reconstruct (optically) a three-dimensional object. In theoretical physics, holography is exemplified in the celebrated formula of Beckenstein–Hawking for the entropy of a black hole. On the one hand, we know from Boltzmann's formula that the entropy of a statistical mechanical system is a measure of the density of states of the system. The black-hole entropy formula says that the entropy is a black hole is proportional to the area of the horizon. In simple terms, the horizon of the black hole is the surface within which light can no longer escape the gravitational attraction of the black hole. The entropy formula is holographic because it tells us that the degrees of freedom of a three-dimensional object like a black hole is determined from the properties of a two-dimensional system: the

horizon, just like with the optical hologram. The "Holographic Principle" roughly states that any theory of quantum gravity, i.e., a theory which can explain the microscopic origin of the entropy of the black hole, must be able to explain the entropy formula and hence be holographic. The Cauchy Integral Formula is holographic in the sense that an analytic function in the plane (which is two-dimensional) is determined by its behaviour on contours (which are one-dimensional).

Notice that by equation (2.32), we have that

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

whence we will have proven the Cauchy Integral Formula if we can show that

$$\oint_\Gamma \frac{f(z) - f(z_0)}{z - z_0}\, dz = 0 \ .$$

As a first step in proving this result, let us use the Cauchy Integral Theorem to conclude that the above integral can be computed along a small circle C_r of radius r about z_0 without changing its value:

$$\oint_\Gamma \frac{f(z) - f(z_0)}{z - z_0}\, dz = \oint_{C_r} \frac{f(z) - f(z_0)}{z - z_0}\, dz \ .$$

Moreover since the radius of the circle does not matter, we are free to take the limit in which the radius goes to zero, so that:

$$\oint_\Gamma \frac{f(z) - f(z_0)}{z - z_0}\, dz = \lim_{r\to 0} \oint_{C_r} \frac{f(z) - f(z_0)}{z - z_0}\, dz \ .$$

Let us parametrise C_r by $z(t) = z_0 + r\exp(2\pi\, i\, t)$ for $t \in [0, 1]$. Then

$$\begin{aligned} \oint_{C_r} \frac{f(z) - f(z_0)}{z - z_0}\, dz &= \int_0^1 \frac{f(z) - f(z_0)}{re^{2\pi\, i\, t}}\, 2\pi\, i\, re^{2\pi\, i\, t}\, dt \\ &= 2\pi\, i \int_0^1 \left(f(z) - f(z_0)\right)\, dt \ . \end{aligned}$$

Let us estimate the integral. Using (2.24) we find

$$\left| \int_0^1 \left(f(z) - f(z_0)\right)\, dt \right| \leq \int_0^1 |f(z) - f(z_0)|\, dt \leq \max_{|z-z_0|=r} |f(z) - f(z_0)| \ .$$

Because f is continuous at z_0—that is, $f(z) \to f(z_0)$ as $z \to z_0$—the limit as $r \to 0$ of $|f(z) - f(z_0)|$ is zero, whence

$$\lim_{r\to 0} \oint_{C_r} \frac{f(z) - f(z_0)}{z - z_0}\, dz = 0 \ .$$

Formally, continuity of f at z_0 says that given any $\varepsilon > 0$ there is a $\delta > 0$ such that $|f(z) - f(z_0)| < \varepsilon$ whenever $|z - z_0| < \delta$. Since we are interested in the limit $r \to 0$, we can always take δ small enough so that $|f(z) - f(z_0)|$ is smaller than any ε. Therefore, $\lim_{r\to 0} |f(z) - f(z_0)| = 0$.

Now let us do something "deep." We will change notation in the Cauchy Integral Formula (2.33) and rewrite it as

$$f(z) = \frac{1}{2\pi\, i} \oint_\Gamma \frac{f(\zeta)}{\zeta - z}\, d\zeta \ .$$

All we have done is change the name of the variable of integration (Shakespeare's Theorem again!); but as a result we have obtained an integral representation of an analytic function which suggests a way to take its derivative simply by sneaking the derivative inside the integral:

$$\begin{aligned} f'(z) &\stackrel{?}{=} \frac{1}{2\pi\, i} \oint_\Gamma \frac{f(\zeta)}{(\zeta - z)^2}\, d\zeta \\ f''(z) &\stackrel{?}{=} \frac{2}{2\pi\, i} \oint_\Gamma \frac{f(\zeta)}{(\zeta - z)^3}\, d\zeta \\ &\vdots \\ f^{(n)}(z) &\stackrel{?}{=} \frac{n!}{2\pi\, i} \oint_\Gamma \frac{f(\zeta)}{(\zeta - z)^{n+1}}\, d\zeta \ . \end{aligned}$$

Of course such manipulations have to be justified, and we will see that indeed this is correct. Given that we are going to spend the effort in justifying this procedure, let us at least get something more out of it.

Integral representation for analytic functions

We already have at our disposal quite a number of analytic functions: rational functions, exponential and related functions, logarithm and complex powers. To some extent these are complex versions of functions with which we are familiar from real calculus. In this section we will learn of yet another way of constructing analytic functions. Functions constructed in this way usually do not have names, since anonymity is the fate which befalls most functions. But by the same token, this means that the method below is a powerful way to construct new analytic functions, or to determine that a function is analytic.

Let g be a function which is continuous in some contour Γ which need *not* be closed. Let z be any complex number not contained in Γ, and define the following function:

$$G(z) = \int_\Gamma \frac{g(\zeta)}{\zeta - z}\, d\zeta \ . \tag{2.34}$$

We claim that $G(z)$ is analytic except possible on Γ, and

$$G'(z) = \int_\Gamma \frac{g(\zeta)}{(\zeta - z)^2}\, d\zeta \ . \tag{2.35}$$

This generalises the above discussion in two important ways: g need *not* be analytic (just continuous) and the contour need not be closed.

To see if $G(z)$ is analytic we need to investigate whether the derivative $G'(z)$ exists and is well-defined. By definition,

$$\begin{aligned} G'(z) &= \lim_{\Delta z \to 0} \frac{G(z+\Delta z) - G(z)}{\Delta z} \\ &= \lim_{\Delta z \to 0} \frac{1}{\Delta z} \int_\Gamma \left(\frac{g(\zeta)}{\zeta - z - \Delta z} - \frac{g(\zeta)}{\zeta - z} \right) d\zeta \\ &= \lim_{\Delta z \to 0} \frac{1}{\Delta z} \int_\Gamma \frac{g(\zeta)\Delta z}{(\zeta - z - \Delta z)(\zeta - z)}\, d\zeta \\ &= \lim_{\Delta z \to 0} \int_\Gamma \frac{g(\zeta)}{(\zeta - z - \Delta z)(\zeta - z)}\, d\zeta \ . \end{aligned}$$

Again, we would be done if we could simply take the limit inside the integral:

$$G'(z) \stackrel{?}{=} \int_\Gamma \lim_{\Delta z \to 0} \frac{g(\zeta)}{(\zeta - z - \Delta z)(\zeta - z)}\, d\zeta = \int_\Gamma \frac{g(\zeta)}{(\zeta - z)^2}\, d\zeta \ .$$

This can be justified (see below), so we are allowed to do so and recover what we were after. The formula (2.34) defines an **integral representation** for the analytic function $G(z)$.

Let us show that one can take the limit inside the integral, so that

$$\lim_{\Delta z \to 0} \int_\Gamma \frac{g(\zeta)}{(\zeta - z - \Delta z)(\zeta - z)}\, d\zeta = \int_\Gamma \frac{g(\zeta)}{(\zeta - z)^2}\, d\zeta \ .$$

Equivalently we would like to show that in the limit $\Delta z \to 0$, the difference

$$\int_\Gamma \ \frac{g(\zeta)}{(\zeta - z - \Delta z)(\zeta - z)} - \frac{g(\zeta)}{(\zeta - z)^2} \ \ d\zeta$$

vanishes. We can rewrite this difference as

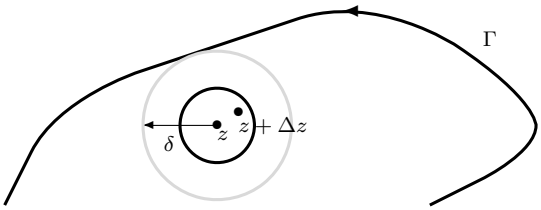
$$\Delta z \int_\Gamma \frac{g(\zeta)}{(\zeta - z - \Delta z)(\zeta - z)^2}\, d\zeta \ ,$$

which we would like to vanish as $\Delta z \to 0$. By equation (2.24), we have that

$$\begin{aligned} \int_\Gamma \frac{g(\zeta)}{(\zeta - z - \Delta z)(\zeta - z)^2}\, d\zeta \ &\le \int_\Gamma \ \frac{g(\zeta)}{(\zeta - z - \Delta z)(\zeta - z)^2} \ \ |d\zeta| \\ &= \max_{\zeta \in \Gamma} \ \frac{g(\zeta)}{(\zeta - z - \Delta z)(\zeta - z)^2} \ \ \ell(\Gamma) \ , \end{aligned}$$

where we have used equation (2.27) for the length $\ell(\Gamma)$ of the contour.

Since $g(\zeta)$ is continuous on Γ, $|g(\zeta)|$ is bounded there: $|g(\zeta)| \leq M$, for some positive real M.



Because z is not on Γ, any point ζ on Γ is at least a certain distance δ from z: $|\zeta - z| \geq \delta > 0$, as shown in the above figure. Now by the triangle inequality (2.1),

$$|\zeta - z| = |\zeta - z - \Delta z + \Delta z| \leq |\zeta - z - \Delta z| + |\Delta z| \ ,$$

whence

$$|\zeta - z - \Delta z| \geq |\zeta - z| - |\Delta z| \ .$$

Since we are taking the limit $\Delta z \to 0$, we can choose $|\Delta z| \leq \frac{1}{2}\delta$ so that

$$|\zeta - z - \Delta z| \geq \delta - \tfrac{1}{2}\delta = \tfrac{1}{2}\delta \ .$$

Therefore putting it all together we find that

$$\int_\Gamma \frac{g(\zeta)}{(\zeta - z - \Delta z)(\zeta - z)^2}\, d\zeta \ \leq \frac{2M\ell(\Gamma)}{\delta^3} \ .$$

Therefore

$$\lim_{\Delta z \to 0} \ \Delta z \int_\Gamma \frac{g(\zeta)}{(\zeta - z - \Delta z)(\zeta - z)^2}\, d\zeta \ \leq \lim_{\Delta z \to 0} |\Delta z| \frac{2M\ell(\Gamma)}{\delta^3} = 0 \ .$$

This is as good a place as any to mention another way of writing the triangle inequality (2.1), which is sometimes more useful and which was used above:

$$|z + w| \geq |z| - |w| \ . \tag{2.36}$$

To obtain the second version of the triangle inequality from the first we simply make the following substitution: $z_1 + z_2 = z$, and $z_2 = -w$, so that $z_1 = z + w$. Then we find from the (2.1), that $|z| \leq |z + w| + |-w| = |z + w| + |w|$, which is can be rewritten as (2.36).

The same argument shows that if we define

$$H(z) = \int_\Gamma \frac{g(\zeta)}{(\zeta - z)^n}\, d\zeta \ , \tag{2.37}$$

where n is a positive integer, then H is analytic and its derivative is given by

$$H'(z) = n \int_\Gamma \frac{g(\zeta)}{(\zeta - z)^{n+1}}\, d\zeta \ . \tag{2.38}$$

The generalised Cauchy Integral Formula

This has as an important consequence: if f is analytic in a neighbourhood of z_0, then *so are all its derivatives* $f^{(n)}$. To prove this simply notice that if f is analytic in a neighbourhood of z_0, there is some $\varepsilon > 0$ such that f is analytic in and on the circle C of radius ε centred at z_0; that is, the closed disk $|\zeta - z_0| \leq \varepsilon$. Therefore for any z in the interior of the circle—that is, such that $|z - z_0| < \varepsilon$—we have the Cauchy Integral representation

$$f(z) = \frac{1}{2\pi\, i} \oint_C \frac{f(\zeta)}{\zeta - z}\, d\zeta \ .$$

But this integral representation is of the form (2.34), whence its derivative is given by the analogue of equation (2.35):

$$f'(z) = \frac{1}{2\pi\, i} \oint_C \frac{f(\zeta)}{(\zeta - z)^2}\, d\zeta \ .$$

But this is of the general form (2.37) (with $n = 2$), whence by the above results, $f'(z)$ is an analytic function and its derivative is given by the analogue of (2.38):

$$f''(z) = \frac{2}{2\pi\, i} \oint_C \frac{f(\zeta)}{(\zeta - z)^3}\, d\zeta \ ,$$

which again follows the pattern (2.37). Continuing in this fashion we deduce that f', f'', ... are analytic in the open ε-disk about z_0.

In summary, *an analytic function is infinitely differentiable*, its derivatives being given by the **generalised Cauchy Integral Formula**:

$$\boxed{f^{(n)}(z) = \frac{n!}{2\pi\, i} \oint_\Gamma \frac{f(\zeta)}{(\zeta - z)^{n+1}}\, d\zeta \ .} \tag{2.39}$$

Notice that if we put $n = 0$ in this formula, define $0! = 1$ and understand the zeroth derivative $f^{(0)}$ as the function f itself, then this is precisely the Cauchy Integral Formula.

Infinite differentiability of harmonic functions.

The generalised Cauchy Integral Formula can also be turned around in order to compute contour integrals. Hence if f is analytic in and on a positively oriented loop Γ, and if z_0 is a point in the interior of Γ, then

$$\boxed{\oint_\Gamma \frac{f(z)}{(z - z_0)^{n+1}}\, dz = \frac{2\pi\, i}{n!}\, f^{(n)}(z_0) \ .} \tag{2.40}$$

For example, let us compute the following contour integral

$$\oint_\Gamma \frac{e^{5z}}{z^3}\,dz\ ,$$

where Γ is the positively oriented unit circle $|z| = 1$. This integral is of the form (2.40) with $n = 2$, $f(z) = e^{5z}$, which is entire and hence, certainly analytic in and on the contour, and with $z_0 = 0$, which lies in the interior of the contour. Therefore by (2.40) we have

$$\oint_\Gamma \frac{e^{5z}}{z^3}\,dz = 2\pi\,i\,\frac{1}{2!}\,\frac{d^2}{dz^2}\left(e^{5z}\right)\Big|_{z=0} = 2\pi\,i\,\frac{1}{2!}\,25 = 25\pi\,i\ .$$

Let us consider a more complicated example. Let us compute the contour integral

$$\int_\Gamma \frac{2z+1}{z(z-1)^2}\,dz\ ,$$

where Γ is the contour depicted in Figure 2.7. Two things prevent us from applying the generalised Cauchy Integral Formula: the contour is not a loop—indeed it is not simple—and the integrand is not of the form $g(z)/(z-z_0)^n$ where $g(z)$ is analytic inside the contour. This last problem could be solved by rewriting the integrand using partial fractions:

$$\frac{2z+1}{z(z-1)^2} = \frac{3}{(z-1)^2} - \frac{1}{z-1} + \frac{1}{z}\ . \tag{2.41}$$

However we are still faced with a contour which is not simple.

Figure 2.7: The contour Γ and an equivalent pair of contours $\{\Gamma_0, \Gamma_1\}$.

This problem can be circumvented by noticing that the smooth contour Γ can be written as a piecewise smooth contour with two smooth components: both starting and ending at the point of self-intersection of Γ. The first such contour is the left lobe of Γ, which is a negatively oriented loop about $z = 0$, and the second is the right lobe of Γ, which is a positively oriented loop about $z = 1$. Because the integrand is analytic everywhere but at $z = 0$ and $z = 1$, the Cauchy Integral Theorem tells us that we get the same result by

integrating around the circular contours Γ_0 and Γ_1 in Figure 2.7. In other words,

$$\int_\Gamma \frac{2z+1}{z(z-1)^2}\,dz = \oint_{\Gamma_0} \frac{2z+1}{z(z-1)^2}\,dz + \oint_{\Gamma_1} \frac{2z+1}{z(z-1)^2}\,dz\ .$$

We can now evaluate this in either of two ways. Using the partial fraction decomposition (2.41) of the integrand, one finds

$$\oint_{\Gamma_0} \frac{2z+1}{z(z-1)^2}\,dz = \oint_{\Gamma_0} \frac{1}{z}\,dz = -\oint_{-\Gamma_0} \frac{1}{z}\,dz = -2\pi\,i\ ,$$
$$\oint_{\Gamma_1} \frac{2z+1}{z(z-1)^2}\,dz = \oint_{\Gamma_1} \frac{3}{(z-1)^2}\,dz - \oint_{\Gamma_1} \frac{1}{z-1}\,dz = 0 - 2\pi\,i = -2\pi\,i\ ;$$

whence

$$\int_\Gamma \frac{2z+1}{z(z-1)^2}\,dz = -4\pi\,i\ .$$

Alternatively we notice that

$$\oint_{\Gamma_0} \frac{2z+1}{z(z-1)^2}\,dz = \oint_{\Gamma_0} \frac{\frac{2z+1}{(z-1)^2}}{z}\,dz = -2\pi\,i\ ,$$

where we have used the fact that $\frac{2z+1}{(z-1)^2}$ is analytic in and on Γ_0 and the Cauchy Integral Formula after taking into account that Γ_0 is negatively oriented. Similarly, one has

$$\oint_{\Gamma_1} \frac{2z+1}{z(z-1)^2}\,dz = \oint_{\Gamma_1} \frac{\frac{2z+1}{z}}{(z-1)^2}\,dz = 2\pi\,i\ \frac{d}{dz}\left(\frac{2z+1}{z}\right)\bigg|_{z=1} = -2\pi\,i\ ,$$

where we have used that $\frac{2z+1}{z}$ is analytic in and on Γ_1, and the generalised Cauchy Integral formula (with $n=1$). Therefore again

$$\int_\Gamma \frac{2z+1}{z(z-1)^2}\,dz = -4\pi\,i\ .$$

Morera's Theorem

Finally we discuss a converse of the Cauchy Integral Theorem, known as Morera's Theorem. Suppose that f is continuous in a domain D and has an antiderivative F in D. This means that F is analytic, and by what we have just shown, so is $f(z) = F'(z)$. Therefore we have just shown that if $f(z)$ is continuous with an antiderivative, then f is analytic. Now from the

path independence lemma, f has an antiderivative if and only if all its loop integrals in D vanish:

$$\oint_\Gamma f(z)dz = 0 \ .$$

Therefore we arrive at **Morera's Theorem** which states that: if $f(z)$ is continuous in D and all the loop integrals of $f(z)$ in D vanish, then f is analytic. This theorem will be of use in Section 2.3.

2.2.6 Liouville's Theorem and its applications

The generalised Cauchy Integral Formula is one of the cornerstones of complex analysis, as it has a number of very useful corollaries. An immediate application of the generalised Cauchy Integral Formula is the so-called *Cauchy estimates* for the derivatives of an analytic function. These estimates will play an important role in the remainder of this section.

Suppose that $f(z)$ is analytic in some domain D containing a circle C of radius R centred about z_0. Suppose moreover that $|f(z)| \leq M$ for all z on the circle C. We can then use the generalised Cauchy Integral Formula (2.39) to obtain a bound for the derivatives of f at z_0:

$$|f^{(n)}(z_0)| = \left| \frac{n!}{2\pi\, i} \oint_C \frac{f(z)}{(z-z_0)^{n+1}}\, dz \right| \leq \frac{n!}{2\pi} \oint_C \frac{|f(z)|}{|z-z_0|^{n+1}}\, |dz| \ ,$$

where we have used (2.28) to arrive at the inequality. On the circle, $|z-z_0| = R$ and $|f(z)| \leq M$, whence

$$|f^{(n)}(z_0)| \leq \frac{n!}{2\pi} \frac{M}{R^{n+1}} \oint_C |dz| \ ,$$

which, using that the length of the contour is $2\pi\, R$, can be rewritten neatly as

$$\boxed{|f^{(n)}(z_0)| \leq \frac{n!\, M}{R^n} \ .} \tag{2.42}$$

This inequality is known as the **Cauchy estimate**.

As an immediate corollary of this estimate suppose that f is analytic in whole complex plane (i.e., that f is an entire function) and that it is bounded, so that $|f(z)| \leq M$ for all z. Then from the Cauchy estimate, at *any* point z_0 in the complex plane, its derivative is bounded by $|f'(z_0)| \leq M/R$. But because the function is entire, we can take R as large as we wish. Now given any number $\varepsilon > 0$, however small, there is always an R large enough for which $M/R < \varepsilon$, so that $|f'(z_0)| < \varepsilon$. Therefore $|f'(z_0)| = 0$, whence

$f'(z_0) = 0$. Since this is true for all z_0 in the complex plane, we have proven **Liouville's theorem**:

a bounded entire function is constant.

This does not violate our experience since the only entire functions we have met are polynomials and the exponential and functions we can make out of them by multiplication, linear combinations and compositions, and these functions are all clearly not bounded.

Indeed, suppose that $P(z)$ is a polynomial of order N; that is,

$$P(z) = z^N + a_{N-1}z^{N-1} + \cdots + a_1 z + a_0 \ .$$

Then intuitively, for large z we expect that $P(z)$ should go as z^N, since the largest power dominates the other ones. The precise statement, to be proven below, is that there exists $R > 0$ large enough such that for $|z| \geq R$, $|P(z)| \geq c|z|^N$, where $0 < c < 1$ depends on R in such a way that as R tends to ∞, c tends to 1.

Let $P(z)$ be the above polynomial and let $A \geq 1$ denote the largest of the moduli of coefficients of the polynomial: $A = \max\{|a_0|, |a_1|, \ldots, |a_{N-1}|, 1\}$. Then let us rewrite the polynomial as $P(z) = z^N \ 1 + a_{N-1}/z + \cdots a_0/z^N$. Now by the triangle inequality (2.36),

$$1 + \frac{a_{N-1}}{z} + \cdots + \frac{a_1}{z^{N-1}} + \frac{a_0}{z^N} \ \geq 1 - \ \frac{a_{N-1}}{z} + \cdots + \frac{a_1}{z^{N-1}} + \frac{a_0}{z^N} \quad .$$

Using the triangle inequality again,

$$\begin{aligned} \frac{a_{N-1}}{z} + \cdots + \frac{a_1}{z^{N-1}} + \frac{a_0}{z^N} \ &\leq \ \frac{a_{N-1}}{z} \ + \cdots + \ \frac{a_1}{z^{N-1}} \ + \ \frac{a_0}{z^N} \\ &\leq \frac{A}{|z|} + \cdots + \frac{A}{|z|^{N-1}} + \frac{A}{|z|^N} \ . \end{aligned}$$

Now take $|z| \geq 1$ so that $|z|^N \geq |z|^{N-1} \geq \cdots \geq |z|$. Then,

$$\frac{a_{N-1}}{z} + \cdots + \frac{a_1}{z^{N-1}} + \frac{a_0}{z^N} \ \leq \frac{NA}{|z|} \ .$$

Therefore,

$$1 + \frac{a_{N-1}}{z} + \cdots + \frac{a_1}{z^{N-1}} + \frac{a_0}{z^N} \ \geq 1 - \frac{NA}{|z|} \ .$$

Hence if we take z such that $|z| \geq R \geq NA \geq 1$, then

$$1 + \frac{a_{N-1}}{z} + \cdots + \frac{a_1}{z^{N-1}} + \frac{a_0}{z^N} \ \geq 1 - \frac{NA}{R} = \frac{R - NA}{R} \ .$$

Finally then,

$$|P(z)| = |z|^N \ 1 + \frac{a_{N-1}}{z} + \cdots + \frac{a_1}{z^{N-1}} + \frac{a_0}{z^N} \ \geq \frac{R - NA}{R}|z|^N \ .$$

Hence $c = (R - NA)/R < 1$ and as $R \to \infty$, $c \to 1$.

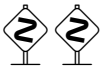
We are now able to prove the **Fundamental Theorem of Algebra** which states that

every nonconstant polynomial has at least one zero.

Indeed, let $P(z)$ be a polynomial and suppose that it does not have any zeros. Then $1/P(z)$ is an entire function. If we manage to prove that this function is bounded, then we can use Liouville's theorem and conclude that $1/P(z)$, and hence $P(z)$, would have to be constant. So let us try to prove that it is bounded. Without loss of generality we can assume that the polynomial has the form $P(z) = z^N + a_{N-1}z^{N-1} + \cdots + a_1 z + a_0$ for some N. Let R be such that $|z| \geq R$, $|P(z)| \geq c|z|^N$, where $0 < c < 1$. Then, for $|z| \geq R$,

$$\left|\frac{1}{P(z)}\right| = \frac{1}{|P(z)|} \leq \frac{1}{c|z|^N} \leq \frac{1}{cR^N} \ .$$

While for $|z| \leq R$, then the function $1/P(z)$, being continuous, is bounded in this disk by some $M = \max_{|z|\leq R} 1/|P(z)|$. Therefore $1/|P(z)|$ is bounded above for all z by the largest of M and $1/(cR^N)$. Hence $1/P(z)$ is bounded.



I have always found this proof of the Fundamental Theorem of Algebra quite remarkable. It is compelling evidence in favour of the vision of mathematics as a coherent whole, that a purely algebraic statement like the Fundamental Theorem of Algebra can be proven in a relatively elementary fashion using complex analysis. I hope that as physicists we can be forgiven the vanity of thinking that this unity of mathematics stems from it being the language of nature.

2.3 Series expansions for analytic functions

This section ushers in the second half of this part of the course. The purpose of this section is to learn about the series representations for analytic functions. We will see that every function analytic in a disk can be approximated by polynomials: the partial sums of its Taylor series. Similarly every function analytic in a punctured disk can be described by a Laurent series, a generalisation of the notion of a power series, where we also allow for negative powers. This will allow us to discuss the different types of singularities that an analytic function can have. This section is organised as follows: we start with a study of sequences and series of complex numbers and of complex functions and of different notions of convergence and methods of establishing convergence. We will then show that a function analytic in the neighbourhood of a point can be approximated there by a power series: its Taylor series. We will then discuss power series and prove that every power series converges to an analytic function in its domain of convergence, and in fact is

the Taylor series of that function. Therefore the power series representation of an analytic function is unique. We then introduce Laurent series: which allows us to represent analytic functions around an isolated singularity. We also prove that they are unique in a sense. We end the section with a discussion of the different isolated singularities which an analytic function can have.

2.3.1 Sequences and Series

In this section we discuss sequences and series and the rudiments of the theory of convergence. This is necessary groundwork to be able to discuss the Taylor and Laurent series representations for analytic functions.

Sequences

By a **sequence** we mean an infinite set $\{z_0, z_1, z_2, z_3, \dots\}$ of complex numbers. It is often denoted $\{z_n\}$ where the index is understood to run over the non-negative integers. Intuitively, a sequence $\{z_n\}$ converges to a complex number z if as n increases, z_n remains ever closer to z. A precise definition is the following. A sequence $\{z_n\}$ is said to **converge** to z (written $z_n \to z$ or $\lim_{n\to\infty} z_n = z$) if given any $\varepsilon > 0$, there exists an integer N, which may depend on ε, such that for all $n \geq N$, $|z_n - z| < \varepsilon$. In other words, the "tail" of the sequence remains arbitrary close to z provided we go sufficiently far into it. A sequence which converges to some point is said to be **convergent**. Convergence is clearly a property only of the tail of the sequence, in the sense that two sequences which differ only in the first N terms (any finite N) but are identical afterwards will have the same convergence properties.

For example, the sequence $\{z_n = 1/n\}$ clearly converges to 0: $|z_n| = 1/n$ and we can make this as small as we like by taking n as large as needed.

A sequence $\{z_n\}$ is said to satisfy the **Cauchy criterion** (or be a **Cauchy sequence**) if it satisfies the following property: given any $\varepsilon > 0$ there exists N (again, depending on ε) such that $|z_n - z_m| < \varepsilon$ for all $n, m \geq N$. This criterion simply requires that the elements in the sequence remain ever closer to each other, not that they should converge to any point. Clearly, if a sequences converges it is Cauchy: simply notice that adding and subtracting z,

$$|z_n - z_m| = |(z_n - z) - (z_m - z)| \leq |z_n - z| + |z_m - z|$$

by the triangle inequality (2.1). Hence if we want z_n and z_m to remain within ε of each other for n, m larger than some N, we need just choose N such that $|z_n - z| < \varepsilon/2$ for all $n \geq N$.

What is a relatively deep result, is that every Cauchy sequence is convergent. This is essentially the fact that the complex numbers are **complete**. To prove this requires a more careful axiomatisation of the real number system than we have time for.

Series

By a **series** we mean a formal sum

$$c_0 + c_1 + c_2 + \cdots + c_j + \cdots$$

of complex numbers, c_j, called the **coefficients**. We say *formal* since just because we can write something down does not mean it makes any sense: it does not make much sense to add an infinite number of terms. What does make sense is the following: define the n-**th partial sum**

$$S_n \equiv \sum_{j=0}^{n} c_j = c_0 + c_1 + \cdots + c_{n-1} + c_n \ .$$

This defines a sequence $\{S_n\}$. Then we can analyse the limit as $n \to \infty$ of this sequence. If one exists, say $S_n \to S$, then we say that the series **converges** to or **sums** to S, and we write

$$S = \sum_{j=0}^{\infty} c_j \ .$$

Otherwise we say that the series is **divergent**. Applying the Cauchy criterion to the sequence of partial sums, we see that a necessary condition for the convergence of a series is that the sequence of coefficients converge to 0. Indeed, if $\{S_n\}$ is convergent, it is Cauchy, whence given any $\varepsilon > 0$, there exists N such that for all $n, m \geq N$, $|S_n - S_m| < \varepsilon$. Taking $m = n - 1$, we see that

$$\left| \sum_{j=0}^{n} c_j - \sum_{j=0}^{n-1} c_j \right| = |c_n| < \varepsilon \ ,$$

for every $n \geq N$. Therefore the sequence $\{c_j\}$ converges to 0. We can summarise this as follows

$$\boxed{\text{If } \sum_{j=0}^{\infty} c_j \text{ converges, then } \lim_{j\to\infty} c_j = 0 \ .}$$

This is a necessary criterion for the convergence of a series, so it can be used to conclude that a series is divergent, but not to conclude that it is

convergent. For example, consider the series

$$\sum_{j=0}^{\infty} \frac{j}{2j+1} \ . \tag{2.43}$$

It is clearly divergent because $j/(2j+1) \to \frac{1}{2}$. On the other hand consider the series (we start at $j=1$ for obvious reasons)

$$\sum_{j=1}^{\infty} \frac{1}{j} \ . \tag{2.44}$$

Now the coefficients do converge to zero, but this series is actually divergent. One way to see this is to notice that for every $n \geq 1$,

$$\sum_{j=1}^{n} \frac{1}{j} = \sum_{j=1}^{n} \int_{j}^{j+1} \frac{dx}{j} > \sum_{j=1}^{n} \int_{j}^{j+1} \frac{dx}{x} = \int_{1}^{n+1} \frac{dx}{x} = \log(n+1) \ ,$$

and $\lim_{n\to\infty} \log(n+1) = \infty$. On the other hand, the series

$$\sum_{j=1}^{\infty} \frac{1}{j^2}$$

does converge. One can argue in a similar style. Notice that for $j \geq 2$,

$$\frac{1}{j^2} = \int_{j-1}^{j} \frac{dx}{j^2} < \int_{j-1}^{j} \frac{dx}{x^2} = \frac{1}{j(j-1)} \ .$$

Hence, for all $n \geq 2$,

$$\sum_{j=1}^{n} \frac{1}{j^2} = 1 + \sum_{j=2}^{n} \frac{1}{j^2} < 1 + \sum_{j=2}^{n} \int_{j-1}^{j} \frac{dx}{x^2} = 1 + \int_{1}^{n} \frac{dx}{x^2} = 2 - \frac{1}{n} \ ,$$

so that in the limit,

$$\sum_{j=1}^{\infty} \frac{1}{j^2} < 2 \ .$$

Indeed, we will be able to compute this sum very easily using contour integration and it will turn out that $\sum_{j=1}^{\infty} \frac{1}{j^2} = \frac{\pi^2}{6} \simeq 1.6449341$. Similarly, one can show in the same way that the series

$$\sum_{j=1}^{\infty} \frac{1}{j^p}$$

converges for any $p > 1$. In fact, p can be any real number.

Establishing convergence

There are two useful tests for establishing the convergence of a series. The first one is known as the **Comparison Test**: Suppose that $\sum_{j=0}^{\infty} M_j$ is a convergent series whose coefficients are non-negative real numbers: $M_j \geq 0$. Let $\sum_{j=0}^{\infty} c_j$ be such that $|c_j| \leq M_j$ for all sufficiently large j. Then $\sum_{j=0}^{\infty} c_j$ also converges.

Prove the Comparison Test.

Of course, in order to apply this test we need to have some examples of convergent series to compare with. We have already seen the series $\sum_{j=1}^{\infty} 1/j^p$, for $p > 1$, but perhaps the most useful series we will come across is the **geometric series** $\sum_{j=0}^{\infty} c^j$, where c is some complex number. To investigate the convergence of this series, simply notice that $|c^j| = |c|^j$ and hence the coefficient sequence $\{c^j\}$ converges to 0 if and only if $|c| < 1$. Thus we let $|c| < 1$ from now on. We proceed as follows:

$$(1-c)S_n = (1-c)(1 + c + \cdots + c^n) = 1 - c^{n+1} \ ,$$

whence

$$S_n = \frac{1-c^{n+1}}{1-c} \qquad \text{or} \qquad S_n - \frac{1}{1-c} = -\frac{c^{n+1}}{1-c} \ .$$

Therefore taking the modulus, we see that

$$\left| S_n - \frac{1}{1-c} \right| = \frac{|c|^{n+1}}{|1-c|} \ ,$$

which converges to 0 as $n \to \infty$ since $|c| < 1$. Therefore

$$\boxed{\sum_{j=0}^{\infty} c^j = \frac{1}{1-c} \qquad \text{if } |c| < 1.} \tag{2.45}$$

As an example, let us consider the following series

$$\sum_{j=0}^{\infty} \frac{3+2i}{(j+1)^j} \ . \tag{2.46}$$

Its coefficient sequence converges to zero. Notice also that

$$\left| \frac{3+2i}{(j+1)^j} \right| = \frac{\sqrt{13}}{(j+1)^j} < \frac{4}{(j+1)^j} \ .$$

Hence for $j \geq 3$,

$$\left| \frac{3+2i}{(j+1)^j} \right| < \frac{1}{2^j} \ .$$

But since $\frac{1}{2} < 1$, the geometric series

$$\sum_{j=0}^{\infty} \frac{1}{2^j} = 2$$

converges. Hence by the comparison test, the original series (2.46) converges as well.

A further convergence criterion is the **Ratio Test**: Let $\sum_{j=0}^{\infty} c_j$ be such that the limit

$$L \equiv \lim_{j \to \infty} \left| \frac{c_{j+1}}{c_j} \right|$$

exists. Then if $L < 1$ the series converges, and if $L > 1$ the series diverges. (Alas, if $L = 1$ we cannot conclude anything.)

Prove the Ratio Test.

The Ratio Test does not contradict our experience so far: for the geometric series $L = |c|$, and we certainly needed $|c| < 1$ for convergence. Moreover in this case $L \geq 1$ implies divergence. Similarly, the series (2.44) has $L = 1$, so that the test tells us nothing. The same goes for the series (2.43). Notice that there are series for which the Ratio Test cannot even be applied, since the limit L may not exist.

Sequences and series of functions: uniform convergence

Our primary interest in series and sequences being the construction of analytic functions, let us now turn our attention to the important case of sequences and series of *functions*. Consider a sequence $\{f_n\}$ whose elements are functions $f_n(z)$ defined on some domain in the complex plane. For a fixed point z we can study the sequence of complex numbers $\{f_n(z)\}$ and analyse its convergence. If it does converge, let us call the limit $f(z)$; that is, $f_n(z) \to f(z)$. This procedure defines a function f for those z such that the sequence $\{f_n(z)\}$ converges. If this is the case we say that the sequence $\{f_n\}$ converges **pointwise** to f. Now suppose that each f_n is continuous (or analytic) will f be continuous (or analytic)? It turns out that pointwise convergence is too weak in order to guarantee that the limit function shares some of these properties of the f_n.

For instance, it is easy to cook up a pointwise limit of analytic functions which is not even continuous. Consider the functions $f_n(z) = \exp(-nz^2)$. Clearly these functions are analytic for each n. Let us now consider the functions restricted to the real axis: $z = x$, and consider the limit function $f(x)$. For all n, $f_n(0) = 1$, whence in the limit $f(0) = 1$. On the other hand, let $x \neq 0$. Then given any $\varepsilon > 0$, however small, there will be N such that $\exp(-nx^2) < \varepsilon$ for $n \geq N$. Hence

$$f(x) = \begin{cases} 1 & \text{for } x = 0; \\ 0 & \text{otherwise.} \end{cases}$$

In other words, the limit function has a discontinuity at the origin. Continuity would require $f(0) = 0$. To understand what is going on here, notice that to make $f_n(x) < \varepsilon$ we require that

$$e^{-nx^2} < \varepsilon \qquad \Longrightarrow \qquad n > \frac{\log(1/\varepsilon)}{x^2} \ ,$$

as can be easily seen by taking the logarithm of both sides of the first inequality. Hence as x becomes smaller, the value of n has to be larger and larger to the extent that in the limit as $x \to 0$, there is no n for which this is the case.

The above "post mortem" analysis prompts the following definition. A sequence of functions $\{f_n\}$ is said to **converge to a function f uniformly in a subset** U if given any $\varepsilon > 0$ there exists an N such that for all $n \geq N$,

$$|f_n(z) - f(z)| < \varepsilon \qquad \text{for } \textit{all } z \in U.$$

In other words, N can depend on ε but *not* on z.

Similarly one says that a series of functions

$$\sum_{j=0}^{\infty} f_j(z) \ ,$$

converges pointwise or uniformly if the sequence of partial sums does.

To show that this definition takes care of the kind of pathologies encountered above, let us first of all prove that the uniform limit of continuous functions is again continuous. Indeed, let $\{f_n(z)\}$ be a sequence of functions which are continuous at z_0, and let it converge to a function $f(z)$ uniformly in a neighbourhood of z_0. We claim that $f(z)$ is continuous at z_0. This means that given any $\varepsilon > 0$, there exists $\delta > 0$ such that $|f(z) - f(z_0)| < \varepsilon$

whenever $|z - z_0| < \delta$. To prove this we will employ a device known as the $\varepsilon/3$ *trick*. Let us rewrite $|f(z) - f(z_0)|$ as follows

$$\begin{aligned} |f(z) - f(z_0)| &= |f(z) - f_n(z) + f_n(z) - f_n(z_0) + f_n(z_0) - f(z_0)| \\ &\leq |f(z) - f_n(z)| + |f_n(z) - f_n(z_0)| + |f_n(z_0) - f(z_0)| \ , \end{aligned}$$

by the triangle inequality. Now, because $f_n(z) \to f(z)$ uniformly, we can choose n above so large that $|f(z) - f_n(z)| < \varepsilon/3$ for all z, so in particular for $z = z_0$. Similarly, because $f_n(z)$ is continuous at z_0, there exists δ such that $|f_n(z) - f_n(z_0)| < \varepsilon/3$ whenever $|z - z_0| < \delta$. Therefore,

$$|f(z) - f(z_0)| < \varepsilon/3 + \varepsilon/3 + \varepsilon/3 = \varepsilon \ .$$

In other words, we have shown that

the uniform limit of continuous functions is continuous.

Similarly we will see that the uniform limit of analytic functions is analytic. Uniform convergence is sufficiently strong to allow us to manipulate sequences of functions naively and yet sufficiently weak to allow for many examples. For instance we will see that if a series converges uniformly to a function, then the series can be differentiated and integrated termwise and it will converge to the derivative or integral of the limit function.

In practice, the way one checks that a sequence $\{f_n\}$ of functions converges uniformly in U to a function f is to write

$$f_n(z) = f(z) + R_n(z)$$

and then to see whether the remainder $R_n(z)$ can be made arbitrarily small for some large enough n independently of z in U. Let us see this for the geometric series:

$$\sum_{j=0}^{\infty} z^j \ . \tag{2.47}$$

The partial sums are the functions

$$f_n(z) = \sum_{j=0}^{n} z^j = 1 + z + z^2 + \cdots + z^n = \frac{1 - z^{n+1}}{1 - z} \ .$$

We claim that this geometric series converges uniformly to the function $1/(1-z)$ on every closed disk $|z| \leq R$ with $R < 1$. Indeed, we have the following estimate for the remainder:

$$\left| f_n(z) - \frac{1}{1-z} \right| = \frac{|z|^{n+1}}{|1-z|} \leq \frac{R^{n+1}}{|1-z|} \ .$$

Now, using the triangle inequality (2.36),

$$|z-1| = |1-z| \geq 1 - |z| \qquad \text{whence} \qquad \frac{1}{|1-z|} \leq \frac{1}{1-|z|} \leq \frac{1}{1-R} \ .$$

In other words,

$$\left| f_n(z) - \frac{1}{1-z} \right| = \frac{|z|^{n+1}}{|1-z|} \leq \frac{R^{n+1}}{1-R} \ .$$

This bound is independent of z and can be made as small as desired since $R < 1$, whence the convergence is uniform.

Another way to check for uniform convergence is the **Weierstrass M-test**, which generalises the Comparison Test. Suppose that $\sum_{j=0}^{\infty} M_j$ is a convergent series with real non-negative terms $M_j \geq 0$. Suppose further that for all z in some subset U of the complex plane and for all sufficiently large j, $|f_j(z)| \leq M_j$. Then the series $\sum_{j=0}^{\infty} f_j(z)$ converges uniformly in U. (Notice that the Comparison Test is obtained as a special case, when $f_j(z)$ are constant functions.)

Proof of the Weierstrass M-test.

Using the Weierstrass M-test we can prove the uniform convergence of the geometric series on any closed disk $|z| \leq R < 1$. Indeed, notice that $|z^j| = |z|^j \leq R^j$ and that since $R < 1$, the geometric series $\sum_{j=0}^{\infty} R^j$ converges.

2.3.2 Taylor series

In this section we will prove the remarkable result that a function analytic in the neighbourhood of a point can be approximated by a sequence of polynomials, namely by its Taylor series. Moreover we will see that convergence is uniform inside the largest open disk over which the function is analytic.

The Taylor series of a function is the result of successive approximations of the function by polynomials. Suppose that $f(z)$ is analytic in a neighbourhood of z_0. Then as we saw in Section 2.2.5 f is infinitely differentiable around z_0. Let us then write down a polynomial function f_n such that it agrees with f at z_0 up to an including its n-th derivative. In other words, $f_n^{(j)}(z_0) = f^{(j)}(z_0)$ for $j = 0, 1, \ldots, n$. The polynomial function of least order which satisfies this condition is

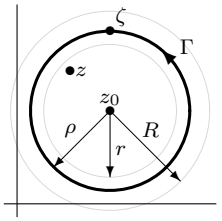
$$f_n(z) = f(z_0) + f'(z_0)(z - z_0) + \frac{f''(z_0)}{2}(z - z_0)^2 + \cdots + \frac{f^{(n)}(z_0)}{n!}(z - z_0)^n \ .$$

The sequence $\{f_n\}$, if it converges, does so to the **Taylor series around** z_0 of the function f:

$$\sum_{j=0}^{\infty} \frac{f^{(j)}(z_0)}{j!} (z - z_0)^j \ . \tag{2.48}$$

(If $z_0 = 0$ this series is also called the **Maclaurin series** of f.)

We will now prove the following important result: Let $f(z)$ be analytic in the disk $|z - z_0| < R$ centred at z_0. Then the Taylor series for f around z_0 converges to $f(z)$ for all z in the disk and moreover the convergence is uniform on any closed subdisk $|z - z_0| \leq r < R$.



The proof uses the generalised Cauchy Integral Formula with an appropriate choice of contour, as shown in the diagram. Let Γ denote the positively oriented circle centred at z_0 with radius ρ where $r < \rho < R$. By hypothesis, f is analytic in and on the contour Γ, whence for any z satisfying $|z - z_0| \leq r$, we have the Cauchy Integral Formula:

$$f(z) = \frac{1}{2\pi\, i} \oint_\Gamma \frac{f(\zeta)}{\zeta - z}\, d\zeta \ .$$

Now we rewrite the integrand:

$$\frac{1}{\zeta - z} = \frac{1}{(\zeta - z_0) - (z - z_0)} = \frac{1}{\zeta - z_0} \frac{1}{1 - \frac{z - z_0}{\zeta - z_0}} \ ,$$

and use the geometric series to write

$$\frac{1}{1 - \frac{z - z_0}{\zeta - z_0}} = \sum_{j=0}^{\infty} \left(\frac{z - z_0}{\zeta - z_0} \right)^j \ ,$$

which is valid because $|z - z_0| = r < \rho = |\zeta - z_0|$. Putting it all together, we have

$$\frac{1}{\zeta - z} = \sum_{j=0}^{\infty} \frac{(z - z_0)^j}{(\zeta - z_0)^{j+1}} \ .$$

Inserting it into the Cauchy Integral Formula,

$$f(z) = \frac{1}{2\pi\, i} \oint_\Gamma \sum_{j=0}^{\infty} \frac{f(\zeta)}{(\zeta - z_0)^{j+1}} (z - z_0)^j\, d\zeta \ .$$

Now we would be tempted to interchange the order of the integral and the summation and arrive at

$$f(z) \stackrel{?}{=} \sum_{j=0}^{\infty} \left[\frac{1}{2\pi\, i} \oint_\Gamma \frac{f(\zeta)}{(\zeta - z_0)^{j+1}}\, d\zeta \right] (z - z_0)^j$$
$$= \sum_{j=0}^{\infty} \frac{f^{(j)}(z_0)}{j!}\, (z - z_0)^j \ ,$$

where we have used the generalised Cauchy Integral Formula. This manipulation turns out to be allowed, but doing it this way we do not see the uniform convergence. This is done with more care below.

Let us prove the Taylor series theorem carefully. It is not hard, but it takes a bit more bookkeeping. Rather than using the geometric series in its entirety, let us use its n-th partial sum:

$$\frac{1}{1 - \frac{z-z_0}{\zeta - z_0}} = \sum_{j=0}^{n} \left(\frac{z - z_0}{\zeta - z_0} \right)^j + \frac{\left(\frac{z-z_0}{\zeta - z_0} \right)^{n+1}}{1 - \frac{z-z_0}{\zeta - z_0}} \ ,$$

whence

$$\frac{1}{\zeta - z} = \sum_{j=0}^{n} \frac{(z - z_0)^j}{(\zeta - z_0)^{j+1}} + \frac{\left(\frac{z-z_0}{\zeta - z_0} \right)^{n+1}}{\zeta - z} \ .$$

Into the Cauchy Integral Formula, we have

$$f(z) = \frac{1}{2\pi\, i} \oint_\Gamma f(\zeta) \left[\sum_{j=0}^{n} \frac{(z - z_0)^j}{(\zeta - z_0)^{j+1}} + \frac{\left(\frac{z-z_0}{\zeta - z_0} \right)^{n+1}}{\zeta - z} \right] d\zeta \ .$$

Now this is only a finite sum, so by linearity we can integrate it term by term. Using the generalised Cauchy Integral Formula we have

$$f(z) = \sum_{j=0}^{n} \frac{f^{(j)}(z_0)}{j!}\, (z - z_0)^j + R_n(z) \ ,$$

where

$$R_n(z) \equiv \frac{1}{2\pi\, i} \oint_\Gamma \frac{f(\zeta)}{\zeta - z} \left(\frac{z - z_0}{\zeta - z_0} \right)^{n+1} d\zeta \ .$$

In other words,

$$f(z) - \sum_{j=0}^{n} \frac{f^{(j)}(z_0)}{j!}\, (z - z_0)^j = R_n(z) \ ,$$

whence in order to prove uniform convergence of the Taylor series, we only have to show that we can make $|R_n(z)|$ as small as desired for all z by simply taking n sufficiently large. Let us estimate $|R_n(z)|$. Using (2.28)

$$|R_n(z)| \leq \frac{1}{2\pi} \oint_\Gamma \frac{|f(\zeta)|}{|\zeta - z|} \left| \frac{z - z_0}{\zeta - z_0} \right|^{n+1} |d\zeta| \ .$$

We now use that $|z - z_0| \leq r$, $|\zeta - z_0| = \rho$, $|f(\zeta)| \leq M$ for some M, $\ell(\Gamma) = 2\pi\, \rho$, and the triangle inequality (2.36),

$$|\zeta - z| = |(\zeta - z_0) - (z - z_0)| \geq |\zeta - z_0| - |z - z_0| \geq \rho - r \ ,$$

whence

$$\frac{1}{|\zeta - z|} \leq \frac{1}{\rho - r} \ .$$

Therefore,

$$|R_n(z)| \leq \frac{\rho\, M}{\rho - r} \left(\frac{r}{\rho} \right)^{n+1} \ .$$

This is what we wanted, because the right-hand side does not depend on z and can be made as small as desired by taking n large, since $r/\rho < 1$. This proves uniform convergence of the Taylor series.

Notice that this result implies that *the Taylor series will converge to $f(z)$ everywhere inside the largest open disk, centred at z_0, over which f is analytic*.

As an example, let us compute the Taylor series for the functions $\operatorname{Log} z$ around $z_0 = 1$ and also $1/(1-z)$ around $z_0 = 0$. The derivatives of the principal branch of the logarithm are:

$$\frac{d^j \operatorname{Log} z}{dz^j} = (-1)^{j+1}(j-1)!\frac{1}{z^j} \ .$$

Evaluating at $z = 1$ and constructing the Taylor series, we have

$$\operatorname{Log} z = \sum_{j=1}^{\infty} \frac{(-1)^{j+1}}{j} (z-1)^j \ .$$

This series is valid for $|z - 1| < 1$ which is the largest open disk centred at $z = 1$ over which $\operatorname{Log} z$ is analytic, as seen in Figure 2.8. Similarly,

Figure 2.8: Analyticity disks for the Taylor series of $\operatorname{Log} z$ and $1/(1-z)$.

$$\frac{d^j}{dz^j} \frac{1}{1-z} = \frac{j!}{(1-z)^{j+1}} \ ,$$

whence evaluating at $z = 0$ and building the Taylor series we find the geometric series

$$\frac{1}{1-z} = \sum_{j=0}^{\infty} z^j \ ,$$

which is valid for $|z| < 1$ since that is the largest open disk around the origin over which $1/(1-z)$ is analytic, as seen in Figure 2.8. *Now notice something remarkable.* We have two a priori different series representations for the function $1/(1-z)$ around the origin: one is the Taylor series and another is the geometric series. Yet we have shown that these series are the same. This is not a coincidence and we will see in Section 2.3.3 that series representations for analytic functions are unique: they are all essentially Taylor series.

Basic properties of Taylor series

Taking the derivative of the Taylor series for $\operatorname{Log} z$ about $z_0 = 1$ term by term, we find the series

$$\sum_{j=1}^{\infty} \frac{(-1)^{j+1}}{\qquad}(z-1)^{j-1} = \sum_{j=0}^{\infty}(-1)^j\,(z-1)^j = \sum_{j=0}^{\infty}(1-z)^j\ .$$

This is a geometric series which for $|z-1| < 1$ converges to

$$\frac{1}{1-(1-z)} = \frac{1}{z}\ ,$$

which is precisely the derivative of $\operatorname{Log} z$. This might not seem at all remarkable, but it is. There is no reason *a priori* why the termwise differentiation of an infinite series which converges to a function $f(z)$, should converge to the derivative $f'(z)$ of the function. This is because there are two limits involved: the limit in the definition of the derivative and the one which we take to approach the function $f(z)$, and we know from previous experience that the order in which one takes limits matters in general. On the other hand, what we have just seen is that for the case of the $\operatorname{Log} z$ function, these two limits commute; that is, they can be taken in any order. It turns out that this is not just a property of $\operatorname{Log} z$ but indeed of any analytic function.

To see this recall that we saw in Section 2.2.5 that if a function $f(z)$ is analytic in a disk $|z - z_0| < R$, then so are all its derivatives. In particular $f(z)$ and $f'(z)$ have Taylor series in the disk which converge uniformly on any closed subdisk. The Taylor series for $f'(z)$ is given by equation (2.48) applied to f' instead of f:

$$\sum_{j=0}^{\infty} \frac{(f')^{(j)}(z_0)}{j!}\,(z-z_0)^j\ .$$

But notice that the j-th derivative of f' is just the $(j+1)$-st derivative of f:

$(f')^{(j)} = f^{(j+1)}$. Therefore we can rewrite the above Taylor series as

$$\sum_{j=0}^{\infty} \frac{f^{(j+1)}(z_0)}{j!} (z - z_0)^j \ . \tag{2.49}$$

On the other hand, differentiating the Taylor series (2.48) for f termwise, we get

$$\begin{aligned} \sum_{j=0}^{\infty} \frac{f^{(j)}(z_0)}{j!} \, j \, (z - z_0)^{j-1} &= \sum_{j=1}^{\infty} \frac{f^{(j)}(z_0)}{(j-1)!} (z - z_0)^{j-1} \\ &= \sum_{k=0}^{\infty} \frac{f^{(k+1)}(z_0)}{k!} (z - z_0)^k \ , \end{aligned}$$

where we have reindexed the last sum by introducing $k = j - 1$. Finally, Shakespeare's Theorem tells us that this last series is the same as the one in equation (2.49). In other words, we have proven that if $f(z)$ is analytic around z_0, the Taylor series for $f'(z)$ around z_0 is obtained by termwise differentiation of the Taylor series for $f(z)$ around z_0.

Similarly one can show that Taylor series have additional properties. Let $f(z)$ and $g(z)$ be analytic around z_0. That means that there is some disk $|z - z_0| < R$ in which the two functions are analytic. Then as shown in Section 2.1.4, $\alpha f(z)$, for α any complex number, and $f(z) + g(z)$ are also analytic in the disk. Then one can show

- The Taylor series for $\alpha f(z)$ is the series obtained by multiplying each term in the Taylor series for $f(z)$ by α:

$$\sum_{j=0}^{\infty} \frac{\alpha \, f^{(j)}(z_0)}{j!} (z - z_0)^j \ .$$

- The Taylor series of $f(z) + g(z)$ is the series obtained by adding the terms for the Taylor series of $f(z)$ and $g(z)$:

$$\sum_{j=0}^{\infty} \frac{f^{(j)}(z_0) + g^{(j)}(z_0)}{j!} (z - z_0)^j \ .$$

These results follow from equations (2.9) and (2.10).

Finally, let $f(z)$ and $g(z)$ be analytic in a disk $|z - z_0| < R$ around z_0. We also saw in Section 2.1.4 that their product $f(z)g(z)$ is analytic there. Therefore it has a Taylor series which converges uniformly in any closed

subdisk. What is the relation between this series and the Taylor series for $f(z)$ and $g(z)$? Let us compute the first couple of terms. We have that the first few derivatives of fg are

$$(fg)(z_0) = f(z_0)g(z_0) \qquad (fg)'(z_0) = f'(z_0)g(z_0) + f(z_0)g'(z_0)$$
$$(fg)''(z_0) = f''(z_0)g(z_0) + 2f'(z_0)g'(z_0) + f(z_0)g''(z_0) \ ,$$

so that the first few terms of the Taylor series for fg are

$$f(z_0)g(z_0) + (f'(z_0)g(z_0) + f(z_0)g'(z_0))\ (z - z_0) + \frac{f''(z_0)g(z_0) + 2f'(z_0)g'(z_0) + f(z_0)g''(z_0)}{2}\,(z - z_0)^2 + \cdots$$

Notice that this can be rewritten as follows:

$$\left(f(z_0) + f'(z_0)(z - z_0) + \frac{f''(z_0)}{2}(z - z_0)^2 + \cdots\right) \times \left(g(z_0) + g'(z_0)(z - z_0) + \frac{g''(z_0)}{2}(z - z_0)^2 + \cdots\right) \ ,$$

which looks like the product of the first few terms in the Taylor series of f and g. Appearances do not lie in this case and one can show that the Taylor series for the product fg of any two analytic functions is the product of their Taylor series, provided one defines the product of the Taylor series appropriately.

Let us see this. To save some writing let me write the Taylor series for $f(z)$ as $\sum_{j=0}^{\infty} a_j(z - z_0)^j$ and for $g(z)$ as $\sum_{j=0}^{\infty} b_j(z - z_0)^j$. In other words, I have introduced abbreviations $a_j = f^{(j)}(z_0)/j!$ and $b_j = g^{(j)}(z_0)/j!$. The **Cauchy product** of these two series is defined by multiplying the series formally and collecting terms with the same power of $z - z_0$. In other words,

$$\left(\sum_{j=0}^{\infty} a_j(z - z_0)^j\right) \times \left(\sum_{j=0}^{\infty} b_j(z - z_0)^j\right) = \sum_{j=0} c_j(z - z_0)^j \ ,$$

where

$$c_j = \sum_{\substack{k,\ell=0 \\ k+\ell=j}}^{\infty} a_k b_\ell = \sum_{k=0}^{j} a_k b_{j-k} = \sum_{k=0}^{j} \frac{f^{(k)}(z_0)}{k!}\,\frac{g^{(j-k)}(z_0)}{(j-k)!} \ .$$

On the other hand, the Taylor series for fg can be written an

$$\sum_{j=0}^{\infty} \frac{(fg)^{(j)}(z_0)}{j!}\,(z - z_0)^j \ ,$$

where one can use the **generalised Leibniz rule** to obtain

$$(fg)^{(j)}(z_0) = \sum_{k=0}^{j} \binom{j}{k} f^{(k)}(z_0)g^{(j-k)}(z_0) \ ,$$

where $\binom{j}{k}$ is the binomial coefficient

$$\binom{j}{k} = \frac{j!}{k!(j-k)!} \, .$$

Therefore the Taylor series for fg can be written as

$$\sum_{j=0}^{\infty} \frac{1}{j!} \sum_{k=0}^{j} \binom{j}{k} f^{(k)}(z_0) g^{(j-k)}(z_0) \, (z-z_0)^j$$

$$= \sum_{j=0}^{\infty} \sum_{k=0}^{j} \frac{1}{k!(j-k)!} f^{(k)}(z_0) g^{(j-k)}(z_0) \, (z-z_0)^j = \sum_{j=0}^{\infty} c_j (z-z_0)^j \, ,$$

with the c_j being the same as above.

2.3.3 Power series

Taylor series are examples of a more general type of series, called power series, whose study is the purpose of this section. We will see that power series are basically always the Taylor series of some analytic function. This shows that series representations of analytic functions are in some sense unique, so that if we manage to cook up, by whatever means, a power series converging to a function in some disk, we know that this series will be its Taylor series of the function around the centre of the disk.

By a **power series** around z_0 we mean a series of the form

$$\sum_{j=0}^{\infty} a_j \, (z-z_0)^j \, ,$$

and where $\{a_j\}$ are known as the **coefficients** of the power series. A power series is clearly determined by its coefficients and by the point z_0. Given a power series one can ask many questions: For which z does it converge? Is the convergence uniform? Will it converge to an analytic function? Will the power series be a Taylor series?

We start the section with the following result, which we will state without proof. It says that to any power series $\sum_{j=0}^{\infty} a_j \, (z-z_0)^j$ one can associate a number $0 \leq R \leq \infty$, called the **radius of convergence**, depending only on the coefficients $\{a_j\}$, such that the series converges in the disk $|z-z_0| < R$, uniformly on any closed subdisk, and the series diverges in $|z-z_0| > R$.

Introduce lim sup, root test and the proof of this theorem.

One can actually give a formula for the number R in terms of the coefficients $\{a_j\}$ but we will not do so here in general. Instead we will give a

formula which is valid only in those cases when the Ratio Test can be used. Recall that the Ratio Test says that if the limit

$$L \equiv \lim_{j\to\infty} \left| \frac{c_{j+1}}{c_j} \right| \tag{2.50}$$

exists, then the series $\sum_{j=0}^{\infty} c_j$ converges for $L < 1$ and diverges for $L > 1$. In the case of a power series, we have

$$L = \lim_{j\to\infty} \left| \frac{a_{j+1}(z-z_0)^{j+1}}{a_j(z-z_0)^j} \right| = \lim_{j\to\infty} \left| \frac{a_{j+1}}{a_j} \right| |z-z_0| \ .$$

Therefore convergence is guaranteed if $L < 1$, which is equivalent to

$$|z-z_0| < \lim_{j\to\infty} \left| \frac{a_j}{a_{j+1}} \right|$$

and divergence is guaranteed for $L > 1$, which is equivalent to

$$|z-z_0| > \lim_{j\to\infty} \left| \frac{a_j}{a_{j+1}} \right| \ .$$

Therefore if the limit (2.50) exists, we have that the radius of convergence is given by

$$\boxed{R = \lim_{j\to\infty} \left| \frac{a_j}{a_{j+1}} \right| \ .} \tag{2.51}$$

Notice that this agrees with our experience with the geometric series (2.47), which is clearly a power series around the origin. Since all the coefficients are equal, the limit exists and $R = 1$, which is precisely the radius of convergence we had established previously.

Power series are Taylor series

We are now going to prove the main result of this section: that a power series is the Taylor series of the functions it approximates. This is a very useful result, because it says that in order to compute the Taylor series of a function it is enough to produce any power series which converges to that function. The proof will follow two steps. The first is to show that a power series converges to an analytic function and the second step will use the Cauchy Integral formula to relate the coefficients of the power series with those of the Taylor series. The first step will itself require two preliminary results, which we state in some more generality.

Suppose that $\{f_n\}$ is a sequence of continuous functions which converges uniformly to a function $f(z)$ in the closed disk $|z - z_0| \leq R$. Let Γ be any contour (not necessarily closed) inside the disk, and let ℓ be the length of the contour. Then we claim that the sequence $\int_\Gamma f_n(z)\,dz$ converges to the integral $\int_\Gamma f(z)\,dz$. To see this, let $\varepsilon > 0$. Then because of uniform convergence, there exists N depending only on ε such that for all $n \geq N$, one has $|f(z) - f_n(z)| < \varepsilon/\ell$ for all z in the disk. Then

$$\begin{aligned}\left|\int_\Gamma f(z)\,dz - \int_\Gamma f_n(z)\,dz\right| &= \left|\int_\Gamma \left(f(z) - f_n(z)\right)\,dz\right| \\ &\leq \max_{z\in\Gamma} |f(z) - f_n(z)|\,\ell \qquad \text{(using (2.28))} \\ &< (\varepsilon/\ell)\ell = \varepsilon\ .\end{aligned}$$

Now suppose that the sequence $\{f_n\}$ is the sequence of partial sums of some infinite series of functions. Then the above result says that one can integrate the series termwise, since for any partial sum, the integral of the sum is the sum of the integrals. In other words, when integrating an infinite series which converges uniformly in some region U along any contour in U, we can interchange the order of the summation and the integration.

Now suppose that the functions $\{f_n\}$ are not just continuous but actually analytic, and let Γ be any loop; that is, a closed simple contour. Then by the Cauchy Integral Theorem, $\oint_\Gamma f_n(z)\,dz = 0$, whence by what we have just shown

$$\oint_\Gamma f(z)\,dz = \lim_{n\to\infty} \oint_\Gamma f_n(z)\,dz = 0\ .$$

Therefore by Morera's theorem, $f(z)$ is also analytic. Therefore we have shown that

the uniform limit of analytic functions is analytic.

In particular, let $\sum_{j=0}^\infty a_j(z - z_0)^j$ be a power series with circle of convergence $|z - z_0| = R > 0$. Since each of the partial sums, being a polynomial function, is analytic in the disk (in fact, in the whole plane), the limit is also analytic in the disk. In other words, a power series converges to an analytic function inside its disk of convergence.

Now that we know that $\sum_{j=0}^\infty a_j(z - z_0)^j$ defines an analytic function, call it $f(z)$, in its disk of convergence, we can compute its Taylor series and compare it with the original series. The Taylor series of $f(z)$ around z_0 has coefficients given by the generalised Cauchy Integral Formula:

$$\frac{f^{(j)}(z_0)}{j!} = \frac{1}{2\pi\,i} \oint_\Gamma \frac{f(z)}{(z - z_0)^{j+1}}\,dz\ ,$$

where Γ is any positively oriented loop inside the disk of convergence of the power series which contains the point z_0 in its interior. Because the power series converges uniformly, we can now substitute the power series for $f(z)$ inside the integral and compute the integral termwise:

$$\begin{aligned}\frac{f^{(j)}(z_0)}{j!} &= \frac{1}{2\pi\, i}\oint_\Gamma \frac{f(z)}{(z-z_0)^{j+1}}\,dz \\ &= \sum_{k=0}^{\infty}\frac{1}{2\pi\, i}\oint_\Gamma a_k\,\frac{(z-z_0)^k}{(z-z_0)^{j+1}}\,dz \\ &= \sum_{k=0}^{\infty} a_k\,\frac{1}{2\pi\, i}\oint_\Gamma (z-z_0)^{k-j-1}\,dz\ .\end{aligned}$$

But now, from the generalised Cauchy Integral Formula,

$$\oint_\Gamma (z-z_0)^{k-j-1}\,dz = \begin{cases} 2\pi\, i & \text{if } j=k, \text{ and} \\ 0 & \text{otherwise.}\end{cases} \tag{2.52}$$

Therefore, only one term contributes to the $\sum_k$, namely the term with $k=j$, and hence we see that

$$\frac{f^{(j)}(z_0)}{j!} = a_j\ .$$

In other words, the power series *is* the Taylor series. Said differently, any power series is the Taylor series of a function analytic in the disk of convergence $|z-z_0|<R$.

For example, let us compute the Taylor series of the function

$$\frac{1}{(z-1)(z-2)}$$

in the disk $|z|<1$. This is the largest disk centred at the origin where we could hope to find a convergent power series for this function, since it has singularities at $z=1$ and $z=2$. The naive solution to this problem would be to take derivatives and evaluate them at the origin and build the Taylor series this way. However from our discussion above, it is enough to exhibit any power series which converges to this function in the specified region. We use partial fractions to rewrite the function as a sum of simple fractions:

$$\frac{1}{(z-1)(z-2)} = \frac{1}{1-z} - \frac{1}{2-z}\ .$$

Now we use geometric series for each of them. For the first fraction we have

$$\frac{1}{1-z} = \sum_{j=0}^{\infty} z^j \qquad \text{valid for } |z|<1;$$

whereas for the second fraction we have

$$\frac{-1}{2-z} = \frac{-1/2}{1-(z/2)} = -\frac{1}{2}\sum_{j=0}^{\infty}\frac{z^j}{2^j} = -\sum_{j=0}^{\infty}\frac{z^j}{2^{j+1}} \ ,$$

which is valid for $|z| < 2$, which contains the region of interest. Therefore, putting the two series together,

$$\frac{1}{(z-1)(z-2)} = \sum_{j=0}^{\infty}\left(1 - \frac{1}{2^{j+1}}\right) z^j \ , \qquad \text{for } |z| < 1.$$

2.3.4 Laurent series

In the previous section we saw that any function which is analytic in some neighbourhood of a point z_0 can be approximated by a power series (its Taylor series) about that point. How about a function which has a "mild" singularity at z_0? For example, how about a function of the form $g(z)/(z-z_0)$? Might we not expect to be able to approximate it by some sort of power series? It certainly could not be a power series of the type we have been discussing because these series are analytic at z_0. There is, however, a simple yet useful generalisation of the notion of power series which can handle these cases. These series are known as Laurent series and consist of a sum of two power series.

A **Laurent series** about the point z_0 is a sum of two power series one consisting of positive powers of $z - z_0$ and the other of negative powers:

$$\sum_{j=0}^{\infty} a_j (z-z_0)^j + \sum_{j=1}^{\infty} a_{-j}(z-z_0)^{-j} \ .$$

Laurent series are often abbreviated as

$$\sum_{j=-\infty}^{\infty} a_j (z-z_0)^j,$$

but we should keep in mind that this is only an abbreviation: conceptually a Laurent series is the sum of two independent power series.

A Laurent series is said to converge if each of the power series converges. The first series, being a power series in $z - z_0$ converges inside some circle of convergence $|z - z_0| = R$, for some $0 \le R \le \infty$. The second series, however, is a power series in $w = 1/(z - z_0)$. Hence it will converge inside a circle of convergence $|w| = R'$; that is, for $|w| < R'$. If we let $R' = 1/r$, then this

condition translates into $|z - z_0| > r$. In other words, such a Laurent series will converge in an annulus: $r < |z - z_0| < R$. (Of course for this to make sense, we need $r < R$. If this is not the case, then the Laurent series does not converge anywhere.)

It turns out that the results which are valid for Taylor series have generalisations for Laurent series. The first main result that we will prove is that any function analytic in an open annulus $r < |z - z_0| < R$ centred at z_0 has a Laurent series around z_0 which converges to it everywhere inside the annulus and uniformly on closed sub-annuli $r < R_1 \leq |z - z_0| \leq R_2 < R$. Moreover the coefficients of the Laurent series are given by

$$a_j = \frac{1}{2\pi\, i} \oint_\Gamma \frac{f(z)}{(z - z_0)^{j+1}}\, dz \ , \qquad \text{for } j = 0, \pm 1, \pm 2, \ldots,$$

where Γ is any positively oriented loop lying in the annulus and containing z_0 in its interior.

Notice that this result generalises the result proven in Section 2.3.2 for functions analytic in the disk. Indeed, if $f(z)$ were analytic in $|z - z_0| < R$, then by the Cauchy Integral Theorem and the above formula for a_j, it would follow that that $a_{-j} = 0$ for $j = 1, 2, \ldots$, and hence that the Laurent series is the Taylor series. Notice also that the Laurent series is a nontrivial generalisation of the Taylor series in that the coefficients a_{-j} for $j = 1, 2, \ldots$ are not just simply derivatives of the function, but rather require contour integration.

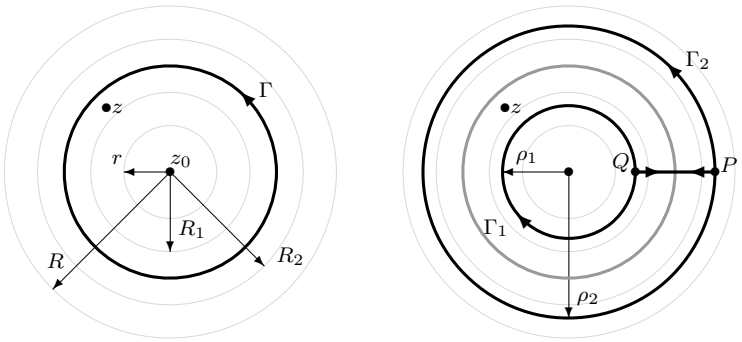


Figure 2.9: Contours Γ, Γ_1 and Γ_2.

In order to follow the logic of the proof, it will be convenient to keep Figure 2.9 in mind. The left-hand picture shows the annuli $r < R_1 \leq |z - z_0| \leq R_2 < R$ and the contour Γ. The right-hand picture shows the equivalent

contours Γ_1 and Γ_2, circles with radii ρ_1 and ρ_2 satisfying the inequalities $r < \rho_1 < R_1$ and $R_2 < \rho_2 < R$.

Consider the closed contour C, starting and ending at the point P in the Figure, and defined as follows: follow Γ_2 all the way around until P again, then go to Q via the 'bridge' between the two circles, then all the way along Γ_1 until Q, then back to P along the 'bridge.' This contour encircles the point z once in the positive sense, hence by the Cauchy Integral Formula we have that

$$f(z) = \frac{1}{2\pi\, i} \oint_C \frac{f(\zeta)}{\zeta - z}\, d\zeta \ .$$

On the other hand, because the 'bridge' is traversed twice in opposite directions, their contribution to the integral cancels and we are left with

$$f(z) = \frac{1}{2\pi\, i} \oint_{\Gamma_1} \frac{f(\zeta)}{\zeta - z}\, d\zeta + \frac{1}{2\pi\, i} \oint_{\Gamma_2} \frac{f(\zeta)}{\zeta - z}\, d\zeta \ .$$

We now treat each integral at a time.

The integral along Γ_2 can be treated *mutatis mutandis* as we did the similar integral in the proof of the Taylor series theorem in Section 2.3.2. We simply quote the result:

$$\frac{1}{2\pi\, i} \oint_{\Gamma_2} \frac{f(\zeta)}{\zeta - z}\, d\zeta = \sum_{j=0}^{\infty} a_j (z - z_0)^j \ ,$$

where

$$a_j = \frac{1}{2\pi\, i} \oint_{\Gamma_2} \frac{f(\zeta)}{(\zeta - z_0)^{j+1}}\, d\zeta = \frac{f^{(j)}(z_0)}{j!} \ . \tag{2.53}$$

Moreover the series converges uniformly in the closed disk $|z - z_0| \leq R_2$, as was shown in that section.

The integral along Γ_1 can be treated along similar lines, except that because $|z - z_0| > |\zeta - z_0|$, we must expand the integrand differently. We will be brief, since the idea is very much the same as what was done for the Taylor series. We start by rewriting $1/(\zeta - z)$ appropriately:

$$\frac{1}{\zeta - z} = \frac{1}{(\zeta - z_0) - (z - z_0)} = -\frac{1}{z - z_0}\, \frac{1}{1 - \frac{\zeta - z_0}{z - z_0}} \ .$$

Let us write this now as a geometric series:

$$\frac{1}{\zeta - z} = -\frac{1}{z - z_0} \left[\sum_{j=0}^{n} \left(\frac{\zeta - z_0}{z - z_0} \right)^j + \frac{\left(\frac{\zeta - z_0}{z - z_0} \right)^{n+1}}{1 - \frac{\zeta - z_0}{z - z_0}} \right] \ ;$$

whence

$$\frac{1}{2\pi\, i}\oint_{\Gamma_1}\frac{f(\zeta)}{\zeta - z}\,d\zeta = \sum_{j=1}^{m+1} a_{-j}(z - z_0)^{-j} + S_n(z)\ ,$$

where

$$a_{-j} = -\frac{1}{2\pi\, i}\oint_{\Gamma_1}\frac{f(\zeta)}{(\zeta - z_0)^{-j+1}}\,d\zeta\ , \tag{2.54}$$

and where

$$S_n(z) = \frac{1}{2\pi\, i}\oint_{\Gamma_1}\frac{f(\zeta)}{\zeta - z}\,\frac{(\zeta - z_0)^{n+1}}{(z - z_0)^{n+1}}\,d\zeta\ .$$

Now, for ζ in Γ_1 we have that $|\zeta - z_0| = \rho_1$ and from the triangle inequality (2.36), that $|\zeta - z| \geq R_1 - \rho_1$. We also note that $|z - z_0| \geq R_1$. Furthermore, $f(\zeta)$, being continuous, is bounded so that $|f(\zeta)| \leq M$ for some M and all ζ on Γ_1. Therefore using (2.28) and the above inequalities,

$$|S_n(z)| \leq \frac{M\,\rho_1}{R_1 - \rho_1}\left(\frac{\rho_1}{R_1}\right)^{n+1}\ ,$$

which is independent of z and, because $\rho_1 < R_1$, can be made arbitrarily small by choosing n large. Hence $S_n(z) \to 0$ as $n \to \infty$ uniformly in $|z - z_0| \geq R_1$, and

$$\frac{1}{2\pi\, i}\oint_{\Gamma_1}\frac{f(\zeta)}{\zeta - z}\,d\zeta = \sum_{j=1}^{\infty} a_{-j}(z - z_0)^{-j}\ ,$$

where the a_{-j} are still given by (2.54). In other words,

$$\frac{1}{2\pi\, i}\oint_{\Gamma_1}\frac{f(\zeta)}{\zeta - z}\,d\zeta = \sum_{j=1}^{\infty} a_{-j}(z - z_0)^{-j}\ ,$$

and the series converges uniformly to the integral for $|z - z_0| \leq R_1$. In summary, we have that proven that $f(z)$ is approximated by the Laurent series

$$f(z) = \sum_{j=-\infty}^{\infty} a_j(z - z_0)^j\ ,$$

everywhere on $r < |z - z_0| < R$ and uniformly on any closed sub-annulus, where the coefficients a_j are given by (2.53) for $j \geq 0$ and by (2.54) for $j < 0$.

We are almost done, except that in the statement of the theorem the coefficients a_j are given by contour integrals along Γ and what we have shown is that they are given by contour integrals along Γ_1 or Γ_2. But notice that the integrand in (2.53) is analytic in the domain bounded by the contours Γ and Γ_2; and similarly for the integrand in (2.54) in the region bounded by

the contours Γ and Γ_1. Therefore we can deform the contours Γ_1 and Γ_2 to $-\Gamma$ and Γ respectively, in the integrals

$$a_{-j} = -\frac{1}{2\pi\, i} \oint_{\Gamma_1} \frac{f(\zeta)}{(\zeta - z_0)^{-j+1}}\, d\zeta = \frac{1}{2\pi\, i} \oint_{\Gamma} \frac{f(\zeta)}{(\zeta - z_0)^{-j+1}}\, d\zeta$$
$$a_j = \frac{1}{2\pi\, i} \oint_{\Gamma_2} \frac{f(\zeta)}{(\zeta - z_0)^{j+1}}\, d\zeta = \frac{1}{2\pi\, i} \oint_{\Gamma} \frac{f(\zeta)}{(\zeta - z_0)^{j+1}}\, d\zeta \ ,$$

which proves the theorem.

Laurent series are unique

We saw in Section 2.3.3 that any power series is the Taylor series of the analytic function it converges to. In other words, the power series representation of an analytic function is unique (in the domain of convergence of the series, of course). Since Laurent series are generalisations of the Taylor series and agree with them when the function is analytic not just in the annulus but in fact in the whole disk, we might expect that the same is true and that the Laurent series representation of a function analytic in an annulus should also be unique. This turns out to be true and the proof follows basically from that of the uniqueness of the power series.

More precisely, one has the following result. Let

$$\sum_{j=0}^{\infty} c_j (z - z_0)^j \qquad \text{and} \qquad \sum_{j=1}^{\infty} c_{-j} (z - z_0)^{-j}$$

be any two power series converging in $|z - z_0| < R$ and $|z - z_0| > r$, respectively, with $R > r$. Then there is a function $f(z)$ analytic in the annulus $r < |z - z_0| < R$, such that

$$\sum_{j=0}^{\infty} c_j (z - z_0)^j + \sum_{j=1}^{\infty} c_{-j} (z - z_0)^{-j}$$

is its Laurent series. We shall omit the proof, except to notice that this follows from the uniqueness of the power series applied to each of the series in turn.

Do this in detail.

This is a very useful result because it says that no matter how we obtain the power series, their sum is guaranteed to be the Laurent series of the analytic function in question. Let us illustrate this in order to compute the Laurent series of some functions.

For example, let us compute the Laurent series of the rational function $(z^2-2z+3)(z-2)$ in the region $|z-1|>1$. Let us first rewrite the numerator as a power series in $(z-1)$:

$$z^2-2z+3=(z-1)^2+2\ .$$

Now we do the same with the denominator:

$$\frac{1}{z-2}=\frac{1}{(z-1)-1}=\frac{1}{z-1}\,\frac{1}{1-\frac{1}{z-1}}\ ,$$

where we have already left it in a form which suggests that we try a geometric series in $1/(z-1)$, which converges in the specified region $|z-1|>1$. Indeed, we have that in this region,

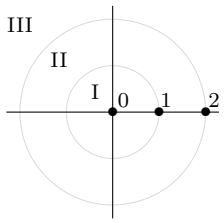
$$\frac{1}{z-1}\,\frac{1}{1-\frac{1}{z-1}}=\frac{1}{z-1}\sum_{j=0}^{\infty}\frac{1}{(z-1)^j}=\sum_{j=0}^{\infty}\frac{1}{(z-1)^{j+1}}\ .$$

Putting the two series together,

$$\begin{aligned}\frac{z^2-2z+3}{z-2}&=\big((z-1)^2+2\big)\sum_{j=0}^{\infty}\frac{1}{(z-1)^{j+1}}\\&=(z-1)+1+\sum_{j=0}^{\infty}\frac{3}{(z-1)^{j+1}}\ .\end{aligned}$$

By the uniqueness of the Laurent series, this is *the* Laurent series for the function in the specified region.

As a final example, consider the function $1/(z-1)(z-2)$. Let us find its Laurent expansions in the regions: $|z|<1$, $1<|z|<2$ and $|z|>2$, which we have labelled I, II and III in the figure. We start by decomposing the function into partial fractions:



$$\frac{1}{(z-1)(z-2)}=\frac{1}{z-2}-\frac{1}{z-1}\ .$$

In region I, we have the following geometric series:

$$-\frac{1}{z-1}=\frac{1}{1-z}=\sum_{j=0}^{\infty}z^j\quad\text{valid for }|z|<1;\text{ and}$$

$$\frac{1}{z-2}=\frac{-\frac{1}{2}}{1-(z/2)}=-\tfrac{1}{2}\sum_{j=0}^{\infty}\Big(\frac{z}{2}\Big)^j=\sum_{j=0}^{\infty}\frac{-1}{2^{j+1}}\,z^j\quad\text{valid for }|z|<2.$$

Therefore in their common region of convergence, namely region I, we have that

$$\frac{1}{(z-1)(z-2)} = \sum_{j=0}^{\infty} \left(1 - \frac{1}{2^{j+1}}\right) z^j \ .$$

In region II, the first of the geometric series above is not valid, but the second one is. Because in region II, $|z| > 1$, this means that $|1/z| < 1$, whence we should try and use a geometric series in $1/z$. This is easy:

$$-\frac{1}{z-1} = -\frac{1}{z}\frac{1}{1-(1/z)} = -\frac{1}{z}\sum_{j=0}^{\infty}\left(\frac{1}{z}\right)^j = \sum_{j=0}^{\infty}\frac{-1}{z^{j+1}} \quad \text{valid for } |z| > 1.$$

Therefore in region II we have that

$$\frac{1}{(z-1)(z-2)} = \sum_{j=0}^{\infty}\frac{-1}{z^{j+1}} + \sum_{j=0}^{\infty}\frac{-1}{2^{j+1}}\, z^j \ .$$

Finally in region III, we have that $|z| > 2$, so that we will have to find another series converging to $1/(z-2)$ in this region. Again, since now $|2/z| < 1$ we should try to use a geometric series in $2/z$. This is once again easy:

$$\frac{1}{z-2} = \frac{1}{z}\frac{1}{1-(2/z)} = \frac{1}{z}\sum_{j=0}^{\infty}\left(\frac{2}{z}\right)^j = \sum_{j=0}^{\infty}\frac{2^j}{z^{j+1}} \quad \text{valid for } |z| > 2.$$

Therefore in region III we have that

$$\frac{1}{(z-1)(z-2)} = \sum_{j=0}^{\infty}\left(-1+2^j\right)\ \frac{1}{z^{j+1}} \ .$$

Again by the uniqueness of the Laurent series, we know that these are *the* Laurent series for the function in the specified regions.

2.3.5 Zeros and Singularities

As a consequence of the existence of power and Laurent series representations for analytic functions we are able to characterise the possible singularities that an analytic function can have, and this is the purpose of this section.

A point z_0 is said to be a **singularity** for a function $f(z)$, if f ceases to be analytic at z_0. Singularities can come in two types. One says that a a point z_0 is an **isolated singularity** for a function $f(z)$, if f is analytic in some punctured disk around the singularity; that is, in $0 < |z - z_0| < R$ for some

$R > 0$. We have of course already encountered isolated singularities; e.g., the function $1/(z - z_0)$ has an isolated singularity at z_0. In fact, we will see below that the singularities of a rational function are always isolated. Singularities need not be isolated, of course. For example, any point $-x$ in the non-positive real axis is a singularity for the principal branch $\operatorname{Log} z$ of the logarithm function which is *not* isolated, since any disk around $-x$, however small, will contain other singularities. In this section we will concentrate on isolated singularities. We will see that there are three types of isolated singularities, distinguished by the behaviour of the function as it approaches the singularity. Before doing so we will discuss the singularities of rational functions. As these occur at the zeros of the denominators, we will start by discussing zeros.

Zeros of analytic functions

Let $f(z)$ be analytic in a neighbourhood of a point z_0. This means that there is an open disk $|z - z_0| < R$ in which f is analytic. We say that z_0 is a **zero** of f if $f(z_0) = 0$. More precisely we say that z_0 is a **zero of order** m, for $m = 1, 2, \ldots$, if

$$f(z_0) = f'(z_0) = f''(z_0) = \cdots = f^{(m-1)}(z_0) = 0 \qquad \textit{but } f^{(m)}(z_0) \neq 0.$$

(A zero of order $m = 1$ is often called a **simple zero**.) Because $f(z)$ is analytic in the disk $|z - z_0| < R$, it has a power series representation there: namely the Taylor series:

$$f(z) = \sum_{j=0}^{\infty} \frac{f^{(j)}(z_0)}{j!} (z - z_0)^j \ .$$

But because z_0 is a zero of order m, the first m terms in the Taylor series vanish, whence

$$f(z) = \sum_{j=m}^{\infty} \frac{f^{(j)}(z_0)}{j!} (z - z_0)^j = (z - z_0)^m g(z) \ ,$$

where $g(z)$ has a power series representation

$$g(z) = \sum_{j=0}^{\infty} \frac{f^{(j+m)}(z_0)}{(j+m)!} (z - z_0)^j$$

in the disk, whence it is analytic there and moreover, by hypothesis, $g(z_0) = f^{(m)}(z_0)/m! \neq 0$. It follows from this that the zeros of an analytic function

are isolated. Because $g(z)$ is analytic, and hence continuous, in the disk $|z - z_0| < R$ and $g(z_0) \neq 0$, it means that there is a disk $|z - z_0| < \varepsilon < R$ in which $g(z) \neq 0$, and hence neither is $f(z) = (z - z_0)^m \, g(z)$ zero there.

Now let $P(z)/Q(z)$ be a rational function. Its singularities will be the zeroes of $Q(z)$ and we have just seen that these are isolated, whence the singularities of a rational function are isolated.

Isolated singularities

Now let z_0 be an isolated singularity for a function $f(z)$. This means that f is analytic in some punctured disk $0 < |z - z_0| < R$, for some $R > 0$. The punctured disk is a degenerate case of an open annulus $r < |z - z_0| < R$, corresponding to $r = 0$. By the results of the previous section, we know that $f(z)$ has a Laurent series representation there. We can distinguish three types of singularities depending on the Laurent expansion:

$$f(z) = \sum_{j=-\infty}^{\infty} a_j (z - z_0)^j \ .$$

Let us pay close attention to the negative powers in the Laurent expansion: we can either have no negative powers—that is, $a_j = 0$ for all $j < 0$; a finite number of negative powers—that is, $a_j = 0$ for all but a finite number of $j < 0$; or an infinite number of negative powers—that is, $a_j \neq 0$ for an infinite number of $j < 0$. This trichotomy underlies the following definitions:

- We say that z_0 is a **removable singularity** of f, if the Laurent expansion of f around z_0 has no negative powers; that is,

$$f(z) = \sum_{j=0}^{\infty} a_j (z - z_0)^j \ .$$

- We say that z_0 is a **pole of order** m for f, if the Laurent expansion of f around z_0 has a_j for all $j < -m$ and $a_{-m} \neq 0$; that is powers; that is,

$$f(z) = \frac{a_{-m}}{(z - z_0)^m} + \cdots + a_0 + a_1 (z - z_0) + \cdots \qquad \text{with } a_{-m} \neq 0.$$

  A pole of order $m = 1$ is often called a **simple pole**.

- Finally we say that z_0 is an **essential singularity** of f if the Laurent expansion of f around z_0 has an infinite number of nonzero terms with negative powers of $(z - z_0)$.

The different types of isolated singularities can be characterised by the way the function behaves in the neighbourhood of the singularity. For a removable singularity the function is clearly bounded as $z \to z_0$, since the power series representation

$$f(z) = \sum_{j=0}^{\infty} a_j (z - z_0)^j = a_0 + a_1(z - z_0) + \cdots$$

certainly has a well-defined limit as $z \to z_0$: namely, a_0. This is *not* the same thing as saying that $f(z_0) = a_0$. If this were the case, then the function would not have a singularity at z_0, but it would be analytic there as well. Therefore, removable singularities are due to f being incorrectly or "peculiarly" defined at z_0. For example, consider the following bizarre-looking function:

$$f(z) = \begin{cases} e^z & \text{for } z \neq 0; \\ 26 & \text{at } z = 0. \end{cases}$$

This function is clearly analytic in the punctured plane $|z| > 0$, since it agrees with the exponential function there, which is an entire function. This means that in the punctured plane, $f(z)$ has a power series representation which agrees with the Taylor series of the exponential function:

$$f(z) = \sum_{j=0}^{\infty} \frac{1}{j!}\, z^j \ .$$

However this series has the limit 1 as $z \to 0$, which is the value of the exponential for $z = 0$, and this does not agree with the value of f there. Hence the function has a singularity, but one which is easy to cure: we simply redefine f at the origin so that $f(z) = \exp(z)$ throughout the complex plane. Other examples of removable singularities are

$$\frac{\sin z}{z} = \frac{1}{z}\left(z - \frac{z^3}{3!} + \frac{z^5}{5!} - \cdots\right) = 1 - \frac{z^2}{3!} + \frac{z^4}{5!} - \cdots \ ; \tag{2.55}$$

and

$$\frac{z^2 - 1}{z - 1} = \frac{1}{z - 1}\,\Big((z-1)^2 + 2(z-1)\Big) = (z - 1) + 2 \ .$$

Of course in this last example we could have simply noticed that $z^2 - 1 = (z - 1)(z + 1)$ and simplified the rational function to $z + 1 = (z - 1) + 2$. In summary, at a removable singularity the function is bounded and can be redefined at the singularity so that the new function is analytic there, in effect removing the singularity.

In contrast, a pole is a true singularity for the function f. Indeed, around a pole z_0 of order m, the Laurent series for f looks like

$$f(z) = \frac{1}{(z-z_0)^m}\, h(z) \ ,$$

where $h(z)$ has a series expansion around z_0 given by

$$h(z) = \sum_{j=0}^{\infty} a_{j-m}(z-z_0)^j = a_{-m} + a_{-m+1}(z-z_0) + \cdots$$

This means that $h(z)$ has at most a removable singularity at z_0. We have already seen many examples of functions with poles throughout these lectures, so we will not give more examples. Let us however pause to discuss the singularities of a rational function.

Let $f(z) = P(z)/Q(z)$ be a rational function. Then we claim that $f(z)$ has either a pole or a removable singularities at the zeros of $Q(z)$. Let us be a little bit more precise. Suppose that z_0 is a zero of $Q(z_0)$, and assume that it is a zero of order m. This means that

$$Q(z) = (z-z_0)^m\, q(z) \ ,$$

where $q(z)$ is an analytic function around z_0 and such that $q(z_0) \neq 0$. If z_0 is *not* a zero of $P(z)$, then z_0 is a pole of f of order m. If z_0 is a zero of order k of $P(z)$, then we have that

$$P(z) = (z-z_0)^k\, p(z) \ ,$$

where $p(z)$ is analytic and $p(z_0) \neq 0$. Therefore we have that

$$f(z) = \frac{(z-z_0)^k\, p(z)}{(z-z_0)^m\, q(z)} = \frac{1}{(z-z_0)^{m-k}}\, \frac{p(z)}{q(z)} \ ;$$

whence $f(z)$ has a pole of order $m-k$ if $m > k$ and has a removable singularity otherwise.

How about essential singularities? A result known as **Picard's Theorem** says that a function takes all possible values (with the possible exception of one) in any neighbourhood of an essential singularity. This is a deep result in complex analysis and one we will not even attempt to prove. Let us however verify this for the function $f(z) = \exp(1/z)$. This function is analytic in the punctured plane $|z| > 0$ since the exponential function is entire. For any finite w we have seen that the exponential function has a power series expansion:

$$e^w = \sum_{j=0}^{\infty} \frac{1}{j!}\, w^j \ .$$

Therefore for $|z| > 0$, we have that

$$e^{1/z} = \sum_{j=0}^{\infty} \frac{1}{j!} \frac{1}{z^j} \ ,$$

whence $z_0 = 0$ is an essential singularity. According to Picard's theorem, the function $\exp(1/z)$ takes every possible value (except possibly one) in *any* neighbourhood of the origin. Clearly, the value 0 is never attainable, but we can easily check that any other value is obtained. Let $c \neq 0$ be any nonzero complex number, and let us solve for those z such that $\exp(1/z) = c$. The multiple-valuedness of the logarithm says that there are infinitely many such z, satisfying:

$$\frac{1}{z} = \log(c) = \mathrm{Log}\,|c| + i\,\mathrm{Arg}(c) + 2\pi\,i\,k \ ,$$

for $k = 0, \pm 1 \ \pm 2, \ldots$, whose moduli are given by

$$|z| = \frac{\mathrm{Log}\,|c| - i\,\mathrm{Arg}(c) - 2\pi\,i\,k}{(\mathrm{Log}\,|c|)^2 + (\mathrm{Arg}(c) + 2\pi\,k)^2} \ ,$$

which can be as small as desired by taking k as large as necessary. Therefore in any neighbourhood of the origin, there are an infinite number of points for which the function $\exp(1/z)$ takes as value a given nonzero complex number.

2.4 The residue calculus and its applications

We now start the final section of this part of the course. It is the culmination of a lot of hard work and formalism but one which is worth the effort and the time spent developing the necessary vocabulary. In this section we will study the theory of residues. The theory itself is very simple and is basically a matter of applying what we have learned already in the appropriate way. Most of the sections are applications of the theory to the computation of real integrals and infinite sums. These are problems which are simple to state in the context of real calculus but whose solutions (at least the elementary ones) take us to the complex plane. In a sense they provide the simplest instance of a celebrated phrase by the French mathematician Hadamard, who said that the shortest path between two real truths often passes by a complex domain.

2.4.1 The Cauchy Residue Theorem

Let us consider the behaviour of an analytic function around an isolated singularity. To be precise let z_0 be an isolated singularity for an analytic

function $f(z)$. The function is analytic in some punctured disk $0 < |z - z_0| < R$, for some $R > 0$, and has a Laurent series there of the form

$$f(z) = \sum_{j=-\infty}^{\infty} a_j (z - z_0)^j \ .$$

Consider the contour integral of the function $f(z)$ along a positively oriented loop Γ contained in the punctured disk and having the singularity z_0 in its interior. Because the Laurent series converges uniformly, we can integrate the series term by term:

$$\oint_\Gamma f(z)\, dz = \sum_{j=-\infty}^{\infty} a_j \oint_\Gamma (z - z_0)^j\, dz \ .$$

From the (generalised) Cauchy Integral Formula or simply by deforming the contour to a circle of radius $\rho < R$, we have that (c.f., equation (2.52))

$$\oint_\Gamma (z - z_0)^j\, dz = \begin{cases} 2\pi\, i & \text{for } j = -1, \text{ and} \\ 0 & \text{otherwise;} \end{cases}$$

whence only the $j = -1$ term contributes to the sum, so that

$$\oint_\Gamma f(z)\, dz = 2\pi\, i\, a_{-1} \ .$$

This singles out the coefficient a_{-1} in the Laurent series, and hence we give it a special name. We say that a_{-1} is the **residue** of f at z_0, and we write this as $\mathrm{Res}(f; z_0)$ or simply as $\mathrm{Res}(z_0)$ when f is understood.

For example, consider the function $z \exp(1/z)$. This function has an essential singularity at the origin and is analytic everywhere else. The residue can be computed from the Laurent series:

$$z e^{1/z} = z \sum_{j=0}^{\infty} \frac{1}{j!}\, \frac{1}{z^j} = \sum_{j=0}^{\infty} \frac{1}{j!}\, \frac{1}{z^{j-1}} = z + 1 + \frac{1}{2z} + \cdots \ ,$$

whence the residue is given by $\mathrm{Res}(0) = \frac{1}{2}$.

It is often not necessary to calculate the Laurent expansion in order to extract the residue of a function at a singularity. For example, the residue of a function at a removable singularity vanishes, since there are no negative powers in the Laurent expansion. On the other hand, if the singularity happens to be a pole, we will see that the residue can be computed by differentiation.

Suppose, for simplicity, that $f(z)$ has a simple pole at z_0. Then the Laurent series of $f(z)$ around z_0 has the form

$$f(z) = \frac{a_{-1}}{z - z_0} + a_0 + a_1(z - z_0) + \cdots \ ,$$

whence the residue can be computed by

$$\begin{aligned} \operatorname{Res}(f; z_0) &= \lim_{z \to z_0} (z - z_0)\, f(z) \\ &= \lim_{z \to z_0} \left(a_{-1} + a_0(z - z_0) + a_1(z - z_0)^2 + \cdots \right) \\ &= a_{-1} + 0 \ . \end{aligned}$$

For example, the function $f(z) = e^z / z(z+1)$ has simple poles at $z = 0$ and $z = -1$; therefore,

$$\begin{aligned} \operatorname{Res}(f; 0) &= \lim_{z \to 0} z\, f(z) = \lim_{z \to 0} \frac{e^z}{z+1} = 1 \\ \operatorname{Res}(f; -1) &= \lim_{z \to -1} (z+1)\, f(z) = \lim_{z \to -1} \frac{e^z}{z} = -\frac{1}{e} \ . \end{aligned}$$

Suppose that $f(z) = P(z)/Q(z)$ where P and Q are analytic at z_0 and Q has a simple zero at z_0 whereas $P(z_0) \neq 0$. Clearly f has a simple pole at z_0, whence the residue is given by

$$\operatorname{Res}(f; z_0) = \lim_{z \to z_0} (z - z_0) \frac{P(z)}{Q(z)} = \lim_{z \to z_0} \frac{P(z)}{\frac{Q(z) - Q(z_0)}{z - z_0}} = \frac{P(z_0)}{Q'(z_0)} \ ,$$

where we have used that $Q(z_0) = 0$ and the definition of the derivative, which exists since Q is analytic at z_0.

We can use this to compute the residues at each singularity of the function $f(z) = \cot z$. Since $\cot z = \cos z / \sin z$, the singularities occur at the zeros of the sine function: $z = n\pi$, $n = 0, \pm 1, \pm 2, \ldots$. These zeros are simple because $\sin'(n\pi) = \cos(n\pi) = (-1)^n \neq 0$. Therefore we can apply the above formula to deduce that

$$\operatorname{Res}(f; n\pi) = \left. \frac{\cos z}{(\sin z)'} \right|_{z = n\pi} = \frac{\cos(n\pi)}{\cos(n\pi)} = 1 \ .$$

This result will be crucial for the applications concerning infinite series later on in this section.

Now suppose that f has a pole of order m at z_0. The Laurent expansion is then

$$f(z) = \frac{a_{-m}}{(z - z_0)^m} + \cdots + \frac{a_{-1}}{z - z_0} + a_0 + a_1(z - z_0) + \cdots$$

Let us multiply this by $(z-z_0)^m$ to obtain

$$(z-z_0)^m\, f(z) = a_{-m} + \cdots + a_{-1}(z-z_0)^{m-1} + a_0(z-z_0)^m + \cdots \ ,$$

whence taking $m-1$ derivatives, we have

$$\frac{d^{m-1}}{dz^{m-1}}\left[(z-z_0)^m\, f(z)\right] = (m-1)!\, a_{-1} + m!\, a_0(z-z_0) + \cdots \ .$$

Finally if we evaluate this at $z=z_0$, we obtain $(m-1)!\, a_{-1}$, which then gives a formula for the residue of f at a pole of order m:

$$\boxed{\mathrm{Res}(f;z_0) = \lim_{z\to z_0} \frac{1}{(m-1)!}\frac{d^{m-1}}{dz^{m-1}}\left[(z-z_0)^m\, f(z)\right] \ .} \tag{2.56}$$

For example, let us compute the residues of the function

$$f(z) = \frac{\cos z}{z^2(z-\pi)^3} \ .$$

This function has a pole of order 2 at the origin and a pole of order 3 at $z=\pi$. Therefore, applying the above formula, we find

$$\begin{aligned}
\mathrm{Res}(f;0) &= \lim_{z\to 0}\frac{1}{1!}\frac{d}{dz}\left[z^2 f(z)\right] \\
&= \lim_{z\to 0}\frac{d}{dz}\left[\frac{\cos z}{(z-\pi)^3}\right] \\
&= \lim_{z\to 0}\left[\frac{-\sin z}{(z-\pi)^3} - \frac{3\cos z}{(z-\pi)^4}\right] \\
&= -\frac{3}{\pi^4} \ , \\
\mathrm{Res}(f;\pi) &= \lim_{z\to \pi}\frac{1}{2!}\frac{d^2}{dz^2}\left[(z-\pi)^3 f(z)\right] \\
&= \lim_{z\to \pi}\frac{1}{2}\frac{d^2}{dz^2}\left[\frac{\cos z}{z^2}\right] \\
&= \lim_{z\to \pi}\frac{1}{2}\left[\frac{6\cos z}{z^4} + \frac{4\sin z}{z^3} - \frac{\cos z}{z^2}\right] \\
&= -\frac{6-\pi^2}{2\pi^4} \ .
\end{aligned}$$

We are now ready to state the main result of this section, which concerns the formula for the integral of a function $f(z)$ which is analytic on a

positively-oriented loop Γ and has only a finite number of isolated singularities $\{z_k\}$ in the interior of the loop. Because of the analyticity of the function, and using a contour deformation argument, we can express the integral of $f(z)$ along Γ as the sum of the integrals of $f(z)$ along positively-oriented loops Γ_k, each one encircling one of the isolated singularities. But we have just seen that the integral along each of these loops is given by $2\pi\,i$ times the residue of the function at the singularity. In other words, we have

$$\oint_\Gamma f(z)\,dz = \sum_k \oint_{\Gamma_k} f(z)\,dz = \sum_k 2\pi\,i\ \mathrm{Res}(f;z_k)\ .$$

In other words, we arrive at the **Cauchy Residue Theorem**, which states that the integral of $f(z)$ along Γ is equal to $2\pi\,i$ times the sum of the residues of the singularities in the interior of the contour:

$$\boxed{\oint_\Gamma f(z)\,dz = 2\pi\,i \sum_{\substack{\text{singularities}\\ z_k\in \mathrm{Int}\,\Gamma}} \mathrm{Res}(f;z_k)\ .}$$

For example, let us compute the integral

$$\oint_\Gamma \frac{1-2z}{z(z-1)(z-3)}\,dz\ ,$$

along the positively oriented circle of radius 2: $|z|=2$. The integrand $f(z)$ has simple poles at $z=0$, $z=1$ and $z=2$, but only the first two lie in the interior of the contour. Thus by the residue theorem,

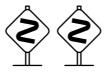
$$\oint_\Gamma \frac{1-2z}{z(z-1)(z-3)}\,dz = 2\pi\,i\,[\mathrm{Res}(f;0)+\mathrm{Res}(f;1)]\ ,$$

and

$$\begin{aligned}
\mathrm{Res}(f;0) &= \lim_{z\to 0} z\,f(z)\\
&= \lim_{z\to 0} \frac{(1-2z)}{(z-1)(z-3)}\\
&= \frac{1}{3}\ ,\\
\mathrm{Res}(f;1) &= \lim_{z\to 1}(z-1)\,f(z)\\
&= \lim_{z\to 1} \frac{(1-2z)}{z(z-3)}\\
&= \frac{1}{2}\ ;
\end{aligned}$$

so that

$$\oint_\Gamma \frac{1-2z}{z(z-1)(z-3)}\,dz = 2\pi\,i\left(\frac{1}{3}+\frac{1}{2}\right) = \frac{5\pi\,i}{3}\ .$$



Notice something curious. Computing the residue at $z=3$, we find,

$$\begin{aligned}\mathrm{Res}(f;3) &= \lim_{z\to 3}(z-3)\,f(z)\\ &= \lim_{z\to 3}\frac{(1-2z)}{z(z-1)}\\ &= \frac{-5}{6}\ ;\end{aligned}$$

whence the sum of all three residues is 0. This can be explained by introducing the Riemann sphere model for the extended complex plane, and thus noticing that a contour which would encompass all three singularities can be deformed to surround the point at infinity but in the opposite sense. Since the integrand is analytic at infinity, the Cauchy Integral Theorem says that the integral is zero, but (up to factors) this is equal to the sum of the residues.

2.4.2 Application: trigonometric integrals

The first of the applications of the residue theorem is to the computation of trigonometric integrals of the form

$$\int_0^{2\pi} R(\cos\theta,\sin\theta)\,d\theta\ ,$$

where R is a rational function of its arguments and such that it is finite in the range $0\le\theta\le 2\pi$. We want to turn this into a complex contour integral so that we can apply the residue theorem. One way to do this is the following. Consider the contour Γ parametrised by $z=\exp(i\theta)$ for $\theta\in[0,2\pi]$: this is the unit circle traversed once in the positive sense. On this contour, we have $z=\cos\theta+i\,\sin\theta$ and $1/z=\cos\theta-i\,\sin\theta$. Therefore we can solve for $\cos\theta$ and $\sin\theta$ in terms of z and $1/z$ as follows:

$$\cos\theta=\frac{1}{2}\left(z+\frac{1}{z}\right)\qquad\text{and}\qquad\sin\theta=\frac{1}{2i}\left(z-\frac{1}{z}\right)\ .$$

Similarly, $dz=d\exp(i\theta)=iz\,d\theta$, whence $d\theta=\frac{dz}{iz}$. Putting it all together we have that

$$\int_0^{2\pi} R(\cos\theta,\sin\theta)\,d\theta=\oint_\Gamma\frac{1}{iz}R\left(\frac{z+\frac{1}{z}}{2},\frac{z-\frac{1}{z}}{2i}\right)\,dz\ ,$$

which is the contour integral of a rational function of z, and hence can be computed using the residue theorem:

$$\int_0^{2\pi} R(\cos\theta, \sin\theta)\, d\theta = 2\pi \sum_{\substack{\text{singularities} \\ |z_k|<1}} \text{Res}(f; z_k) \ , \tag{2.57}$$

where $f(z)$ is the rational function

$$f(z) \equiv \frac{1}{z} R\left(\frac{z + \frac{1}{z}}{2}, \frac{z - \frac{1}{z}}{2i} \right) \ . \tag{2.58}$$

As an example, let us compute the integral

$$I = \int_0^{2\pi} \frac{(\sin\theta)^2}{5 + 4\cos\theta}\, d\theta \ .$$

First of all notice that the denominator never vanishes, so that we can go ahead. The rational function $f(z)$ given in (2.58) is

$$f(z) = \frac{1}{z} \frac{\left(\frac{1}{2i}\left(z - \frac{1}{z}\right)\right)^2}{5 + 4\frac{1}{2}\left(z + \frac{1}{z}\right)} = -\frac{1}{4} \frac{(z^2-1)^2}{z^2(2z^2+5z+2)} = -\frac{1}{8} \frac{(z^2-1)^2}{z^2(z+\frac{1}{2})(z+2)} \ ,$$

whence it has a double pole at $z = 0$ and single poles at $z = -\frac{1}{2}$ and $z = -2$. Of these, only the poles at $z = 0$ and $z = -\frac{1}{2}$ lie inside the unit disk, whence

$$I = 2\pi \ \left[\text{Res}(f; 0) + \text{Res}(f; -\tfrac{1}{2}) \right] \ .$$

Let us compute the residues. The singularity at $z = 0$ is a pole of order 2, whence by equation (2.56), we have

$$\text{Res}(f; 0) = \lim_{z\to 0} \frac{d}{dz} \left[-\frac{1}{8} \frac{(z^2-1)^2}{(z+\frac{1}{2})(z+2)} \right] = \frac{5}{16} \ .$$

The pole at $z = -\frac{1}{2}$ is simple, so that its residue is even simpler to compute:

$$\text{Res}(f; -\mathbb{1}/2) = \lim_{z\to -\frac{1}{2}} \left[-\frac{1}{8} \frac{(z^2-1)^2}{z^2(z+2)} \right] = -\frac{3}{16} \ .$$

Therefore, the integral becomes

$$I = 2\pi \left(\frac{5}{16} - \frac{3}{16} \right) = \frac{\pi}{4} \ .$$

As a mild check on our result, we notice that it is real whence it is not obviously wrong.

Let us do another example:

$$I = \int_0^\pi \frac{d\theta}{2 - \cos\theta} \ .$$

This time the integral is only over $[0, \pi]$, so that we cannot immediately use the residue theorem. However in this case we notice that because $\cos(2\pi - \theta) = \cos\theta$, we have that

$$\int_\pi^{2\pi} \frac{d\theta}{2 - \cos\theta} = \int_\pi^0 \frac{d(2\pi - \theta)}{2 - \cos(2\pi - \theta)} = -\int_\pi^0 \frac{d\theta}{2 - \cos\theta} = \int_0^\pi \frac{d\theta}{2 - \cos\theta} \ .$$

Therefore,

$$I = \tfrac{1}{2} \int_0^{2\pi} \frac{d\theta}{2 - \cos\theta} \ ,$$

which using equation (2.57) and paying close attention to the factor of $\frac{1}{2}$, becomes π times the sum of the residues of the function

$$f(z) = \frac{1}{z} \frac{1}{2 - \frac{1}{2}(z + \frac{1}{z})}$$

lying inside the unit disk. After a little bit of algebra, we find that

$$f(z) = -\frac{2}{z^2 - 4z + 1} = -\frac{2}{(z - 2 + \sqrt{3})(z - 2 - \sqrt{3})} \ .$$

Of the two simple poles of this function only the one at $z = 2 - \sqrt{3}$ lies inside the unit disk, hence

$$\mathrm{Res}(f; 2 - \sqrt{3}) = \lim_{z \to 2-\sqrt{3}} \frac{-2}{z - 2 - \sqrt{3}} = \frac{1}{\sqrt{3}} \ ,$$

and thus the integral becomes

$$I = \frac{\pi}{\sqrt{3}} \ .$$

2.4.3 Application: improper integrals

In this section we consider improper integrals of rational functions and of products of rational and trigonometric functions.

Let $f(x)$ be a function of a real variable, which is continuous in $0 \leq x < \infty$. Then by the improper integral $\int_0^\infty f(x)\,dx$, we mean the limit

$$\int_0^\infty f(x)\,dx \equiv \lim_{R\to\infty} \int_0^R f(x)\,dx \ ,$$

if such a limit exists. Similarly, if $f(x)$ is continuous in $-\infty < x \leq 0$, then the improper integral $\int_{-\infty}^0 f(x)\,dx$ is defined by the limit

$$\int_{-\infty}^0 f(x)\,dx \equiv \lim_{r\to-\infty} \int_r^0 f(x)\,dx \ ,$$

again provided that it exists. If $f(x)$ is continuous on the whole real line and both of the above limits exists, we define

$$\int_{-\infty}^\infty f(x)\,dx \equiv \lim_{\substack{R\to\infty \\ r\to-\infty}} \int_r^R f(x)\,dx \ . \tag{2.59}$$

If such limits exist, then we get the same result by symmetric integration:

$$\int_{-\infty}^\infty f(x)\,dx = \lim_{\rho\to\infty} \int_{-\rho}^\rho f(x)\,dx \ . \tag{2.60}$$

Notice however that the symmetric integral may exist even if the improper integral (2.59) does not. For example consider the function $f(x) = x$. Clearly the integrals $\int_0^\infty x\,dx$ and $\int_{-\infty}^0 x\,dx$ do not exist, yet because x is an odd function, $\int_{-\rho}^\rho x dx = 0$ for all ρ, whence the limit is 0. In cases like this we say that equation (2.60) defines the **Cauchy principal value** of the integral, and we denote this by

$$\text{p. v.} \int_{-\infty}^\infty f(x)\,dx \equiv \lim_{\rho\to\infty} \int_{-\rho}^\rho f(x)\,dx \ .$$

We stress to point out that whenever the improper integral (2.59) exists it agrees with its principal value (2.60).

Improper integrals of rational functions over $(-\infty, \infty)$

Let us consider as an example the improper integral

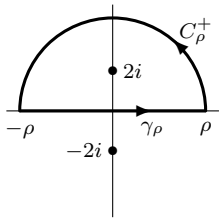
$$I = \text{p. v.} \int_{-\infty}^\infty \frac{dx}{x^2+4} = \lim_{\rho\to\infty} \int_{-\rho}^\rho \frac{dx}{x^2+4} \ .$$

The integral for finite ρ can be interpreted as the complex integral of the function $f(z) = 1/(z^2+4)$,

$$\int_{\gamma_\rho} \frac{dz}{z^2+4} \ ,$$

where γ_ρ is the straight line segment on the real axis: $y = 0$ and $-\rho \le x \le \rho$. In order to use the residue theorem we need to *close the contour*; that is, we must produce a closed contour along which we can apply the residue theorem. Ot f course, in so doing we are introducing a further integral, and the success of the method depends on whether the extra integral is computable. We will see that in this case, the extra integral, if chosen judiciously, vanishes.

Let us therefore complete the contour γ_ρ to a closed contour. One suggestion is to consider the semicircular contour C_ρ^+ in the upper half plane, parametrised by $z(t) = \rho \exp(it)$, for $t \in [0, \pi]$. Let Γ_ρ be the composition of both contours: it is a closed contour as shown in the figure. Then, according to the residue theorem,



$$\int_{\Gamma_\rho} \frac{dz}{z^2+4} = \int_{\gamma_\rho} \frac{dz}{z^2+4} + \int_{C_\rho^+} \frac{dz}{z^2+4} = 2\pi\, i \sum_{\substack{\text{singularities} \\ z_k \in \text{Int}\,\Gamma_\rho}} \text{Res}(f; z_k) \ ;$$

whence

$$\int_{\gamma_\rho} \frac{dz}{z^2+4} = 2\pi\, i \sum_{\substack{\text{singularities} \\ z_k \in \text{Int}\,\Gamma_\rho}} \text{Res}(f; z_k) - \int_{C_\rho^+} \frac{dz}{z^2+4} \ .$$

We will now argue that the integral along C_ρ^+ vanishes in the limit $\rho \to \infty$. Of course, this is done using (2.28):

$$\left| \int_{C_\rho^+} \frac{dz}{z^2+4} \right| \le \int_{C_\rho^+} \frac{|dz|}{|z^2+4|} \ . \tag{2.61}$$

Using the triangle inequality (2.36), we have that on C_ρ^+,

$$|z^2+4| \ge |z^2| - 4 = |z|^2 - 4 = \rho^2 - 4 \ ,$$

whence

$$\frac{1}{|z^2+4|} \le \frac{1}{\rho^2-4} \ .$$

Plugging this into (2.61), and taking into account that the length of the semicircle C_ρ^+ is $\pi\rho$,

$$\left| \int_{C_\rho^+} \frac{dz}{z^2+4} \right| \le \frac{\pi\rho}{\rho^2-4} \to 0 \qquad \text{as } \rho \to \infty.$$

Therefore in the limit,

$$\int_{\gamma_\rho} \frac{dz}{z^2+4} = 2\pi\, i \sum_{\substack{\text{singularities} \\ z_k \in \text{Int}\,\Gamma_\rho}} \text{Res}(f; z_k) \ .$$

The function $f(z)$ has poles at $z = \pm 2i$, of which only the one at $z = 2i$ lies inside the closed contour Γ_ρ, for large ρ. Computing the residue there, we find from (2.56) that

$$\text{Res}(f; 2i) = \lim_{z \to 2i} \left[\frac{1}{z+2i} \right] = \frac{1}{4i} \ ,$$

and hence the integral is given by

$$I = 2\pi\, i\, \frac{1}{4i} = \frac{\pi}{2} \ .$$

There is no reason why we chose to close the contour using the top semicircle C_ρ^+ instead of using the bottom semicircle C_ρ^- parametrised by $z(t) = \rho \exp(it)$ for $t \in [\pi, 2\pi]$. The same argument shows that in the limit $\rho \to \infty$ the integral along C_ρ^- vanishes. It is now the pole at $-2i$ that we have to take into account, and one has that $\text{Res}(f; -2i) = -1/4i$. Notice however that the closed contour is negatively-oriented, which produces an extra $-$ sign from the residue formula, in such a way that the final answer is again

$$I = -2\pi\, i\, \frac{-1}{4i} = \frac{\pi}{2} \ .$$

The technique employed in the calculation of the above integral can be applied in more general situations. All that we require is for the integral along the large semicircle C_ρ^+ to vanish and this translates into a condition on the behaviour of the integrand for large $|z|$.

We will now show the following general result. Let $R(x) = P(x)/Q(x)$ be a rational function of a real variable satisfying the following two criteria:

- $Q(x) \neq 0$; and
- $\deg Q - \deg P \geq 2$.

Then the improper integral of $R(x)$ along the real line is given by considering the residues of the complex rational function $R(z)$ at its singularities in the upper half-plane. Being a rational function the only singularities are either removable or poles, and only these latter ones contribute to the residue. In summary,

$$\boxed{\text{p. v.} \int_{-\infty}^{\infty} R(x)\, dx = 2\pi\, i \sum_{\substack{\text{poles } z_k \\ \text{Im}(z_k) > 0}} \text{Res}(R; z_k) \ .} \tag{2.62}$$

The proof of this relation follows the same steps as in the computation of the integral I above. The trick is to close the contour using the upper semicircle C_ρ^+ and then argue that the integral along the semicircle vanishes. This is guaranteed by the the behaviour of $R(z)$ for large $|z|$.

Let us do this in detail. The integral to be computed is

$$I = \text{p. v.} \int_{-\infty}^{\infty} R(x)\, dx = \lim_{\rho\to\infty} \int_{-\rho}^{\rho} R(x)\, dx = \lim_{\rho\to\infty} \int_{\gamma_\rho} R(z)\, dz \ .$$

Closing the contour with C_ρ^+ to Γ_ρ, we have

$$\int_{\gamma_\rho} R(z)\, dz = \oint_{\Gamma_\rho} R(z)\, dz - \int_{C_\rho^+} R(z)\, dz \ .$$

The first integral in the right-hand side can be easily dispatched using the residue theorem. In the limit $\rho \to \infty$, one finds

$$\lim_{\rho\to\infty} \oint_{\Gamma_\rho} R(z)\, dz = 2\pi\, i \sum_{\substack{\text{poles } z_k \\ \text{Im}(z_k)>0}} \text{Res}(R; z_k) \ .$$

All that remains then is to show that the second integral vanishes in the limit $\rho \to \infty$. We can estimate it using (2.28) as usual:

$$\int_{C_\rho^+} R(z)\, dz \ \leq \int_{C_\rho^+} |R(z)|\, |dz| = \int_{C_\rho^+} \frac{|P(z)|}{|Q(z)|}\, |dz| \ .$$

Let the degree of the polynomial $P(z)$ be p and that of $Q(z)$ be q, where by hypothesis we have that $q - p \geq 2$. Recall from our discussion in Section 2.2.6 that for large $|z|$ a polynomial $P(z)$ of degree N behaves like $|P(z)| \sim c|z|^N$ for some c. Similar considerations in this case show that the rational function $R(z) = P(z)/Q(z)$ with $q = \deg Q > \deg P = p$ obeys

$$|R(z)| \leq \frac{c}{|z|^{q-p}} \ ,$$

for some constant c independent of $|z|$. Using this into the estimate of the integral along C_ρ^+, and using that the semicircle has length $\pi\rho$,

$$\int_{C_\rho^+} R(z)\, dz \ \leq \frac{c\pi\rho}{\rho^{q-p}} \ .$$

Since $q - p \geq 2$, we have that this goes to zero in the limit $\rho \to \infty$, as desired.

As an example, let us compute the following integral

$$I = \text{p. v.} \int_{-\infty}^{\infty} \frac{x^2}{(x^2+1)^2}\, dx \ .$$

The integrand is rational and obeys the two criteria above: it is always finite and the degree of the denominator is 4 whereas that of the numerator is 2,

whence $4 - 2 \geq 2$. In order to compute the integral it is enough to compute the residues of the rational function

$$f(z) = \frac{z^2}{(z^2+1)^2} \ ,$$

at the poles in the upper half-plane. This function has poles of order 2 at the points $z = \pm i$, of which only $z = +i$ is in the upper half-plane, hence from (2.62) we have

$$I = 2\pi\, i \lim_{z\to i} \frac{d}{dz} \left[\frac{z^2}{(z+i)^2}\right] = 2\pi\, i \lim_{z\to i} \left[\frac{2i\, z}{(z+i)^3}\right] = 2\pi\, i\, \frac{-i}{4} = \frac{\pi}{2} \ .$$

Improper integrals of rational and trigonometric functions

The next type of integrals which can be be handled by the method of residues are of the kind

$$\text{p. v.} \int_{-\infty}^{\infty} R(x)\cos(ax)\, dx \qquad \text{and} \qquad \text{p. v.} \int_{-\infty}^{\infty} R(x)\sin(ax)\, dx \ ,$$

where $R(x)$ is a rational function which is continuous everywhere in the real line (except maybe at the zeros of $\cos(ax)$ and $\sin(ax)$, depending on the integral), and where a is a nonzero real number.

As an example, consider the integral

$$I = \text{p. v.} \int_{-\infty}^{\infty} \frac{\cos(3x)}{x^2+4}\, dx = \lim_{\rho\to\infty} \int_{-\rho}^{\rho} \frac{\cos(3x)}{x^2+4}\, dx \ .$$

From the discussion in the previous section, we are tempted to try to express the integral over $[-\rho, \rho]$ as a complex contour integral, close the contour and use the residue theorem. Notice however that we cannot use the function $\cos(3z)/(z^2+4)$ because $|\cos(3z)|$ is not bounded for large values of $|\operatorname{Im}(z)|$. Instead we notice that we can write the integral as the real part of a complex integral $I = \operatorname{Re}(I_0)$, where

$$I_0 = \lim_{\rho\to\infty} \int_{-\rho}^{\rho} \frac{e^{i3x}}{x^2+4}\, dx \ .$$

Therefore let us consider the integral

$$\int_{-\rho}^{\rho} \frac{e^{i3x}}{x^2+4}\, dx = \int_{\gamma_\rho} \frac{e^{i3z}}{z^2+4}\, dz \ ,$$

where γ_ρ is the line segment on the real axis from $-\rho$ to ρ. We would like to close this contour to be able to use the residue theorem, and in such a way that the integral vanishes on the extra segment that we must add to close it. Let us consider the upper semicircle C_ρ^+. There we have that

$$\left|\frac{e^{i3z}}{z^2+4}\right| = \frac{e^{-3\,\mathrm{Im}(z)}}{|z^2+4|} \leq \frac{e^{-3\,\mathrm{Im}(z)}}{\rho^2-4} \ ,$$

where to reach the inequality we used (2.36) as was done above. The function $e^{-3\,\mathrm{Im}(z)}$ is bounded above by 1 in the *upper* half-plane, and in particular along C_ρ^+, hence we have that on the semicircle,

$$\left|\frac{e^{i3z}}{z^2+4}\right| \leq \frac{1}{\rho^2-4} \ .$$

Therefore the integral along the semicircle is bounded above by

$$\left|\int_{C_\rho^+} \frac{e^{i3z}}{z^2+4}\,dz\right| \leq \frac{\pi\rho}{\rho^2-4} \to 0 \qquad \text{as } \rho\to\infty.$$

Therefore we can use the residue theorem to express I_0 in terms of the residues of the function $f(z) = \exp(i3z)/(z^2+4)$ at the poles in the upper half-plane. This function has simple poles at $z = \pm 2i$, but only $z = 2i$ lies in the upper half-plane, whence

$$I_0 = 2\pi\,i\ \mathrm{Res}(f;2i) = 2\pi\,i\ \lim_{z\to 2i}\left[\frac{e^{i3z}}{z+2i}\right] = 2\pi\,i\,\frac{e^{-6}}{4i} = \frac{\pi}{2e^6} \ ,$$

which is already real. (One could have seen this because the imaginary part is the integral of $\sin(3x)/(x^2+4)$ which is an odd function and hence integrates to zero under symmetric integration.) Therefore,

$$I = \mathrm{Re}(I_0) = \frac{\pi}{2e^6} \ .$$

Suppose instead that we had wanted to compute the integral

$$\mathrm{p.\,v.}\int_{-\infty}^{\infty} \frac{e^{-i3x}}{x^2+4}\,dx \ .$$

Of course, now we could do it because this is the complex conjugate of the integral we have just computed, but let us assume that we had not yet done the other integral. We would follow the same steps as before, but notice that now,

$$\left|\frac{e^{-i3z}}{z^2+4}\right| = \frac{e^{3\,\mathrm{Im}(z)}}{|z^2+4|} \ ,$$

which is no longer bounded in the upper half-plane. In this case we would be forced the close the contour using the lower semicircle C_ρ^-, keeping in mind that the closed contour is now negatively oriented. The lesson to learn from this is that there is some choice in how to close to contour and that one has to exercise this choice judiciously for the calculation to work out.

This method of course generalises to compute integrals of the form

$$\text{p. v.} \int_{-\infty}^{\infty} R(x) e^{iax}\, dx \tag{2.63}$$

where a is real. Surprisingly the conditions on the rational function $R(x)$ are now slightly weaker. Indeed, we have the following general result.

Let $R(x) = P(x)/Q(x)$ be a rational function satisfying the following conditions:

- $Q(x) \neq 0$,[3] and
- $\deg Q - \deg P \geq 1$.

Then the improper integral (2.63) is given by considering the residues of the function $f(z) = R(z)e^{iaz}$ at its singularities in the upper (if $a > 0$) or lower (if $a < 0$) half-planes. These singularities are either removable or poles, and again only the poles contribute to the residues. In summary,

$$\boxed{\text{p. v.} \int_{-\infty}^{\infty} R(x) e^{iax}\, dx = \begin{cases} 2\pi\, i \sum_{\substack{\text{poles } z_k \\ \text{Im}(z_k)>0}} \text{Res}(f; z_k) & \text{if } a > 0; \\ -2\pi\, i \sum_{\substack{\text{poles } z_k \\ \text{Im}(z_k)<0}} \text{Res}(f; z_k) & \text{if } a < 0. \end{cases}} \tag{2.64}$$

This result is similar to (2.62) with two important differences. The first is that we have to choose the contour appropriately depending on the integrand; that is, depending on the sign of a. The second one is that the condition on the rational function is less restrictive than before: now we simply demand that the degree of Q be greater than the degree of P. This will therefore require a more refined estimate of the integral along the semicircle, which goes by the name of the **Jordan lemma**, which states that

$$\lim_{\rho\to\infty} \int_{C_\rho^+} e^{iaz} \frac{P(z)}{Q(z)}\, dz = 0\ ,$$

whenever $a > 0$ and $\deg Q > \deg P$. Of course an analogous result holds for $a < 0$ and along C_ρ^-.

[3]This could in principle be relaxed provided the zeros of Q at most gave rise to removable singularities in the integrand.

Let us prove this lemma. Parametrise the semicircle C_ρ^+ by $z(t) = \rho \exp(it)$ for $t \in [0, \pi]$. Then by (2.25)

$$\int_{C_\rho^+} e^{iaz} \frac{P(z)}{Q(z)}\, dz = \int_0^\pi e^{ia\rho e^{it}} \frac{P(\rho e^{it})}{Q(\rho e^{it})}\, \rho\, i\, dt\ .$$

Let us now estimate the integrand term by term. First we have that

$$e^{ia\rho e^{it}} \ = \ e^{ia\rho(\cos t + i \sin t)} \ = e^{-a\rho \sin t}\ .$$

Similarly, since $\deg Q - \deg P \geq 1$, we have that

$$\frac{P(\rho e^{it})}{Q(\rho e^{it})} \ \leq \frac{c}{\rho}\ ,$$

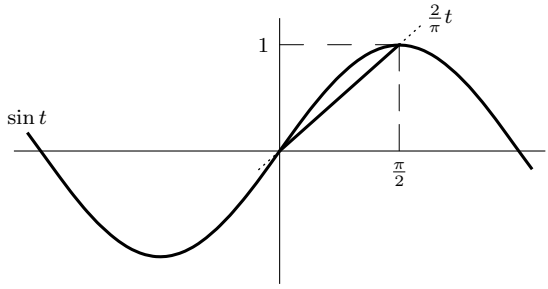
for ρ large, for some $c > 0$. Now using (2.24) on the t-integral together with the above (in)equalities,

$$\int_{C_\rho^+} e^{iaz} \frac{P(z)}{Q(z)}\, dz \ = \ \int_0^\pi e^{ia\rho e^{it}} \frac{P(\rho e^{it})}{Q(\rho e^{it})}\, \rho\, i\, dt \ \leq c \int_0^\pi e^{-a\rho \sin t}\, dt\ .$$

We need to show that this latter integral goes to zero in the limit $\rho \to \infty$. First of all notice that $\sin t = \sin(\pi - t)$ for $t \in [0, \pi]$, whence

$$\int_0^\pi e^{-a\rho \sin t}\, dt = 2 \int_0^{\pi/2} e^{-a\rho \sin t}\, dt\ .$$

Next notice that for $t \in [0, \pi/2]$, $\sin t \geq 2t/\pi$. This can be seen pictorially as in the following picture, which displays the function $\sin t$ in the range $t \in [-\pi, \pi]$ and the function $2t/\pi$ in the range $t \in [0, \pi/2]$ and makes the inequality manifest.



Therefore,

$$\int_0^{\pi/2} e^{-a\rho \sin t}\, dt \leq \int_0^\pi e^{-2a\rho\, t/\pi}\, dt = \frac{\pi}{2a\rho}\ \ 1 - e^{-2a\rho\pi} \quad .$$

Putting this all together, we see that

$$\int_{C_\rho^+} e^{iaz} \frac{P(z)}{Q(z)}\, dz \ \leq \frac{c\pi}{a\rho}\ \ 1 - e^{-2a\rho\pi} \quad ,$$

which clearly goes to 0 in the limit $\rho \to \infty$, proving the lemma.

As an example, let us compute the integral

$$I = \text{p. v.} \int_{-\infty}^{\infty} \frac{x \sin x}{1 + x^2}\, dx\ .$$

This is the imaginary part of the integral

$$I_0 = \text{p. v.} \int_{-\infty}^{\infty} \frac{x e^{ix}}{1+x^2}\, dx \ ,$$

which satisfies the conditions which permit the use of (2.64) with $a = 1$ and $R(z) = z/(1+z^2)$. This rational function has simple poles for $z = \pm i$, but only $z = i$ lies in the upper half-plane. According to (2.64) then, and letting $f(z) = R(z)e^{iz}$, we have

$$I_0 = 2\pi\, i\, \text{Res}(f; i) = 2\pi\, i \lim_{z\to i} \left[\frac{z\, e^{iz}}{z+i}\right] = 2\pi\, i \frac{ie^{-1}}{2i} = \frac{i\pi}{e} \quad \Longrightarrow \quad I = \frac{\pi}{e} \ .$$

Improper integrals of rational functions on $(0, \infty)$

The next type of integrals which can be tackled using the residue theorem are integrals of rational functions but over the half line; that is, integrals of the form:

$$\int_0^{\infty} R(x)\, dx \ ,$$

where $R(x)$ is continuous for $x \geq 0$. Of course, if $R(x)$ were an even function, i.e., $R(-x) = R(x)$, then we would have $\int_0^{\infty} R(x)\, dx = \frac{1}{2}\int_{-\infty}^{\infty} R(x)\, dx$, and we could use the method discussed previously. However for more general integrands, this does not work and we have to do something different.

The following general result is true. Let $R(x) = P(x)/Q(x)$ be a rational function of a real variable satisfying the following two conditions

- $Q(x) \neq 0$; and
- $\deg Q - \deg P \geq 2$.

Further let $f(z) = \log(z)\, R(z)$ with the branch of the logarithm chosen to be analytic at the poles $\{z_k\}$ of R; for example, we can choose the branch $\text{Log}_0(z)$ which has the cut along the positive real axis, since $Q(x)$ has no zeros there. Then,

$$\boxed{\int_0^{\infty} R(x)\, dx = -\sum_{\text{poles } z_k} \text{Res}(f; z_k) \ , \quad \text{for } f(z) = \log(z)\, R(z).} \tag{2.65}$$

The details.

This same method can be applied to integrals of the form

$$\int_a^\infty R(x)\,dx\ ,$$

where the rational function $R(x) = P(x)/Q(x)$ satisfies the same conditions as above except that now $Q(x) \neq 0$ only for $x \geq a$. In this case we must consider the function $f(z) = \log(z-a)\,R(z)$.

Details?

Similarly, since $\int_0^a = \int_0^\infty - \int_a^\infty$, we can use this method to compute *indefinite* integrals of rational functions.

2.4.4 Application: improper integrals with poles

Suppose that we want to compute the principal value integral

$$I = \text{p. v.}\int_{-\infty}^{\infty} \frac{\sin x}{x}\,dx\ .$$

This integral should converge: the singularity at $x = 0$ is removable, as we saw in equation (2.55), so that the integrand is continuous for all x, and the rational function $1/x$ satisfies the conditions of the Jordan Lemma. Following the ideas in the previous section, we would be write

$$I = \text{Im}(I_0) \qquad \text{where} \qquad I_0 = \text{p. v.}\int_{-\infty}^{\infty} \frac{e^{ix}}{x}\,dx\ , \tag{2.66}$$

and compute I_0. However notice that now the integrand of I_0 has a pole at $x = 0$. Until now we have always assumed that integrands have no poles along the contour, so the methods developed until now are not immediately applicable to perform the above integral. We therefore need to make sense out of integrals whose integrands are not continuous everywhere in the region of integration.

Let $f(x)$ be a function of a real variable, which is continuous in the interval $[a, b]$ except for a discontinuity at some point c, $a < c < b$. Then the **improper integrals** of f over the intervals $[a, c]$, $[c, b]$ and $[a, b]$ are defined by

$$\int_a^c f(x)\,dx \equiv \lim_{r \searrow 0} \int_a^{c-r} f(x)\,dx\ ,$$
$$\int_c^b f(x)\,dx \equiv \lim_{s \searrow 0} \int_{c+s}^b f(x)\,dx\ ,$$

and

$$\int_a^b f(x)\,dx \equiv \lim_{r\searrow 0}\int_a^{c-r} f(x)\,dx + \lim_{s\searrow 0}\int_{c+s}^b f(x)\,dx\ , \tag{2.67}$$

provided the appropriate limit(s) exist. We have used the notation $r \searrow 0$ to mean that r approaches 0 from above; that is, $r > 0$ as we take the limit. As an example, consider the function $1/\sqrt{x}$ integrated on $[0, 1]$:

$$\int_0^1 \frac{dx}{\sqrt{x}} = \lim_{s\searrow 0}\int_s^1 \frac{dx}{\sqrt{x}} = \lim_{s\searrow 0} 2\sqrt{x}\Big|_s^1 = \lim_{s\searrow 0}\left[2 - 2\sqrt{s}\right] = 2\ .$$

If the limits in (2.67) exist, then we can calculate the integral using symmetric integration, which defines the **principal value** of the integral,

$$\text{p. v.}\int_a^b f(x)\,dx \equiv \lim_{r\searrow 0}\left[\int_a^{c-r} f(x)\,dx + \int_{c+r}^b f(x)\,dx\right]\ .$$

However the principal value integral may exist even when the improper integral does not. Take, for instance,

$$\begin{aligned}\text{p. v.}\int_1^4 \frac{dx}{x-2} &= \lim_{r\searrow 0}\left[\int_1^{2-r} + \int_{2+r}^4\right]\frac{dx}{x-2}\\ &= \lim_{r\searrow 0}\left[\text{Log}\,|x-2|\Big|_1^{2-r} + \text{Log}\,|x-2|\Big|_{2+r}^4\right]\\ &= \lim_{r\searrow 0}\left[\text{Log}\,r + \text{Log}\,2 - \text{Log}\,r\right] = \text{Log}\,2\ ,\end{aligned}$$

whereas it is clear that the improper integral $\int_1^4 \frac{dx}{x-2}$ does not exist.

When the function $f(x)$ is continuous everywhere in the real line except at the point c we define the principal value integral by

$$\text{p. v.}\int_{-\infty}^{\infty} f(x)\,dx \equiv \lim_{\substack{\rho\to\infty\\ r\searrow 0}}\left[\int_{-\rho}^{c-r} f(x)\,dx + \int_{c+r}^{\rho} f(x)\,dx\right]\ , \tag{2.68}$$

provided the limits $\rho \to \infty$ and $r \searrow 0$ exist independently. In the case of several discontinuities $\{c_i\}$ we extend the definition of the improper integral in the obvious way: excising a small symmetric interval $(c_i - r_i, c_i + r_i)$ about each discontinuity and then taking the limits $r_i \searrow 0$ and, if applicable, $\rho \to \infty$.

It turns out that principal value integrals of this type can often be evaluated using the residue theorem. The residue theorem applies to closed contours, so in computing a principal value integral we need to close the contour, not just ρ to $-\rho$ as in the previous session, but also $c - r$ to $c + r$.

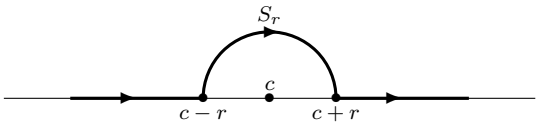


Figure 2.10: Closing the contour around a singularity.

One way to do this is to consider a small semicircle S_r of radius r around the singular point c, as in Figure 2.10.

Because we are interested in the limit $r \searrow 0$, we will have to consider the integral

$$\lim_{r \searrow 0} \int_{S_r} f(z)\, dz \ .$$

When the singularity of $f(z)$ at $z = c$ is a simple pole, this integral can be evaluated using the following result, which we state in some generality.

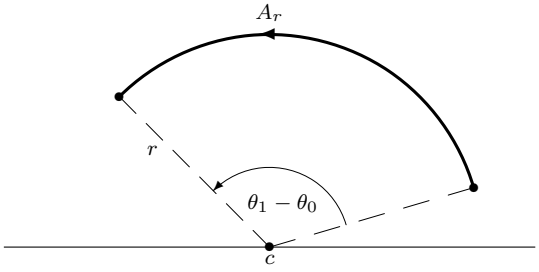


Figure 2.11: A small circular arc.

Let $f(z)$ have a simple pole at $z = c$ and let A_r be the circular arc in Figure 2.11, parametrised by $z(\theta) = c + r\exp(i\theta)$ with $\theta_0 \le \theta \le \theta_1$. Then

$$\lim_{r \searrow 0} \int_{A_r} f(z)\, dz = i\,(\theta_1 - \theta_0)\ \mathrm{Res}(f; c) \ .$$

Therefore for the semicircle S_r in Figure 2.10, we have

$$\lim_{r \searrow 0} \int_{S_r} f(z)\, dz = -i\,\pi\ \mathrm{Res}(f; c) \ . \tag{2.69}$$

Let us prove this result. Since $f(z)$ has a simple pole at c, its Laurent expansion in a punctured disk $0 < |z - c| < R$ has the form

$$f(z) = \frac{a_{-1}}{z - c} + \sum_{k=0}^{\infty} a_k (z - c)^k \ ,$$

where

$$g(z) \equiv \sum_{k=0}^{\infty} a_k (z-c)^k$$

defines an analytic function in the disk $|z-c| < R$. Now let $0 < r < R$ and consider the integral

$$\int_{A_r} f(z)\, dz = a_{-1} \int_{A_r} \frac{dz}{z-c} + \int_{A_r} g(z)\, dz \ .$$

Because $g(z)$ is analytic it is in particular bounded on some neighbourhood of c, so that $|g(z)| \le M$ for some M and all $|z-c| < R$. Then we can estimate its integral by using (2.28):

$$\int_{A_r} g(z)\, dz \ \le \int_{A_r} |g(z)|\, |dz| \le M\ell(A_r) = Mr(\theta_1 - \theta_0) \ ,$$

whence

$$\lim_{r \searrow 0} \int_{A_r} g(z)\, dz = 0 \ .$$

On the other hand,

$$\int_{A_r} \frac{dz}{z-c} = \int_{\theta_0}^{\theta_1} \frac{rie^{i\theta}}{re^{i\theta}}\, d\theta = i \int_{\theta_0}^{\theta_1} d\theta = i\ (\theta_1 - \theta_0) \ .$$

Therefore

$$\lim_{r \searrow 0} \int_{A_r} f(z)\, dz = i\,(\theta_1 - \theta_0) a_{-1} + 0 = i\,(\theta_1 - \theta_0)\ \mathrm{Res}(f; c) \ .$$

Having discussed the basic theory, let us go back to the original problem: the computation of the integral I_0 given in (2.66):

$$I_0 = \lim_{\substack{\rho \to \infty \\ r \searrow 0}} \left[\int_{-\rho}^{-r} \frac{e^{ix}}{x}\, dx + \int_{r}^{\rho} \frac{e^{ix}}{x}\, dx \right] \ ,$$

which for finite ρ and nonzero r can be understood as a contour integral in the complex plane along the subset of the real axis consisting of the intervals $[-\rho, -r]$ and $[r, \rho]$. In order to use the residue theorem we must close this contour. The Jordan lemma forces us to join ρ and $-\rho$ via a large semicircle C_ρ^+ of radius ρ in the *upper* half-plane. In order to join $-r$ and r we choose a small semicircle S_r also in the upper half-plane. The resulting closed contour is depicted in Figure 2.12.

Because the function is analytic on and inside the contour, the Cauchy Integral Theorem says that the contour integral vanishes. Splitting this contour integral into its different pieces, we have that

$$\left[\int_{-\rho}^{-r} + \int_{S_r} + \int_{r}^{\rho} + \int_{C_\rho^+} \right] \frac{e^{iz}}{z}\, dz = 0 \ ,$$

which remains true in the limits $\rho \to \infty$ and $r \searrow 0$. By the Jordan lemma, the integral along C_ρ^+ vanishes in the limit $\rho \to \infty$, whence, using (2.69),

$$I_0 = -\lim_{r \searrow 0} \int_{S_r} \frac{e^{iz}}{z}\, dz = \lim_{r \searrow 0} \int_{-S_r} \frac{e^{iz}}{z}\, dz = i\pi\ \mathrm{Res}(0) = i\pi \ ,$$

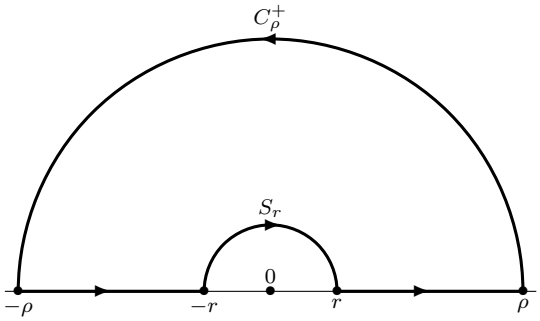


Figure 2.12: The contour in the calculation of I_0 in (2.66).

since the residue of e^{iz}/z at $z = 0$ is equal to 1. Therefore, we have that

$$\text{p. v.} \int_{-\infty}^{\infty} \frac{\sin x}{x}\, dx = \text{Im}(i\pi) = \pi \ .$$

There are plenty of other integrals which can be calculated using the residue theorem; e.g., integrals involving multi-valued functions. We will not have time to discuss them all, but the lesson to take home from this cursory introduction to residue techniques is that when faced with a real integral, one should automatically think of this as a parametrisation of a contour integral in the complex plane, where we have at our disposal the powerful tools of complex analysis.

2.4.5 Application: infinite series

The final section of this part of the course is a beautiful application of the theory of residues to the computation of infinite sums.

How can one use contour integration in order to calculate sums like the following one:

$$\sum_{n=1}^{\infty} \frac{1}{n^2} \ ? \tag{2.70}$$

The idea is to exhibit this sum as part of the right-hand side of the Cauchy Residue Theorem. For this we need a function $F(z)$ which has only simple poles at the integers and whose residue is 1 there. We already met a function which has an infinite number of poles which are integrally spaced: the function $\cot z$ has simple poles for $z = n\pi$, $n = 0, \pm 1, \pm 2, \ldots$ with residues equal to 1. Therefore the function $F(z) = \pi \cot(\pi z)$ has simple poles at $z = n$, n an integer, and the residue is still 1:

$$\text{Res}(F; n) = \lim_{z \to n} \frac{\pi \cos(\pi z)}{(\sin(\pi z))'} = \lim_{z \to n} \frac{\pi \cos(\pi z)}{\pi \cos(\pi z)} = 1 \ .$$

Now let $R(z) = P(z)/Q(z)$ be any rational function such that $\deg Q - \deg P \geq 2$. Consider the function $f(z) = \pi \cot(\pi z) R(z)$ and let us integrate this along the contour Γ_N, for N a positive integer, defined as the positively oriented square with vertices $(N+\frac{1}{2})(1+i)$, $(N+\frac{1}{2})(-1+i)$, $(N+\frac{1}{2})(-1-i)$ and $(N+\frac{1}{2})(1-i)$, as shown in Figure 2.13. Notice that the contour misses the poles of $\pi \cot(\pi z)$. Assuming that N is taken to be large enough, and since $R(z)$ has a finite number of poles, one can also guarantee that the contour will miss the poles of $R(z)$.

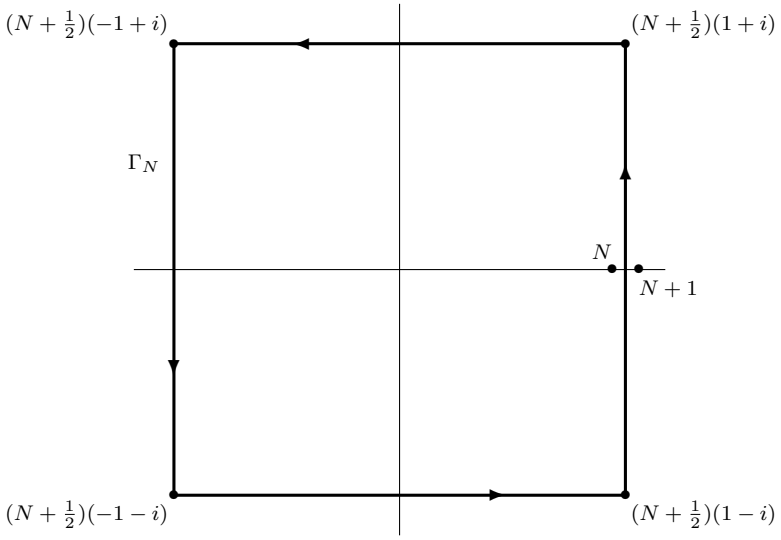


Figure 2.13: The contour Γ_N.

Let us compute the integral of the function $f(z)$ along this contour,

$$\int_{\Gamma_N} \pi \cot(\pi z) R(z)\, dz \ ,$$

in two ways. On the one hand we can use the residue theorem to say that the integral will be $(2\pi i)$ times the sum of the residues of the poles of $f(z)$. These poles are of two types: the poles of $R(z)$ and the poles of $\pi \cot(\pi z)$, which occur at the integers. Let us assume for simplicity that $R(z)$ has no poles at integer values of z, so that the poles of $R(z)$ and $\pi \cot(\pi z)$ do not coincide. Therefore we see that

$$\int_{\Gamma_N} \pi \cot(\pi z) R(z)\, dz = 2\pi\, i \left(\sum_{n=-N}^{N} \mathrm{Res}(f; n) + \sum_{\substack{\text{poles } z_k \text{ of } R \\ \text{inside } \Gamma_N}} \mathrm{Res}(f; z_k) \right) \ .$$

The residue of $f(z)$ at $z-n$ is easy to compute. Since by assumption $R(z)$ is analytic there and $\pi\cot(\pi z)$ has a simple pole with residue 1, we see that around $z=n$, we have

$$f(z) = R(z)\pi\cot(\pi z) = R(z)\left(\frac{1}{z-n} + \cdots\right) = \frac{R(z)}{z-n} + h(z)\ ,$$

where $h(z)$ is analytic at $z=n$. Therefore,

$$\mathrm{Res}(f;n) = \lim_{z\to n}\left[(z-n)f(z)\right] = R(n) + 0\ ,$$

and as a result,

$$\int_{\Gamma_N} \pi\cot(\pi z)R(z)\,dz = 2\pi\,i\left(\sum_{n=-N}^{N} R(n) + \sum_{\substack{\text{poles } z_k \text{ of } R\\ \text{inside } \Gamma_N}} \mathrm{Res}(f;z_k)\right)\ . \quad (2.71)$$

On the other hand we can estimate the integral for large enough N as follows. First of all because of the condition on $R(z)$, we have that for large $|z|$,

$$|R(z)| \le \frac{C}{|z|^2}\ .$$

Similarly, it can be shown that the function $\pi\cot(\pi z)$ is bounded along the contour, so that $|\pi\cot(\pi z)| \le K$ for some K independent of N.

Indeed, notice that

$$|\cot(\pi z)| = \frac{\cos(\pi z)}{\sin(\pi z)} = \frac{e^{i\pi z}+e^{-i\pi z}}{e^{i\pi z}-e^{-i\pi z}} = \frac{1+e^{-2i\pi z}}{1-e^{-2i\pi z}}\ .$$

Therefore along the segment of the contour parametrised by $z(t) = (N+\frac{1}{2}) + it$ for $t\in[-N-\frac{1}{2}, N+\frac{1}{2}]$, we have that

$$\begin{aligned}|\cot(\pi z(t))| &= \frac{1+e^{i2\pi((N+\frac{1}{2})+it)}}{1-e^{i2\pi((N+\frac{1}{2})+it)}}\\ &= \frac{1-e^{\pi(2N+1)\,t}}{1+e^{\pi(2N+1)\,t}} < 1\ ;\end{aligned}$$

whereas along the segments of the contour parametrised by $z(t) = t - i(N+\frac{1}{2})$ for $t\in[-N-\frac{1}{2}, N+\frac{1}{2}]$, we have that

$$\begin{aligned}|\cot(\pi z(t))| &= \frac{1+e^{-i\pi(2N+1)(t-i)}}{1-e^{-i\pi(2N+1)(t-i)}}\\ &= \frac{1+e^{2\pi t}e^{-\pi(2N+1)}}{1-e^{2\pi t}e^{-\pi(2N+1)}}\\ &\le \frac{1+e^{-\pi(2N+1)}}{1-e^{-\pi(2N+1)}}\end{aligned}$$

where we have used the triangle inequalities (2.1) on the numerator and (2.36) on the denominator. But

$$\frac{1+e^{-\pi(2N+1)}}{1-e^{-\pi(2N+1)}}$$

is maximised for $N=0$, whence it is bounded.

Since the length of the contour Γ_N is given by $4(2N+1)$, equation (2.28) gives the following estimate for the integral

$$\left|\int_{\Gamma_N} \pi\cot(\pi z)R(z)\,dz\right| \leq \frac{Kc}{(N+\frac{1}{2})^2}\,4(2N+1)\ ,$$

which vanishes in the limit $N\to\infty$. Therefore, taking the limit $N\to\infty$ of equation (2.71), and using that the left-hand side vanishes, one finds

$$\sum_{n=-\infty}^{\infty} R(n) = -\sum_{\text{poles } z_k \text{ of } R} \text{Res}(f;z_k)\ .$$

More generally, if $R(z)$ does have some poles for integer values of z, then we have to take care not to over-count these poles in the sum of the residues. We will count them as poles of $R(z)$ and not as poles of $\pi\cot(\pi z)$, and the same argument as above yields the general formula:

$$\boxed{\sum_{\substack{n=-\infty\\ n\neq z_k}}^{\infty} R(n) = -\sum_{\substack{\text{poles}\\ z_k \text{ of } R}} \text{Res}(f;z_k)\ ,\quad \text{for } f(z)=\pi\cot(\pi z)\,R(z).} \tag{2.72}$$

Let us compute then the sum (2.70). Notice that

$$\sum_{n=1}^{\infty}\frac{1}{n^2} = \tfrac{1}{2}\sum_{\substack{n=-\infty\\ n\neq 0}}^{\infty}\frac{1}{n^2}\ .$$

The function $R(z)=1/z^2$ has a double pole at $z=0$, hence by (2.72)

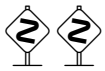
$$\sum_{n=1}^{\infty}\frac{1}{n^2} = -\tfrac{1}{2}\lim_{z\to 0}\frac{d}{dz}\left[\pi\cot(\pi z)\right]\ .$$

Now, the Laurent expansion of $\pi\cot(\pi z)$ around $z=0$ is given by

$$\frac{1}{z} - \frac{\pi^2\,z}{3} - \frac{\pi^4\,z^3}{45} + O(z^5)\ , \tag{2.73}$$

whence

$$\sum_{n=1}^{\infty}\frac{1}{n^2} = -\tfrac{1}{2}\left[-\frac{\pi^2}{3}\right] = \frac{\pi^2}{6}\ .$$



This sum has an interesting history. Its computation was an open problem in the 18th century for quite some time. It was known that the series was convergent (proven in fact by one of the Bernoullis) but it was up to Euler to calculate it. His "proof" is elementary and quite clever. Start with the Taylor series for the sine function:

$$\sin x = x \quad 1 - \frac{x^2}{3!} + \frac{x^4}{5!} - \cdots \quad ,$$

and treat the expression in parenthesis as an algebraic equation in x^2. Its solutions are known: $n^2\pi^2$ for $n = 1, 2, 3, \ldots$. Suppose we could factorise the expression in parenthesis:

$$1 - \frac{x^2}{\pi^2} \quad 1 - \frac{x^2}{(2\pi)^2} \quad 1 - \frac{x^2}{(3\pi)^2} \quad \cdots$$
$$= 1 - x^2 \quad \frac{1}{\pi^2} + \frac{1}{(2\pi)^2} + \frac{1}{(3\pi)^2} + \cdots \quad + O(x^4) \ .$$

Therefore, comparing the coefficient of x^2, we see that

$$\frac{1}{3!} = \frac{1}{\pi^2} + \frac{1}{(2\pi)^2} + \frac{1}{(3\pi)^2} + \cdots = \sum_{n=1}^{\infty} \frac{1}{(n\pi)^2} \ ,$$

which upon multiplication by π^2 yields the sum.

Similarly, we can compute the sum

$$\sum_{n=1}^{\infty} \frac{1}{n^4} = -\tfrac{1}{2} \operatorname{Res}(f; 0) \ ,$$

where $f(z) = \pi \cot(\pi z)/z^4$, whose Laurent series about $z = 0$ is can be read off from (2.73) above:

$$\frac{1}{z^5} - \frac{\pi^2}{3z^3} - \frac{\pi^4}{45\, z} + O(z) \ ,$$

whence

$$\sum_{n=1}^{\infty} \frac{1}{n^4} = \frac{\pi^4}{90} \ .$$

Infinite alternating sums

The techniques above can be extended to the computation of infinite alternating sums of the form

$$\sum_{n=-\infty}^{\infty} (-1)^n R(n) \ ,$$

where $R(z) = P(z)/Q(z)$ is a rational function with $\deg Q - \deg P \geq 2$. Now what is needed is a function $G(z)$ which has a simple pole at $z = n$, for n

an integer, and whose residue there is $(-1)^n$. We claim that this function is $\pi \csc(\pi z)$. Indeed, the Laurent expansion about $z = 0$ is given by

$$\pi \csc(\pi z) = \frac{1}{z} + \frac{\pi^2 z}{6} + \frac{7\,\pi^4 z^3}{360} + O(z^5) \; ; \tag{2.74}$$

whence its residue at 0 is 1. Because of the periodicity $\csc(\pi(z + 2k)) = \csc(\pi z + 2k\pi) = \csc(\pi z)$ for any integer k, this is also the residue about every even integer. Now from the periodicity $\csc(\pi(z + 1)) = \csc(\pi z + \pi) = -\csc(\pi z)$, we notice that the residue at every odd integer is -1. Therefore we conclude that for $G(z) = \pi \csc(\pi z)$, $\mathrm{Res}(G; n) = (-1)^n$.

The trigonometric identity

$$(\csc(\pi z))^2 = 1 + (\cot(\pi z))^2 \; ,$$

implies that $\csc(\pi z)$ is also bounded along the contour Γ_N, with a bound which is independent of N just like for $\cot(\pi z)$. Just as was done above for the cotangent function, we can now prove that the integral of the function $f(z) = \pi \csc(\pi z) R(z)$ along Γ_N vanishes in the limit $N \to 0$. This proof is virtually identical to the one given above. Therefore we can conclude that

$$\boxed{\sum_{\substack{n=-\infty \\ n \neq z_k}}^{\infty} (-1)^n R(n) = -\sum_{\substack{\text{poles} \\ z_k \text{ of } R}} \mathrm{Res}(f; z_k) \; , \quad \text{for } f(z) = \pi \csc(\pi z)\, R(z).} \tag{2.75}$$

As an example, let us compute the alternating sums

$$S_1 = \sum_{n=1}^{\infty} \frac{(-1)^n}{n^2} \qquad \text{and} \qquad S_2 = \sum_{n=1}^{\infty} \frac{(-1)^n}{n^4} \; .$$

For the first sum we have that

$$S_1 = -\tfrac{1}{2}\,\mathrm{Res}(f; 0) \; ,$$

where $f(z) = \pi \csc(\pi z)/z^2$, whose Laurent expansion about $z = 0$ can be read off from (2.74):

$$f(z) = \frac{1}{z^3} + \frac{\pi^2}{6\,z} + \frac{7\,\pi^4 z}{360} + O(z^3) \; ,$$

whence the residue is $\pi^2/6$ and the sum

$$S_1 = -\frac{\pi^2}{12} \; .$$

For the second sum we also have that

$$S_2 = -\tfrac{1}{2}\,\mathrm{Res}(f;0)\ ,$$

where the function $f(z) = \pi\csc(\pi z)/z^4$ has now a Laurent series

$$f(z) = \frac{1}{z^5} + \frac{\pi^2}{6\,z^3} + \frac{7\,\pi^4}{360\,z} + O(z^1)\ ,$$

whence the residue is $7\pi^4/360$ and the sum

$$S_2 = -\frac{7\,\pi^4}{720}\ .$$

Sums involving binomial coefficients

There are other types of sums which can also be performed or at least estimated using residue techniques, particularly sums whose coefficients are related to the binomial coefficients, as in $\sum_{n=1}^{\infty}\binom{2n}{n}R(n)$. By definition, the binomial coefficient $\binom{n}{k}$ is the coefficient of z^k in the binomial expansion of $(1+z)^n$. In other words, using the residue theorem,

$$\binom{n}{k} = \frac{1}{2\pi i}\oint_\Gamma \frac{(1+z)^n}{z^{k+1}}\,dz\ ,$$

where Γ is any positively oriented loop surrounding the origin.

Suppose that we wish to compute the sum

$$S = \sum_{n=0}^{\infty}\binom{2n}{n}\frac{1}{5^n}\ .$$

We can substitute the integral representation for the binomial coefficient,

$$S = \sum_{m=0}^{\infty}\left[\frac{1}{2\pi i}\int_\Gamma \frac{(1+z)^{2n}}{z^{n+1}}\,dz\right]\frac{1}{5^n} = \frac{1}{2\pi i}\sum_{n=0}^{\infty}\int_\Gamma \frac{(1+z)^{2n}}{(5z)^n}\,\frac{dz}{z}\ .$$

Now provided that we choose Γ inside the domain of convergence of the series $\sum_{n=0}^{\infty}\frac{(1+z)^{2n}}{(5z)^n}$ then we would obtain that by uniform convergence, the integral of the sum is the sum of the termwise integrals. Being a geometric series, its convergence is uniform in the region

$$\left|\frac{(1+z)^2}{5z}\right| < 1\ ,$$

so choose the contour Γ inside this region. For definiteness we can choose the unit circle, since on the unit circle:

$$\left|\frac{(1+z)^2}{5z}\right| \leq \frac{4}{5} \ .$$

In this case, we can interchange the order of the summation and the integration:

$$S = \frac{1}{2\pi i}\int_{|z|=1}\sum_{n=0}^{\infty}\frac{(1+z)^{2n}}{(5z)^n}\,\frac{dz}{z} = \frac{5}{2\pi i}\int_{|z|=1}\frac{1}{3z-1-z^2}\,dz \ .$$

Now the integral can be performed using the residue theorem. The integrand has simple poles at $(3\pm\sqrt{5})/2$ of which only the $(3-\sqrt{5})/2$ lies inside the contour. Therefore,

$$S = 5\,\mathrm{Res}\left(f; \frac{3-\sqrt{5}}{2}\right) \qquad \text{where} \qquad f(z) = \frac{1}{3z-1-z^2} \ .$$

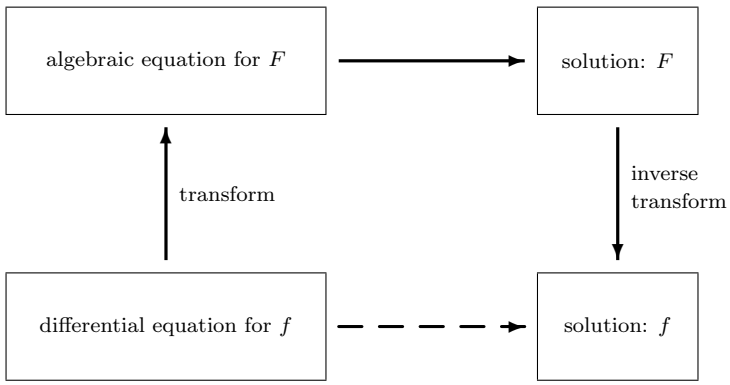
Computing the residue, we find $\mathrm{Res}((3-\sqrt{5})/2) = 1/\sqrt{5}$, whence $S = \sqrt{5}$.

Chapter 3

Integral Transforms

This part of the course introduces two extremely powerful methods to solving differential equations: the Fourier and the Laplace transforms. Beside its practical use, the Fourier transform is also of fundamental importance in quantum mechanics, providing the correspondence between the position and momentum representations of the Heisenberg commutation relations.

An integral transform is useful if it allows one to turn a complicated problem into a simpler one. The transforms we will be studying in this part of the course are mostly useful to solve differential and, to a lesser extent, integral equations. The idea behind a transform is very simple. To be definite suppose that we want to solve a differential equation, with unknown function f. One first applies the transform to the differential equation to turn it into an equation one can solve easily: often an algebraic equation for the transform F of f. One then solves this equation for F and finally applies the inverse transform to find f. This circle (or square!) of ideas can be represented diagrammatically as follows:



We would like to follow the dashed line, but this is often very difficult.

Therefore we follow the solid line instead: it may seem a longer path, but it has the advantage of being straightforward. After all, what is the purpose of developing formalism if not to reduce the solution of complicated problems to a set of simple rules which even a machine could follow?

We will start by reviewing Fourier series in the context of one particular example: the vibrating string. This will have the added benefit of introducing the method of separation of variables in order to solve partial differential equations. In the limit as the vibrating string becomes infinitely long, the Fourier series naturally gives rise to the Fourier integral transform, which we will apply to find steady-state solutions to differential equations. In particular we will apply this to the one-dimensional wave equation. In order to deal with transient solutions of differential equations, we will introduce the Laplace transform. This will then be applied, among other problems, to the solution of initial value problems.

3.1 Fourier series

In this section we will discuss the Fourier expansion of periodic functions of a real variable. As a practical application, we start with the study of the vibrating string, where the Fourier series makes a natural appearance.

3.1.1 The vibrating string

Consider a string of length L which is clamped at both ends. Let x denote the position along the string: such that the two ends of the string are at $x = 0$ and $x = L$, respectively. The string has tension T and a uniform mass density μ, and it is allowed to vibrate. If we think of the string as being composed of an infinite number of infinitesimal masses, we model the vibrations by a function $\psi(x, t)$ which describes the vertical displacement at time t of the mass at position x. It can be shown that for small vertical displacements, $\psi(x, t)$ obeys the following equation:

$$T \frac{\partial^2 \psi(x,t)}{\partial x^2} = \mu \frac{\partial^2 \psi(x,t)}{\partial t^2} \ ,$$

which can be recognised as the **one-dimensional wave equation**

$$\boxed{\frac{\partial^2}{\partial x^2}\psi(x,t) = \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\psi(x,t) \ ,} \tag{3.1}$$

where $c = \sqrt{T/\mu}$ is the wave velocity. This is a partial differential equation which needs for its solution to be supplemented by boundary conditions for

x and initial conditions for t. Because the string is clamped at both ends, the boundary conditions are

$$\psi(0,t) = \psi(L,t) = 0 \ , \qquad \text{for all } t. \tag{3.2}$$

As initial conditions we specify that at $t = 0$,

$$\left.\frac{\partial\psi(x,t)}{\partial t}\right|_{t=0} = 0 \qquad \text{and} \qquad \psi(x,0) = f(x) \ , \qquad \text{for all } x, \tag{3.3}$$

where f is a continuous function which, for consistency with the boundary conditions (3.2), must satisfy $f(0) = f(L) = 0$. In other words, the string is released from rest from an initial shape given by the function f.

This is not the only type of initial conditions that could be imposed. For example, in the case of, say, a piano string, it would be much more sensible to consider an initial condition in which the string is horizontal so that $\psi(x,0) = 0$, but such that it is given a blow at $t = 0$, which means that $\frac{\partial\psi(x,t)}{\partial t}|_{t=0} = g(x)$ for some function g. More generally still, we could consider mixed initial conditions in which $\psi(x,0) = f(x)$ and $\frac{\partial\psi(x,t)}{\partial t}|_{t=0} = g(x)$. These different initial conditions can be analysed in roughly the same way.

We will solve the wave equation by the method of **separation of variables**. This consists of choosing as an Ansatz for $\psi(x,t)$ the product of two functions, one depending only on x and the other only on t: $\psi(x,t) = u(x)\,v(t)$. We do not actually expect the solution to be of this form; but because, as we will review below, the equation is linear and one can use the principle of superposition to construct the desired solution out of decomposable solutions of this type. At any rate, inserting this Ansatz into (3.1), we have

$$u''(x)\,v(t) = \frac{1}{c^2}\,u(x)v''(t) \ ,$$

where we are using primes to denote derivatives with respect to the variable on which the function depends: $u'(x) = du/dx$ and $v'(t) = dv/dt$. We now divide both sides of the equation by $u(x)\,v(t)$, and obtain

$$\frac{u''(x)}{u(x)} = \frac{1}{c^2}\,\frac{v''(t)}{v(t)} \ .$$

Now comes the reason that this method works, so pay close attention. Notice that the right-hand side does not depend on x, and that the left-hand side does not depend on t. Since they are equal, both sides have to be equal to a constant which, with some foresight, we choose to call $-\lambda^2$, as it will be a negative number in the case of interest. The equation therefore breaks up into two ordinary differential equations:

$$u''(x) = -\lambda^2\,u(x) \qquad \text{and} \qquad v''(t) = -\lambda^2\,c^2\,v(t) \ .$$

The boundary conditions say that $u(0) = u(L) = 0$.

Let us consider the first equation. It has three types of solutions depending on whether λ is nonzero real, nonzero imaginary or zero. (Notice that $-\lambda^2$ has to be real, so that these are the only possibilities.) If $\lambda = 0$, then the solution is $u(x) = a + b\,x$. The boundary condition $u(0) = 0$ means that $a = 0$, but the boundary condition $u(L) = 0$ then means that $b = 0$, whence $u(x) = 0$ for all x. Clearly this is a very uninteresting solution. Let us consider λ imaginary. Then the solution is now $a\exp(|\lambda|\,x) + b\exp(-|\lambda|\,x)$. Again the boundary conditions force $a = b = 0$. Therefore we are left with the possibility of λ real. Then the solution is

$$u(x) = a\cos\lambda x + b\sin\lambda x\ .$$

The boundary condition $u(0) = 0$ forces $a = 0$. Finally the boundary condition $u(L) = 0$ implies that

$$\sin\lambda L = 0 \qquad \Longrightarrow \qquad \lambda = \frac{n\,\pi}{L} \qquad \text{for } n \text{ an integer.}$$

Actually $n = 0$ is an uninteresting solution, and because of the fact that the sine is an odd function, negative values of n give rise to the same solution (up to a sign) as positive values of n. In other words, all nontrivial distinct solution are given (up to a constant multiple) by

$$u_n(x) \equiv \sin\lambda_n\,x\ , \quad \text{with } \lambda_n = \frac{n\,\pi}{L} \quad \text{and where } n = 1, 2, 3, \cdots . \tag{3.4}$$

Let us now solve for $v(t)$. Its equation is

$$v''(t) = -\lambda^2 c^2 v(t)\ ,$$

whence

$$v(t) = a\,\cos\lambda ct + b\,\sin\lambda ct\ .$$

The first of the two initial conditions (3.3) says that $v'(0) = 0$ whence $b = 0$. Therefore for any positive integer n, the function

$$\psi_n(x,t) = \sin\lambda_n x\,\cos\lambda_n ct\ , \qquad \text{with } \lambda_n = \frac{n\pi}{L}\ ,$$

satisfies the wave equation (3.1) subject to the boundary conditions (3.2) and to the first of the initial conditions (3.3).

Now notice something important: the wave equation (3.1) is **linear**; that is, if $\psi(x,t)$ and $\phi(x,t)$ are solutions of the wave equation, so is any linear combination $\alpha\,\psi(x,t) + \beta\,\phi(x,t)$ where α and β are constants.

Clearly then, any linear combination of the $\psi_n(x,t)$ will also be a solution. In other words, the most general solution subject to the boundary conditions (3.2) and the first of the initial conditions in (3.3) is given by a linear combination

$$\psi(x,t) = \sum_{n=1}^{\infty} b_n \sin \lambda_n x \, \cos \lambda_n c t \ .$$

Of course, this expression is formal as it stands: it is an infinite sum which does not necessarily make sense, unless we chose the coefficients $\{b_n\}$ in such a way that the series converges, and that the convergence is such that we can differentiate the series termwise at least twice.

We can now finally impose the second of the initial conditions (3.3):

$$\psi(x,0) = \sum_{n=1}^{\infty} b_n \sin \lambda_n x = f(x) \ . \tag{3.5}$$

At first sight this seems hopeless: can any function $f(x)$ be represented as a series of this form? The Bernoullis, who were the first to get this far, thought that this was not the case and that in some sense the solution was only valid for special kinds of functions for which such a series expansion is possible. It took Euler to realise that, in a certain sense, all functions $f(x)$ with $f(0) = f(L) = 0$, can be expanded in this way. He did this by showing how the coefficients $\{b_n\}$ are determined by the function $f(x)$.

To do so let us argue as follows. Let n and m be positive integers and consider the functions $u_n(x)$ and $u_m(x)$ defined in (3.4). These functions satisfy the differential equations:

$$u_n''(x) = -\lambda_n^2 \, u_n(x) \qquad \text{and} \qquad u_m''(x) = -\lambda_m^2 \, u_m(x) \ .$$

Let us multiply the first equation by $u_m(x)$ and the second equation by $u_n(x)$ and subtract one from the other to obtain

$$u_n''(x) \, u_m(x) - u_n(x) \, u_m''(x) = (\lambda_m^2 - \lambda_n^2) \, u_n(x) \, u_m(x) \ .$$

We notice that the left-hand side of the equation is a total derivative

$$u_n''(x) \, u_m(x) - u_n(x) \, u_m''(x) = \left(u_n'(x) \, u_m(x) - u_n(x) \, u_m'(x)\right)' \ ,$$

whence integrating both sides of the equation from $x = 0$ to $x = L$, we obtain

$$(\lambda_m^2 - \lambda_n^2) \int_0^L u_m(x) \, u_n(x) \, dx = \left(u_n'(x) \, u_m(x) - u_n(x) \, u_m'(x)\right)\Big|_0^L = 0 \ ,$$

since $u_n(0) = u_n(L) = 0$ and the same for u_m. Therefore we see that unless $\lambda_n^2 = \lambda_m^2$, which is equivalent to $n = m$ (since n, m are positive integers), the integral $\int_0^L u_m(x)\, u_n(x)\, dx$ vanishes. On the other hand, if $m = n$, we have that

$$\int_0^L u_n(x)^2\, dx = \int_0^L \left(\sin \frac{n\pi x}{L}\right)^2 dx = \int_0^L \left(\frac{1}{2} - \frac{1}{2}\cos \frac{2n\pi x}{L}\right) dx = \frac{L}{2}\ .$$

Therefore, in summary, we have the **orthogonality property** of the functions $u_m(x)$:

$$\int_0^L u_m(x)\, u_n(x)\, dx = \begin{cases} \frac{L}{2}\ , & \text{if } n = m, \text{ and} \\ 0\ , & \text{otherwise.} \end{cases} \tag{3.6}$$

Let us now go back to the solution of the remaining initial condition (3.5). This condition can be rewritten as

$$f(x) = \sum_{n=1}^{\infty} b_n\, u_n(x) = \sum_{n=1}^{\infty} b_n\, \sin \frac{n\pi x}{L}\ . \tag{3.7}$$

Let us multiply both sides by $u_m(x)$ and integrate from $x = 0$ to $x = L$:

$$\int_0^L f(x)\, u_m(x)\, dx = \sum_{n=1}^{\infty} b_n \int_0^L u_n(x)\, u_m(x)\, dx\ ,$$

where we have interchanged the order of integration and summation with impunity.[1] Using the orthogonality relation (3.6) we see that of all the terms in the right-hand side, only the term with $n = m$ contributes to the sum, whence

$$\int_0^L f(x)\, u_m(x)\, dx = b_m\, \frac{L}{2}\ ,$$

or in other words,

$$b_m = \frac{2}{L} \int_0^L f(x)\, u_m(x)\, dx\ , \tag{3.8}$$

a formula due to Euler. Finally, the solution of the wave equation (3.1) with boundary conditions (3.2) and initial conditions (3.3) is

$$\psi(x,t) = \sum_{n=1}^{\infty} b_n\, \sin \frac{n\pi x}{L}\, \cos \frac{n\pi ct}{L}\ , \tag{3.9}$$

[1] This would have to be justified, but in this part of the course we will be much more cavalier about these things. The amount of material that would have to be introduced to be able to justify this procedure is too much for a course at this level and of this length.

where

$$b_n = \frac{2}{L}\int_0^L f(x)\,\sin\frac{n\pi x}{L}\,dx\ .$$

Inserting this expression into the solution (3.9), we find that

$$\begin{aligned}\psi(x,t) &= \sum_{n=1}^{\infty}\left[\frac{2}{L}\int_0^L f(y)\,\sin\frac{n\pi y}{L}\,dy\right]\sin\frac{n\pi x}{L}\,\cos\frac{n\pi ct}{L}\\ &= \int_0^L\left[\frac{2}{L}\sum_{n=1}^{\infty}\sin\frac{n\pi y}{L}\,\sin\frac{n\pi x}{L}\,\cos\frac{n\pi ct}{L}\right]f(y)\,dy\\ &= \int_0^L K(x,y;t)\,f(y)\,dy\ ,\end{aligned}$$

where the **propagator** $K(x,y,t)$ is (formally) defined by

$$K(x,y;t) \equiv \sum_{n=1}^{\infty}\frac{2}{L}\,\sin\frac{n\pi y}{L}\,\sin\frac{n\pi x}{L}\,\cos\frac{n\pi ct}{L}\ .$$

To understand why it is called a propagator, notice that

$$\psi(x,t) = \int_0^L K(x,y;t)\,\psi(y,0)\,dy\ ,$$

so that one can obtain $\psi(x,t)$ from its value at $t=0$ simply by multiplying by $K(x,y;t)$ and integrating; hence $K(x,y;t)$ allows us to *propagate* the configuration at $t=0$ to any other time t.

Actually, the attentive reader will have noticed that we never showed that the series $\sum_{n=1}^{\infty} b_n u_n(x)$, with b_n given by the Euler formula (3.8) converges to $f(x)$. In fact, it is possible to show that it does, but the convergence is not necessarily pointwise (and certainly not uniform). We state without proof the following result:

$$\lim_{N\to\infty}\int_0^L\left(f(x) - \sum_{n=1}^{N} b_n\,u_n(x)\right)^2\,dx = 0\ . \tag{3.10}$$

In other words, the function

$$h(x) \equiv f(x) - \sum_{n=1}^{\infty} b_n\,u_n(x)$$

has the property that the integral

$$\int_0^L h(x)^2\,dx = 0\ .$$

This however does not mean that $h(x) = 0$, but only that it is zero **almost everywhere**.

To understand this notice consider the (discontinuous) function

$$h(x) = \begin{cases} 1 \ , & \text{for } x = x_0, \text{ and} \\ 0 \ , & \text{otherwise.} \end{cases}$$

Then it is clear that the improper integral

$$\int_0^L h(x)^2 \, dx = \lim_{r,s \searrow 0} \int_0^{x_0 - r} + \int_{x_0 + s}^L h(x)^2 \, dx = 0 \ .$$

The same would happen if $h(x)$ were zero but at a *finite* number of points.

Of course, if $h(x)$ were continuous *and* zero almost everywhere, it would have to be identically zero. In this case the convergence of the series (3.7) would be pointwise. This is the case if $f(x)$ is itself continuous.

Expanding (3.10), we find that

$$\int_0^L f(x)^2 \, dx - 2 \sum_{n=1}^{\infty} b_n \int_0^L f(x) \, u_n(x) \, dx + \sum_{n,m=1}^{\infty} b_n \, b_m \int_0^L u_n(x) \, u_m(x) \, dx = 0 \ .$$

Using (3.8) and (3.6) we can simplify this a little

$$\int_0^L f(x)^2 \, dx - 2 \sum_{n=1}^{\infty} \frac{L}{2} \, b_n^2 + \sum_{n=1}^{\infty} \frac{L}{2} \, b_n^2 = 0 \ ,$$

whence

$$\sum_{n=1}^{\infty} b_n^2 = \frac{2}{L} \int_0^L f(x)^2 \, dx \ .$$

Since $f(x)^2$ is continuous, it is integrable, and hence the right-hand side is finite, whence the series $\sum_{n=1}^{\infty} b_n^2$ also converges. In particular, it means that $\lim_{n \to \infty} b_n = 0$.

3.1.2 The Fourier series of a periodic function

We have seen above that a continuous function $f(x)$ defined on the interval $[0, L]$ and vanishing at the boundary, $f(0) = f(L) = 0$, can be expanded in terms of the functions $u_n(x) = \sin(n\pi x/L)$. In this section we will generalise this and consider similar expansions for periodic functions.

To be precise let $f(x)$ be a complex-valued function of a real variable which is periodic with period L: $f(x+L) = f(x)$ for all x. Periodicity means that $f(x)$ is uniquely determined by its behaviour within a period. In other words, if we know f in the interval $[0, L]$ then we know $f(x)$ everywhere. Said differently, any function defined on $[0, L]$, obeying $f(0) = f(L)$ can be extended to the whole real line as a periodic function. More generally, the interval $[0, L]$ can be substituted by any one period $[x_0, x_0 + L]$, for some x_0 with the property that $f(x_0) = f(x_0+L)$. This is not a useless generalisation: it will be important when we discuss the case of $f(x)$ being a discontinuous function. The strength of the Fourier expansion is that it treats discontinuous functions (at least those with a finite number of discontinuities in any one period) as easily as it treats continuous functions. The reason is, as we stated briefly above, that the convergence of the series is not pointwise but rather in the sense (3.10), which simply means that it converges pointwise almost everywhere.

The functions $e_n(x) \equiv \exp(i2\pi nx/L)$ are periodic with period L, since $e_n(x + L) = e_n(x) \exp(i2\pi n) = e_n(x)$. Therefore we could try to expand

$$\boxed{f(x) = \sum_{n=-\infty}^{\infty} c_n \, e_n(x) \ ,} \tag{3.11}$$

for some complex coefficients $\{c_n\}$. This series is known as a **trigonometric** of **Fourier series** of the periodic function f, and the $\{c_n\}$ are called the **Fourier coefficients**. Under complex conjugation, the exponentials $e_n(x)$ satisfy $e_n(x)^* = e_{-n}(x)$, and also the following orthogonality property:

$$\int_0^L e_m(x)^* \, e_n(x) \, dx = \int_0^L e^{i2\pi(n-m)x/L} \, dx = \begin{cases} L \ , & \text{if } n = m, \text{ and} \\ 0 \ , & \text{otherwise.} \end{cases}$$

Therefore if we multiply both sides of (3.11) by $e_m(x)^*$ and integrate, we find the following formula for the Fourier coefficients:

$$c_m = \frac{1}{L} \int_0^L e_m(x)^* \, f(x) \, dx \ .$$

It is important to realise that the exponential functions $e_n(x)$ satisfy the orthogonality relation for any one period, not necessarily $[0, L]$:

$$\int\limits_{\substack{\text{one} \\ \text{period}}} e_m(x)^* \, e_n(x) \, dx = \begin{cases} L \ , & \text{if } n = m, \text{ and} \\ 0 \ , & \text{otherwise;} \end{cases} \tag{3.12}$$

whence the Fourier coefficients can be obtained by integrating over any one period:

$$\boxed{c_m = \frac{1}{L} \int\limits_{\substack{\text{one} \\ \text{period}}} e_m(x)^* \, f(x) \, dx \ .} \tag{3.13}$$

Again we can state without proof that the series converges pointwise almost everywhere within any one period, in the sense that

$$\lim_{N \to \infty} \int\limits_{\substack{\text{one} \\ \text{period}}} \left| f(x) - \sum_{n=-N}^{N} c_n \, e_n(x) \right|^2 dx = 0 \ ,$$

whenever the $\{c_n\}$ are given by (3.13).

There is one special case where the series converges pointwise and uniformly. Let $g(z)$ be a function which is analytic in an open annulus containing the unit circle $|z| = 1$. We saw in Section 2.3.4 that such a function is approximated uniformly by a Laurent series of the form

$$g(z) = \sum_{n=-\infty}^{\infty} b_n \, z^n \ ,$$

where the $\{b_n\}$ are given by equations (2.53) and (2.54). Evaluating this on the unit circle $z = e^{i\theta}$, we have that

$$g(e^{i\theta}) = \sum_{n=-\infty}^{\infty} b_n \, e^{in\theta} \ , \tag{3.14}$$

and the coefficients $\{b_n\}$ are given by

$$b_n = \frac{1}{2\pi} \int_0^{2\pi} g(e^{i\theta}) \, e^{-in\theta} \, d\theta \ , \tag{3.15}$$

which agrees precisely with the Fourier series of the function $g(e^{i\theta})$ which is periodic with period 2π. We can rescale this by defining $\theta = 2\pi x/L$ where x is periodic with period L. Let $f(x) \equiv g(\exp(i2\pi x/L))$, which is now periodic with period L. Then the Laurent series (3.14) becomes the Fourier series (3.11) where the Laurent coefficients (3.15) are now give by the Fourier coefficients (3.13).

Some examples

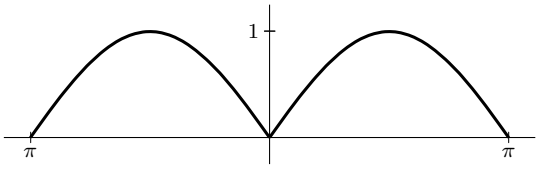


Figure 3.1: Plot of $|\sin x|$ for $x \in [-\pi, \pi]$.

Let us now compute some examples of Fourier series. The first example is the function $f(x) = |\sin x|$. A graph of this function shows that it is periodic with period π, as seen in Figure 3.1. We therefore try an expansion of the form

$$|\sin x| = \sum_{n=-\infty}^{\infty} c_n \, e^{i2nx} \ ,$$

where the coefficients $\{c_n\}$ are given by

$$c_n = \frac{1}{\pi} \int_0^{\pi} |\sin x| \, e^{-i2nx} \, dx = \frac{1}{\pi} \int_0^{\pi} \sin x \, e^{-i2nx} \, dx \ .$$

We can expand $\sin x$ into exponentials to obtain

$$\begin{aligned} c_n &= \frac{1}{2\pi i} \int_0^{\pi} \left(e^{ix} - e^{-ix} \right) \, e^{-i2nx} \, dx \\ &= \frac{1}{2\pi i} \left[\int_0^{\pi} e^{-i(2n-1)x} \, dx - \int_0^{\pi} e^{-i(2n+1)x} \, dx \right] \\ &= \frac{1}{2\pi i} \left[\frac{i}{2n-1} \left(e^{-i(2n-1)\pi} - 1 \right) - \frac{i}{2n+1} \left(e^{-i(2n+1)\pi} - 1 \right) \right] \\ &= \frac{1}{2\pi i} (-2i) \left[\frac{1}{2n-1} - \frac{1}{2n+1} \right] \\ &= -\frac{2}{\pi} \, \frac{1}{4n^2 - 1} \ . \end{aligned}$$

Therefore,

$$|\sin x| = \sum_{n=-\infty}^{\infty} -\frac{2}{\pi} \, \frac{1}{4n^2 - 1} \, e^{i2nx} = \frac{2}{\pi} - \sum_{n=1}^{\infty} \frac{4}{\pi} \, \frac{1}{4n^2 - 1} \cos 2nx \ .$$

Notice that this can be used in order to compute infinite sums. Evaluating this at $x = 0$, we have that

$$\sum_{n=1}^{\infty} \frac{1}{4n^2 - 1} = \tfrac{1}{2} \ ,$$

whereas evaluating this at $x = \pi/2$, we have that

$$\sum_{n=1}^{\infty} \frac{(-1)^n}{4n^2 - 1} = \frac{2 - \pi}{4} \ .$$

Of course, we could have summed these series using the residue theorem, as explained in Section 2.4.5.

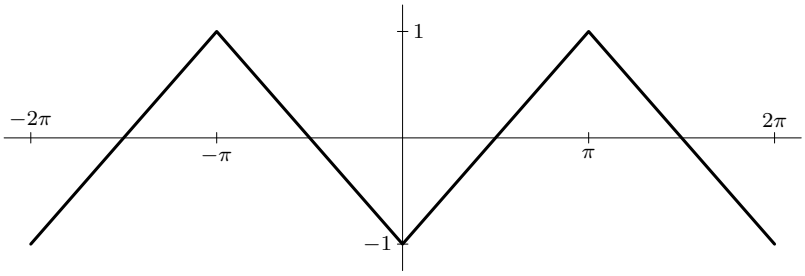


Figure 3.2: Plot of $f(x)$ for $x \in [-2\pi, 2\pi]$.

As a second example, let us consider the function $f(x)$ defined in the interval $[-\pi, \pi]$ by

$$f(x) = \begin{cases} -1 - \frac{2}{\pi}\, x \ , & \text{if } -\pi \le x \le 0, \text{ and} \\ -1 + \frac{2}{\pi}\, x \ , & \text{if } 0 \le x \le \pi. \end{cases} \tag{3.16}$$

and extended periodically to the whole real line. A plot of this function for $x \in [-2\pi, 2\pi]$ is shown in Figure 3.2. It is clear from the picture that $f(x)$ has periodicity 2π, whence we expect a Fourier series of the form

$$f(x) = \sum_{n=-\infty}^{\infty} c_n\, e^{inx} \ ,$$

where the coefficients are given by

$$\begin{aligned} c_n &= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x)\, e^{-inx}\, dx \\ &= \frac{1}{2\pi} \left[\int_{-\pi}^{0} (-1 - \frac{2}{\pi}\, x)\, e^{-inx}\, dx + \int_{0}^{\pi} (-1 + \frac{2}{\pi}\, x)\, e^{-inx}\, dx \right] \\ &= \frac{1}{2\pi} \int_{0}^{\pi} (-1 + \frac{2}{\pi}\, x) \, \left[e^{inx} + e^{-inx} \right] \, dx \\ &= -\frac{1}{\pi} \int_{0}^{\pi} \cos(nx)\, dx + \frac{2}{\pi^2} \int_{0}^{\pi} x\, \cos(nx)\, dx \ . \end{aligned}$$

We must distinguish between $n = 0$ and $n \neq 0$. Performing the elementary integrals for both of these cases, we arrive at

$$c_n = \begin{cases} \frac{2}{\pi^2\, n^2} \left[(-1)^n - 1 \right] \ , & \text{for } n \neq 0, \text{ and} \\ 0 \ , & \text{for } n = 0. \end{cases} \tag{3.17}$$

Therefore we have that

$$\begin{aligned} f(x) &= \sum_{\substack{n=-\infty \\ n\neq 0}}^{\infty} \frac{2}{\pi^2\, n^2}\, [(-1)^n - 1]\; e^{inx} \\ &= \sum_{n=1}^{\infty} \frac{4}{\pi^2\, n^2}\, [(-1)^n - 1]\, \cos nx \\ &= \sum_{\substack{n=1 \\ n \text{ odd}}}^{\infty} -\frac{8}{\pi^2\, n^2}\, \cos nx \\ &= \sum_{\ell=0}^{\infty} \frac{8}{\pi^2}\, \frac{-1}{(2\ell+1)^2}\, \cos(2\ell+1)\, x \ . \end{aligned}$$

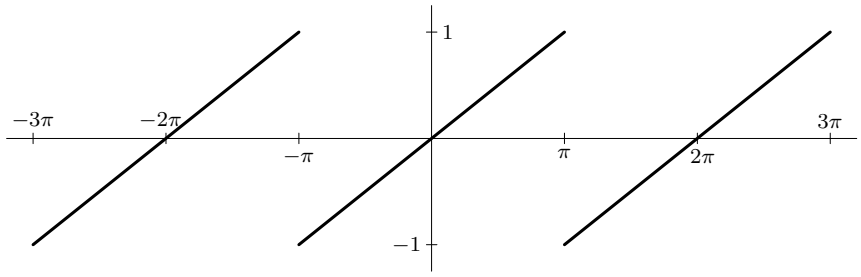


Figure 3.3: Plot of $g(x)$ for $x \in [-3\pi, 3\pi]$.

Finally we consider the case of a discontinuous function:

$$g(x) = \frac{x}{\pi}\ , \qquad \text{where } x \in [-\pi, \pi],$$

and extended periodically to the whole real line. The function has period 2π, and so we expect a series expansion of the form

$$g(x) = \sum_{n=-\infty}^{\infty} c_n\, e^{inx}\ ,$$

where the Fourier coefficients are given by

$$\begin{aligned} c_n &= \frac{1}{2\pi} \int_{-\pi}^{\pi} g(x)\, e^{-inx}\, dx \\ &= \frac{1}{2\pi^2} \int_{-\pi}^{\pi} x\, e^{-inx}\, dx\ . \end{aligned}$$

We must distinguish the cases $n = 0$ and $n \neq 0$. In either case we can perform the elementary integrals to arrive at

$$c_n = \begin{cases} 0 \ , & \text{if } n = 0, \\ \frac{i}{\pi\, n}\, (-1)^n \ , & \text{otherwise.} \end{cases}$$

Therefore,

$$g(x) = \sum_{\substack{n=-\infty \\ n\neq 0}}^{\infty} \frac{i}{\pi\, n}\, (-1)^n e^{inx} = \sum_{n=1}^{\infty} \frac{-2}{\pi\, n}\, (-1)^n \, \sin n\, x \ .$$

Now notice something curious: the function $g(x)$ is discontinuous at $x = (2\ell + 1)\pi$. Evaluating the series at such values of x we see that because $\sin n(2\ell + 1)\pi = 0$, the series sums to zero for these values. In other words, $g(x)$ is only equal to the Fourier series at those values x where $g(x)$ is continuous. At the values where $g(x)$ is discontinuous, the Fourier series can be shown to converge to the mean of the left and right limits of the function: in this case, $\lim_{x\searrow \pi} g(x) = -1$ and $\lim_{x\nearrow \pi} g(x) = 1$, and the average is 0, in agreement with what we just saw.

3.1.3 Some properties of the Fourier series

In this section we explore some general properties of the Fourier series of a complex periodic function $f(x)$ with period L.

Let us start with the following observation. If $f(x)$ is real, then the Fourier coefficients obey $c_n^* = c_{-n}$. This follows from the following. Taking the complex conjugate of the Fourier series for $f(x)$, we have

$$f(x)^* = \left(\sum_{n=-\infty}^{\infty} c_n\, e_n(x) \right)^* = \sum_{n=-\infty}^{\infty} c_n^*\, e_{-n}(x) \ ,$$

where we have used that $e_n(x)^* = e_{-n}(x)$. Since $f(x)$ is real, $f(x) = f(x)^*$ for all x, whence

$$\sum_{n=-\infty}^{\infty} c_n\, e_n(x) = \sum_{n=-\infty}^{\infty} c_n^*\, e_{-n}(x) = \sum_{n=-\infty}^{\infty} c_{-n}^*\, e_n(x) \ .$$

Multiplying both sides of the equation by $e_m^*(x)$, integrating over one period and using the orthogonality relation (3.12), we find that $c_m = c_{-m}^*$.

Fourier sine and cosine series

Suppose that $f(x)$ is periodic and also even, so that $f(-x) = f(x)$. Then this means that

$$f(x) = \tfrac{1}{2}\left[f(x) + f(-x)\right] \ .$$

If we substitute its Fourier series

$$f(x) = \sum_{n=-\infty}^{\infty} c_n \, e_n(x) \ ,$$

we see that

$$f(x) = \sum_{n=-\infty}^{\infty} \tfrac{1}{2}\, c_n \left[e_n(x) + e_n(-x)\right] = \sum_{n=-\infty}^{\infty} c_n \, \cos\lambda_n x \ ,$$

where $\lambda_n = 2\pi n/L$. Now we use the fact that $\cos\lambda_{-n}x = \cos\lambda_n x$ to rewrite the series as

$$f(x) = c_0 + \sum_{n=1}^{\infty} \left[c_n + c_{-n}\right] \cos\lambda_n \, x = \tfrac{1}{2}a_0 + \sum_{n=1}^{\infty} a_n \, \cos\lambda_n \, x \ ,$$

where $a_n \equiv [c_n + c_{-n}]$. Using (3.13) we find the following expression for the $\{a_n\}$:

$$a_n = \frac{2}{L} \int\limits_{\substack{\text{one} \\ \text{period}}} \cos\lambda_n x \, f(x) \, dx \ .$$

The above expression for $f(x)$ as a sum of cosines is known as a **Fourier cosine series** and the $\{a_n\}$ are the **Fourier cosine coefficients**.

Similarly, one can consider the Fourier series of an odd periodic function $f(-x) = -f(x)$. Now we have that

$$f(x) = \tfrac{1}{2}\left[f(x) - f(-x)\right] \ ,$$

which, when we substitute its Fourier series, becomes

$$f(x) = \sum_{n=-\infty}^{\infty} \tfrac{1}{2}\, c_n \left[e_n(x) - e_n(-x)\right] = \sum_{n=-\infty}^{\infty} i \, c_n \, \sin\lambda_n x \ .$$

Now we use the fact that $\sin\lambda_{-n}x = -\sin\lambda_n x$, and that $\lambda_0 = 0$, to rewrite the series as

$$f(x) = \sum_{n=1}^{\infty} i \ \left[c_n - c_{-n}\right] \sin\lambda_n x = \sum_{n=1}^{\infty} b_n \, \sin\lambda_n x \ ,$$

where $b_n \equiv i\,[c_n - c_{-n}]$. Using (3.13) we find the following expression for the $\{b_n\}$:

$$b_n = \frac{2}{L} \int\limits_{\substack{\text{one} \\ \text{period}}} \sin \lambda_n x\, f(x)\, dx \ .$$

The above expression for $f(x)$ as a sum of sines is known as a **Fourier sine series** and the $\{b_n\}$ are the **Fourier sine coefficients**.

Any function can be decomposed into the sum of an odd and an even function and this is reflected in the fact that the complex exponential $e_n(x)$ can be decomposed into a sum of a cosine and a sine: $e_n(x) = \cos \lambda_n x + i \sin \lambda_n x$. Therefore for $f(x)$ periodic, we have

$$\begin{aligned} f(x) = \sum_{n=-\infty}^{\infty} c_n\, e_n(x) = \sum_{n=-\infty}^{\infty} c_n\, [\cos \lambda_n x + i\, \sin \lambda_n x] \\ = \tfrac{1}{2} a_0 + \sum_{n=1}^{\infty} a_n\, \cos \lambda_n x + \sum_{n=1}^{\infty} b_n\, \sin \lambda_n x \ , \end{aligned}$$

where the first two terms comprise a Fourier cosine series and the last term is a Fourier sine series.

Parseval's identity

Let $f(x)$ be a complex periodic function and let us compute the following integral

$$\|f\|^2 \equiv \frac{1}{L} \int\limits_{\substack{\text{one} \\ \text{period}}} |f(x)|^2\, dx \ ,$$

using the Fourier series.

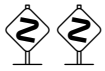
$$\|f\|^2 = \frac{1}{L} \int\limits_{\substack{\text{one} \\ \text{period}}} \left| \sum_{n=-\infty}^{\infty} c_n\, e_n(x) \right|^2 dx \ .$$

Expanding the right-hand side and interchanging the order of integration and summation, we have

$$\|f\|^2 = \frac{1}{L} \sum_{n,m=-\infty}^{\infty} c_n^*\, c_m \int\limits_{\substack{\text{one} \\ \text{period}}} e_n(x)^*\, e_m(x)\, dx = \sum_{n=-\infty}^{\infty} |c_n|^2 \ ,$$

where we have used the orthogonality relation (3.12). In other words, we have derived **Parseval's identity**:

$$\boxed{\sum_{n=-\infty}^{\infty} |c_n|^2 = \|f\|^2 \ .} \tag{3.18}$$



Explain the Fourier series as setting up an isometry between L^2 and ℓ^2.

The Dirac delta "function"

Let us insert the expression (3.13) for the Fourier coefficients back into the Fourier series (3.11) for a periodic function $f(x)$:

$$f(x) = \sum_{n=-\infty}^{\infty} \left[\frac{1}{L} \int\limits_{\substack{\text{one} \\ \text{period}}} e_n(y)^* \, f(y) \, dy \right] \, e_n(x) \ .$$

Interchanging the order of summation and integration, we find

$$f(x) = \int\limits_{\substack{\text{one} \\ \text{period}}} \left[\sum_{n=-\infty}^{\infty} \frac{1}{L} \, e_n(y)^* \, e_n(x) \right] \, f(y) \, dy = \int\limits_{\substack{\text{one} \\ \text{period}}} \delta(x-y) \, f(y) \, dy \ ,$$

where we have introduced the **Dirac delta "function"**

$$\boxed{\delta(x-y) \equiv \sum_{n=-\infty}^{\infty} \frac{1}{L} \, e_n(y)^* \, e_n(x) = \sum_{n=-\infty}^{\infty} \frac{1}{L} \, e^{i(x-y)2\pi n/L} \ .} \tag{3.19}$$

Despite its name, the delta function is *not* a function, even though it is a limit of functions. Instead it is a **distribution**. Distributions are only well-defined when integrated against sufficiently well-behaved functions known as **test functions**. The delta function is the distribution defined by the condition:

$$\int\limits_{\substack{\text{one} \\ \text{period}}} \delta(x-y) \, f(y) \, dy = f(x) \ .$$

In particular,

$$\int\limits_{\substack{\text{one} \\ \text{period}}} \delta(y) \, dy = 1 \ ,$$

hence it depends on the region of integration. This is clear from the above expression which has an explicit dependence on the period L. In the following section, we will see another delta functions adapted to a different region of integration: the whole real line.

3.1.4 Application: steady-state response

We now come to one of the main applications of the Fourier series: finding steady-state solutions to differential equations.

Consider a system governed by a differential equation

$$\frac{d^2\phi(t)}{dt^2} + a_1 \frac{d\phi(t)}{dt} + a_0\, \phi(t) = e^{i\omega t} \ .$$

The function $\phi(t)$ can be understood as the response of the system which is being driven by a sinusoidal force $e^{i\omega t}$. After sufficient time has elapsed, or assuming that we have been driving the system in this fashion for an infinitely long time, say, for all $t < 0$, a realistic system will be in a so-called **steady state**: in which $\phi(t) = A(\omega)e^{i\omega t}$. The reason is that energy dissipates in a realistic system due to damping or friction, so that in the absence of the driving term, the system will tend to lose all its energy: so that $\phi(t) \to 0$ in the limit as $t \to \infty$. To find the steady-state response of the above system one then substitutes $\phi(t) = A(\omega)e^{i\omega\, t}$ in the equation and solves for $A(\omega)$:

$$\frac{d^2\phi(t)}{dt^2} + a_1 \frac{d\phi(t)}{dt} + a_0\, \phi(t) = A(\omega)\left(-\omega^2 + i\, a_1\, \omega + a_0\right)\, e^{i\omega t} = e^{i\omega t} \ ,$$

whence

$$A(\omega) = \frac{1}{-\omega^2 + i\, a_1\, \omega + a_0} \ .$$

In practice, one would like however to analyse the steady-state response of a system which is being driven not by a simple sinusoidal function but by a general periodic function $f(t)$, with period T. This suggests that we expand the driving force in terms of a Fourier series:

$$f(t) = \sum_{n=-\infty}^{\infty} c_n\, e^{i2\pi nt/T} \ ,$$

where the coefficients are given by

$$c_n = \frac{1}{T} \int\limits_{\substack{\text{one} \\ \text{period}}} f(t)\, e^{-i2\pi nt/T}\, dt \ .$$

Above we found the steady-state response of the system for the sinusoidal forces $\exp(i2\pi nt/T)$, namely

$$\phi_n(t) = \frac{1}{-4\pi^2 n^2/T^2 + i\, a_1\, 2\pi\, n/T + a_0}\, e^{i2\pi nt/T} \ .$$

Because the equation is linear, we see that the response to a force which is a linear combination of simple sinusoids will be the same linear combination of the responses to the simple sinusoidal forces. Assuming that this can be extended to infinite linear combinations[2], we see that since $\phi_n(t)$ solves the differential equation for the driving force $\exp(i2\pi nt/T)$, then the series

$$\phi(t) = \sum_{n=-\infty}^{\infty} c_n\, \phi_n(t)$$

solves the differential equation for the driving force

$$f(t) = \sum_{n=-\infty}^{\infty} c_n e^{i2\pi nt/T} \ .$$

As an example, let us consider the differential equation

$$\frac{d^2\phi(t)}{dt^2} + 2\,\frac{d\phi(t)}{dt} + 2\,\phi(t) = f(t) \ ,$$

where $f(t)$ is the periodic function defined in (3.16). This function has period $T = 2\pi$ and according to what was said above above, the solution of this equation is

$$\phi(t) = \sum_{n=-\infty}^{\infty} \frac{c_n}{-n^2 + 2i\, n + 2}\, e^{int} \ ,$$

where the coefficients c_n are given in (3.17). Explicitly, we have

$$\phi(t) = \sum_{\substack{n=-\infty \\ n\neq 0}}^{\infty} \frac{2\,((-1)^n - 1)}{\pi^2\, n^2}\, \frac{e^{int}}{-n^2 + 2i\, n + 2} \ .$$

[2]Fourier series, since they contain an infinite number of terms, are limiting cases of linear combinations and strictly speaking we would have to justify that, for example, the derivative of the series is the series of termwise derivatives. This would follow if the series were uniformly convergent, for example. In the absence of general theorems, which will be the case in this course, one has to justify this *a posteriori*.

We would now have to check that $\phi(t)$ is twice differentiable. It could not be differentiable three times because, by the defining equation, the second derivative is given by

$$\frac{d^2\phi(t)}{dt^2} = f(t) - 2\,\frac{d\phi(t)}{dt} - 2\,\phi(t)\ ,$$

and $f(t)$ is not differentiable. The twice-differentiability of $\phi(t)$ follows from the uniform convergence of the above series for $\phi(t)$. To see this we apply the Weierstrass M-test:

$$\left|\frac{2\,((-1)^n-1)}{\pi^2\,n^2}\,\frac{e^{int}}{-n^2+2i\,n+2}\right| \le \frac{8}{\pi^2\,n^4}\ ,$$

and the series

$$\sum_{\substack{n=-\infty\\ n\neq 0}}^{\infty}\frac{8}{\pi^2\,n^4}$$

is absolutely convergent. Every time we take a derivative with respect to t, we bring down a factor of $i\,n$, hence we see that the series for $\phi(t)$ can be legitimately differentiated termwise only twice, since the series

$$\sum_{\substack{n=-\infty\\ n\neq 0}}^{\infty}\frac{8}{\pi^2\,n^3} \qquad \text{and} \sum_{\substack{n=-\infty\\ n\neq 0}}^{\infty}\frac{8}{\pi^2\,n^2}$$

are still absolutely convergent, but the series

$$\sum_{\substack{n=-\infty\\ n\neq 0}}^{\infty}\frac{8}{\pi^2\,n}$$

is not.

Green's functions

Let us return for a moment to the general second order differential equation treated above:

$$\frac{d^2\phi(t)}{dt^2} + a_1\,\frac{d\phi(t)}{dt} + a_0\,\phi(t) = f(t)\ , \tag{3.20}$$

where $f(t)$ is periodic with period T and can be expanded in a Fourier series

$$f(t) = \sum_{n=-\infty}^{\infty} c_n\,e^{i2\pi nt/T}\ .$$

Then as we have just seen, the solution is given by

$$\phi(t) = \sum_{n=-\infty}^{\infty} \frac{c_n \, e^{i2\pi nt/T}}{-4\pi^2 n^2/T^2 + i\, a_1 \, 2\pi \, n/T + a_0} \ ,$$

where the coefficients c_n are given by

$$c_n = \frac{1}{T} \int\limits_{\substack{\text{one} \\ \text{period}}} f(t) \, e^{-i2\pi nt/T} \, dt \ .$$

Inserting this back into the solution, we find

$$\phi(t) = \sum_{n=-\infty}^{\infty} \left[\frac{1}{T} \int\limits_{\substack{\text{one} \\ \text{period}}} f(\tau) \, e^{-i2\pi n\tau/T} \, d\tau \right] \frac{e^{i2\pi nt/T}}{-4\pi^2 n^2/T^2 + i\, a_1 \, 2\pi \, n/T + a_0}$$

Interchanging the order of summation and integration,

$$\phi(t) = \int\limits_{\substack{\text{one} \\ \text{period}}} \left[\frac{1}{T} \sum_{n=-\infty}^{\infty} \frac{e^{i2\pi n(t-\tau)/T}}{-4\pi^2 n^2/T^2 + i\, a_1 \, 2\pi \, n/T + a_0} \right] f(\tau) \, d\tau \ ,$$

which we can write as

$$\phi(t) = \int\limits_{\substack{\text{one} \\ \text{period}}} G(t-\tau) \, f(\tau) \, d\tau \ , \tag{3.21}$$

where

$$G(t) \equiv \sum_{n=-\infty}^{\infty} \frac{T \, e^{i2\pi nt/T}}{-4\pi^2 n^2 + i\, a_1 \, 2\pi \, nT + a_0 \, T^2}$$

is the **Green's function** for the above equation. It is defined (formally) as the solution of the differential equation

$$\frac{d^2 G(t)}{dt^2} + a_1 \frac{dG(t)}{dt} + a_0 \, G(t) = \delta(t) \ , \tag{3.22}$$

where

$$\delta(t) = \frac{1}{T} \sum_{n=-\infty}^{\infty} e^{i2\pi nt/T}$$

is the Dirac delta "function." In other words, the Green's function is the response of the system to a delta function. It should be clear that if $G(t)$ satisfies (3.22) then $\phi(t)$ given by (3.21) satisfies the original equation (3.20):

$$\begin{aligned}\frac{d^2\phi(t)}{dt^2} + a_1\,\frac{d\phi(t)}{dt} + a_0\,\phi(t) \\ &= \int\limits_{\substack{\text{one}\\ \text{period}}} \left(\frac{d^2G(t-\tau)}{dt^2} + a_1\,\frac{dG(t-\tau)}{dt} + a_0\,G(t-\tau)\right) f(\tau)\,d\tau \\ &= \int\limits_{\substack{\text{one}\\ \text{period}}} \delta(t-\tau)\,f(\tau)\,d\tau = f(t)\ .\end{aligned}$$

3.2 The Fourier transform

In the previous section we have seen how to expand a periodic function as a trigonometric series. This can be thought of as a decomposition of a periodic function in terms of elementary modes, each of which has a definite frequency allowed by the periodicity. If the function has period L, then the frequencies must be integer multiples of the fundamental frequency $k = 2\pi/L$. In this section we would like to establish a similar decomposition for functions which are not periodic. A non-periodic function can be thought of as a periodic function in the limit $L \to \infty$. Clearly, the larger L is, the less frequently the function repeats, until in the limit $L \to \infty$ the function does not repeat at all. In the limit $L \to \infty$ the allowed frequencies become a continuum and the Fourier sum goes over to a Fourier integral. In this section we will discuss this integral as well as some of its basic properties, and apply it to a variety of situations: solution of the wave equation and steady-state solutions to differential equations. As in the previous section we will omit most of the analytic details which are necessary to justify the cavalier operations we will be performing.

3.2.1 The Fourier integral

Consider a function $f(x)$ defined on the real line. If $f(x)$ were periodic with period L, say, we could try to expand $f(x)$ in a Fourier series converging to it almost everywhere within each period

$$f(x) = \sum_{n=-\infty}^{\infty} c_n\,e^{i2\pi nx/L}\ ,$$

where the coefficients $\{c_n\}$ are given by

$$c_n = \frac{1}{L} \int_{-L/2}^{L/2} f(x)\, e^{-i2\pi nx/L}\, dx \ , \tag{3.23}$$

where we have chosen the period to be $[-L/2, L/2]$ for convenience in what follows. Even if $f(x)$ is *not* periodic, we can still define a function

$$f_L(x) = \sum_{n=-\infty}^{\infty} c_n\, e^{i2\pi nx/L} \ , \tag{3.24}$$

with the same $\{c_n\}$ as above. By construction, this function $f_L(x)$ is periodic with period L and moreover agrees with $f(x)$ for almost all $x \in [-L/2, L/2]$. Then it is clear that as we make L larger and larger, then $f_L(x)$ and $f(x)$ agree (almost everywhere) on a larger and larger subset of the real line. One should expect that in the limit $L \to \infty$, $f_L(x)$ should converge to $f(x)$ in some sense. The task ahead is to find reasonable expressions for the limit $L \to \infty$ of the expression (3.24) of $f_L(x)$ and of the coefficients (3.23).

The continuum limit in detail.

This prompts us to define the **Fourier (integral) transform** of the function $f(x)$ as

$$\boxed{\mathcal{F}\{f\}(k) \equiv \hat{f}(k) = \frac{1}{2\pi} \int_{-\infty}^{\infty} f(x)\, e^{-ikx}\, dx \ ,} \tag{3.25}$$

provided that the integral exists. Not every function $f(x)$ has a Fourier transform. A sufficient condition is that it be **square-integrable**; that is, so that the following integral converges:

$$\|f\|^2 \equiv \int_{-\infty}^{\infty} |f(x)|^2\, dx \ .$$

If in addition of being square-integrable, the function is continuous, then one also has the **inversion formula**

$$\boxed{f(x) = \int_{-\infty}^{\infty} \hat{f}(k)\, e^{ikx}\, dk \ .} \tag{3.26}$$

More generally, one has the **Fourier inversion theorem**, which states that if $f(x)$ is square-integrable, then the Fourier transform $\hat{f}(k)$ exists and moreover

$$\int_{-\infty}^{\infty} \hat{f}(k)\, e^{ikx}\, dk = \begin{cases} f(x), & \text{if } f \text{ is continuous at } x, \text{ and} \\ \frac{1}{2}\left[\lim_{y \nearrow x} + \lim_{y \searrow x}\right] f(y), & \text{otherwise.} \end{cases}$$

In other words, at a point of discontinuity, the inverse transform produces the average of the left and right limiting values of the function f. This was also the case with the Fourier series. In any case, assuming that the function $f(x)$ is such that its points of discontinuity are isolated, then the inverse transform will agree with $f(x)$ everywhere but at the discontinuities.

Some examples

Before discussing any general properties of the Fourier transform, let us compute some examples.

Let $f(x) = 1/(4+x^2)$. This function is clearly square-integrable. Indeed, the integral

$$\|f\|^2 = \int_{-\infty}^{\infty} \frac{1}{(4+x^2)^2}\, dx$$

can be computed using the residue theorem as we did in Section 2.4.3. We will not do the calculation in detail, but simply remark that $\|f\|^2 = \pi/16$. Therefore its Fourier transform exists:

$$\hat{f}(k) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{e^{-ikx}}{4+x^2}\, dx\ .$$

We can compute this integral using the residue theorem. According to equation (2.64), we have that for $k < 0$, we pick up the residues of the poles in the upper half-plane, whereas for $k > 0$ we pick up the poles in the lower half-plane. The function $\exp(ikz)/(4+z^2)$ has simple poles at $z = \pm 2i$. Therefore we have

$$\begin{aligned} \hat{f}(k) &= \begin{cases} \frac{1}{2\pi}\, 2\pi i\ \mathrm{Res}(2i)\ , & \text{if } k \le 0, \text{ and} \\ \frac{1}{2\pi}\, (-2\pi i)\ \mathrm{Res}(-2i)\ , & \text{if } k \ge 0; \end{cases} \\ &= \begin{cases} \frac{1}{4}\, e^{2k}\ , & \text{if } k \le 0, \text{ and} \\ \frac{1}{4}\, e^{-2k}\ , & \text{if } k \ge 0; \end{cases} \\ &= \frac{1}{4}\, e^{-2|k|}\ . \end{aligned}$$

We can also verify the inversion formula. Indeed,

$$\begin{aligned} f(x) &\stackrel{?}{=} \int_{-\infty}^{\infty} \frac{1}{4}\, e^{-2|k|}\, e^{ikx}\, dk \\ &= \int_{-\infty}^{0} \frac{1}{4} e^{2k}\, e^{ikx}\, dk + \int_{0}^{\infty} \frac{1}{4} e^{-2k}\, e^{ikx}\, dk \\ &= \int_{0}^{\infty} \frac{1}{4} e^{-2k-ikx}\, dk + \int_{0}^{\infty} \frac{1}{4} e^{-2k+ikx}\, dk \\ &= \frac{1}{4}\left[\frac{1}{2+ix} + \frac{1}{2-ix}\right] \\ &= \frac{1}{4+x^2}\ . \end{aligned}$$

As our second example consider the Fourier transform of a **pulse**:

$$f(x) = \begin{cases} 1\ , & \text{for } |x| < \pi, \text{ and} \\ 0\ , & \text{otherwise.} \end{cases} \tag{3.27}$$

It is clearly square-integrable, with $\|f\|^2 = 2\pi$. Its Fourier transform is given by

$$\hat{f}(k) = \frac{1}{2\pi} \int_{-\infty}^{\infty} f(x)\, e^{-ikx}\, dx = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-ikx}\, dx = \frac{\sin \pi k}{\pi k}\ .$$

We will not verify the inversion formula in this example. If we were to do this we would be able to evaluate the integral in the inversion formula for $x \neq \pm\pi$ and we would obtain $f(x)$ for those values. The residue methods fail at the discontinuities $x = \pm\pi$, and one has to appeal to more advanced methods we will not discuss in this course.

Finally consider the Fourier transform of a **finite wave train**:

$$f(x) = \begin{cases} \sin x\ , & \text{for } |x| \leq 6\pi; \text{ and} \\ 0\ , & \text{otherwise.} \end{cases}$$

This function is clearly square-integrable, since

$$\|f\|^2 = \int_{-6\pi}^{6\pi} (\sin x)^2\, dx = 6\pi\ .$$

Its Fourier transform is given by

$$\begin{aligned}
\hat{f}(k) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} f(x)\, e^{-ikx}\, dx \\
&= \frac{1}{2\pi}\int_{-6\pi}^{6\pi} \sin x\, e^{-ikx}\, dx \\
&= \frac{1}{4\pi i}\int_{-6\pi}^{6\pi} \left(e^{ix} - e^{-ix}\right)\, e^{-ikx}\, dx \\
&= \frac{1}{4\pi i}\int_{-6\pi}^{6\pi} \left(e^{i(1-k)x} - e^{-i(1+k)x}\right)\, dx \\
&= \frac{1}{4\pi i}\left[\frac{i}{k-1}\left(e^{-ik6\pi} - e^{ik6\pi}\right) - \frac{i}{1+k}\left(e^{-ik6\pi} - e^{ik6\pi}\right)\right] \\
&= \frac{i\sin 6\pi k}{2\pi}\left[\frac{1}{1-k} + \frac{1}{1+k}\right] \\
&= \frac{i\sin 6\pi k}{\pi(1-k^2)}\ .
\end{aligned}$$

We will not verify the inversion formula for this transform; although in this case the formula holds for all x since the original function is continuous.

3.2.2 Some properties of the Fourier transform

In this section we will discuss some basic properties of the Fourier transform. All the basic properties of Fourier series extend in some way to the Fourier integral. Although we will not discuss all of them, it would be an instructive exercise nevertheless to try and guess and prove the extensions by yourself.

The first basic property is that if $\hat{f}(k)$ is the Fourier transform of $f(x)$, then $\hat{f}(-k)^*$ is the Fourier transform of $f(x)^*$. This follows simply by taking the complex conjugate of the Fourier integral (3.25):

$$\hat{f}(k)^* = \frac{1}{2\pi}\int_{-\infty}^{\infty} f(x)^*\, e^{ikx}\, dx = \mathcal{F}\left\{f(x)^*\right\}(-k)\ ,$$

whence $\hat{f}(-k)^* = \mathcal{F}\left\{f(x)^*\right\}(k)$. Therefore we conclude that if $f(x)$ is real, then $\hat{f}(k)^* = \hat{f}(-k)$.

Suppose that $f'(x) = \frac{df(x)}{dx}$ is also square-integrable. Its Fourier transform is given by

$$\mathcal{F}\left\{f'(x)\right\}(k) = \frac{1}{2\pi}\int_{-\infty}^{\infty} f'(x)\, e^{-ikx}\, dx\ .$$

Let us integrate by parts:

$$\mathcal{F}\left\{f'(x)\right\}(k) = \frac{1}{2\pi}\int_{-\infty}^{\infty}(ik)\, f(x)\, e^{-ikx}\, dx \ ,$$

where we have dropped the boundary terms since $f(x)$ is square-integrable and hence vanishes in the limit $|x| \to \infty$. In other words,

$$\mathcal{F}\left\{f'(x)\right\}(k) = ik\, \mathcal{F}\left\{f(x)\right\}(k) \ . \tag{3.28}$$

More generally, if the n-th derivative $f^{(n)}(x)$ is square-integrable, then

$$\boxed{\mathcal{F}\left\{f^{(n)}(x)\right\}(k) = (ik)^n\, \mathcal{F}\left\{f(x)\right\}(k) \ .} \tag{3.29}$$

This is one of the most useful properties of the Fourier transform, since it will allow us to turn differential equations into algebraic equations.

Another version of the Dirac delta function

Let $f(x)$ be a continuous square-integrable function. In this case, the Fourier inversion theorem says that the inversion formula is valid, so that

$$f(x) = \int_{-\infty}^{\infty} \hat{f}(k)\, e^{ikx}\, dk \ .$$

If we insert the definition of the Fourier transform $\hat{f}(k)$ in this equation, we obtain

$$f(x) = \int_{-\infty}^{\infty}\left[\frac{1}{2\pi}\int_{-\infty}^{\infty} f(y)\, e^{-iky}\, dy\right] e^{ikx}\, dk \ .$$

If f is in addition sufficiently well-behaved[3] we can exchange the order of integrations to obtain

$$f(x) = \int_{-\infty}^{\infty}\left[\frac{1}{2\pi}\int_{-\infty}^{\infty} e^{ik(x-y)}\, dk\right] f(y)\, dy = \int_{-\infty}^{\infty}\delta(x-y)\, f(y)\, dy \ ,$$

where we have introduced the **Dirac delta function**

$$\delta(x) \equiv \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{ikx}\, dk \ .$$

[3]Technically, it is enough that f belong to the **Schwarz class**, consisting of those infinitely differentiable functions which decay, together with all its derivatives, sufficiently fast at infinity.

Notice that we can also write this as

$$\boxed{\delta(x) \equiv \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-ikx}\, dk\ ,} \tag{3.30}$$

which makes it clear that it is the Fourier transform of the constant function $f(x) = 1$. Of course, this function is not square-integrable, so this statement is purely formal. We should not expect anything better because the Dirac delta function is not a function. This version of the Dirac delta function is adapted to the integral over the whole real line, as opposed to the one defined by equation (3.19), which is adapted to a finite interval.

Parseval's identity revisited

Another result from Fourier series which extends in some fashion to the Fourier integral transform is the one in equation (3.18). We will first attempt to show that the Fourier transform of a square-integrable function is itself square-integrable. Let us compute

$$\begin{aligned}
\|\hat{f}\|^2 &= \int_{-\infty}^{\infty} |\hat{f}(k)|^2\, dk \\
&= \int_{-\infty}^{\infty} \left| \frac{1}{2\pi} \int_{-\infty}^{\infty} f(x)\, e^{-ikx}\, dx \right|^2 dk \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{1}{4\pi^2} f(x)\, f(y)^*\, e^{-ikx}\, e^{iky}\, dx\, dy\, dk\ .
\end{aligned}$$

Being somewhat cavalier, let us interchange the order of integration so that we do the k-integral first. Recognising the result as $2\pi\delta(x-y)$, with $\delta(x-y)$ the delta function of (3.30), we can simplify this to

$$\|\hat{f}\|^2 = \int_{-\infty}^{\infty} \frac{1}{2\pi} f(x)\, f(x)^*\, dx = \frac{1}{2\pi} \int_{-\infty}^{\infty} |f(x)|^2\, dx = \frac{1}{2\pi} \|f\|^2\ .$$

Therefore since $\|f\|^2$ is finite, so is $\|\hat{f}\|^2$, and moreover their norms are related by **Parseval's identity**:

$$\boxed{\|\hat{f}\|^2 = \frac{1}{2\pi} \|f\|^2\ ,} \tag{3.31}$$

which is the integral version of equation (3.18).

For many applications this factor of $1/2\pi$ is a nuisance and one redefines the Fourier transform so that

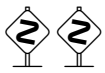
$$\hat{\mathcal{F}}\{f\}(k) \equiv \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x)\, e^{-ikx}\, dx\ ,$$

and the inversion formula is more symmetrical

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \hat{\mathcal{F}}\{f\}(k)\, e^{ikx}\, dk \ .$$

In this case, Parseval's identity becomes simply

$$\|\hat{\mathcal{F}}\{f\}\|^2 = \|f\|^2 \ .$$



One should mention that the Fourier transform is an isometry from L^2 to L^2.

3.2.3 Application: one-dimensional wave equation

Let us now illustrate the use of the Fourier transform to solve partial differential equations by considering the one-dimensional wave equation (3.1) again. This time, however, we are not imposing the boundary conditions (3.2) for x. Instead we may impose that at each moment in time t, $\psi(x,t)$ is square-integrable, which is roughly equivalent to saying that the wave has a finite amount of energy. As initial conditions we will again impose (3.3), where $f(x)$ is a square-integrable function.

We will analyse this problem by taking the Fourier transform of the wave equation. From equation (3.29) with $n = 2$ we have that

$$\mathcal{F}\left\{\frac{\partial^2}{\partial x^2}\psi(x,t)\right\} = -k^2\,\hat{\psi}(k,t) \ ,$$

where

$$\hat{\psi}(k,t) \equiv \frac{1}{2\pi}\int_{-\infty}^{\infty} \psi(x,t)\, e^{-ikx}\, dx \ ,$$

is the Fourier transform of $\psi(x,t)$. Similarly, taking the derivative inside the integral,

$$\mathcal{F}\left\{\frac{\partial^2}{\partial t^2}\psi(x,t)\right\} = \frac{\partial^2}{\partial t^2}\hat{\psi}(k,t) \ .$$

Therefore the wave equation becomes

$$\frac{1}{c^2}\,\frac{\partial^2}{\partial t^2}\hat{\psi}(k,t) = -k^2\hat{\psi}(k,t) \ .$$

The most general solution is given by a linear combination of two sinusoids:

$$\hat{\psi}(k,t) = \hat{a}(k)\,\cos kct + \hat{b}(k)\,\sin kct \ ,$$

where the "constants" $\hat{a}$ and $\hat{b}$ can still depend on k. The first of the initial conditions (3.3) implies that

$$\left.\frac{\partial\hat{\psi}(k,t)}{\partial t}\right|_{t=0} = 0 \ ,$$

whence we have that $\hat{b}(k) = 0$. Using the inversion formula (3.26), we can write

$$\psi(x,t) = \int_{-\infty}^{\infty} \hat{a}(k)\, \cos kct\, e^{ikx}\, dk \ .$$

Evaluating at $t = 0$, we have that

$$\psi(x,t) = f(x) = \int_{-\infty}^{\infty} \hat{a}(k)\, e^{ikx}\, dk \ ,$$

whence comparing with the inversion formula (3.26), we see that $\hat{a}(k) = \hat{f}(k)$, so that

$$\psi(x,t) = \int_{-\infty}^{\infty} \hat{f}(k)\, \cos kct\, e^{ikx}\, dk \ , \tag{3.32}$$

where

$$\hat{f}(k) = \frac{1}{2\pi} \int_{-\infty}^{\infty} f(x)\, e^{-ikx}\, dx \ .$$

Inserting back this expression into the solution (3.32) and interchanging the order of integration, we have

$$\begin{aligned}
\psi(x,t) &= \int_{-\infty}^{\infty} \left[\frac{1}{2\pi} \int_{-\infty}^{\infty} f(y)\, e^{-iky}\, dy \right] \cos kct\, e^{ikx}\, dk \\
&= \int_{-\infty}^{\infty} \left[\frac{1}{2\pi} \int_{-\infty}^{\infty} \cos kct\, e^{ik(x-y)}\, dk \right] f(y)\, dy \\
&= \int_{-\infty}^{\infty} K(x-y,t)\, f(y)\, dy \ ,
\end{aligned}$$

where we have introduced the **propagator** $K(x,t)$ defined by

$$K(x,t) \equiv \frac{1}{2\pi} \int_{-\infty}^{\infty} \cos kct\, e^{ikx}\, dk \ .$$

Notice that $K(x,t)$ clearly satisfies the wave equation (3.1):

$$\frac{\partial^2}{\partial x^2} K(x,t) = \frac{1}{c^2}\, \frac{\partial^2}{\partial t^2} K(x,t) \ ,$$

with initial conditions

$$\left. \frac{\partial K(x,t)}{\partial t} \right|_{t=0} = 0 \ ,$$

and

$$K(x,0) = \delta(x) \ ,$$

according to (3.30).

3.2.4 Application: steady-state response

Another useful application of the Fourier transform is to solve for the steady-state solutions of linear ordinary differential equations. Suppose that we have a system governed by a differential equation

$$\frac{d^2\phi(t)}{dt^2} + a_1 \frac{d\phi(t)}{dt} + a_0\,\phi(t) = f(t) \ ,$$

where $f(t)$ is some driving term. We saw that when $f(t)$ is periodic we can use the method of Fourier series in order to solve for $\phi(t)$. If $f(t)$ is not periodic, then it makes sense that we try and use the Fourier integral transform. Let us define the Fourier transform $\hat{\phi}(\omega)$ of $\phi(t)$ by

$$\mathcal{F}\left\{\phi(t)\right\}(\omega) \equiv \hat{\phi}(\omega) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \phi(t)\, e^{-i\omega t}\, dt \ .$$

Similarly, let $\hat{f}(\omega)$ denote the Fourier transform of $f(t)$. Then we can take the Fourier transform of the differential equation and we obtain an algebraic equation for $\hat{\phi}(\omega)$:

$$-\omega^2\,\hat{\phi}(\omega) + i\,a_1\omega\,\hat{\phi}(\omega) + a_0\,\hat{\phi}(\omega) = \hat{f}(\omega) \ ,$$

which can be readily solved to yield

$$\hat{\phi}(\omega) = \frac{1}{-\omega^2 + i\,a_1\omega + a_0}\,\hat{f}(\omega) \ .$$

Now we can transform back via the inversion formula

$$\phi(t) = \int_{-\infty}^{\infty} \hat{\phi}(\omega)\, e^{i\omega t}\, d\omega = \int_{-\infty}^{\infty} \frac{e^{i\omega t}}{-\omega^2 + i\,a_1\omega + a_0}\,\hat{f}(\omega)\, d\omega \ .$$

Using the definition of $\hat{f}(\omega)$, we have

$$\phi(t) = \int_{-\infty}^{\infty} \frac{e^{i\omega t}}{-\omega^2 + i\,a_1\omega + a_0}\,\left[\frac{1}{2\pi}\int_{-\infty}^{\infty} f(\tau)\, e^{-i\tau\omega}\, d\tau\right]\, d\omega \ .$$

If, as we have been doing without justification in this part of the course, we interchange the order of integration, we obtain

$$\phi(t) = \int_{-\infty}^{\infty} G(t-\tau)\, f(\tau)\, d\tau \ ,$$

where we have introduced the **Green's function** $G(t)$, defined by

$$G(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{e^{i\omega t}}{-\omega^2 + i\,a_1\omega + a_0}\, d\omega \ .$$

Notice that as in the case of the Fourier series, $G(t)$ satisfies the equation

$$\frac{d^2G(t)}{dt^2} + a_1 \frac{dG(t)}{dt} + a_0\, G(t) = \delta(t) \ ,$$

so that it is the response of the system to a delta function input.

As a concrete illustration of the method, let us find a steady-state solution to the following differential equation:

$$\frac{d^2\phi(t)}{dt^2} + 2\,\frac{d\phi(t)}{dt} + 2\,\phi(t) = f(t) \ ,$$

where $f(t)$ is the pulse defined in (3.27). Let us first compute the Green's function for this system:

$$G(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{e^{i\omega t}}{-\omega^2 + 2i\,\omega + 2}\, d\omega \ .$$

We can compute this using the residue theorem and, in particular, equation (2.64). The integrand has simple poles at $i \pm 1$, which lie in the upper half-plane. Therefore it follows immediately from equation (2.64), that $G(t) = 0$ for $t \leq 0$. For $t > 0$, we have that

$$G(t) = \frac{1}{2\pi}\, 2\pi\, i\, \left[\mathrm{Res}(i+1) + \mathrm{Res}(i-1)\right] \ .$$

We compute the residues to be

$$\mathrm{Res}(i+1) = -\frac{e^{-t+it}}{2} \qquad \text{and} \qquad \mathrm{Res}(i-1) = \frac{e^{-t-it}}{2} \ ,$$

whence for $t > 0$, we have

$$G(t) = -\,e^{-t}\,\sin t \ .$$

In summary, the Green's function for this system is

$$G(t) = \begin{cases} 0 \ , & \text{for } t < 0, \text{ and} \\ -\,e^{-t}\,\sin t \ , & \text{for } t \geq 0. \end{cases}$$

Notice that although it is continuous at $t = 0$, its first derivative is not continuous there, and hence the second derivative does not exist at $t = 0$. This is to be expected, since the second derivative of $G(t)$ at $t = 0$ is related

to the delta function, which is not a function. In any case, we can now integrate this against the pulse $f(t)$ to find the solution:

$$\begin{aligned}\phi(t) &= \int_{-\infty}^{\infty} G(t-\tau)\, f(\tau)\, d\tau \\ &= \int_{-\pi}^{\pi} G(t-\tau)\, d\tau \ .\end{aligned}$$

Taking into account that $G(t) = 0$ for $t < 0$, we are forced to distinguish between three epochs: $t < -\pi$, $-\pi \leq t \leq \pi$, and $t > \pi$, corresponding to the time before the pulse, during the pulse and after the pulse. We can perform the integral in each of these three epochs with the following results:

$$\phi(t) = \begin{cases} 0 \ , & \text{for } t < -\pi, \\ -\frac{1}{2} - \frac{1}{2}e^{-\pi - t}\,(\cos t + \sin t) \ , & \text{for } t \in [-\pi, \pi], \text{ and} \\ e^{-t} \sinh \pi (\cos t + \sin t) \ , & \text{for } t > \pi. \end{cases}$$

Notice that before the pulse the system is at rest, and that after the pulse the response dies off exponentially. This is as we expect for a steady-state response to an input of finite duration.

3.3 The Laplace transform

In the previous section we introduced the Fourier transform as a tool to find steady-state solutions to differential equations. These solutions can be interpreted as the response of a system which has been driven for such a long time that any transient solutions have died out. In many systems, however, one is also interested in the transient solutions, and in any case, mathematically one usually finds the most general solution of the differential equation. The Laplace transform will allow us to do this. In many ways the Laplace transform is reminiscent of the Fourier transform, with the important difference that it incorporates in a natural way the initial conditions.

3.3.1 The Heaviside D-calculus

Let us start by presenting the D-calculus introduced by Heaviside. The justification for this method is the Laplace transform. An example should suffice to illustrate the method, but first we need to introduce a little bit of notation.

Differential operators

The result of taking the derivative of a function is another function: for example, $d/dt\,(t^n) = nt^{n-1}$ or $d/dt\,\sin t = \cos t$. Therefore we can think of the derivative as some sort of machine to which one feeds a function as input and gets another function in return. Such machines are generally called **operators**. It is convenient to introduce symbols for operators and, in the case of the derivative operator, it is customary to call it D. Therefore, if f is a function, Df is the function one obtains by having D act on f. A function is defined by specifying its values at every point t. In the case of Df we have

$$Df(t) = \frac{df(t)}{dt} \ .$$

Operators can be composed. For example we can consider D^2 to be the operator which acting on a function f gives $D^2 f = D(Df)$, or

$$D^2 f(t) = D(Df)(t) = \frac{dDf(t)}{dt} = \frac{d^2 f(t)}{dt^2} \ .$$

Therefore D^2 is the second derivative. Operators can be multiplied by functions, and in particular, by constants. If a is a constant, the operator aD is defined by

$$(aD)f(t) \equiv a\,Df(t) = a\,\frac{df(t)}{dt} \ .$$

Similarly, if $g(t)$ is a function, then the operator gD is defined by

$$(gD)f(t) \equiv g(t)\,Df(t) = g(t)\,\frac{df(t)}{dt} \ .$$

Operators can also be added: if g and h are functions, then the expression $gD^2 + hD$ is an operator, defined by

$$(gD^2 + hD)f(t) = g(t)\,\frac{d^2 f(t)}{dt^2} + h(t)\,\frac{df(t)}{dt} \ .$$

In other words, linear combinations of operators are again operators. Operators which are formed by linear combinations with function coefficients of D and its powers are known as **differential operators**. A very important property shared by all differential operators is that they are **linear**. Let us consider the derivative operator D, and let $f(t)$ and $g(t)$ be functions. Then,

$$D(f+g)(t) = \frac{d(f(t)+g(t))}{dt} = \frac{df(t)}{dt} + \frac{dg(t)}{dt} = Df(t) + Dg(t) \ .$$

In other words, $D(f+g) = Df + Dg$. Similarly, it is easy to see that this is still true for any power of D and for any linear combination of powers of D. In summary, differential operators are linear.

The highest power of D which occurs in a differential operator is called the **order** of the differential operator. This agrees with the nomenclature used for differential equations. In fact, a second order ordinary differential equation, like this one

$$a(t)\,\frac{d^2f(t)}{dt^2} + b(t)\,\frac{df(t)}{dt} + c(t)\,f(t) = h(t)\ ,$$

can be rewritten as an operator equation $Kf(t) = h(t)$, where we have introduced the second order differential operator $K = a\,D^2 + b\,D + c$.

An example

Suppose we want to solve the following differential equation

$$\frac{d^2f(t)}{dt^2} + 3\,\frac{df(t)}{dt} + 2\,f(t) = e^{it}\ . \tag{3.33}$$

We first write it down as an operator equation:

$$(D^2 + 3D + 2)\,f(t) = e^{it}\ .$$

Next we will manipulate the operator formally as if D were a variable and not an operator:

$$D^2 + 3D + 2 = (D+2)(D+1)\ ;$$

whence formally

$$f(t) = \frac{1}{(D+2)(D+1)}\,e^{it} = \left[\frac{1}{D+1} - \frac{1}{D+2}\right]\,e^{it}\ , \tag{3.34}$$

where we have used a partial fraction expansion: remember we are treating D as if it were a variable z, say. Now we do something even more suspect and expand each of the simple fractions using a geometric series:

$$\frac{1}{D+1} = \sum_{j=0}^{\infty}(-1)^j D^j \qquad \text{and} \qquad \frac{1}{D+2} = \sum_{j=0}^{\infty}(-1)^j \frac{1}{2^{j+1}}\,D^j\ .$$

Now notice that $D\,e^{it} = i\,e^{it}$; hence

$$\frac{1}{D+1}\,e^{it} = \sum_{j=0}^{\infty}(-1)^j D^j\,e^{it} = \sum_{j=0}^{\infty}(-1)^j i^j\,e^{it} = \frac{1}{i+1}\,e^{it}\ ,$$

$$\frac{1}{D+2}\,e^{it} = \sum_{j=0}^{\infty}(-1)^j \frac{1}{2^{j+1}}\,D^j\,e^{it} = \sum_{j=0}^{\infty}(-1)^j \frac{i^j}{2^{j+1}}\,e^{it} = \frac{1}{i+2}\,e^{it}\ .$$

Therefore into equation (3.34), we obtain

$$f(t) = \left[\frac{1}{i+1} - \frac{1}{i+2}\right] e^{it} = \frac{1-3i}{10}\, e^{it} \ ,$$

which can be checked to obey equation (3.33) by direct substitution.

Of course this is only a **particular solution** to the differential equation (3.33). In order to obtain the most general solution we have to add to it the **complementary solution**, which is the most general solution of the **associated homogeneous equation**:

$$Kf(t) = (D+1)(D+2)f(t) = \frac{d^2 f(t)}{dt^2} + 3\,\frac{df(t)}{dt} + 2\,f(t) = 0 \ .$$

The reason for this is that if $g(t)$ solves the equation $Kg(t) = 0$, and $Kf(t) = e^{it}$, then, by linearity, $K(f+g)(t) = Kf(t) + Kg(t) = e^{it} + 0 = e^{it}$. To find the complementary solution, notice that

$$(D+1)(D+2)f(t) = 0$$

has two kinds of solutions:

$$(D+1)f_1(t) = 0 \qquad \text{and} \qquad (D+2)f_2(t) = 0 \ .$$

These first order equations can be read off immediately:

$$f_1(t) = a\, e^{-t} \qquad \text{and} \qquad f_2(t) = b\, e^{-2t} \ ,$$

where the constants a and b are to be determined from the initial conditions: $f(0)$ and $f'(0)$, say. In summary, we have the following general solution to the differential equation (3.33):

$$f(t) = \frac{1-3i}{10}\, e^{it} + a\, e^{-t} + b\, e^{-2t} \ , \tag{3.35}$$

which can be checked explicitly to solve the differential equation (3.33). Notice that the first term corresponds to the steady-state response and the last two terms are transient.

3.3.2 The Laplace transform

The D-calculus might seem a little suspect, but it can be justified by the use of the **Laplace transform**, which we define as follows

$$\boxed{\mathcal{L}\left\{f\right\}(s) \equiv F(s) = \int_0^\infty f(t)\, e^{-st}\, dt \ ,} \tag{3.36}$$

provided that the integral exists. This might restrict the values of s for which the transform exists.

A function $f(t)$ is said to be of **exponential order** if there exist real constants M and α for which

$$|f(t)| \leq M\, e^{\alpha t} \ .$$

It is not hard to see that if $f(t)$ is of exponential order, then the Laplace transform $F(s)$ of $f(t)$ exists provided that $\mathrm{Re}(s) > \alpha$.

To see this let us estimate the integral

$$|F(s)| = \int_0^\infty f(t)\, e^{-st}\, dt \ \leq \int_0^\infty |f(t)|\, |e^{-st}|\, dt \leq \int_0^\infty M\, e^{\alpha t}\, e^{-\,\mathrm{Re}(s)t}\, dt \ .$$

Provided that $\mathrm{Re}(s) > \alpha$, this integral exists and

$$|F(s)| \leq \frac{M}{\mathrm{Re}(s) - \alpha} \ .$$

Notice that in particular, in the limit $\mathrm{Re}(s) \to \infty$, $F(s) \to 0$. This can be proven in more generality: so that if a function $F(s)$ does not approach 0 in the limit $\mathrm{Re}(s) \to \infty$, it cannot be the Laplace transform of any function $f(t)$.

We postpone a more complete discussion of the properties of the Laplace transform until later, but for now let us note the few properties we will need to justify the D-calculus solution of the differential equation (3.33) above.

The first important property is that the Laplace transform is **linear**. Clearly, if $f(t)$ and $g(t)$ are functions whose Laplace transforms $F(s)$ and $G(s)$ exist, then for those values of s for which *both* $F(s)$ and $G(s)$ exist, we have that

$$\mathcal{L}\left\{f + g\right\}(s) = \mathcal{L}\left\{f\right\}(s) + \mathcal{L}\left\{g\right\}(s) = F(s) + G(s) \ .$$

Next let us consider the function $f(t) = \exp(at)$ where a is some complex number. This function is of exponential order, so that its Laplace transform exists provided that $\mathrm{Re}(s) > \mathrm{Re}(a)$. This being the case, we have that

$$\mathcal{L}\left\{e^{at}\right\}(s) = \int_0^\infty e^{at}\, e^{-st}\, dt = \frac{1}{s - a} \ . \tag{3.37}$$

Suppose now that $f(t)$ is a differentiable function. Let us try to compute the Laplace transform of its derivative $f'(t)$. By definition,

$$\mathcal{L}\left\{f'\right\}(s) = \int_0^\infty f'(t)\, e^{-st}\, dt \ ,$$

which can be integrated by parts to obtain

$$\mathcal{L}\{f'\}(s) = \int_0^\infty s\, f(t)\, e^{-st}\, dt + f(t)\, e^{-st}\Big|_0^\infty$$
$$= s\,\mathcal{L}\{f\}(s) - f(0) + \lim_{t\to\infty} f(t)\, e^{-st} \ .$$

Provided the last term is zero, which might imply conditions on f and/or s, we have that

$$\boxed{\mathcal{L}\{f'\}(s) = s\,\mathcal{L}\{f\}(s) - f(0) \ .} \tag{3.38}$$

We can iterate this expression in order to find the Laplace transform of higher derivatives of $f(t)$. For example, the Laplace transform of the second derivative is easy to find by understanding $f''(t)$ as the first derivative of $f'(t)$ and iterating the above formula:

$$\begin{aligned}\mathcal{L}\{f''\}(s) &= \mathcal{L}\{(f')'\}(s) \\ &= s\,\mathcal{L}\{f'\}(s) - f'(0) \\ &= s\,(s\,\mathcal{L}\{f\}(s) - f(0)) - f'(0) \\ &= s^2\,\mathcal{L}\{f\}(s) - s\,f(0) - f'(0) \ ,\end{aligned}$$

provided that $f(t)\,\exp(-st)$ and $f'(t)\,\exp(-st)$ both go to zero in the limit $t \to \infty$.

The D-calculus justified

We are now ready to justify the D-calculus solution of the previous section. This serves also to illustrate how to solve initial value problems using the Laplace transform.

Consider again the differential equation (3.33):

$$\frac{d^2 f(t)}{dt^2} + 3\,\frac{df(t)}{dt} + 2\,f(t) = e^{it} \ ,$$

and let us take the Laplace transform of both sides of the equation. Since the Laplace transform is linear, we can write this as

$$\mathcal{L}\{f''\}(s) + 3\,\mathcal{L}\{f'\}(s) + 2\,\mathcal{L}\{f\}(s) = \mathcal{L}\left\{e^{it}\right\}(s) \ .$$

Letting $F(s)$ denote the Laplace transform of f, we can use equations (3.37) and (3.38) to rewrite this as

$$s^2 F(s) - sf(0) - f'(0) + 3\,(sF(s) - f(0)) + 2F(s) = \frac{1}{s-i} \ ,$$

which can be solved for $F(s)$:

$$F(s) = \frac{1}{s^2+3s+2}\left[\frac{1}{s-i} + (s+3)f(0) + f'(0)\right] .$$

Expanding this out, and factorising $s^2+3s+2 = (s+1)(s+2)$, we have

$$F(s) = \frac{1}{(s-i)(s+1)(s+2)} + \frac{(s+3)f(0)+f'(0)}{(s+1)(s+2)} .$$

We now decompose this into partial fractions:

$$F(s) = \frac{\frac{1}{10}(1-3i)}{s-i} + \frac{2f(0)+f'(0)-\frac{1}{2}(1-i)}{s+1} + \frac{\frac{1}{5}(2-i)-f(0)-f'(0)}{s+2} .$$

Using linearity again and (3.37) we can recognise this as the Laplace transform of the function

$$f(t) = \frac{1-3i}{10}\, e^{it} + \left(2f(0)+f'(0)-\frac{1-i}{2}\right) e^{-t} + \left(\frac{2-i}{5} - f(0) - f'(0)\right) e^{-2t} ,$$

which agrees with (3.35) and moreover displays manifestly the dependence of the coefficients a and b in that expression in terms of the initial conditions.

The inverse Laplace transform

The Laplace transform is applicable to a wide range of initial value problems. The main difficulty stems from inverting the transform, which might be difficult. In practice one resorts to tables of Laplace transforms, like Table 3.1 below; but if this does not work, there is an inversion formula, as for the Fourier transform, which we will state without proof. It says that if $F(s)$ is the Laplace transform of a function $f(t)$, then one can recover the function (except maybe at points of discontinuity) by

$$f(t) = \frac{1}{2\pi i}\int_{-i\infty}^{i\infty} F(s)\, e^{st}\, ds ,$$

where the integral is meant to be a contour integral along the imaginary axis. In other words, parametrising $s = iy$, we have

$$f(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} F(iy)\, e^{iyt}\, dy .$$

It may happen, however, that Laplace transform $F(s)$ does not make sense for $\mathrm{Re}(s) = 0$, because the integral (3.36) does not converge. Suppose instead that there is some positive real number a such that the Laplace transform of $f(t)\, e^{-at}$ does exist for $\mathrm{Re}(s) = 0$. In this case, we can use the inversion formula to obtain

$$f(t)\, e^{-at} = \frac{1}{2\pi i} \int_{-i\infty}^{i\infty} \mathcal{L}\left\{ f(t)\, e^{-at} \right\}(s)\, e^{st}\, ds \ .$$

Using the shift formula (3.43), $\mathcal{L}\left\{ f(t)\, e^{-at} \right\}(s) = F(s+a)$, whence, multiplying by e^{at} on both sides of the inversion formula:

$$f(t) = \frac{1}{2\pi i} \int_{-i\infty}^{i\infty} F(s+a)\, e^{(s+a)t}\, ds \ .$$

Changing variables of integration to $u = s + a$, we have

$$f(t) = \frac{1}{2\pi i} \int_{a-i\infty}^{a+i\infty} F(u)\, e^{ut}\, du \ , \tag{3.39}$$

which can now be interpreted as a contour integral along the line $u = a$. In other words, we can for free shift the original contour of integration to the right until $F(s)$ makes sense on it.

3.3.3 Basic properties of the Laplace transform

We shall now discuss the basic properties of the Laplace transform. We have already seen that it is linear and we computed the transform of a simple exponential function $\exp(at)$ in equation (3.37). From this simple result, we can compute the Laplace transforms of a few simple functions related to the exponential.

Let ω be a real number. From the fact that $\exp(i\omega t) = \cos\omega t + i\, \sin\omega t$, linearity of the Laplace transform implies that

$$\begin{aligned} \mathcal{L}\left\{ e^{i\omega t} \right\}(s) &= \mathcal{L}\left\{ \cos\omega t \right\}(s) + i\, \mathcal{L}\left\{ \sin\omega t \right\}(s) \\ &= \frac{1}{s - i\omega} = \frac{s}{s^2 + \omega^2} + i\, \frac{\omega}{s^2 + \omega^2} \ , \end{aligned}$$

from where we can read off the Laplace transforms of $\cos\omega t$ and $\sin\omega t$. Notice that these expressions are valid for $\mathrm{Re}(s) > 0$, since this is the condition for the existence of the Laplace transform of the exponential.

Similarly, let β be a real number and recall the trigonometric identities (2.16), from where we can deduce that

$$\cosh\beta t = \cos i\beta t \qquad \text{and} \qquad \sinh\beta t = -i \sin i\beta t \ .$$

As a result, we immediately see that the Laplace transforms of the hyperbolic functions are given by

$$\mathcal{L}\left\{\cosh \beta t\right\}(s) = \frac{s}{s^2 - \beta^2} \qquad \text{and} \qquad \mathcal{L}\left\{\sinh \beta t\right\}(s) = \frac{\beta}{s^2 - \beta^2} \ ,$$

where the condition is now $\mathrm{Re}(s) > |\beta|$.

Putting $a = 0$ in (3.37), we see that the Laplace transform of the constant function $f(t) = 1$, is given by

$$\mathcal{L}\left\{1\right\}(s) = \frac{1}{s} \ ,$$

which is valid for $\mathrm{Re}(s) > 0$.

Suppose that $f(t)$ has Laplace transform $F(s)$. Then by taking derivatives with respect to s of the expression (3.36) for $F(s)$, we arrive at

$$\mathcal{L}\left\{t^n\, f(t)\right\}(s) = (-1)^n F^{(n)}(s) \ , \tag{3.40}$$

which is valid for those values of s for which the Laplace transform $F(s)$ of $f(t)$ exists. In particular, if we take $f(t) = 1$, we arrive at

$$\mathcal{L}\left\{t^n\right\}(s) = (-1)^n \frac{d^n}{ds^n} \frac{1}{s} = \frac{n!}{s^{n+1}} \ , \tag{3.41}$$

valid for $\mathrm{Re}(s) > 0$.

How about if n is a negative integer? Let us consider the Laplace transform of $g(t) \equiv f(t)/t$, and let us call it $G(s)$. From equation (3.40) for $n = 1$, we have that

$$\mathcal{L}\left\{f(t)\right\}(s) = \mathcal{L}\left\{tg(t)\right\}(s) = -G'(s) \ ,$$

so that $G(s)$ in an antiderivative for $-F(s)$; that is,

$$G(s) = -\int_a^s F(\sigma)\, d\sigma \ .$$

If we demand that $G(s)$ vanishes in the limit $s \to \infty$, then we must choose $a = \infty$, and hence

$$\mathcal{L}\left\{f(t)/t\right\}(s) = \int_s^\infty F(\sigma)\, d\sigma \ . \tag{3.42}$$

Another important property of the Laplace transform is the **shifting formula**:

$$\mathcal{L}\left\{e^{at} f(t)\right\}(s) = \mathcal{L}\left\{f(t)\right\}(s-a) = F(s-a) \ , \tag{3.43}$$

which is evident from the definition (3.36) of the Laplace transform. Related to this property is the following. Given a function $f(t)$, let τ be a positive

real constant, and introduce the notion of the **delayed function** $f_\tau(t)$, defined by

$$f_\tau(t) = \begin{cases} f(t-\tau)\ , & \text{for } t \geq \tau, \text{ and} \\ 0\ , & \text{otherwise.} \end{cases} \tag{3.44}$$

In other words, the delayed function is the same as the original function, but it has been translated in time by τ, hence the name. The Laplace transform of the delayed function is given by

$$\begin{aligned} \mathcal{L}\left\{f_\tau\right\}(s) &= \int_0^\infty f_\tau(t)\, e^{-st}\, dt \\ &= \int_\tau^\infty f(t-\tau)\, e^{-st}\, dt \\ &= e^{-s\tau} \int_0^\infty f(u)\, e^{-su}\, du \\ &= e^{-s\tau} \mathcal{L}\left\{f\right\}(s)\ , \end{aligned}$$

where we have changed the variable of integration from t to $u = t - \tau$. In other words,

$$\mathcal{L}\left\{f_\tau\right\}(s) = e^{-s\tau} F(s)\ . \tag{3.45}$$

Although the delta function $\delta(t-\tau)$ is not a function, we can nevertheless attempt to compute its Laplace transform:

$$\mathcal{L}\left\{\delta(t-\tau)\right\}(s) = \int_0^\infty \delta(t-\tau)\, e^{-st}\, dt = \begin{cases} e^{-s\tau}\ , & \text{if } \tau \geq 0, \text{ and} \\ 0\ , & \text{otherwise.} \end{cases}$$

Introducing the **Heaviside step function** $\theta(t)$, defined as

$$\theta(t) = \begin{cases} 1\ , & \text{for } t \geq 0, \text{ and} \\ 0\ , & \text{for } t < 0; \end{cases}$$

we see that

$$\mathcal{L}\left\{\delta(t-\tau)\right\}(s) = \theta(\tau)\, e^{-s\tau}\ .$$

Finally let us consider the Laplace transforms of integrals and derivatives of functions. In the previous section we derived equation (3.38) for the Laplace transform of the derivative $f'(t)$ of a function $f(t)$. Iterating this expression we can find a formula for the Laplace transform of the n-th derivative of a function:

$$\mathcal{L}\left\{f^{(n)}\right\}(s) = s^n\, F(s) - \sum_{k=0}^{n-1} s^{n-1-k}\, f^{(k)}(0)\ , \tag{3.46}$$

where by $f^{(0)}$ we mean the original function f. This formula is valid whenever $\lim_{t\to\infty} f^{(n)}(t)\exp(-st) = 0$. How about integration? Consider the function

$$g(t) = \int_0^t f(\tau)\, d\tau \ .$$

What is its Laplace transform? We know that since $g(t)$ is an antiderivative for f, $g'(t) = f(t)$ and moreover, from the definition, that $g(0) = 0$. Therefore we can compute the Laplace transform of $f(t) = g'(t)$, in two ways. On the one hand it is simply $F(s)$, but using (3.38) we can write

$$\mathcal{L}\left\{f\right\}(s) = \mathcal{L}\left\{g'\right\}(s) = s\mathcal{L}\left\{g\right\}(s) - g(0) = s\mathcal{L}\left\{g\right\}(s) \ ,$$

whence

$$\mathcal{L}\left\{\int_0^t f(\tau)\, d\tau\right\}(s) = \frac{F(s)}{s} \ .$$

These properties are summarised in Table 3.1.

3.3.4 Application: stability and the damped oscillator

In this section we will use the Laplace transform to characterise the notion of stability of a dynamical system which is governed by a linear ordinary differential equation.

Many systems are governed by differential equations of the form

$$K\, f(t) = u(t) \ , \tag{3.47}$$

where K is an n-th order differential operator which we will take, for simplicity, to have constant coefficients and such that the coefficient of the term of highest degree is 1 ; that is,

$$K = D^n + a_{n-1}\, D^{n-1} + \cdots + a_1\, D + a_0 \ .$$

The differential equation (3.47) describes the **output response** $f(t)$ of the system to an **input** $u(t)$. For the purposes of this section we will say that a system is **stable** if in the absence of any input all solutions are transient; that is,

$$\lim_{t\to\infty} f(t) = 0 \ ,$$

regardless the initial conditions.

Often one extends the notion of stability to systems for which $f(t)$ remains bounded as $t \to \infty$: for example, if the solutions oscillate; but we will not do this here. In any case, the method we will employ extends trivially to this weaker notion of stability.

Function	Transform	Conditions
$f(t)$	$F(s)$	convergence
e^{at}	$\dfrac{1}{s-a}$	$\mathrm{Re}(s) > \mathrm{Re}(a)$
$\cos \omega t$	$\dfrac{s}{s^2+\omega^2}$	$\omega \in \mathbb{R}$ and $\mathrm{Re}(s) > 0$
$\sin \omega t$	$\dfrac{\omega}{s^2+\omega^2}$	$\omega \in \mathbb{R}$ and $\mathrm{Re}(s) > 0$
$\cosh \beta t$	$\dfrac{s}{s^2-\beta^2}$	$\beta \in \mathbb{R}$ and $\mathrm{Re}(s) > \lvert\beta\rvert$
$\sinh \beta t$	$\dfrac{\beta}{s^2-\beta^2}$	$\beta \in \mathbb{R}$ and $\mathrm{Re}(s) > \lvert\beta\rvert$
t^n	$\dfrac{n!}{s^{n+1}}$	$n = 0, 1, \ldots$ and $\mathrm{Re}(s) > 0$
$e^{at} f(t)$	$F(s-a)$	convergence
$t^n f(t)$	$(-1)^n F^{(n)}(s)$	same as for $F(s)$
$\dfrac{f(t)}{t}$	$\displaystyle\int_s^\infty F(\sigma)\, d\sigma$	same as for $F(s)$
$f_\tau(t)$	$e^{-s\tau} F(s)$	$\tau > 0$ and same as for $F(s)$
$\delta(t-\tau)$	$\theta(t-\tau)\, e^{-s\tau}$	none
$f^{(n)}(t)$	$s^n F(s) - \displaystyle\sum_{k=0}^{n-1} s^{n-1-k} f^{(k)}(0)$	$\displaystyle\lim_{t\to\infty} f^{(k)}(t) e^{-st} = 0$
$\displaystyle\int_0^t f(\tau)\, d\tau$	$\dfrac{F(s)}{s}$	same as for $F(s)$

Table 3.1: Some Laplace transforms

Stability can be analysed using the Laplace transform. In order to see this let us take the Laplace transform of the equation (3.47). Letting $F(s)$ and $U(s)$ denote the Laplace transforms of $f(t)$ and $u(t)$ respectively, we

have

$$(s^n + a_{n-1}\, s^{n-1} + \cdots + a_1\, s + a_0)F(s) = U(s) + P(s)\ , \tag{3.48}$$

where $P(s)$ is a polynomial in s of order at most $n-1$ depending on the initial conditions: $f^{(k)}(0)$ for $k = 0, 1, \ldots, n-1$. In fact, a little bit of algebra using equation (3.46) shows that

$$P(s) = \sum_{i=0}^{n-1} p_i\, s^i\ , \qquad \text{with } p_i = \sum_{j=0}^{n-i-1} a_{j+i+1} f^{(j)}(0)\ ,$$

with the conventions that $a_n = 1$. We will not need its explicit expression, however. We can solve for $F(s)$ in the transformed equation (3.48):

$$F(s) = \frac{U(s)}{s^n + \cdots + a_0} + \frac{P(s)}{s^n + \cdots + a_0}\ .$$

Notice that the first term in the right-hand side of the equation depends on the input, whereas the second term depends on the initial conditions. Moreover the common denominator depends only on the differential operator K; that is, it is intrinsic to the system. It is convenient to define the function

$$H(s) = \frac{1}{s^n + \cdots + a_0}\ .$$

It is called the **transfer function** of the system and it encodes a great deal of information about the qualitative dynamics of the system. In particular we can will be able to characterise the stability of the system by studying the poles of the transfer function in the complex s-plane.

Let us start with the case of a first order equation:

$$(D + a_0)f(t) = u(t)\ .$$

Taking the Laplace transform and solving for the Laplace transform $F(s)$ of $f(t)$ we have

$$F(s) = \frac{U(s)}{s + a_0} + \frac{f(0)}{s + a_0}\ .$$

In the absence of any input ($u = 0$), the solution of this equation is given by

$$f(t) = f(0)\, e^{-a_0\, t}\ .$$

This solution is transient provided that $\mathrm{Re}(a_0) > 0$. This is equivalent to saying that the pole $-a_0$ of the transfer function $1/(s + a_0)$ lies in the left half of the plane.

Let us now consider a second order equation:

$$(D^2 + a_1 D + a_0) f(t) = u(t) \ .$$

Taking the Laplace transform and solving for $F(s)$, we find

$$F(s) = H(s)\, U(s) + H(s)\ [(s + a_1)\, f(0) + f'(0)] \ , \tag{3.49}$$

where the transfer function is given by

$$H(s) = \frac{1}{s^2 + a_1\, s + a_0} \ .$$

The poles of $H(s)$ occur at the zeros of $s^2 + a_1\, s + a_0$. Two possibilities can occur: the zeros are simple and distinct: $s_\pm$ say, or there is one double zero at s_0. In either case we will decompose the right-hand side of the transformed equation (3.49) with $u(t)$ and hence $U(s)$ set to zero, into partial fractions. In the case of distinct zeros, we have

$$F(s) = \frac{A_1}{s - s_+} + \frac{A_2}{s - s_-} \ ,$$

where A_1 and A_2 are constants depending on $f(0)$ and $f'(0)$. The transform is trivial to invert:

$$f(t) = A_1\, e^{s_+\, t} + A_2\, e^{s_-\, t} \ ,$$

which is transient for all A_1 and A_2 if and only if $\mathrm{Re}(s_\pm) < 0$; in other words, if and only if the poles of the transfer function lie in the left side of the plane. On the other hand, if the zero is double, then we have

$$F(s) = \frac{B_1}{s - s_0} + \frac{B_2}{(s - s_0)^2} \ ,$$

where again B_1 and B_2 are constants depending on $f(0)$ and $f'(0)$. We can invert the transform and find that

$$f(t) = B_1\, e^{s_0\, t} + B_2\, t\, e^{s_0\, t} \ ,$$

which is transient for all B_1 and B_2 if and only if $\mathrm{Re}(s_0) < 0$; so that s_0 lies in the left side of the plane.

In fact this is a general result: a system is stable if all the poles of the transfer function lie in the left side of the plane. A formal proof of this statement is not hard, but takes some bookkeeping, so we will leave it as an exercise for the industrious reader.

Notice that if one relaxes the condition that the solutions should be transient for *all* initial conditions, then it may happen that for certain types of initial conditions non-transient solutions have a zero coefficient. The system may therefore seem stable, but only because of the special choice of initial conditions.

The damped harmonic oscillator

Stability is not the only property of a system that can be detected by studying the poles of the transfer function. With some experience one can detect change in the qualitative behaviour of a system by studying the poles. A simple example is provided by the damped harmonic oscillator.

This system is defined by two parameters μ and ω, both positive real numbers. The differential equation which governs this system is

$$(D^2 + 2\mu\, D + \omega^2)\, f(t) = u(t)\ .$$

The transfer function is

$$H(s) = \frac{1}{s^2 + 2\mu\, s + \omega^2}\ ,$$

which has poles at

$$s_\pm = -\mu \pm \sqrt{\mu^2 - \omega^2}\ .$$

We must distinguish three separate cases:

(a) (**overdamped**) $\mu > \omega$
In this case the poles are real and negative:

$$s_\pm = -\mu \left(1 \mp \sqrt{1 - \frac{\omega^2}{\mu^2}}\right)\ .$$

(b) (**critically damped**) $\mu = \omega$
In this case there is a double pole, real and negative: $s_+ = s_- = -\mu$.

(c) (**underdamped**) $\mu < \omega$
In this case the poles are complex:

$$s_\pm = -\mu \pm i\omega\sqrt{1 - \frac{\mu^2}{\omega^2}}\ .$$

Hence provided that μ is positive, the system is stable.

Suppose that we start with the system being overdamped so that the ratio $\varrho \equiv \omega/\mu$ is less than 1: $\varrho < 1$. As we increase ϱ either by increasing ω or decreasing μ, the poles of the transfer function, which start in the negative real axis, start moving towards each other, coinciding when $\varrho = 1$. If we continue increasing ϱ so that it becomes greater than 1, the poles move vertically away from each other keeping their real parts constant. It is the transition from real to complex poles which offers the most drastic qualitative change in the behaviour of the system.

3.3.5 Application: convolution and the tautochrone

In this section we discuss a beautiful application of the Laplace transform. We also take the opportunity to discuss the convolution of two functions.

The convolution

Suppose that $f(t)$ and $g(t)$ are two functions with Laplace transforms $F(s)$ and $G(s)$. Consider the product $F(s)\,G(s)$. Is this the Laplace transform of any function? It turns out it is! To see this let us write the product $F(s)\,G(s)$ explicitly:

$$F(s)\,G(s) = \left(\int_0^\infty f(u)\,e^{-su}\,du\right)\,\left(\int_0^\infty g(v)\,e^{-sv}\,dv\right)\ .$$

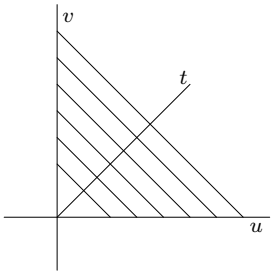
We can think of this as a double integral in the positive quadrant of the (u,v)-plane:

$$F(s)\,G(s) = \iint e^{-s(u+v)}\,f(u)\,g(v)\,du\,dv\ . \tag{3.50}$$

If this were the Laplace transform of anything, it would have to be of the form

$$F(s)\,G(s) \stackrel{?}{=} \int_0^\infty h(t)\,e^{-st}\,dt\ . \tag{3.51}$$

Comparing the two equations we are prompted to define $t = u + v$. In the positive quadrant in the (u,v)-axis, t runs from 0 to ∞: lines of constant t having slope -1. Therefore we see that integrating (u,v) in the positive quadrant is the same as integrating (t,v) where t runs from 0 to ∞ and for every t, v runs from 0 to t:



$$\iint k(u,v)\,du\,dv = \int_0^\infty dt\,\int_0^t k(t-v,v)\,dv$$

In other words, we can rewrite equation (3.50) as

$$F(s)\,G(s) = \int_0^\infty e^{-st}\,\int_0^t f(t-v)\,g(v)\,dv\ .$$

Comparing with equation (3.51), we see that this equation is true provided that

$$h(t) = \int_0^t f(t-v)\, g(v)\, dv \ .$$

This means that $h(t)$ is the **convolution** of f and g. The convolution is often denoted $f \star g$:

$$\boxed{(f \star g)(t) \equiv \int_0^t f(t-\tau)\, g(\tau)\, d\tau \ ,} \tag{3.52}$$

and it is characterised by the **convolution theorem**:

$$\boxed{\mathcal{L}\left\{f \star g\right\}(s) = F(s)\, G(s) \ .} \tag{3.53}$$

Notice that $f \star g = g \star f$. This is clear from the fact that $F(s)\, G(s) = G(s)\, F(s)$, but can also be checked directly by making a change of variables $\tau = t - \sigma$ in the integral in (3.52).

Abel's mechanical problem and the tautochrone

As an amusing application of the convolution theorem for the Laplace transform, let us consider Abel's mechanical problem. In short, the problem can be described as follows. Consider a bead of mass m which can slide down a wire frame under the influence of gravity but without any friction. Suppose that the bead is dropped from rest from a height h. Let $\tau(h)$ denote the time it takes to slide down to the ground. If one knows the shape of the wire it is a simple matter to determine the function $\tau(h)$, and we will do so below. Abel's mechanical problem is the inverse: given the function $\tau(h)$ determine the shape of the wire. As we will see below, this leads to an integral equation which has to be solved. In general integral equations are difficult to solve, but in this particular case, the integral is in the form of a convolution, whence its Laplace transform factorises. It is precisely this feature which makes the problem solvable.

To see what I mean, consider the following integral equation for the unknown function $f(t)$:

$$f(t) = 1 + \int_0^t f(t-\tau)\, \sin\tau\, d\tau \ . \tag{3.54}$$

We can recognise the integral as the convolution of the functions $f(t)$ and $\sin t$, whence taking the Laplace transform of both sides of the equation, we have

$$F(s) = \frac{1}{s} + F(s)\frac{1}{s^2+1} \ ,$$

which we can immediately solve for $F(s)$:

$$F(s) = \frac{1}{s} + \frac{1}{s^3} \ ,$$

which is the Laplace transform of the function

$$f(t) = 1 + \tfrac{1}{2}\, t^2 \ .$$

One can verify directly that this function obeys the original integral equation (3.54).

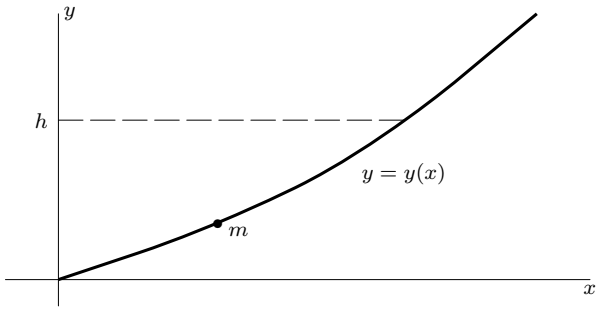


Figure 3.4: Abel's mechanical problem

In order to set up Abel's mechanical problem, it will prove convenient to keep Figure 3.4 in mind. We will assume that the wire has no torsion, so that the motion of the bead happens in one plane: the (x, y) plane with y the vertical displacement and x the horizontal displacement. We choose our axes in such a way that wire touches the ground at the origin of the plane: $(0, 0)$. The shape of the wire is given by a function $y = y(x)$, with $y(0) = 0$. Let ℓ denote the length *along* the wire from the origin to the point $(x, y = y(x))$ on the wire. We drop the bead from rest from a height h. Because there is no friction, energy is conserved. The kinetic energy of the bead at any time t after being dropped is given by

$$T = \tfrac{1}{2}\, m \left(\frac{d\ell}{dt}\right)^2 \ ,$$

whereas the potential energy is given by

$$V = -mg\,(h - y) \ .$$

Conservation of energy says that $T + V$ is a constant. To compute this constant, let us evaluate this at the moment the bead is dropped, $t = 0$. Because it is dropped from rest, $d\ell/dt = 0$ at $t = 0$, and hence $T = 0$.

Since at $t = 0$, $y = h$ the potential energy also vanishes and we have that $T + V = 0$. This identity can be rewritten as

$$\tfrac{1}{2}\, m \left(\frac{d\ell}{dt}\right)^2 = mg\,(h - y)\ ,$$

from which we can find a formula for $d\ell/dt$:

$$\frac{d\ell}{dt} = -\sqrt{2g\,(h - y)}\ , \tag{3.55}$$

where we have chosen the negative sign for the square root, because as the bead falls, ℓ decreases. Now, the length element along the wire is given by

$$d\ell = \sqrt{dx^2 + dy^2}\ ,$$

where dx and dy are not independent since we have a relation $y = y(x)$. Inverting this relation gives x as a function of y and we can use this to write

$$d\ell = \sqrt{1 + \left(\frac{dx}{dy}\right)^2}\, dy \equiv f(y)\, dy\ , \tag{3.56}$$

which defines the function $f(y)$. Clearly, $f(y)$ encodes the information about the shape of the wire: knowing $f(y)$ for all y allows us to solve for the dependence of x on y and viceversa. Indeed, suppose that $f(y)$ is known, then solving for dx/dy, we have that

$$\frac{dx}{dy} = \sqrt{f(y)^2 - 1}\ ,$$

from where we have

$$dx = \sqrt{f(y)^2 - 1}\, dy\ , \tag{3.57}$$

which can then be integrated to find x as a function of y, and by inverting this, y as a function of x.

Let us rewrite equation (3.55) as

$$dt = -\frac{1}{\sqrt{2g\,(h - y)}}\, d\ell\ .$$

and insert equation (3.56) in this equation, to obtain

$$dt = -\frac{f(y)}{\sqrt{2g\,(h - y)}}\, dy\ .$$

Finally we integrate this along the trajectory of the bead, as it falls from $y = h$ at $t = 0$ until $y = 0$ at $t = \tau(h)$:

$$\int_0^{\tau(h)} dt = -\int_h^0 \frac{f(y)}{\sqrt{2g\,(h-y)}}\,dy\ ,$$

whence

$$\boxed{\tau(h) = \frac{1}{\sqrt{2g}}\int_0^h f(y)\,\frac{1}{\sqrt{h-y}}\,dy\ .} \tag{3.58}$$

This formula gives us how long it takes for the bead to fall along the wire from a height h: so if we know the shape of the wire, and hence $f(y)$, we can compute $\tau(h)$ just by integrating. On the other hand, suppose that we are given $\tau(h)$ and we want to solve for the shape of the wire. This means solving equation (3.58) for $f(y)$ and then finding $y = y(x)$ from $f(y)$. The latter half of the problem is a first order differential equation, but the former half is an integral equation. In general this problem would be quite difficult, but because we notice that the integral in the right hand side is in the form of a convolution, we can try to solve this by using the Laplace transform.

Before doing so, however, let us check that we have not made a mistake, by testing the integral expression for $\tau(h)$ in a some cases where we know the answer. Suppose, for instance, that the wire is completely vertical. This means that $dx/dy = 0$, whence $f(y) = 1$. In this case, equation (3.58) simplifies enormously, and we get

$$\tau(h) = \frac{1}{\sqrt{2g}}\int_0^h \frac{dy}{\sqrt{h-y}} = \sqrt{\frac{2h}{g}}\ ,$$

as expected from elementary newtonian mechanics. Similarly, if the wire is inclined θ degrees from the horizontal, so that $y(x) = \tan\theta\, x$. Then $dx/dy = \cot\theta$, and hence $f(y)$ is given by

$$f(y) = \sqrt{1 + \left(\frac{dx}{dy}\right)^2} = \sqrt{1 + (\cot\theta)^2} = \csc\theta\ .$$

Therefore, the time taken to fall is simply $\csc\theta$ times the vertical time of fall:

$$\tau(h) = \csc\theta\,\sqrt{\frac{2h}{g}}\ ,$$

which, since $\csc(\pi/2) = 1$, agrees with the previous result.

Let us now take the Laplace transform of both sides of equation (3.58), thinking of them both as functions of h; in other words, the Laplace transform $F(s)$ of a function $g(h)$ is given by

$$G(s) = \int_0^\infty e^{-sh}\, g(h)\, dh \ .$$

(This is Shakespeare's theorem yet again!) Applying this to equation (3.58), we find

$$T(s) = \frac{1}{\sqrt{2g}} F(s)\, \mathcal{L}\left\{\frac{1}{\sqrt{h}}\right\}(s) \ ,$$

where $T(s)$ is the Laplace transform of the function τ, and $F(s)$ is the Laplace transform of the function f. The Laplace transform of the function $1/\sqrt{h}$ was worked out in the problems and the result is:

$$\mathcal{L}\left\{\frac{1}{\sqrt{h}}\right\}(s) = \sqrt{\frac{\pi}{s}} \ . \tag{3.59}$$

We can then solve for $F(s)$ in terms of $T(s)$ as follows:

$$F(s) = \sqrt{\frac{2g}{\pi}}\, \sqrt{s}\, T(s) \ , \tag{3.60}$$

which can in principle be inverted to solve for f, either from Table 3.1 or, if all else fails, from the inversion formula (3.39).

Let us apply this to solving for the shape that the wire must have for it to have the curious property that no matter what height we drop the bead from, it will take the same amount of time to fall to the ground. Such a shape is known as the *tautochrone*. Clearly, the tautochrone is such that $\tau(h) = \tau$ is constant, whence its Laplace transform is $T(s) = \tau/s$. Into equation (3.60), we get

$$F(s) = \sqrt{\frac{2g}{\pi}}\, \sqrt{s}\, \frac{\tau}{s} = \sqrt{2g}\, \frac{\tau}{\pi} \sqrt{\frac{\pi}{s}} \ ,$$

where we have rewritten it in a way that makes it easy to invert. From equation (3.59) we immediately see that

$$f(y) = \sqrt{2g}\, \frac{\tau}{\pi}\, \frac{1}{\sqrt{y}} \ .$$

To reconstruct the formula for the shape of the wire, we apply equation (3.57) to obtain

$$dx = \sqrt{\frac{2g\tau^2}{\pi^2}\, \frac{1}{y} - 1}\, dy \ ,$$

which can be integrated to

$$x = \int_0^y \sqrt{\frac{2g\tau^2}{\pi^2}\,\frac{1}{y} - 1}\, dy = \int_0^y \frac{\sqrt{\frac{2g\tau^2}{\pi^2} - y}}{\sqrt{y}}\, dy \ . \tag{3.61}$$

Notice that the constant of integration is fixed to 0 since the wire is such that when $x = 0$, $y = 0$. This integral can be performed by a trigonometric substitution. First of all let us define

$$b \equiv \frac{2g\tau^2}{\pi^2} \ ,$$

and let $y = b\,(\sin\phi)^2$, so that

$$dy = 2b\,\sin\phi\,\cos\phi\, d\phi \ .$$

Into the integral in (3.61), we find

$$x = \int_0^{\phi(y)} 2b\,(\cos\phi)^2\, d\phi = \frac{b}{2}\,\left[2\phi(y) + \sin 2\phi(y)\right] \ ,$$

where

$$y = b\,(\sin\phi(y))^2 = \frac{b}{2}\,(1 - \cos 2\phi(y)) \ .$$

If we define $a = b/2$ and $\theta = 2\phi(y)$, we have the following parametric representation for the curve in the (x, y) plane defining the wire:

$$x = a(\theta + \sin\theta) \qquad \text{and} \qquad y = a(1 - \cos\theta) \ .$$

This curve is called a **cycloid**. It is the curve traced by a point in the rim of a circle of radius a rolling upside down without sliding along the line $y = a$, as shown in the Figure 3.5.

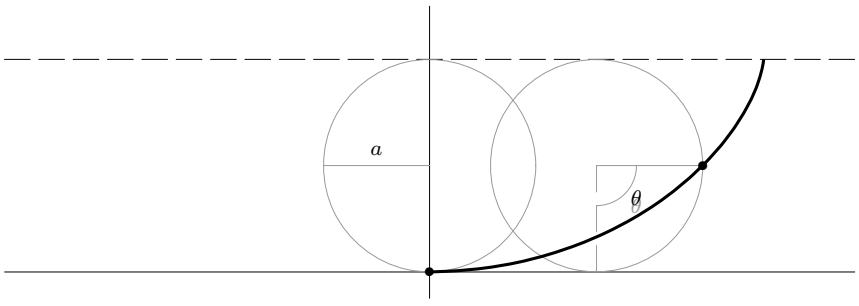


Figure 3.5: The cycloid

The cycloid also has another interesting property: it is the *brachistochrone*, namely the shape of the wire for which the time $\tau(h)$ is minimised. Although the proof is not hard, we will not do it here.

3.3.6 The Gamma and Zeta functions

This section falls outside the main scope of these notes, but since it allows a glimpse at some of the deepest and most beautiful aspects of mathematics, I could not resist the temptation to include it.

It is possible to consider the Laplace transform of complex powers t^z, with z some complex number. We see that

$$\mathcal{L}\left\{t^z\right\}(s) = \int_0^\infty t^z\, e^{-st}\, dt = \frac{1}{s^{z+1}} \int_0^\infty u^z\, e^{-u}\, du \ ,$$

where we have changed the variable of integration from t to $u = st$. Let us introduce the **Euler Gamma function**

$$\boxed{\Gamma(z) \equiv \int_0^\infty t^{z-1}\, e^{-t}\, dt \ ,} \tag{3.62}$$

which converges for $\mathrm{Re}(z) > 0$. Then we have that

$$\mathcal{L}\left\{t^z\right\}(s) = \frac{\Gamma(z+1)}{s^{z+1}} \ .$$

Comparing with equation (3.41), we see that $\Gamma(n+1) = n!$, whence we can think of the Gamma function as a way to define the factorial of a complex number. Although the integral representation (3.62) is only defined for $\mathrm{Re}(z) > 0$ it is possible to extend $\Gamma(z)$ to a holomorphic function with only isolated singularities in the whole complex plane: simple poles at the nonpositive integers.

To see this notice that for $\mathrm{Re}(z) > 0$, we can derive a recursion relation for $\Gamma(z)$ extending the well-known $n! = n\,(n-1)!$ for positive integers n. Consider

$$\Gamma(z+1) = \int_0^\infty t^z\, e^{-t}\, dt \ .$$

Integrating by parts,

$$\Gamma(z+1) = \int_0^\infty z\, t^{z-1}\, e^{-t}\, dt - t^z\, e^{-t}\Big|_0^\infty = z\,\Gamma(z) + \lim_{t\to 0} t^z\, e^{-st} \ .$$

Provided that $\mathrm{Re}(z) > 0$, the boundary term vanishes and we have

$$\Gamma(z+1) = z\,\Gamma(z) \ . \tag{3.63}$$

Turning this equation around, we have that

$$\Gamma(z) = \frac{\Gamma(z+1)}{z} \ .$$

Since $\Gamma(1) = 1$, which incidentally justifies the usual claim that $0! = 1$, we see that $\Gamma(z)$ has a simple pole at $z = 0$ with residue 1. Using this recursion relation repeatedly, we see that $\Gamma(z)$ has simple poles at all the nonpositive integers, with residue

$$\mathrm{Res}(\Gamma; -k) = \frac{(-1)^k}{k!} \ ,$$

and these are all the singularities.

The Gamma function is an extremely important function in mathematics, not least of all because it is intimately related to another illustrious function: the **Riemann Zeta function** $\zeta(z)$, defined for $\mathrm{Re}(z) > 1$ by the converging series

$$\zeta(z) = \sum_{n=1}^{\infty} \frac{1}{n^z} \ .$$

To see the relation notice that

$$\int_0^\infty t^{z-1}\, e^{-nt}\, dt = \frac{1}{n^z} \int_0^\infty u^{z-1}\, e^{-u}\, du = \frac{\Gamma(z)}{n^z} \ ,$$

where we have changed variables of integration from t to $u = nt$. Summing both sides of this identity over all positive integers n, we have, on the one hand

$$\sum_{n=1}^{\infty} \frac{\Gamma(z)}{n^z} = \Gamma(z)\, \zeta(z) \ ,$$

and on the other

$$\sum_{n=1}^{\infty} \int_0^\infty t^{z-1}\, e^{-nt}\, dt = \int_0^\infty t^{z-1} \sum_{n=1}^{\infty} e^{-nt}\, dt = \int_0^\infty t^{z-1}\, \frac{1}{e^t - 1}\, dt \ .$$

where we have interchanged the summation inside the integral, and summed the geometric series. (This can be justified, although we will not do so here.) As a result we have the following integral representation for the Zeta function

$$\zeta(z) = \frac{1}{\Gamma(z)} \int_0^\infty \frac{t^{z-1}}{e^t - 1}\, dt \ .$$

The only source of singularities in the integral is the zero of $e^t - 1$ at the origin, so we can split the integral into two as follows:

$$\zeta(z) = \frac{1}{\Gamma(z)} \left[\int_0^1 \frac{t^{z-1}}{e^t - 1}\, dt + \int_1^\infty \frac{t^{z-1}}{e^t - 1}\, dt \right] \ .$$

It is possible to show that $\Gamma(z)$ has no zeros, whence $1/\Gamma(z)$ is entire. Similarly, the second integral $\int_1^\infty$ is also entire since the integrand is continuous

there. Hence the singularity structure of the Zeta function is contained in the first integral. We can do a Laurent expansion of the integrand around $t = 0$:

$$\frac{1}{e^t - 1} = \frac{1}{t} - \frac{1}{2} + \frac{t}{12} + O(t^3) \ ,$$

where only odd powers of t appear after the first. Therefore integrating termwise, which we can do because Laurent series converge uniformly, we have that

$$\int_0^1 \frac{t^{z-1}}{e^t - 1}\, dt = \frac{1}{z-1} - \frac{1}{2}\frac{1}{z} + \frac{1}{12}\frac{1}{z+1} + \cdots \ , \tag{3.64}$$

where the terms which have been omitted are all of the form $a_k/(z+k)$ where k is a positive odd integer. This shows that the integral $\int_0^1$ has simple poles at $z = 1$, $z = 0$, and $z = -k$ with k a positive odd integer. Because the integral is multiplied by $1/\Gamma(z)$, and the Gamma function has simple poles at the nonnegative integers we see immediately that

- $\zeta(z)$ has a simple pole at $z = 1$ with residue $\Gamma(1) = 1$, and is analytic everywhere else; and
- $\zeta(-2n) = 0$ where n is any positive integer: these are the zeros of $1/\Gamma(z)$ which are not cancelled by the poles in (3.64).

The celebrated **Riemann hypothesis** states that all other zeros of $\zeta(z)$ occur in the line $\mathrm{Re}(z) = \frac{1}{2}$. Now that Fermat's Last Theorem has been proven, the Riemann hypothesis remains the most important open problem in mathematics today.

The importance of settling this hypothesis stems from the intimate relationship between the Zeta function and the theory of numbers. The key to this relationship is the following infinite product expansion for $\zeta(z)$, valid for $\mathrm{Re}(z) > 1$:

$$\frac{1}{\zeta(z)} = \prod_{\substack{\text{primes} \\ p}} \left(1 - \frac{1}{p^z}\right) \ ,$$

which follows from the unique factorisation of every positive integer into a product of primes. To see this notice that since, for $\mathrm{Re}(z) > 1$, one has

$$\zeta(z) = 1 + \frac{1}{2^z} + \frac{1}{3^z} + \frac{1}{4^z} + \frac{1}{5^z} + \cdots \ ,$$

then it follows that

$$\frac{1}{2^z}\zeta(z) = \frac{1}{2^z} + \frac{1}{4^z} + \frac{1}{6^z} + \frac{1}{8^z} + \frac{1}{10^z} + \cdots \ ;$$

whence

$$\left(1-\frac{1}{2^z}\right)\zeta(z) = 1+\frac{1}{3^z}+\frac{1}{5^z}+\frac{1}{7^z}+\frac{1}{9^z}+\cdots\ .$$

In other words we have in the right-hand side only those terms $1/n^z$ where n is odd. Similarly,

$$\left(1-\frac{1}{3^z}\right)\left(1-\frac{1}{2^z}\right)\zeta(z) = 1+\frac{1}{5^z}+\frac{1}{7^z}+\frac{1}{11^z}+\frac{1}{13^z}+\cdots\ ,$$

where now we have in the right-hand side only those terms $1/n^z$ where n is not divisible by 2 or by 3. Continuing in this fashion, we have that

$$\prod_{\substack{\text{primes}\\ p}}\left(1-\frac{1}{p^z}\right)\zeta(z) = 1\ .$$

By the way, this shows that $\zeta(z)$ has no zeros for $\mathrm{Re}(z) > 1$.

The Zeta function and its generalisations also play a useful role in physics: particularly in quantum field theory, statistical mechanics, and, of course, in string theory. In fact, together with the heat kernel, introduced in the problems, the (generalised) Zeta function proves invaluable in computing determinants and traces of infinite-dimensional matrices

Areas of spheres

As a minor application of the Gamma function, let us compute the area of a unit sphere in n dimensions, for $n \geq 2$.

What do we mean by a unit sphere in n dimensions? The unit sphere in n dimensions is the set of points in n-dimensional euclidean space which are a unit distance away from the origin. If we let $(x_1, x_2, \ldots, x_n)$ be the coordinates for euclidean space, the unit sphere is the set of points which satisfy the equation

$$\sum_{i=1}^{n} x_i^2 = x_1^2 + x_2^2 + \ldots x_n^2 = 1\ .$$

In $n = 2$ dimensions, the unit "sphere" is a circle, whereas in $n = 3$ dimensions it is the usual sphere of everyday experience. For $n > 3$, the sphere is harder to visualise, but one can still work with it via the algebraic description above.

What do we mean by its area? We mean the $n-1$-dimensional area: so if $n = 2$, we mean the circumference of the circle, and if $n = 3$ we mean

the usual area of everyday experience. Again it gets harder to visualise for $n > 3$, but one can again tackle the problem algebraically as above.

Clearly every point in n-dimensional space lies on some sphere: if it is a distance r away from the origin then, by definition, it lies on the sphere of radius r. There are an uncountable number of spheres in euclidean space, one for every positive real number. All these spheres taken together with the origin (a "sphere" of zero radius) make up all of euclidean space. A simple scaling argument shows that if we double the radius, we multiply the area of the sphere by 2^{n-1}. More generally, the area of the sphere at radius r will be r^{n-1} times the area of the unit sphere. Therefore the volume element in n-dimensions is

$$d^n x = r^{n-1}\, dr\, d\Omega \ ,$$

where $d\Omega$ is the area element of the unit sphere. We will now integrate the function $\exp(-r^2)$ over all of the euclidean space. We can compute this integral in either of two ways. On the one hand,

$$\begin{aligned} I &\equiv \int \cdots \int e^{-r^2}\, d^n x \\ &= \int \cdots \int e^{-x_1^2 - x_2^2 - \cdots - x_n^2}\, dx_1 dx_2 \cdots dx_n \\ &= \left(\int_{-\infty}^{\infty} e^{-x_1^2}\, dx_1 \right) \left(\int_{-\infty}^{\infty} e^{-x_2^2}\, dx_2 \right) \cdots \left(\int_{-\infty}^{\infty} e^{-x_n^2}\, dx_n \right) \\ &= \left(\int_{-\infty}^{\infty} e^{-x^2}\, dx \right)^n \ , \end{aligned}$$

which is computed to give $(\sqrt{\pi})^n$ after using the elementary gaussian result: $\int_{-\infty}^{\infty} \exp(-x^2)\, dx = \sqrt{\pi}$. On the other hand,

$$I = \int \cdots \int e^{-r^2}\, r^{n-1}\, dr\, d\Omega = \left(\int_0^{\infty} e^{-r^2}\, r^{n-1}\, dr \right) \left(\int \cdots \int d\Omega \right) \ .$$

The integral of $d\Omega$ is simply the area A of the unit sphere, which is what we want to calculate. The radial integral can be calculated in terms of the Gamma function after changing the variable of integration from r to $t = r^2$:

$$\int_0^{\infty} e^{-r^2}\, r^{n-1}\, dr = \int_0^{\infty} e^{-t}\, \tfrac{1}{2}\, t^{(n-2)/2}\, dt = \frac{\Gamma(n/2)}{2} \ .$$

Equating both ways to compute the integral, we arrive at the following formula for the area $A(n)$ of the unit sphere in n dimensions:

$$A(n) = \frac{2\,\pi^{n/2}}{\Gamma(n/2)} \ . \tag{3.65}$$

To see that this beautiful formula is not obviously wrong, let us see that it reproduces what we know. For $n = 2$, by the area of the unit sphere we mean the circumference of the unit circle, that is 2π. In $n = 3$, we expect the area of the standard unit sphere and that is 4π. Let us see if our expectations are born out. According to the formula,

$$A(2) = \frac{2\pi}{\Gamma(1)} = 2\pi \ ,$$

as expected. For $n = 3$ the formula says

$$A(3) = \frac{2\,\pi^{3/2}}{\Gamma(3/2)} \ .$$

We can compute the half-integral values of the Gamma function as follows. First we have that

$$\Gamma(\tfrac{1}{2}) = \int_0^\infty t^{-1/2}\, e^{-t}\, dt \ .$$

Changing variables to $t = u^2$, we have

$$\Gamma(\tfrac{1}{2}) = 2\int_0^\infty e^{-u^2}\, du = \int_{-\infty}^\infty e^{-u^2}\, du = \sqrt{\pi} \ .$$

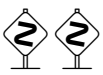
Now using the recursion relation (3.63), we have that

$$\Gamma(k + \tfrac{1}{2}) = \frac{2k-1}{2}\,\frac{2k-3}{2}\,\cdots\,\frac{1}{2}\,\Gamma(\tfrac{1}{2}) = \frac{(2k-1)!!}{2^k}\sqrt{\pi} \ .$$

In particular, $\Gamma(\frac{3}{2}) = \frac{\sqrt{\pi}}{2}$, whence

$$A(3) = \frac{2\pi^{3/2}}{\Gamma(3/2)} = 4\pi \ ,$$

as expected.



How about for $n = 1$? This case is a little special: in one dimension the unit sphere consists of the points ± 1. So that it is a zero-dimensional set. Is there an intrinsic notion of area for a zero-dimensional set? If we evaluate the above formula for $A(n)$ at $n = 1$, we get an answer: $A(1) = 2$, which is *counting* the number of points: in other words, zero-dimensional area is simply the **cardinality** of the set: the number of elements. This is something that perhaps we would not have expected. As someone said once, some formulae are more clever than the people who come up with them.

Now that we trust the formula, we can compute a few more values to learn something new. First let us simplify the formula by evaluating the

Gamma function at the appropriate values. Distinguishing between odd and even dimensions, we find

$$A(n) = \begin{cases} \dfrac{2\,\pi^\ell}{(\ell-1)!}\ , & \text{for } n = 2\ell, \text{ and} \\ \dfrac{2^{\ell+1}\,\pi^\ell}{(2\ell-1)!!}\ , & \text{for } n = 2\ell+1. \end{cases} \tag{3.66}$$

The next few values are

$$A(4) = 2\,\pi^2\ , \quad A(5) = \frac{8\pi^3}{3}\ , \quad A(6) = \pi^3\ , \quad A(7) = \frac{16\pi^3}{15}\ , \quad A(8) = \frac{\pi^4}{3}\ .$$

In case you are wondering whether this is at all useful, it actually comes in handy when normalising electromagnetic fields in higher-dimensional field theories so that they have integral fluxes around charged objects (e.g., branes and black holes).

Let us end with another nice formula. How about the n-dimensional volume $V(n)$ of the unit ball, i.e., the interior of the unit sphere? We can compute this by integrating the areas of the spheres from radius 0 to radius 1. The area of the sphere of radius r will be r^{n-1} times the area of the sphere of unit radius, so that the volume is then

$$V(n) = \int_0^1 A(n) r^{n-1}\, dr = \frac{A(n)}{n}\ .$$

Using the formula (3.65), we see that

$$V(n) = \frac{2\pi^{n/2}}{n\Gamma(n/2)} = \frac{\pi^{n/2}}{(n/2)\Gamma(n/2)} = \frac{\pi^{n/2}}{\Gamma((n+2)/2)}\ ,$$

where we have used the recursion formula (3.63). Because the the unit sphere is inscribed inside the cube of length 2, the ratio of the volume of the unit ball to that of the cube circumscribing it is given by

$$\varrho(n) = \frac{V(n)}{2^n} = \frac{\pi^{n/2}}{2^n\,\Gamma(\frac{n}{2}+1)}\ .$$

If we were to plot this as a function of n we notice that it starts at 1 for $n = 1$ and then decreases quite fast, so that the ball takes up less and less of the volume of he cube which circumscribes it.