Mathematical Methods PHYS50007

Lectured by Dmitry Turaev*
assisted by Henry Chiu, Lifan Xuan, and Stuart Patching

Spring Term 2021

1 Lecture 1: Analytic Functions of a Complex Variable

1.1 Definitions and Basic Properties

A complex number has the form $z = x + iy$. We denote the *real* and *imaginary* parts respectively as $x = \mathrm{Re}(z)$ and $y = \mathrm{Im}(z)$. We may also think of z as a point in the complex plane,which in polar coordinates has *absolute value* $r = |z| = \sqrt{x^2 + y^2}$ and *argument* φ satisfying $x = r\cos\varphi$ and $y = r\sin\varphi$. Note that φ is defined periodically with period 2π and may also be denoted $\varphi = \arg(z)$.

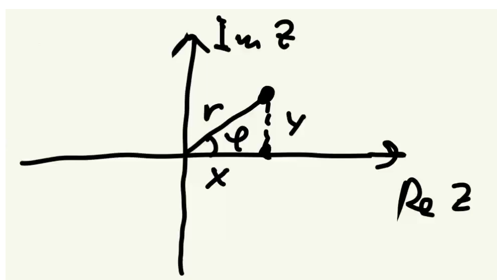


Figure 1: A complex number visualised on the complex plane.

Complex numbers obey the following algebraic properties:

$$\begin{aligned} z_1 + z_2 =& (x_1 + x_2) + i(y_1 + y_2) \\ z_1 z_2 =& (x_1 + iy_1)(x_2 + iy_2) = (x_1x_2 - y_1y_2) + i(y_1x_2 + x_1y_2) \end{aligned}$$

i.e. they have the same algebra as the real numbers, with the convention $i^2 = -1$.

*dturaev@imperial.ac.uk; lecture notes from previous years are available at `https://wwwf.imperial.ac.uk/~dturaev/`

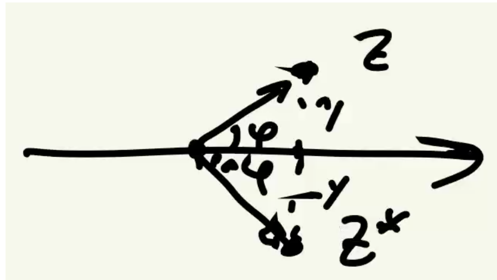


Figure 2: Complex conjugate $z^* = x - iy$ is defined by reflecting in the x-axis.

If $z = x + iy$ then the *complex conjugate* is defined as $z^* = x - iy$, which is the symmetric image of z with respect to the x-axis. Note the following properties of the complex conjugate:

$$\begin{aligned} r(z) =& \sqrt{x^2 + y^2} = r(z^*) \\ \varphi(z) =& -\varphi(z^*) \\ zz^* =& (x + iy)(x - iy) = x^2 + y^2 = |z|^2 \end{aligned}$$

This helps us to obtain a useful formula for the division of complex numbers:

$$\frac{z_1}{z_2} = \frac{z_1 z_2^*}{z_2 z_2^*} = \frac{(x_1 + iy_1)(x_2 - iy_2)}{x_2^2 + y_2^2}$$

We may also do multiplication in polar coordinates:

$$\begin{aligned} z_1 z_2 =& (r_1 \cos\varphi_1 + ir_1 \sin\varphi_1)(r_2 \cos\varphi_2 + ir_2 \sin\varphi_2) \\ =& r_1 r_2 \left[\cos\varphi_1 \cos\varphi_2 - \sin\varphi_1 \sin\varphi_2 + i(\cos\varphi_1 \sin\varphi_2 + \sin\varphi_1 \cos\varphi_2)\right] \\ =& r_1 r_2 \left[\cos(\varphi_1 + \varphi_2) + i \sin(\varphi_1 + \varphi_2)\right] \end{aligned}$$

where we use the double-angle formulae for sin and cos. Then we see that

$$|z_1 z_2| = |z_1| \, |z_2|$$

and

$$\arg(z_1 z_2) = \arg(z_1) + \arg(z_2).$$

Thus, we may take the nth power and, after that, nth root of z as:

$$\begin{aligned} z^n =& r^n (\cos(n\varphi) + i \sin(n\varphi)) \\ \sqrt[n]{z} =& \cos\left(\frac{\varphi + 2\pi k}{n}\right) + i \sin\left(\frac{\varphi + 2\pi k}{n}\right) \qquad \text{for} \qquad k = 0, ..., n-1 \end{aligned}$$

Each non-zero z has n different nth roots and these are related by rotations in the complex plane, see Figure 3.

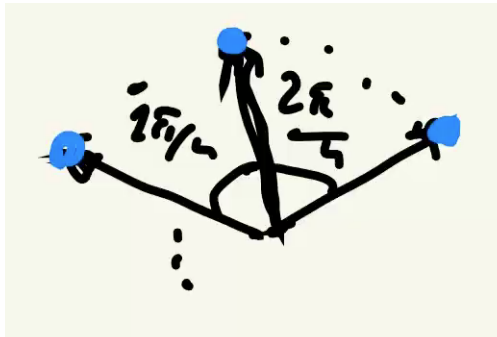


Figure 3: The nth roots are related by rotation through angle $2\pi/n$ in the complex plane.

1.2 Functions of Complex Variables

A standard notation we use for a complex function is

$$f(z) = u(x, y) + iv(x, y)$$

Examples

1. The simplest complex functions are $f(z) = a = \textit{constant}$ and $f(z) = z$. By doing multiplication and addition with these, we obtain

2. Polynomials:

   $$P(z) = a_0 + a_1 z + ... + a_n z^n$$

   Employing division operation, we obtain

3. Rational functions:

   $$R(z) = \frac{P(z)}{Q(z)}$$

   where P and Q are both polynomials.

4. A more advanced example is given by
   Sums of convergent power series:

   $$f(z) = \sum_{n=0}^{\infty} a_n z^n$$

   If there exist $C > 0$ and $\rho > 0$ such that $|a_n| \leq C\rho^{-n}$ then the series is convergent for all z with $|z| < \rho$. The maximal such ρ is called the *radius of convergence* of f. In particular, if the sequence a_n is bounded, then f is convergent at least for $|z| < 1$.
   For example, consider the following function (the sum of geometric progression):

   $$f(z) = \sum_{n=0}^{\infty} z^n$$

In this case $a_n = 1$ for all n, thus the radius of convergence $\rho = 1$. However, we may define an *analytic continuation* of f outside the range $|z| < 1$, as:

$$\tilde{f}(z) = \frac{1}{1-z}$$

which is defined for all $z \neq 1$. So f is undefined for $z > 1$, but for all $z < 1$ we have $f(z) = \tilde{f}(z)$.

Consider also the exponential function:

$$\exp(z) = e^z := \sum_{n=0}^{\infty} \frac{z^n}{n!}$$

Here $a_n = \frac{1}{n!}$ and the radius of convergence is $\rho = \infty$. The coefficients of this series are exactly the same as for the real exponential function, so, because the algebraic operations on real and complex numbers obey the same rules, the exponential function defined on complex numbers obeys the same algebra as the exponential function on real numbers:

$$e^{z_1} e^{z_2} = e^{z_1 + z_2}$$

Hence, we have

$$e^{x+iy} = e^x \cos y + i e^x \sin y$$

once we prove that

$$e^{iy} = \cos y + i \sin y$$

To establish the latter formula, we write

$$e^{iy} = \sum_{n=0}^{\infty} \frac{i^n y^n}{n!} = \sum_{k=0}^{\infty} \frac{(-1)^k y^{2k}}{(2k)!} + i \sum_{k=0}^{\infty} \frac{(-1)^k y^{2k+1}}{(2k+1)!} = \cos y + i \sin y$$

From the formula

$$e^{i\varphi} = \cos\varphi + i \sin\varphi,$$

we see that $e^{i\varphi}$ is 2π-periodic in φ:

$$e^{i\phi + 2\pi k i} = e^{i\varphi} \qquad \text{for} \qquad k \in \mathbb{Z}$$

This periodicity implies that the inverse of the exponential function is *multi-valued*:

$$\ln\left(re^{i\varphi}\right) = \ln r + i\left(\varphi + 2k\pi\right) \qquad \text{for} \qquad k \in \mathbb{Z}$$

We may use the logarithm to define arbitrary complex powers of a complex number z:

$$z^\alpha := e^{\alpha \ln z}$$

This function is also multivalued (when α is not an integer).

5. The previous examples are 'good' functions, in the sense that they are analytic, as we will see later. However, we may also define 'bad' functions of complex variables, such as:

$$\begin{aligned} f(z) &= \mathrm{Re}(z) \\ g(z) &= |z| \\ h(z) &= z^* \end{aligned}$$

   These functions are non-analytic.

The analyticity is the main topic here. While a typical complex function is not analytic, many important functions are analytic and, as a consequence, possess non-trivial and useful properties. The analyticity can be defined in several equivalent ways. We start with the notion of the derivative.

Definition 1.1. The *derivative* of a complex function f at the point z is given by:

$$f'(z) = \lim_{\Delta z \to 0} \frac{f(z + \Delta z) - f(z)}{\Delta z} \tag{1}$$

Note that this limit $\Delta z \to 0$ may be taken in many directions; for the derivative of f to exist, the limit must exist and be the same regardless of the direction in which the limit is taken. In particular, we may write $\Delta z = re^{i\varphi}$, so the limit $\Delta z \to 0$ must be independent of φ, the direction along which Δz approaches zero.

Definition 1.2. A function $f(z)$ is *analytic* at the point z_0 if it has a derivative at all points close to z_0.

Note that $f(z) = |z|^2$ is not analytic at $z = 0$, even though it has a derivative at $z = 0$, because one can show that it does not have a derivative for any $z \neq 0$.

Definition 1.3. A function $f(z)$ is *analytic* in an open region D if the derivative $f'(z)$ exists at every points $z \in D$.

Let us write $f(z) = u(x, y) + iv(x, y)$. In order for f to have a derivative at z, the functions u and v must be differentiable at the corresponding point in the $x - y$ plane; however, the differentiability of u and v does not guarantee the existence of $f'(z)$, i.e., it is a necessary but not sufficient condition.

Theorem 1.4 (Cauchy-Riemann Conditions)**.** *Let $f(z) = u(x, y) + iv(x, y)$. The derivative $f'(z)$ exists at $z = x + iy$ if and only if u and v are differentiable and satisfy the Cauchy-Riemann equations*

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y}, \tag{2a}$$

$$\frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x}. \tag{2b}$$

Proof. Suppose $f'(z)$ exists. Then the limit in eq. (1) exists and is independent of the direction in which the limit is taken. Therefore set $\Delta z = \Delta x$ and we see that:

$$f'(z) = \lim_{\Delta x \to 0} \frac{u(x + \Delta x, y) + iv(x + \Delta x, y) - u(x, y) + iv(x, y)}{\Delta x} = \frac{\partial u}{\partial x} + i\frac{\partial v}{\partial x}$$

Similarly, if we take $\Delta z = i\Delta y$ then we find:

$$f'(z) = \lim_{\Delta x \to 0} \frac{u(x, y + \Delta y) + iv(x, y + \Delta y) - u(x, y) + iv(x, y)}{i\Delta y} = -i\frac{\partial u}{\partial y} + \frac{\partial v}{\partial y}$$

Equating the real and imaginary parts gives eq. (2).

Suppose conversely that u and v satisfy eq. (2). We keep a general form for $\Delta z = \Delta x + i\Delta y$ and then expand:

$$\begin{aligned} f(z+\Delta z) =& u(x+\Delta x, y+\Delta y) + iv(x+\Delta x, y+\Delta y) \\ u(x+\Delta x, y+\Delta y) =& u(x,y) + \frac{\partial u}{\partial x}\Delta x + \frac{\partial u}{\partial y}\Delta y + o\left(\sqrt{\Delta x^2 + \Delta y^2}\right) \\ v(x+\Delta x, y+\Delta y) =& v(x,y) + \frac{\partial v}{\partial x}\Delta x + \frac{\partial v}{\partial y}\Delta y + o\left(\sqrt{\Delta x^2 + \Delta y^2}\right) \\ \Longrightarrow \qquad f(z+\Delta z) - f(z) =& \frac{\partial u}{\partial x}\Delta x + \frac{\partial u}{\partial y}\Delta y + i\left(\frac{\partial v}{\partial x}\Delta x + \frac{\partial v}{\partial y}\Delta y\right) + o\left(|\Delta z|\right). \end{aligned}$$

By eq. (2), the right-hand side of this formula is

$$= \left(\frac{\partial u}{\partial x} + i\frac{\partial v}{\partial x}\right)\Delta z + o\left(|\Delta z|\right)$$

and therefore $f'(z) = \left(\frac{\partial u}{\partial x} + i\frac{\partial v}{\partial x}\right)$ exists. □

Note the word ”differentiable” in the formulation of this theorem. It means more than just the existence of the partial derivatives $\frac{\partial u}{\partial x}$, $\frac{\partial u}{\partial y}$, $\frac{\partial v}{\partial x}$, and $\frac{\partial v}{\partial y}$. Namely, we want that u and v have derivatives along any direction (not just along the x- and y- axes). Or, equivalently, we want to have expansions

$$\begin{aligned} u(x+\Delta x, y+\Delta y) =& u(x,y) + \frac{\partial u}{\partial x}\Delta x + \frac{\partial u}{\partial y}\Delta y + o\left(\sqrt{\Delta x^2 + \Delta y^2}\right) \\ v(x+\Delta x, y+\Delta y) =& v(x,y) + \frac{\partial v}{\partial x}\Delta x + \frac{\partial v}{\partial y}\Delta y + o\left(\sqrt{\Delta x^2 + \Delta y^2}\right) \end{aligned}$$

A sufficient condition for that is the continuity of $\frac{\partial u}{\partial x}$, $\frac{\partial u}{\partial y}$, $\frac{\partial v}{\partial x}$, and $\frac{\partial v}{\partial y}$.

Thus, an equivalent definition of analyticity is this:
the function is analytic in an open region D of the complex plane if its real and imaginary parts are differentiable and satisfy Cauchy-Riemann conditions at every point of D.

2 Lecture 2: Derivatives and integrals

2.1 Derivatives and integrals, Cauchy theorem

We repeat that a function $f(z)$ is *analytic* (or holomorphic) at the point z_0 if it has a derivative at all points z close to z_0. The derivative of f is defined like in the real case:

$$f'(z) = \lim_{\Delta z \to 0} \frac{f(z + \Delta z) - f(z)}{\Delta z},$$

which is independent of the direction along which $\Delta z \to 0$. This independence condition is equivalent to the Cauchy-Riemann condition, i.e. letting $f = u + iv$, we have

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y}, \qquad \frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x}.$$

The derivative satisfies the same rules as in the real case: for any analytic functions f and g, we have

1. $(f + g)' = f' + g'$,
2. $(fg)' = f'g + fg'$,
3. $(\frac{f}{g})' = \frac{f'g - fg'}{g^2}$,
4. (chain rule) for any z in the domain of g, $(f(g(z)))' = f'(g(z))g'(z)$.

Remark If g is the inverse of f, i.e. for any z in the domain of f, $g(f(z)) = z$, then $g'(f(z)) = \frac{1}{f'(z)}$, and $g'(z) = \frac{1}{f'(g(z))}$.

Some examples of how to derive the derivatives are given in what follows.

Example i. Suppose $f(z) = c$ for any $z \in \mathbb{C}$, where c is a constant complex number. Since for any $z, \Delta z \in \mathbb{C}$

$$\frac{f(z) + \Delta z - f(z)}{\Delta z} = \frac{c - c}{\Delta z} = 0,$$

the derivative of f is 0.

ii. Suppose $f(z) = z$ for any $z \in \mathbb{C}$. Since for any $z, \Delta z \in \mathbb{C}$,

$$\frac{f(z) + \Delta z - f(z)}{\Delta z} = \frac{z + \Delta z - z}{\Delta z} = 1,$$

we get $f'(z) = 1$.

iii. Using the product rule, for any $z \in \mathbb{C}$, we get

$$(z^n)' = nz^{n-1}$$

iv. Now, for any $z \in \mathbb{C}$, letting $P(z) = a_0 + a_1 z + ... + a_n z^n$, we get the derivative of P as follows:

$$P'(z) = a_1 + ... + na_n z^{n-1}.$$

v. Since any rational function can be written as $\frac{P}{Q}$, where P and Q are polynomials, and since $(\frac{P}{Q})' = \frac{P'Q - PQ'}{Q^2}$, we obtain that rational functions are analytic everywhere except for the points where $Q(z) = 0$.

vi. Suppose f is a power series, i.e. $f(z) = a_0 + a_1 z + ... + a_n z^n + ...$, then

$$f'(z) = a_1 + 2a_2 z + ... + a_n z^{n-1} + ...$$

Furthermore, given that $|a_n| \le c\rho^{-n}$, which implies that

$$|na_n| \le (\frac{cn}{\rho})\rho^{-(n-1)} \le \tilde{c}\tilde{\rho}^{-(n-1)}, \text{ where } \tilde{\rho} < \rho, \text{ and } \tilde{c} \text{ large enough,}$$

we conclude that f' has the same radius convergence as that of f.

Example. For $e^z = \sum_{i=0}^{\infty} \frac{z^n}{n!}$, its derivative is

$$(e^z)' = 1 + z + ... + \frac{nz^{n-1}}{n!} + ... = \sum_{i=0}^{\infty} \frac{z^n}{n!} = e^z.$$

The inverse function of e^z is denotes as $\ln z$ (recall that $\ln(re^{i\psi}) = \ln r + i(\psi + 2k\pi)$). Using the rule for the derivative of the inverse function, we find

$$(\ln z)' = \frac{1}{e^{\ln z}} = \frac{1}{z}.$$

vii. For any $\alpha \in \mathbb{R}$ and $z \in \mathbb{C}$,

$$(z^\alpha)' = (e^{\alpha \ln z})' = \frac{\alpha}{z} e^{\alpha \ln z} = \alpha z^{\alpha - 1}.$$

viii. Letting ζ be a Riemann Zeta function, i.e. $\zeta(z) = \sum_{n=1}^{\infty} n^{-z}$, then its derivative is:

$$\zeta'(z) = -\sum_{n=1}^{\infty} n^{-z}(\ln n).$$

Holomorphic functions have many wonderful, or even magical properties. For example, let $\mathcal{D}$ be a simply connected open region in $\mathbb{C}$, as in Figure 4.

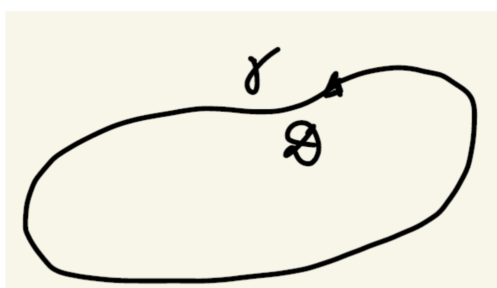


Figure 4: A simply connected region $\mathcal{D}$.

Then, as we will show later, the following two statements are equivalent for a complex function f defined in $\mathcal{D}$:

(a) $f'(z)$ exists for all z in the domain $\mathcal{D}$,

(b) $\int_{z_0}^{z} f(\xi) d\xi$ exists for all z and z_0 in the domain $\mathcal{D}$.

It is indeed, unexpected: we know that for real functions of a real variable the existence of the derivative does not follow from the existence of the integral.
Moreover, property (b) implies, in fact, that for any $z \in \mathcal{D}$,

$$f(z) = \frac{1}{2\pi i} \int_\gamma \frac{f(\xi)}{\xi - z} d\xi. \tag{3}$$

This is *Cauchy formula*, which plays the central role in this theory (here γ is the boundary of $\mathcal{D}$; we move along γ anti-clockwise).
We will explain the notion of the integral along the path γ in a moment. Before that, we just note that this is indeed a magical formula: the values of the analytic function f at any point inside the domain $\mathcal{D}$ are determined by the values of f on the boundary of the domain only!
Magical formulas have magical consequences: using the Cauchy formula, we will also show later that f has derivatives of *all orders* and the Taylor series of f converges to f. Nothing similar happens for real functions of a real variable, they may have first derivative, but not second derivative, or first and second derivative, but no third derivative, etc; or derivatives of all orders may exist, but the Taylor series may diverge, or converge to a wrong function - all these complications disappear for complex-valued functions of complex variables.

In order to properly formulate these results, we need to introduce the notion of an integral over a path in the complex plane. Let $\gamma := \{z = z(t) | \; t \in [0, 1]\}$ be a continuous curve parameterized by a real parameter t. Figure 5 shows a path connecting points $z(0)$ and $z(1)$.

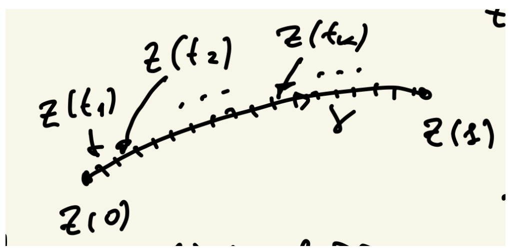


Figure 5: A path γ in $\mathcal{D}$.

We define the integral over γ as:

$$\int_\gamma f(z)dz = \lim_{n\to\infty} \sum_{k=0}^{n-1} f(z(t_k))(z(t_{k+1}) - z(t_k)). \tag{4}$$

Here $\{t_0 = 0, t_1, ..., t_n = 1\}$ is any partition of $[0, 1]$ such that $t_{k+1} - t_k$ tends uniformly to zero as $n \to +\infty$. The main (and the simplest) example is when $t_k = \frac{k}{n}$.

Remark It is important to remember that $(z(t_{k+1}) - z(t_k))$ in (4) is a complex number, not the length $|z(t_{k+1}) - z(t_k)|$.

The path γ is smooth when the function $z(t)$ is smooth (i.e., continuously differentiable). In this case, one uses the following formula for the computation of the integral:

$$\int_\gamma f(z)dz = \int_0^1 f(z(t))z'(t)dt.$$

The integral has the following basic properties.

1. If $\tilde{\gamma}$ is the same curve as γ, just with opposite orientation, (i.e. $\tilde{z}(t) = z(1-t)$), then
$$\int_\gamma f(z)dz = -\int_{\tilde{\gamma}} f(z)dz.$$

2. Let the end point of γ_1 be the starting point of γ_2. Then
$$\int_{\gamma_1 \cup \gamma_2} f(z)dz = \int_{\gamma_1} f(z)dz + \int_{\gamma_2} f(z)dz.$$

Note that Property 1 together with Property 2 implies that if $\gamma = \gamma_1 \cup \gamma_2$ where γ_2 is the same path as γ_1 just with opposite orientation (i.e., γ is a closed path obtained by traversing the same arc twice, forwards and then backwards), then $\int_\gamma f(z)dz = 0$.

We note that the integral over a path γ is well-defined for any continuous function f, which does not need to be analytic. However, we will show in the next lecture that for analytic functions the integral $\int_\gamma f(z)dz$ does not change when we slightly deform γ without moving the end points, so we can write $\int_{z_0}^{z_1} f(\zeta)d\zeta$ without indicating the path connecting the points z_0 and z_1.

3 Lecture 3: Cauchy formula and consequences

We will now prove the claim made in a previous lecture: the innocent looking condition of the existence of the derivative $f'(z)$ at every point of an open region imposes severe restrictions on the behaviour of the function f. Thus,

1. the integral $\int_\gamma f(\zeta)d\zeta$ of the analytic function f is independent of the path of integration, i.e., it is the same for any two paths with the same end points if one path can be continuously deformed to the other within the domain of analyticity of f (without moving the end points); in fact, the path-independence of the integral and, hence the existence of the anti-derivative $\Phi(z) = \int f(z)dz$ such that $\Phi'(z) = f(z)$ gives an equivalent definition of the analyticity;

2. the value of $f(z)$ at any point inside an analyticity domain is completely determined by the values of f at the boundary of the domain: by the *Cauchy formula*

$$f(z) = \frac{1}{2\pi i}\oint_\gamma \frac{f(\xi)}{\xi - z}d\xi; \tag{5}$$

3. the function f has derivatives of *all orders* and is the sum of its Taylor series:

$$f(z) = \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!}(z-a)^n$$

   within the radius of convergence of this power series, which always equals to the distance from a to the nearest singularity of f;

4. the following formula (obtained by differentiating (5)) holds for the derivatives of f:

$$f^n(z) = \frac{n!}{2\pi i}\oint \frac{f(\xi)}{(\xi - z)^{n+1}}d\xi; \tag{6}$$

   this formula implies that

$$|f^{(n)}(z)| \leq n!\frac{M(R)}{R^n}, \tag{7}$$

   where R is any number such that f is analytic everywhere in the disc $|\xi - z| \leq R$ and M is the maximal value of $|f|$ on the boundary of this disc: $M = \max_{|\xi-z|=R}|f(z)|$; hence, the derivatives of an analytic function can be estimated via the absolute value of the function itself.

We will also describe other consequences of the analyticity - maximum principle, Liouville Theorem, and the Fundamental Theorem of Algebra.

3.1 Cauchy's integral theorem

We start with proving the path-independence of the integral $\int f(\xi)dx$

Theorem 3.1. *If γ bounds a simply connected region $\mathcal{D}$, and f is analytic in $\mathcal{D}$, then*

$$\oint_\gamma f(z)dz = 0$$

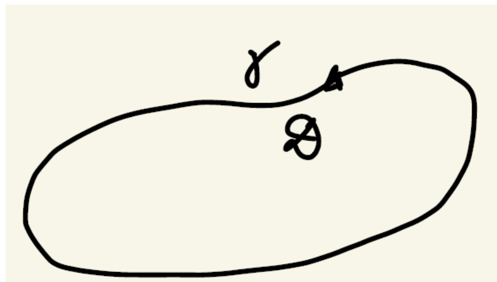


Figure 6: A simply connected region $\mathcal{D}$ is a region with no "holes" inside.

Proof. We prove the theorem for a very special case first: $\mathcal{D}$ is a rectangle, i.e., it is bounded by the segments $\gamma_1 : \{y = y_1, x \in [x_1, x_2]\}$, $\gamma_2 : \{x = x_2, y \in [y_1, y_2]\}$, $\gamma_3 : \{y = y_2, x \in [x_2, x_1]\}$, $\gamma_4 : \{x = x_1, y \in [y_2, y_1]\}$.

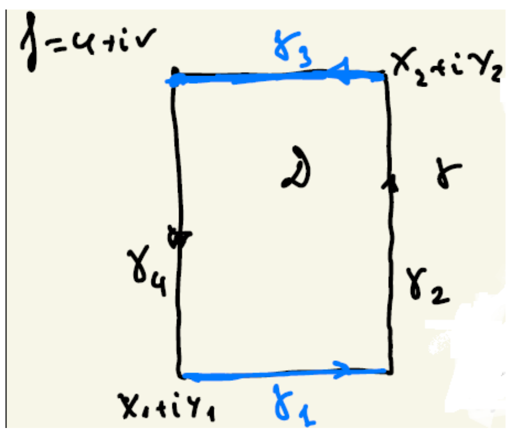


Figure 7: A rectangle.

We need to show that $\int_\gamma f d\xi = \int_{\gamma_1} f d\xi + \int_{\gamma_2} f d\xi + \int_{\gamma_3} f d\xi + \int_{\gamma_4} f d\xi$ vanishes. Note that

$$\int_{\gamma_1} f dz = \int_{x_1}^{x_2} f(x + iy_1)dx, \qquad \int_{\gamma_2} f dz = i \int_{y_1}^{y_2} f(x_2 + iy)dy,$$

$$\int_{\gamma_3} f dz = \int_{x_2}^{x_1} f(x + iy_2)dx, \qquad \int_{\gamma_4} f dz = i \int_{y_2}^{y_1} f(x_1 + iy)dy.$$

Denoting $f(x + iy) = u(x, y) + iv(x, y)$, we obtain

$$Re \int_\gamma f dz = \int_{x_1}^{x_2} [u(x, y_1) - u(x, y_2)]dx + \int_{y_1}^{y_2} [v(x_1, y) - v(x_2, y)]dy$$

$$= -\iint_D \frac{\partial u}{\partial y} dydx - \iint_D \frac{\partial v}{\partial x} dxdy = -\iint_D [\frac{\partial u}{\partial y} + \frac{\partial v}{\partial x}]dxdy.$$

Now note that the Cauchy-Riemann conditions imply that the integrand in the last formula is zero (see (2b)), so

$$Re \int_\gamma f dz = 0.$$

In the same way, condition (2b) (see Theorem 1.4 in Lecture 1) gives

$$Im \int_\gamma f dz = 0,$$

which proves the theorem for the case where D is a rectangle.

Now we can consider the case of a general simply-connected domain D.

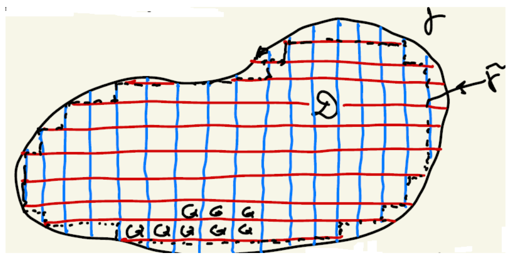


Figure 8: General simply-connected region $\mathcal{D}$ can be approximated by a union $\widetilde{\mathcal{D}}$ of small rectangles such that the boundary γ of D is sufficiently well approximated by the boundary $\tilde{\gamma}$ of $\widetilde{\mathcal{D}}$.

We take a sufficiently fine mesh of vertical and horizontal lines, so that the domain $\mathcal{D}$ is approximated by the union $\widetilde{\mathcal{D}}$ of small rectangles. Let $\tilde{\gamma}$ be the boundary of $\widetilde{\mathcal{D}}$, then

$$\int_{\tilde{\gamma}} f dz = \sum \int_{\tilde{\gamma}_k} f dz,$$

where $\tilde{\gamma}_k$ denotes the boundary of the rectangle number k. This formula is true because $\int_{\gamma_k}$ is, for each k, the sum of four integrals corresponding to the four boundaries of the k rectangle. In the total sum $\sum \int_{\tilde{\gamma}_k}$ each of the subintegrals corresponding to the vertical or horizontal segments which serve as a common boundary of two rectangles enters exactly twice and with different signs (because going around the boundary of each of these rectangles counter-clockwise means that the common boundary of the two rectangles is traversed twice with opposite orientations), i.e., they cancel each other. Hence, the only tersm which are not cancelled in $\sum \int_{\tilde{\gamma}_k}$ are the integrals over the vertical and horizontal segments which constitute the boundary of $\widetilde{\mathcal{D}}$, i.e., exactly the curve $\tilde{\gamma}$.
As we have proved before $\int_{\gamma_k} f(z)dz = 0$ for every rectangular path γ_k. Therefore, $\int_{\tilde{\gamma}} f(z)dz = 0$. We complete the proof of the theorem by noticing that the curves $\tilde{\gamma}$ and γ can be made arbitrarily close to each other (by taking the rectangular mesh finer), so $\int_\gamma f(z)dz$ can be arbitrarily well approximated by $\int_{\tilde{\gamma}} f(z)dz = 0$, i.e., $\int_\gamma f(z)dz = 0$, as claimed. □

This theorem implies
Corollary. *If $f(z)$ is analytic in a simply-connected domain $\mathcal{D}$, then $\int_{\gamma_1} f dz = \int_{\gamma_2} f dz$ for any curves γ_1 and γ_2 in $\mathcal{D}$ which have the same end points.*

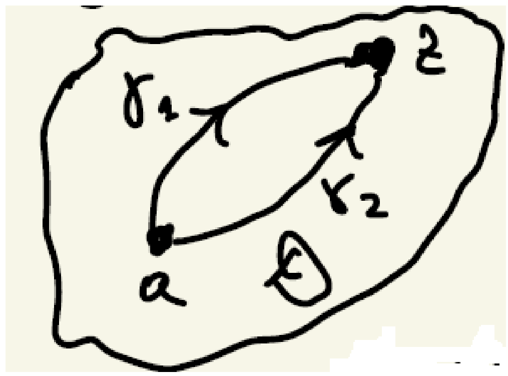


Figure 9: The integral of an analytic function in the simply-connected domain $\mathcal{D}$ can depend only on the end points of the path of integration, and does not depend on the path itself.

To prove this, consider the curve γ obtained by the concatenation of the paths γ_1 and reversed γ_2. We have $\int_\gamma = \int_{\gamma_1} - \int_{\gamma_2}$. The path γ is closed, so $\int_\gamma f(z)dz = 0$ by Theorem 3.1, hence $\int_{\gamma_1} f(z)dz = \int_{\gamma_2} f(z)dz$, as claimed.
In other words, we obtain that for an analytic function in a simply-connected domain $\int f(z)dz$ is a function of the end points of the integration path only, i.e., we have defined a single-valued function

$$\Phi(a, z) = \int_a^z f(\zeta)d\zeta$$

for all a and z in $\mathcal{D}$. Like in the real case, one obtains

$$\frac{d\Phi}{dz} = f(z), \tag{8}$$

so ϕ is the indefinite integral, or the *anti-derivative*, of f. If we replace the initial point a by any other point $b \in \mathcal{D}$, this will result in adding a constant to Φ: indeed, by construction, $\int_b^z f(\zeta)d\zeta = \int_a^b f(\zeta)d\zeta + \int_a^z f(\zeta)d\zeta$, i.e.,

$$\Phi(b, z) = \Phi(a, z) + \Phi(a, b)$$

for all $z \in \mathcal{D}$. So, we do not need to specify the initial point and can simply state that *an analytic function in a simply-connected domain has a well-defined integral*

$$\Phi(z) = \int f(z)dz + C$$

that satisfies the Newton-Leibnitz rule (8) (so, Φ is an analytic function in $\mathcal{D}$). We can also write

$$\Phi(z_1) - \Phi_(z_0) = \int_{z_0}^{z_1} \Phi'(\zeta)d\zeta.$$

This all looks the same as in the real case, however one can show that the analyticity of f (i.e., the existence of the derivative f') is not only a sufficient condition for the existence of the integral Φ, but it is also a necessary condition, and this is not analogous to real functions of a real variable.

Indeed, take any function $f(z)$ (we do not kn ow yet if it is analytic or not) with the property that $\int_\gamma f(z)dz$ depends only on the end points of the path γ. As before, this fact allows us to introduce a single-valued function $\Phi(z) = \int f(z)dz$ and one can check that $\frac{d\Phi}{dz} = f(z)$. Thus, Φ is analytic. We show in the next section that the derivative of any analytic function is also analytic, which gives us the analyticity of f, as it is the derivative of Φ.

Caution: If the domain of analiticity of $f(z)$ is not simply-connected, then $\int f dz$ can be multi-valued. For example, if $f(z) = \frac{1}{z}$, it has a singularity at $z = 0$), so its analyticity domain is $C \setminus \{z = 0\}$ and it is not simply-connected. Then, $\int_1^z f(\zeta)d\zeta = \ln z = \ln r + i\varphi + 2\pi ki$, where k can be any integer. The multi-valuedness of $\int f(z)dz$ implies that whenever we take an integral of f over a closed path which cannot be contracted to a point without hitting the singularity, it does not need to be zero. Thus, $\oint_{|z|=1} \frac{dz}{z} = 2\pi i$, see Figure 10.

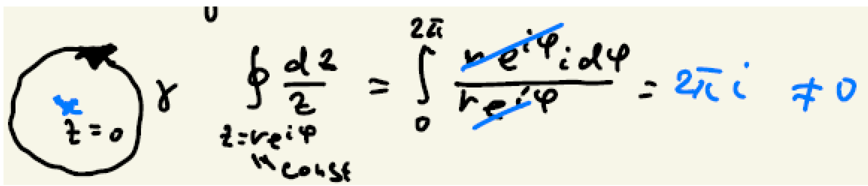


Figure 10: The integral of an analytic function over a closed path surrounding a singularity can be non-zero.

It is easy to generalise Cauchy's integral Theorem 3.1 to the case of domains which are not simply-connected: *the integral over the outer boundary minus the sum of the integrals over the boundaries of the holes equals zero.* An example is given by

Theorem 3.2. *If $f(z)$ is analytic in an annulus bounded by γ_1 and γ_2, then $\int_{\gamma_1} f dz = \int_{\gamma_2} f dz$.*

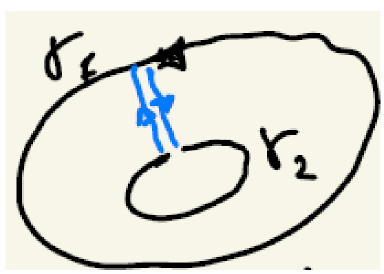


Figure 11: The integral over the outer boundary equals to the integral over the inner boundary.

To see this, we take a cut (shown by blue in Figure 11) connecting the outer boundary γ_1 and the inner boundary γ_2. The closed path starting with the end point of the blue cut, then going along γ_2 clockwise, then along the blue cut towards γ_1, then along γ_1 counter-clockwise, then back along the blue cut, bounds a simply-connected domain (the annulus minus the blue cut). Therefore, the integral of f over this path is zero, according to Theorem 3.1. Since the blue cut is traversed twice in opposite directions, the contributions of the "blue parts" of the path cancel each other, which gives us that the integral over γ_1 taken counter-clockwise plus the integral over γ_2 taken clockwise equals zero, which gives the theorem immediately.

3.2 Cauchy Formula

Our next goal is to prove the Cauchy formula, central for this topic.

Theorem 3.3. *Let f be analytic in a simply-connected region $\mathcal{D}$ bounded by a path γ, then*

$$f(z) = \frac{1}{2\pi i} \oint_\gamma \frac{f(\xi)}{\xi - z} d\xi. \tag{9}$$

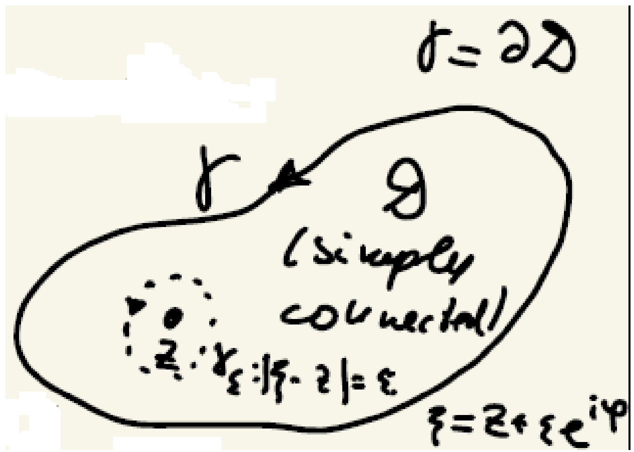


Figure 12: The integral over γ equals to the integral over a small circle around z.

Proof. Since f is analytic in $\mathcal{D}$, it follows that $\frac{f(\xi)}{\xi - z}$ is analytic in $\mathcal{D}\backslash\{\xi = z\}$. Therefore, by Theorem 3.2,

$$\oint_\gamma \frac{f(\xi)}{\xi - z} d\xi = \oint_{\gamma_\varepsilon} \frac{f(\xi)}{\xi - z} d\xi,$$

where γ_ε is a circle of small radius ε with center at the point z. This circle is given by equation $\xi = z + \varepsilon e^{i\varphi}$, where $\varphi \in [0, 2\pi]$. After making this substitution into the integral, we obtain

$$\oint_\gamma \frac{f(\xi)}{\xi - z} d\xi = \oint_{\gamma_\varepsilon} \frac{f(\xi)}{\xi - z} d\xi = \int_0^{2\pi} \frac{f(z + \varepsilon e^{i\varphi})}{\varepsilon e^{i\varphi}} d(\varepsilon e^{i\varphi}) = i \int_0^{2\pi} f(z + \varepsilon e^{i\varphi}) d\varphi = i \int_0^{2\pi} (f(z) + O(\varepsilon)) d\varphi.$$

Now, taking the limit $\varepsilon \to 0$, we obtain formula (9). □

Corollary. *An analytic function $f(z)$ has derivatives of all orders.*
Indeed, $\frac{1}{z - \xi}$ has derivatives of all orders with respect to z when $z \neq \xi$. The latter condition holds true for every $\xi \in \gamma$ in formula (9) (since $z \notin \gamma$), so we may differentiate (9) with respect to z as many times as we want. The result is given by formula (6).
In other words, *the derivative of an analytic function is analytic.* We showed in the previous section that this implies the equivalence of the existence of derivative and the existence of a path-independent integral for functions of complex variables.

The existence of derivatives of all orders allows one to write the Taylor expansion of an analytic function f at any point a of the analyticity domain. We know that for real functions of real variables the Taylor series can diverge, however it is not the case for functions of complex variables.

Theorem 3.4. *The Taylor series* $\displaystyle\sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!}(z-a)^n$ *of an analytic function* f *at any point* a *converges to* f *for all* z *such that* $|z-a| < \rho$*, where the convergence radius* ρ *equals to the distance from* a *to the nearest singularity point of* f*.*

Proof. By Cauchy formula (see (9))

$$f(z) = \frac{1}{2\pi i}\oint_\gamma \frac{f(\xi)}{\xi - z}d\xi = \frac{1}{2\pi i}\oint_\gamma \frac{f(\xi)}{\xi - a - (z-a)}d\xi \overset{u=\frac{z-a}{\xi-a}}{=} \frac{1}{2\pi i}\oint_\gamma \frac{f(\xi)}{(\xi-a)(1-u)}d\xi$$

$$\overset{\text{if } |u|<1}{=} \frac{1}{2\pi i}\oint_\gamma \frac{f(\xi)}{(\xi-a)}\sum_{n=0}^{\infty} u^n d\xi = \frac{1}{2\pi i}\oint_\gamma \frac{f(\xi)}{(\xi-a)}\sum_{n=0}^{\infty}\left(\frac{z-a}{\xi-a}\right)^n d\xi = \sum_{n=0}^{\infty}\left(\frac{1}{2\pi i}\oint_\gamma \frac{f(\xi)d\xi}{(\xi-a)^{n+1}}\right)(z-a)^n$$

$$= f(0) + f'(a)(z-a) + \ldots + \frac{f^{(n)}(a)}{n!}(z-a)^n + \ldots,$$

where we used formula (6) for the derivatives of f. As we see, the Taylor series converges to f if $|u| < 1$, i.e., if $|z-a| < |\xi - a|$ for all $\xi \in \gamma$. In other words, the radius of convergence ρ cannot be smaller than the distance to the integration path γ in the Cauchy formula.

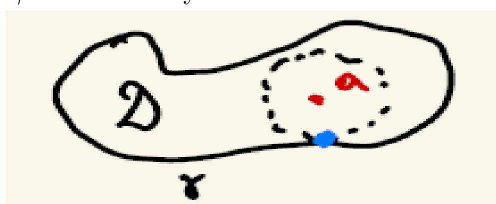
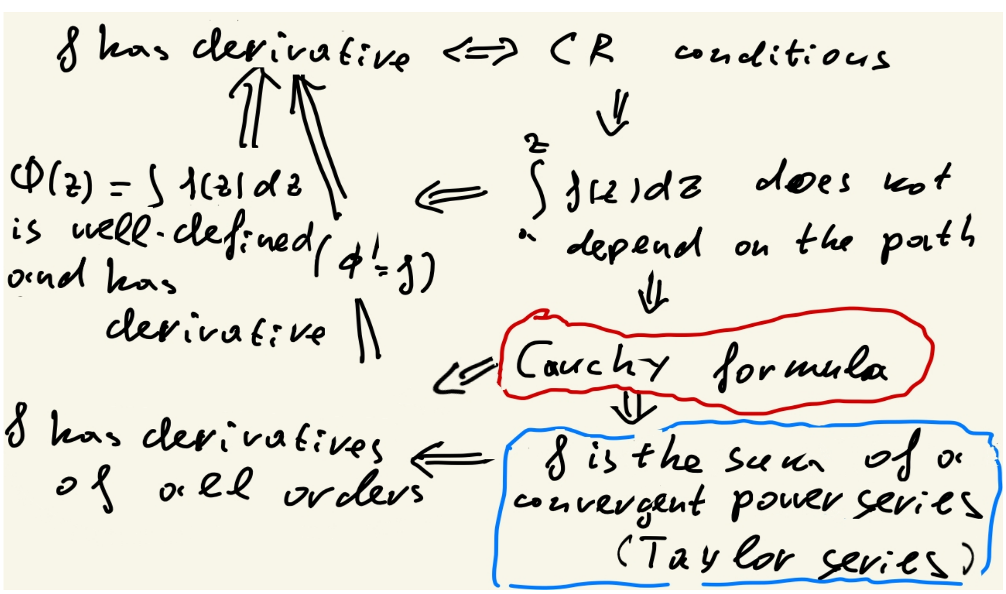


Figure 13: Radius of convergence is the distance to the boundary of the analiticity domain.

The path can be taken as close as we want to the boundary of the analyticity domain of f, hence ρ cannot be smaller than the distance from a to the nearest singularity. Of course, ρ cannot also be larger than this, since the sum of a convergent power series is analytic (non-singular) everywhere within the radius of convergence. □

We can summarise the above by the following diagram of logical relations between different (equivalent) defining properties of analytic functions:



3.3 Estimates on the derivatives of analytic functions

Letting $\gamma = \{|\xi - z| = R\}$ in Cauchy formula (5) and its derivatives (6), we get $\xi = z + Re^{i\varphi}$ and $d\xi = iRe^{i\varphi}d\varphi$, which gives

$$|f^{(n)}(z)| \leq \frac{n!}{2\pi}\int_0^{2\pi}\frac{|f(\xi)|}{R^{n+1}}Rd\varphi \leq n!\frac{M}{R^n},$$

where $M(R)$ is the maximal value of $|f|$ on γ. This proves estimates (7) for the absolute value of the function f (the case $n = 0$) and its derivatives ($n \geq 1$).

If $n = 0$, then we have $|f(z)| \leq \max_{|\xi - z| = R}|f(\xi)|$ for any R such that f is analytic in the disc of radius R around the point z. One can infer from this the following

Maximum Principle: *The absolute value of an analytic function takes its maximum on the boundary of the analyticity domain.*

This implies, in particular,

The Fundamental Theorem of Algebra: Every non-constant polynomial has a root in a complex plane, i.e., there exists $z \in C$ such that P(z)=0.

Indeed, if not, then the function $\frac{1}{P(z)}$ would be analytic everywhere, hence, by Maximum Prnciple, the maximal value of $\frac{1}{|P(z)|}$ will be attained as $z \to \infty$, but this is impossible because $\frac{1}{P(z)} \to 0$ as $z \to \infty$ for any non-constant polynomial.

Another interesting fact is given by

Theorem 3.5 (Liouville theorem)**.** *If an entire (i.e. analytic everywhere in the complex plane) function f is bounded, it is a constant.*

Proof. The boundedness means there exists M such that $|f(z)| < M$ for all z. By (7) with $n = 1$, we obtain

$$|f'(z)| \leq \frac{M}{R}$$

for any R, so

$$f'(z) \equiv 0 \Longrightarrow f(z) = \text{const}.$$

□

4 Lecture 4: Analytic Continuation

It was shown in the previous lecture that the value of an analytic function at any point z inside the analyticity domain are completely determined by its values on the boundary of the domain only (by Cauchy formula). Now our goal is to show somewhat opposite: an information about the local behaviour of an analytic function near any point inside the analyticity domain completely determines the function globally, i.e., everywhere in the domain.

Precisely, we show that an analytic function f defined on some connected domain $\mathcal{D}$ is uniquely defined by any of the following local data:

- The values it takes in a small neighbourhood of any given point $z_0 \in \mathcal{D}$.
- The coefficients of its Taylor expansion around z_0.
- The values it takes at the points of any given converging sequence $z_k \to z_0$.

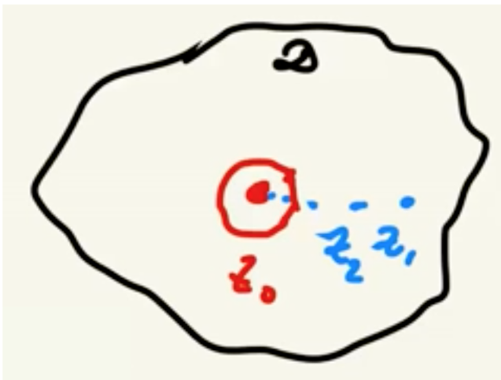


Figure 14: An analytic function is determined by its values in a neighbourhood of $z_0 \in \mathcal{D}$, or by its Taylor coefficients at z_0, or by its values at the points of any sequence converging to z_0.

Note that the first and second statements in this list are equivalent: if we know the function at every point of a however small neighbourhood of z_0, we can compute all the derivatives at this point (by computing ratios of finite differences and taking the limit), i.e., we know all the Taylor coefficients; conversely, the Taylor series of an analytic function f converges to it in some neighborurhood of z_0 (Theorem 3.4), i.e., the Taylor coefficients determine the function f in some small neighbourhood of z_0 uniquely.

The third statement is stronger - it says we do not need to know f everywhere in a neighbourhood of z_0, it is enough to fix the values of f at the points of just one sequence converging to z_0. So, we start with proving that the values of f along any such sequence uniquely determine f everywhere in some neighbourhood of z_0 (Theorem 4.1) and then show that this neighbourhood can be enlarged to cover the entire analyticity domain (Theorem 4.2).

Theorem 4.1. *Let $f(z)$ be analytic at $z = z_0$. Let $z_k \to z_0$ as $k \to \infty$. If g is analytic at z_0 and $g(z_k) = f(z_k)$ for all k then $g(z) = f(z)$ in some neighbourhood of z_0.*

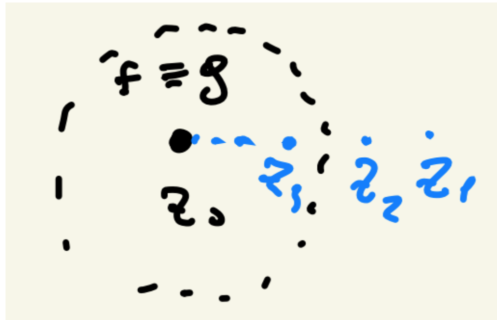


Figure 15: If two analytic functions are equal on some convergent sequence then they are equal everywhere in some neighbourhood of the limit of the sequence.

Proof. It is sufficient to show that $g(z_0) = f(z_0)$, $g'(z_0) = f'(z_0)$, ..., $g^{(n)}(z_0) = f^{(n)}(z_0)$ because these give the Taylor coefficients of f and g around z_0:

$$g(z - z_0) = \sum_{n=0}^{\infty} \frac{g^{(n)}(z_0)}{n!}(z - z_0)^n, \qquad f(z - z_0) = \sum_{n=0}^{\infty} \frac{f^{(n)}(z_0)}{n!}(z - z_0)^n.$$

Therefore, if we show that all the derivatives at z_0 are equal, then we show that f and g have the same Taylor expansion at z_0 and therefore are equal inside the radius of convergence of the Taylor series.

So, we just need to show that the values of f and all its derivatives at z_0 are completely determined by the values of $f(z_k)$ for $k = 1, 2, ...$ with $z_k \to z_0$. Now, since f is analytic it must be continuous. Then by continuity of f we may write:

$$f(z_0) = \lim_{k\to\infty} f(z_k)$$

Similarly, f has a first derivative at z_0, so we may write:

$$\begin{aligned} f'(z_0) &= \lim_{\Delta z\to 0} \frac{f(z_0 + \Delta z) - f(z_0)}{\Delta z} \\ &= \lim_{z_k\to z_0} \frac{f(z_k) - f(z_0)}{z_k - z_0} \end{aligned}$$

where we have set $\Delta z = z_k - z_0$. As we see, the first derivative $f'(z_0)$ is indeed determined uniquely by the values of f on the sequence z_k.

We continue by induction in n (the order of derivative): suppose that $f(z_0)$, $f'(z_0)$, ..., $f^n(z_0)$ have been determined by the values of f on the sequence z_k. Then by the analyticity of f, for all z close to z_0 we have:

$$f(z) = \sum_{r=0}^{n} \frac{f^{(r)}(z_0)}{r!}(z - z_0)^r + \frac{f^{(n+1)}(z_0)}{(n+1)!}(z - z_0)^{n+1} + \mathcal{O}\left((z - z_0)^{n+2}\right)$$

Then we may take $z = z_k$ and re-arrange to get:

$$\begin{aligned} f^{(n+1)}(z_0) =& \frac{(n+1)!}{(z_k - z_0)^{n+1}} \left(f(z_k) - \sum_{r=0}^{n} \frac{f^{(r)}(z_0)}{r!}(z_k - z_0)^r \right) + \mathcal{O}(z_k - z_0) \\ =& \lim_{k\to\infty} \frac{(n+1)!}{(z_k - z_0)^{n+1}} \left(f(z_k) - \sum_{r=0}^{n} \frac{f^{(r)}(z_0)}{r!}(z_k - z_0)^r \right) \end{aligned}$$

Thus, since the right-hand side is determined by the values of f on the sequence z_k (by our induction assumption), so is $f^{(n+1)}(z_0)$. By induction, we can continue this up to infinity, and compute all the derivatives of f at z_0. □

Theorem 4.2. *If f and g are analytic in a connected domain $\mathcal{D}$ and $f \equiv g$ in a neighbourhood of $z_0 \in \mathcal{D}$ then $f \equiv g$ in $\mathcal{D}$.*

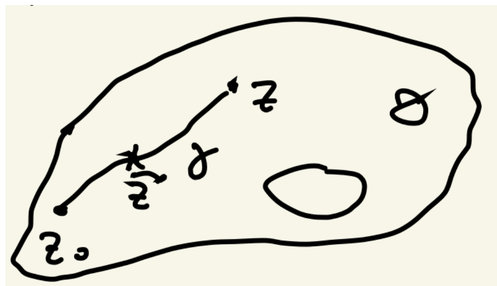


Figure 16: Here, γ is a curve in the domain $\mathcal{D}$ and we suppose that $\tilde{z}$ is the end point of the initial arc of this curve where f and g are equal.

Proof. Let z be an arbitrary point in $\mathcal{D}$ and let γ be a curve in $\mathcal{D}$ connecting z_0 to z. Suppose $\tilde{z}$ is the last point on the curve γ such that $f \equiv g$ on γ between z_0 and $\tilde{z}$. Take any sequence of points z_k which lie on the curve γ between z_0 and $\tilde{z}$ and which converge to $\tilde{z}$. Then, by Theorem 4.1, there exists a neighbourhood of $\tilde{z}$ on which $f \equiv g$. In particular, if $\tilde{z}$ is not equal to z, then f continues to be equal to g on γ after $\tilde{z}$, which contradicts the definition of $\tilde{z}$. Thus we conclude that $\tilde{z} = z$, i.e., $f(z) = g(z)$, and, since z is taken arbitrarily, we conclude that $f \equiv g$ everywhere in $\mathcal{D}$. □

As we see, if we have an analytic function f defined in some neighbourhood of a point z_0, we may try to define the *analytic continuation* of f on a large domain $\mathcal{D}$. The results above guarantee that when such continuation exists, it is unique. One possible way to construct such continuation is to take any point $z \in \mathcal{D}$, connect z_0 and z by a continuous path γ, and cover γ by a finite sequence of open discs B_j of some small radius R and centers $z_0, z_1, ..., z_k = z$ lying on γ such that the distance between the two consecutive points z_j is smaller than R, i.e., $z_{j+1} \in B_j$. We know f in B_0, so we can, theoretically, compute the Taylor coefficients of f at z_1 (as it lies inside B_0). If the convergence radius of this Taylor series is not smaller than R, then we get f defined everywhere in B_1, hence we can build the Taylor series at the point $z_2 \in B_1$, and so on, until we reach the point z after finitely many steps of this procedure.

In this way we define an analytic function f not just at a point z but also in a neighbourhood of the path γ (in the above construction, this neighbourhood is the set $B_0 \cup \ldots \cup B_k$). We say that we have the analytic continuation of f to the point z along the path γ.

We do not know in advance whether this will work or not: a priori, the radius of convergence can get below R and the continuation process would stop. In general, it might happen that any path connecting z_0 and z hits a singularity of f, then no analytic continuation of f to z will be possible. For example, the function $f(z) = 1/z$ cannot be continued from a neighbourhood of $z = 1$ to the point $z = 0$: by Theorem 4.2, if there is an analytic at $z = 0$ function $g(z)$ which coincides with $1/z$ near $z = 1$, then $g(z) = 1/z$ everywhere where $f(z) = 1/z$ is analytic, i.e., for all $z \neq 0$, so $g(0) = \lim_{z\to 0} g(z) = \infty$, a contradiction.

However, if the analytic function f can indeed be continued to some neighbourhood of γ, then the radius of convergence of the Taylor series at any point of γ will be not smaller than the distance from γ to the set of singular points. So, by taking R smaller than this distance, wee will obtain the continuation along γ in a finite number of steps of the above construction with making the Taylor expansion at a sequence of points on γ.

In practice, this continuation algorithm can be difficult to implement. A typical way employed for the analytic continuation is to find some "defining equation" which the function f satisfies near z_0, make the analytic continuation of this equation (whatever it means) outside a small neighbourhood of z_0, and then define f globally as the solution of this equation. This sounds wage and is easier explained in examples:

- the exponential function of real numbers satisfies the equation $f' = f$, with $f(0) = 1$, so we define the exponent of complex numbers as the solution of the same equation; we have $f = f' = f'' = \ldots = f^{(n)}$, for any n, so the n-th coefficient of the Taylor expansion at $z = 0$ equals to $1/n!$, which uniquely defines

$$e^z = \sum_{n=0}^{\infty} \frac{z^n}{n!};$$

- the real logarithm is the inverse of the exponent, so we define $f(z) = \ln z$ as the solution of the equation

$$e^{\ln z} = z,$$

  which gives $\ln z = \ln r + i\varphi$ for $z = re^{i\varphi}$.

Importantly, no matter which trick we use to construct an analytic continuation, an analytic continuation along any two sufficiently close paths yields the same result (just by the definition of the continuation along a path). It follows that the analytic continuation from z_0 to z returns the same value of $f(z)$ for any two paths between z_0 and z which are related by a continuous deformation, provided we do not hit a singularity of f when transforming one path to the other (in particular, the result of analytic continuation does not depend on the path γ when we continue f to a simply-connected domain $\mathcal{D}$).

However, as we saw with the example of $f = 1/z$, the position of the singularities of f (as well as all the other information about f, by Theorem 4.2) is completely determined by the behaviour of f near z_0. In particular, if the Taylor coefficients $a_n = f^{(n)}(z_0)/n!$ computed at $z = z_0$ do not decay super-exponentially with n, then the radius of convergence ρ of the Taylor series is finite, and we know by Theorem 3.4 that somewhere at a distance ρ from z_0 there is at least one singular point z^* of f (a point where f is not analytic). In this case, when we continue f to a point z beyond the radius-ρ disc there exist paths with the same end points z_0 and z which cannot be continuously deformed one to another without hitting the point z^*. Analytic continuation along such paths may give different values for $f(z)$ (see Fig.17).

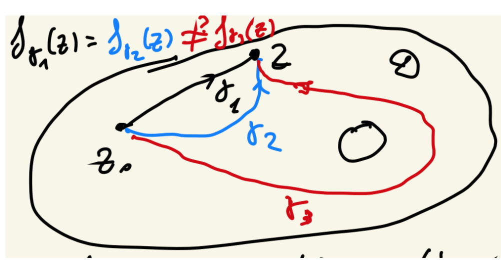


Figure 17: Different paths may be used to define analytic continuations. If the domain is not simply-connected then choosing different paths may lead to different analytic continuations.

Example. Let us consider the logarithm function defined by $\ln z = \ln r + i\varphi$ in some neighbourhood of $z_0 = 1$. Then the value of $\ln i$ depends on the path we take from $z_0 = 1$ to $z = i$.

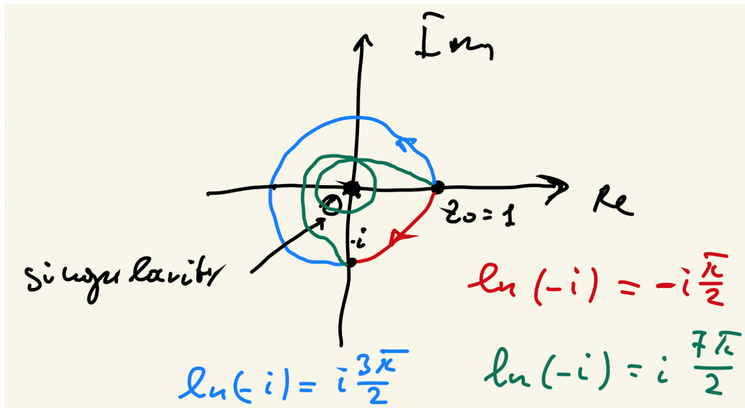


Figure 18: The analytic continuation of the logarithm can take different values, depending on which path is chosen.

Remark. It may appear that the idea of multi-valued analytic continuations contradicts Theorem 4.2. However, the different possible continuations are, in fact, defined on different domains (simply-connected neighbourhoods of the continuation paths). This domains are distinguished by the choice of the so-called branch cuts, which we will discuss in the next lecture.

Complex numbers were invented to help solving problems about real numbers. The main observation of the theory of complex analytic functions is that local data of an analytic function in a neighbourhood of any given point contains all the information about this function anywhere else in the complex plane, however far away from the initial point. In many situations this allows one to replace computations performed in one part of the complex plane by simpler computations at a different part. We will see examples of this approach in the next and following lectures. The powerful idea is that whenever we have to work with a function of real variables which admit an analytic continuation to complex numbers, we continue the function to as large domain as possible in the complex plane, and then perform computations anywhere in this domain in order to obtain (with luck) as much information as possible about the behaviour for real values of the variables.
This method is applied extensively in various branches of physics because many fundamental physical objects and phenomena happened to be described (or approximately modelled) by real-analytic functions, i.e., the functions of a real variable which are given by convergent power series

$$f(x) = \sum_{n=0}^{\infty} a_n (x - x_0)^n.$$

The (unique by Theorem 4.1) analytic continuation to a complex neighbourhood of $z = x_0$ is given by the sum of the same series $f(z) = \sum_{n=0}^{\infty} a_n (z - x_0)^n$. As we mentioned, it can be beneficial to continue the function to a larger domain, but then there can be a price to pay: if the coefficients a_n do not decay super-exponentially with n, then singularities must emerge, the analyticity domain can become not simply-connected, and we may have to deal with a multi-valued function as a result.

5 Lecture 5: Residue formula

As we discussed in the previous lectures, the full information about an analytic function is always encoded in local data. Extracting such information can be very hard, but in some cases effective methods exists. A fruitful approach is the analysis of **singularities**, i.e., the points where the analytic function becomes infinite, or loses continuity, or loses analyticity. There are in general two types of singularities: non-isolated singularities and isolated singularities.

A non-isolated singularity is a point which is a limit point of a sequence of other singularities of the function. For instance, for the function $f(z) = \frac{1}{\sin(\frac{1}{z})}$, let $z_n = \frac{1}{\pi n}$. We have $\sin(1/z_n) = 0$, so $f(z_n) = \infty$, i.e., z_n are singular points of f. The sequence $\{z_n\}$ has a limit point $z^* = 0$, therefore z^* is a non-isolated singularity of f. I am not aware of a systematic theory of non-isolated singularities.

An isolated singularity is a singularity that has no other singularities in its neighbourhood. The isolated singularities consist of

- branching points;
- poles;
- essential singularities.

Definition 5.1. A point z_0 is a branching point of $f(z)$ if f is not single-valued in a neighbourhood of z_0, i.e., for any z close to z_0 the analytic continuation of f from z to the same z along a path γ around z_0 returns a different value of $f(z)$. A way to deal with multi-valued analytic functions is to decompose them into several single-valued branches, by removing from the complex plane the so-called "branch cuts".

Definition 5.2. A branch cut is a line γ such that the multi-valued analytic function becomes a collection of single-valued analytic functions in a complement to γ. Obviously, a branch cut must go through all branching points; often it has an arc which extends to infinity. Note that there is much freedom in the choice of branch cuts, as one can see e.g. in the example below

Example

1. Let $f(z) = lnz = lnr + i\varphi + 2\pi k$ where k is any integer. This is a multi-valued function, because φ is defined modulo 2π. Take the positive real semi-axis $\gamma = \{y = 0, x \geq 0\}$ as a branch cut for f; note that this line goes from the branching point at $z = 0$ to infinity. Different branches of f correspond to different choices of k. To keep every branch continuous in the complement to γ, we set the values of φ to run the interval $0 < \varphi < 2\pi$. Note that the values of φ close to the line γ above and below it approach, respectively, 0 and 2π. Thus, each single-values branch has a discontinuity at the branch cut γ (another way to think of it is that each fixed-k branch is single-valued in the complement to γ and double-valued on γ, with the two values differing by $2\pi i$). In fact any line extending from zero to infinity can serve as a branch cut: one distinguishes different branches by different integers k and just needs to properly define the range of change of φ such that φ would change continuously in the complement to the cut and undergo the $2\pi i$ jump on it. Thus, a natural candidate for another branch cut is the negative real semi-axis $\{y = 0, x \leq 0\}$, which corresponds to the choice $-\pi < \varphi < \pi$.
2. Let $f(z) = \sqrt{z} = \pm\sqrt{r}e^{i\frac{\varphi}{2}}$. A possible branch cut for f is $\{y = 0, x < 0\}$. The range of φ is chosen as $\{-\pi < \varphi < \pi\}$; the branches are distinguished by + or − sign in front of $\sqrt{r}$.

3. Let $f(z) = \sqrt{z^2-1} = z\sqrt{1-\frac{1}{z^2}} = z\sqrt{1-\frac{1}{r^2}e^{-2i\varphi}}$. Take the finite segment $\gamma = \{y=0, -1 \le x \le 1\}$ as a branch cut. The two branches of f correspond to the "plus" and "minus" branches of the square root, as in the previous example. Notice that the real part of $\zeta(z) = 1-\frac{1}{r^2}e^{-2i\varphi}$ equals to $1-\frac{1}{r^2}\cos(2\varphi)$ and is greater than 0 if $r>1$. Therefore, a closed path around the branch cut γ in the z-plane corresponds to a closed path which does not intersect the negative real semi-axis (the branch cut of the square root in the previous example) in the ζ-plane. Hence, going around any closed path in the z-plane which does not intersect γ does not switches the branches of $\sqrt{\zeta(z)}$. On the other hand, crossing γ at a point $z=x$ with $|x|<1$ corresponds to $\zeta(z) = 1-\frac{1}{x^2}$ being real and negative, so we jump from $i\sqrt{1-x^2}$ to $-i\sqrt{1-x^2}$.

As we see, introducing branch cuts cancels the multi-valuedness problem and allows for using the theory developed for single-valued analytic functions. One should be aware, however, that the analyticity domain of the single-valued branches excludes the branch cut and that the single-valued branches become discontinuous at the cuts. Ignoring this could lead to mistakes (for example, when using Cauchy formula, one cannot consider closed paths intersecting branch cuts). On the other hand, an information about the jumps of the function across the branch cut can be useful in computations, see Keyhole Formula in the next lecture.

When an analytic function $f(z)$ is single-valued in a neighbourhood of an isolated singularity z_0, the main toll of the analysis is the *Laurent series*. Namely, it is a theorem (we do not prove it here) that f is a sum of a convergent double-infinite power series

$$f(z) = \sum_{n=-\infty}^{\infty} b_n(z-z_0)^n = \cdots + \frac{b_{-1}}{z-z_0} + b_0 + b_1(z-z_0) + \cdots . \qquad (10)$$

As we see, the Laurent series includes both negative and positive powers of $z-z_0$; it can be represented as the sum of two series, the regular part $\sum_{n=0}^{+\infty} b_n(z-z_))^n$ and the singular part $\sum_{n=1}^{+\infty} b_{-n}(z-z_))^{-n}$ where both series are convergent in a small neighbourhood of $z=z_0$ (with the point z_0 itself taken out for the singular part). The coefficients b_n are uniquely defined by the function f. One can show that

$$b_n = \frac{1}{2\pi i}\oint_{|z-z_0|=\varepsilon} f(z)(z-z_0)^{-n-1}dz,$$

where $\varepsilon > 0$ is small enough (the integral does not depend on ε because of the analyticity). However, this formula is not particularly useful: typically, we are more interested in evaluating the integral in the right-hand side than determining b_n.
Therefore, one resorts to indirect methods of evaluating the Laurent coefficients. For example,

1. Let $f(z) = \frac{1}{z}$. Then $z_0 = 0$ is its singularity, and $b_{-1} = 1$, while $b_n = 0$ for all $n \neq -1$.
2. Let $f(z) = e^{\frac{1}{z}}$. By substituting $\frac{1}{z}$ into the Taylor expansion for the exponent, we obtain

$$f(z) = \cdots + \frac{1}{n!}z^{-n} + \cdots + \frac{1}{z} + 1,$$

   i.e., $b_n = \frac{1}{|n|!}$ for $n \le 0$ and $b_n = 0$ for $n > 0$.

The coefficient b_{-1} plays a special role and is called the *residue* of f at $z = z_0$. The following notation is used

$$b_{-1} = \mathrm{Res}(f, z_0).$$

We now use some simple examples to illustrate how to evaluate b_{-1}:

1. for $f = \frac{3}{z}$, we have $\text{Res}(f, 0) = 3$;
2. for $f = \frac{1}{z^2}$, we have $\text{Res}(f, 0) = 0$;
3. for $f = \cos(\frac{1}{z}) = 1 - \frac{1}{2z^2} + \cdots$, we have $\text{Res}(f, 0) = 0$;
4. for $f = \sin(\frac{1}{z}) = \frac{1}{z} - \frac{1}{6z^3} + \cdots$, we have $\text{Res}(f, 0) = 1$.

Definition 5.3. If the Laurent series of f has infinitely many non-zero coefficients b_n where $n < 0$, then z_0 is call *an essential singularity*. If the Laurent series contains only finitely many non-zero b_n where $n < 0$, then z_0 is called *a pole*.

The pole is the simplest type of singularity. By definition, if z_0 is a pole, then f can be represented as:

$$f(z) = \frac{b_{-N}}{(z - z_0)^{-N}} + \cdots + \frac{b_{-1}}{(z - z_0)^{-1}} + \sum_{n=0}^{\infty} b_n (z - z_0)^n. \tag{11}$$

Notice that $g(z) = f(z)(z - z_0)^N$ is analytic at z_0, so the Laurent coefficients b_k can be computed as Taylor coefficients of the function g at z_0. Thus, in the case of poles, this gives us a regular method of determining the Laurent coefficients.
When $N = 1$, the pole is called simple. In this case, we can write f as

$$f(z) = \frac{b_{-1}}{(z - z_0)^{-1}} + \sum_{n=0}^{\infty} b_n (z - z_0)^n.$$

As we see, $b_{-1} = \lim_{z \to z_0} f(z)(z - z_0)$ for simple poles. If $f(z) = \frac{\phi(z)}{\psi(z)}$, where ϕ and ψ are analytic and $\psi(z_0) = 0$, $\phi(z) \neq 0$, then

$$b_{-1} = \lim_{z \to z_0} \frac{\phi(z)}{\frac{\psi(z)}{z - z_0}} = \frac{\phi(z_0)}{\lim_{z \to z_0} \frac{\psi(z)}{z - z_0}} = \frac{\phi(z_0)}{\psi'(z_0)}.$$

This is a useful formula, so we repeat it:

$$\text{Res}(f, z_0) = \frac{\phi(z_0)}{\psi'(z_0)}. \tag{12}$$

Example Let $f(z) = \frac{1}{z^4 + 1}$. Then the poles of f are the points such that $z^4 + 1 = 0$, which is equivalent to $z = e^{\frac{i\pi}{4} + \frac{i\pi k}{2}} = \pm\frac{\sqrt{2}}{2} \pm \frac{i\sqrt{2}}{2}$. By (12), the residue of f at $z = z_k$ is:

$$\text{Res}(f, z_k) = \frac{1}{4z_k^3} = \frac{z_k}{4z_k^4} = \frac{-z_k}{4}$$

(we use that $z_k^4 = -1$ here).

The next theorem explains why we are interested in the residues of a function:

Theorem 5.4. *For a single-valued analytic function f whose all singularities $\{z_k\}$ inside a domain $\mathcal{D}$ bounded by a closed path γ are isolated and non-branching (either essential singularities, or poles), we have*

$$\oint_\gamma f(z)dz = 2\pi i \sum_k Res(f, z_k), \tag{13}$$

where the integration is taken anti-clockwise.

Proof. Figure 19 shows a domain $\mathcal{D}$ and $\{z_k\}_k$ are the isolated singularities of f.

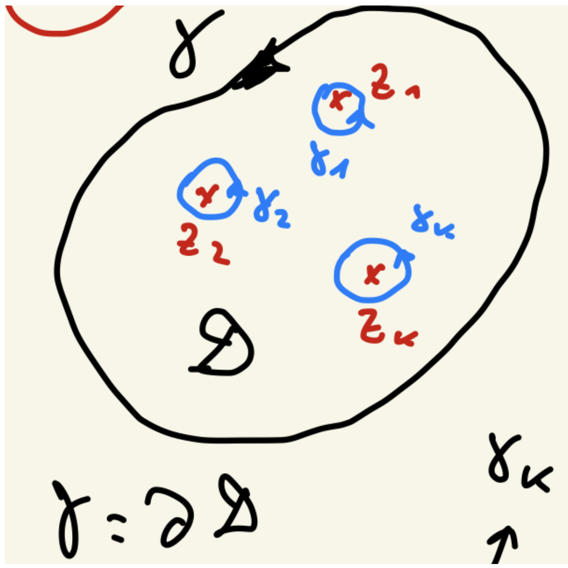


Figure 19: A domain $\mathcal{D}$, bounded by γ, with $\{z_k\}$ as singularities of f. The paths γ_k are small circles of radius ε around z_k.

Consider a small circle around z_k, denoted by

$$\gamma_k := \{z = z_k + \varepsilon e^{i\varphi}|_{\varphi\in[0,2\pi]}\}.$$

Since f is analytic function in $\mathcal{D}\setminus \cup_k \{z_k\}$, we have

$$\oint_\gamma f(z)dz = \sum_k \oint_{\gamma_k} f(z)dz$$

(see comment before Theorem 3.2). Recall that $f(z)$ is the sum of the Laurent series, i.e.,

$$f(z) = \sum_{n=-\infty}^{+\infty} b_n^{(k)}(z - z_k)^n$$

near $z = z_k$. Notice that for the points z on γ_k, we have $z = z_k + \varepsilon e^{i\varphi}$, and thus $dz = i\varepsilon e^{i\varphi} d\varphi$. Therefore,

$\oint_{\gamma_k} f(z)dz$ can be written as:

$$\begin{aligned}
\oint_{\gamma_k} f(z)dz &= \sum_{n=-\infty}^{+\infty} b_n^{(k)} \oint_{\gamma_k} (z - z_k)dz \\
&= \sum_{n=-\infty}^{+\infty} b_n^{(k)} i\varepsilon^{n+1} \int_0^{2\pi} e^{i\varphi(n+1)} d\varphi \\
&= \sum_{n\neq -1} b_n^k i\varepsilon^{n+1} \left. \frac{e^{i\varphi(n+1)}}{n+1} \right|_0^{2\pi} + b_{-1}^{(k)} i \int_0^{2\pi} d\varphi \\
&= 2\pi i \; b_{-1}^k = 2\pi i \; \mathrm{Res}(f, z_k),
\end{aligned}$$

which finishes the proof. □

Example Let $f(z) = \frac{1}{1+z^4}$. We already know that it has four poles: $z_k = \pm\frac{\sqrt{2}}{2} \pm \frac{i\sqrt{2}}{2}$ and the corresponding residues are

$$\mathrm{Res}(f, z_k) = \frac{-z_k}{4}.$$

We want to compute the contour integral of f over different paths, for example, γ_1 and γ_2 as shown in Figure 20; the points z_1 and z_4 are inside γ_1, and z_3 and z_4 are inside γ_2.

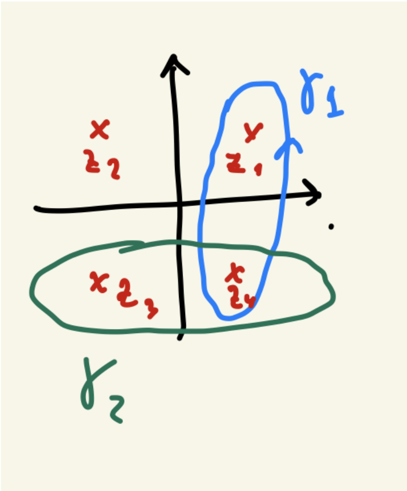


Figure 20: The poles $z_1, ..., z_4$ of f, and two paths γ_1 and γ_2.

We have:

$$\oint_{\gamma_1} f(z)dz = \oint_{\gamma_1} \frac{1}{1+z^4} dz = 2\pi i \left[\mathrm{Res}(f, z_1) + \mathrm{Res}(f, z_4)\right] = -\frac{2\pi i}{4}(z_1 + z_4) = -\pi i \frac{\sqrt{2}}{2},$$

and

$$\oint_{\gamma_2} f(z)dz = \oint_{\gamma_2} \frac{1}{1+z^4} dz = 2\pi i \left[\mathrm{Res}(f, z_3) + \mathrm{Res}(f, z_4)\right] = -\frac{2\pi i}{4}(z_3 + z_4) = \frac{-\pi\sqrt{2}}{2}.$$

6 Lecture 6: Integrals over reals

Space and time are real (not in the sense they really exist, which may be debatable, but in the sense they are described by real numbers). Therefore, the structure and evolution of physical objects are described by functions of real variables. We want to record these functions by means of sufficiently simple formulas - this makes these functions analytic. Hence, we can analytically continue them, in a unique way, to complex variables and use powerful tools of complex analysis for computations.
In our course, we employ this method by applying the residue theorem Theorem 5.4 to the problem of evaluating integrals over the real line.

Theorem 6.1. *Let $f(x)$ be a function defined on the real line, such that f admits an analytic continuation $f(z)$ to the upper half-plane $\{Im(z) \geq 0\}$ with a finite number of non-branching singularities z_k. If $|f(z)| = o\left(\frac{1}{|z|}\right)$ as $|z| \to \infty$ then:*

$$\int_{-\infty}^{\infty} f(x)dx = 2\pi i \sum_{Im(z_k)>0} \text{Res}(f, z_k) \tag{14}$$

Proof. Let us choose γ to be the curve shown in Figure 21. This consists of an interval $(-R, R)$ on the real line, as well as a semi-circle of radius R in the upper half-plane (the semi-circle is given by the equation $\{z = Re^{i\varphi}|_{\varphi\in[0,\pi]}\}$). Since the analytic continuation of f is defined in the upper half-plane, we may integrate along γ, giving:

$$\begin{aligned} \oint_\gamma f(z)dz &= \int_{-R}^{R} f(x)dx + \int_0^\pi iRe^{i\varphi} f(Re^{i\varphi})d\varphi \\ &= 2\pi i \sum_{\substack{\text{Im}(z_k)>0 \\ |z_k|<R}} \text{Res}(f, z_k) \end{aligned} \tag{15}$$

The final equality is given by the residue theorem. Now we take the limit $R \to \infty$, and note that:

$$\left| \int_0^\pi iRe^{i\varphi} f(Re^{i\varphi})d\varphi \right| \leq \pi \max_{\varphi\in(0,\pi)} \left|Rf(Re^{i\varphi})\right| \to 0 \tag{16}$$

because $\left|f(Re^{i\varphi})\right| = o(1/R)$ as $R \to \infty$. We also note that as $R \to \infty$ all of the poles of f will be contained within the area enclosed by γ; thus all of the upper half-plane residues of f are included in the sum in eq. (15) as $R \to \infty$. Thus we obtain eq. (14)

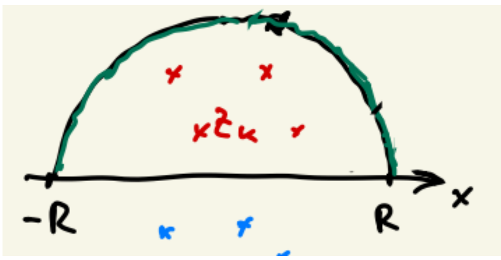


Figure 21: The curve γ consists of the real interval $(-R, R)$ along with a semicircle of radius R in the upper half-plane, traversed anti-clockwise. As we take $R \to \infty$ all of the singularities of f in the upper half-plane will be contained in the area enclosed by γ.

□

Remark

We may make a nearly identical argument in the case in which f may be analytically continued to the lower half-plane. In that case we have:

$$\int_{-\infty}^{\infty} f(x)dx = -2\pi i \sum_{\text{Im}(z_k)<0} \text{Res}(f, z_k) \tag{17}$$

provided $|f(z)| = o\left(\frac{1}{|z|}\right)$ as $|z| \to \infty$. Notice the minus sign here, arising from the fact that we traverse the interval $(-R, R)$ in the opposite direction, as in Figure 22. eq. (14)

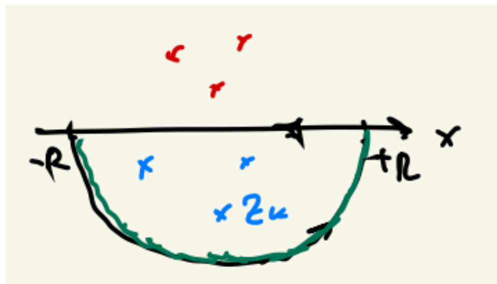


Figure 22: Now if we traverse the semi-circle anti-clockwise, we integrate over the real line in the opposite direction, picking up a minus sign in (17).

Example

Consider the function $f(x) = \frac{1}{1+x^2}$. This has an obvious analytic continuation to the complex plane: $f(z) = \frac{1}{1+z^2}$, with poles at $z_\pm = \pm i$. To compute the residues we write $f(z) = \frac{\phi(z)}{\psi(z)}$ with $\phi(z) = 1$ and $\psi(z) = 1 + z^2$, then the residues are given by:

$$\text{Res}(f, z_\pm) = \frac{\phi(z_\pm)}{\psi'(z_\pm)} = \frac{1}{2z_\pm} = \frac{1}{2i}, -\frac{1}{2i} \tag{18}$$

Moreover, the large-$|z|$ behaviour of $f(z)$ may be bounded above by:

$$|f(z)| \leq \frac{1}{|z|^2} = o\left(\frac{1}{|z|}\right) \tag{19}$$

Thus we may apply eq. (14) to find:

$$\int_{-\infty}^{\infty} f(x)dx = 2\pi i \left(\frac{1}{2i}\right) = \pi \tag{20}$$

We may also use the lower half-plane formula eq. (17):

$$\int_{-\infty}^{\infty} f(x)dx = -2\pi i \left(-\frac{1}{2i}\right) = \pi \tag{21}$$

Finally, we may verify that this is correct by using techniques from real calculus:

$$\int_{-\infty}^{\infty} \frac{1}{1+x^2} dx = \arctan x|_{-\infty}^{\infty} = \frac{\pi}{2} - \left(-\frac{\pi}{2}\right) = \pi. \tag{22}$$

A variation of Theorem 6.1 (with a slightly relaxed condition on the decrease rate at infinity) provides a formula for the so-called oscillating integrals.

Theorem 6.2. *Let f be a real analytic function with an analytic continuation $f(z)$ in the upper half-plane such that $|f(z)| \to 0$ as $z \to \infty$. Then:*

$$\int_{-\infty}^{\infty} f(x)e^{ix}dx = 2\pi i \sum_{Im(z_k)>0} \operatorname{Res}(f(z)e^{iz}, z_k) \tag{23}$$

Proof. As before, we consider an integral around a semi-circle in the upper half-plane:

$$\oint_\gamma f(z)e^{iz}dz = \int_{-R}^{R} f(x)e^{ix}dz + \int_0^\pi f(Re^{i\varphi})e^{iRe^{i\varphi}} iRe^{i\varphi}d\varphi = 2\pi i \sum_{\substack{\operatorname{Im}(z_k)>0 \\ |z_k|<R}} \operatorname{Res}(f(z)e^{iz}, z_k) \tag{24}$$

Now, let us consider the integral around the semi-circle part. We need to show that this converges to 0 for large R. Note that $\left|e^{iz}\right| = \left|e^{iR\cos\varphi - R\sin\varphi}\right| = e^{-R\sin\varphi}$. We have

$$\left|\int_0^\pi f(Re^{i\varphi})e^{iRe^{i\varphi}} iRe^{i\varphi}d\varphi\right| \le R\int_0^\pi \left|f(Re^{i\varphi})\right| e^{-R\sin\varphi}d\varphi \le R \max_{|z|=R,\operatorname{Im}(z)\ge 0} |f(z)| \int_0^\pi e^{-R\sin\varphi}d\varphi.$$

Since $\sin\varphi = \sin(\pi-\varphi)$, we have

$$\int_0^\pi e^{-R\sin\varphi}d\varphi = 2\int_0^{\pi/2} e^{-R\sin\varphi}d\varphi,$$

and since $\sin\varphi > 2\varphi$ for $\varphi \in [0, \pi/2]$, we obtain

$$\int_0^{\pi/2} e^{-R\sin\varphi}d\varphi < \int_0^{\pi/2} e^{-R\varphi/2}d\varphi \le \frac{2}{R}.$$

Thus,

$$\left|\int_0^\pi f(Re^{i\varphi})e^{iRe^{i\varphi}} iRe^{i\varphi}d\varphi\right| \le 4 \max_{|z|=R,\operatorname{Im}(z)\ge 0} |f(z)| \to 0 \quad \text{as} \quad R \to +\infty,$$

which completes the proof of the theorem. □

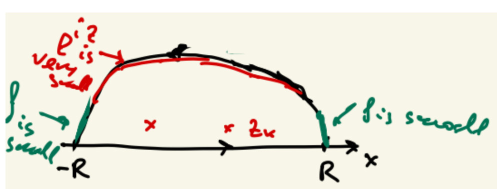


Figure 23: We split the integral up into three parts: two small intervals at the ends of the semicircle; and the rest of the semi-circle, on which $e^{iz} = e^{iR\cos x - R\sin x}$ is very small for large R.

Remark

In the same way, under conditions of Theorem 6.2, we obtain the following formula

$$\int_{-\infty}^{\infty} f(x)e^{i\lambda x}dx = 2\pi i \sum_{\mathrm{Im}(z_k)>0} \mathrm{Res}(f(z)e^{i\lambda z}, z_k) \tag{25}$$

for a constant $\lambda > 0$. Note that this formula is no longer true when $\lambda > 0$: as one can seen in the proof, it is crucial that $e^{i\lambda z} \to 0$ as $z \to 0$ along the positive imaginary axis. This does not hold for $\lambda < 0$, so we must do the integration in the lower half-plan, which gives

$$\int_{-\infty}^{\infty} f(x)e^{i\lambda x}dx = -2\pi i \sum_{\mathrm{Im}(z_k)<0} \mathrm{Res}(f(z)e^{i\lambda z}, z_k) \tag{26}$$

when $\lambda < 0$.

Remark

We may use Theorem 6.2 to calculate two real integrals simultaneously:

$$\begin{aligned} \int_{-\infty}^{\infty} f(x)\cos x\, dx &= \mathrm{Re} \int_{-\infty}^{\infty} f(x)e^{ix}dx, \\ \int_{-\infty}^{\infty} f(x)\sin x\, dx &= \mathrm{Im} \int_{-\infty}^{\infty} f(x)e^{ix}dx. \end{aligned} \tag{27}$$

Note that one cannot simply use an analytic continuation of cos or sin functions to the complex plane: these functions grow to infinity along the imaginary axis both in positive and negative directions which would make invalid the main tool in the proof of Theorem 6.2 (the smallness of the integral along the large semicircle).

Example

Consider the integral:

$$\int_{-\infty}^{\infty} \frac{\cos x}{1+x^2}dx = \mathrm{Re} \int_{-\infty}^{\infty} \frac{e^{ix}}{1+x^2}dx. \tag{28}$$

So we consider the analytic function $f(z) = \frac{e^{iz}}{1+z^2}$, which has poles $z_\pm = \pm i$, with corresponding residues $\frac{e^{iz_\pm}}{2z_\pm} = \frac{e^{-1}}{2i}, \frac{e}{-2i}$. The function f obviously decays as $|z| \to \infty$. Thus we apply Theorem 6.2 to obtain:

$$\int_{-\infty}^{\infty} \frac{e^{ix}}{1+x^2}dx = 2\pi i \left(\frac{e^{-1}}{2i}\right) = \frac{\pi}{e}. \tag{29}$$

Then we conclude:

$$\int_{-\infty}^{\infty} \frac{\cos x}{1+x^2}dx = \frac{\pi}{e}, \tag{30}$$

$$\int_{-\infty}^{\infty} \frac{\sin x}{1+x^2}dx = 0. \tag{31}$$

There is something magical in these formulas: almost out of nothing, just a few lines of simple arithmetics, one obtains absolutely non-intuitive expressions for integrals that look inapproachable, in terms of fundamental constants!

Remark

If z_0 is a simple pole then we may use the formula:

$$\operatorname{Res}(f(z)e^{iz}, z_0) = e^{iz_0}\operatorname{Res}(f(z), z_0). \tag{32}$$

However, if z_0 is not a simple pole, then this formula does not hold. Indeed, the formula $\operatorname{Res}\left(\frac{\phi}{\psi}, z_0\right) = \frac{\phi(z_0)}{\psi'(z_0)}$ give the result for simple poles only.

The next theorem deals with the integrals over the real half-line. Unlike the previous formulas, it takes into account all singularities (not just those in the uper or lower half-plane). It also makes a very effective use of a different idea: a single-valued branch of a multi-valued analytic function has a disconiuity along a branch cut.

Theorem 6.3. *Suppose that f is a real-analytic function such that its analytic continuation obeys:*

$$|f(z)z\ln z| \to 0 \qquad \textit{as} \quad |z|\to\infty \qquad \textit{and as} \quad |z| \to 0. \tag{33}$$

Then we may compute the integral over the positive *real line as:*

$$\int_0^\infty f(x)dx = -\sum_k \operatorname{Res}\left(f(z)\ln z, z_k\right) \tag{34}$$

Proof. We integrate the function $f(z)\ln z$ over the *keyhole* path shown in Figure 24. However, $\ln z$ is a multi-valued function, and so we must define a branch. We take the branch cut to be along the positive real axis, i.e. $\ln z = \ln r + i\varphi$ with $\varphi \in (0, 2\pi)$. Then we have:

$$\begin{aligned}\oint_\gamma f(z)\ln z dz = &-\int_0^{2\pi} f(\varepsilon e^{i\varphi})\ln\left(\varepsilon e^{i\varphi}\right)\varepsilon i e^{i\varphi}d\varphi \\ &+\int_\varepsilon^R f(x)\ln x dx \\ &+\int_0^{2\pi} f(Re^{i\varphi})\ln\left(Re^{i\varphi}\right)Rie^{i\varphi}d\varphi \\ &-\int_\varepsilon^R f(x)\left(\ln x + 2\pi i\right)dx \\ &= 2\pi i\sum_{\varepsilon<|z_k|<R}\operatorname{Res}\left(f(z)\ln z, z_k\right)\end{aligned}$$

Notice that the orange integral picks up an extra $2\pi i$ term due to the fact that the logarithm is evaluated on the lower side of the branch. The logarithmic terms in the purple and orange integrals cancel out, leaving only this additional term. Taking the limits $R\to\infty$ and $\varepsilon\to 0$, the green and blue integrals decay to 0 (by (33)), thus giving exactly eq. (34).

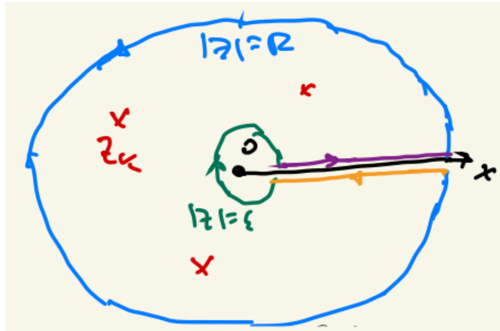


Figure 24: Keyhole contour.

□

Example

Consider the following integral:

$$\int_0^{\infty} \frac{dx}{1+x^2} = \arctan(x)|_0^{+\infty} = \frac{\pi}{2}.$$

Let us check that the Keyhole Formula (34) gives us the same result. Considering $f(z) = \frac{1}{1+z^2}$, we can compute residues:

$$\operatorname{Res}\left(\frac{\ln z}{1+z^2}, z = i\right) = \frac{\ln i}{2i} = \frac{i\pi}{2}\frac{1}{2i} = \frac{\pi}{4}$$
$$\operatorname{Res}\left(\frac{\ln z}{1+z^2}, z = -i\right) = \frac{\ln(-i)}{2i} = \frac{3i\pi}{2}\frac{1}{-2i} = -\frac{3\pi}{4}$$

Note that because of the branch we have chosen $\ln(-i) \neq -i\pi/2$, contrary to what one could easily (and mistakenly) assume. Applying Theorem 6.3 gives us:

$$\int_0^{\infty} \frac{dx}{1+x^2} = -\left[\frac{\pi}{4} - \frac{3\pi}{4}\right] = \frac{\pi}{2},$$

in accordance to Theorem 6.3.

7 Lecture 7: Principle value and half-residues

The integrals considered in the previous lecture diverge when the function f has singularities on the real line, and the formulas we derived are not applicable in this case. However, there are situations when we need to make sense of integrals of analytic functions along the paths going directly through singular points. One way to do it is to consider the so-called *principal value* of the integral.

Suppose we have a function, as shown in Figure 25. The function f goes to infinity both on the right- and left-hand side of its singularity x^*.

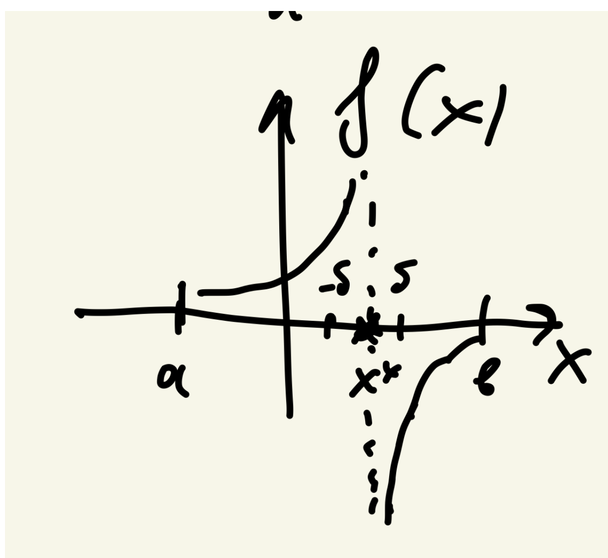


Figure 25: A function $f(x) \sim \frac{1}{x-x^*}$ near a real singular point x^*.

We see that this function diverge when $x \to (x^*)^+$ and when $x \to (x^*)^-$. Therefore we cannot, strictly speaking, integrate f from $a < x^*$ to $b > x^*$. However, we define the principal value of this integral (denoted by v.p.) as follows:

$$\text{v.p.} \int_a^b f(x)dx = \lim_{\delta\to 0}\left(\int_a^{x^*-\delta} f(x)dx + \int_{x^*+\delta}^b f(x)dx\right), \tag{35}$$

where x^* is the singularity. The idea is that we remove a small segment around x^* from the integration path such that x^* is *exactly in the middle of the segment.* When the length of the segment decreases to zero, the integral acquires a large negative contribution on the left and a large positive contribution on the right; the symmetry of the segment ensures that the absolute values of this contributions are, essentially, equal, so they may cancel each other, and the limit (35) can be finite. We explain this in more details in the following example.

Example Let $f(x) = \frac{1}{x}$, and $a < 0$, $b > 0$ are two real numbers. The singularity x^* of f is then 0. The principal value of f is:

$$\begin{aligned}
\text{v.p.} \int_a^b f(x)dx &= \lim_{\delta\to 0}\left(\int_a^{-\delta} \frac{1}{x}dx + \int_\delta^b \frac{1}{x}dx\right) \\
&= \lim_{\delta\to 0}(\ln(|x|)|_a^{-\delta} + (\ln(|x|)|_\delta^b) \\
&= \lim_{\delta\to 0}(\ln\delta - ln|a| + lnb - \ln\delta) = \ln|\frac{b}{a}|.
\end{aligned}$$

In this example, we see that $\ln\delta$ and $-\ln\delta$ (which both tend to infinity) cancel each other. Would the rates, at which the left and right integration limits tend to zero, be different, this cancellation would not happen and the integral would diverge: when computing $\lim_{\delta\to 0,\delta'\to 0}(\int_a^{-\delta}\frac{1}{x}dx + \int_{\delta'}^b \frac{1}{x}dx)$, we would have the term $\ln\delta' - \ln\delta$ which can tend to any limit depending on the ration between δ' and δ.

Next we discuss how to compute the principal value of an integral. Let f be a function of a real variable x and let f be analytic (i.e., it admits an analytic continuation to complex z with small imaginary part) except for one real singularity point at $x = x^*$. Assume $f(x)$ tends to zero sufficiently fast as $\|x\| \to \infty$. Let us integrate f along the path $\gamma = \gamma_1 \cup \gamma_2 \cup \gamma_3$ as shown in Figure 26. More precisely, $\gamma_1 := \{x \le x^* - \delta, y = 0\}$, $\gamma_2 := \{z = x^* + \delta e^{i\varphi}, \quad \varphi \in [\pi, 0]\}$, $\gamma_3 := \{x \ge x^* + \delta, y = 0\}$.

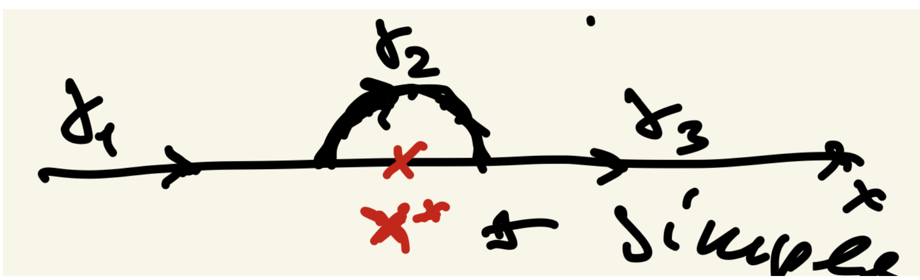


Figure 26: γ_1 and γ_3 are parts of the real line, and γ_2 is a half circle of radius δ centered at x^*, a singularity of the function f.

Theorem 7.1. *Let x^* be a simple pole of f. Then, for all small δ,*

$$\int_\gamma f(x)dx = \text{v.p.} \int_{-\infty}^{+\infty} f(x)dx - \pi i \ \mathit{Res}(f, x^*). \tag{36}$$

Proof. Since x^* is a simple pole, we can write f as follows:

$$f(z) = \frac{\text{Res}(f, x^*)}{z - x^*} + \text{ bounded terms.}$$

Thus, integrating f along γ_2 gives:

$$\int_{\gamma_2} f(z)dz \overset{\delta\to 0}{=} \text{Res}(f, x^*)\int_\pi^0 \frac{d(\delta e^{i\varphi})}{\delta e^{i\varphi}} = \text{Res}(f, x^*)\int_\pi^0 id\varphi = \pi i \ \text{Res}(f, x^*).$$

Since $f(z)$ is analytic at $x \ne x^*$, small changes in the integration path γ do not change the integral, i.e., $\int_\gamma f(x)dx$ does not depend on δ when $\delta > 0$ is small. So, by taking the limit $\delta \to +0$ in (36), we obtain the theorem, by the definition of v.p.. ☐

The integral over γ in formula (36) can be computed by methods discussed in the previous lecture. For example, if f has no singularities in the upper half-plane and tends to zero sufficiently fast as $|z| \to \infty$, then

$$\int_{\gamma_2} f(z)dz = 0,$$

which gives

$$\text{v.p.} \int_{-\infty}^{+\infty} f(x)dx = \pi i \ \text{Res}(f, x^*).$$

Example Let $f(x) = \frac{\sin x}{x}$. This function has no singularities (its Taylor series converges for all complex z), so we can write

$$\int_{-\infty}^{+\infty} f(x)dx = \text{v.p.} \int_{-\infty}^{+\infty} f(x)dx.$$

Thus,

$$\int_{-\infty}^{+\infty} \frac{\sin x}{x} dx = \text{v.p.} \int_{-\infty}^{+\infty} \frac{\sin x}{x} dx = \text{Im} \ (\text{ v.p.} \int_{-\infty}^{+\infty} \frac{e^{ix}}{x} dx).$$

Now, the function $\frac{e^{iz}}{z}$ has a simple pole at $z = 0$. It has no singularities in the upper half-plane and tends to zero fast enough as $|z| \to \infty$ in the upper half-plane (note that this is not true in the lower half-plane!). Thus, like in the proof of Theorem 6.2, we establish that $\int_\gamma \frac{e^{iz}}{z} = 0$ (where γ is the path described in Theorem 7.1). Now, by (36), we obtain

$$\text{v.p.} \int_{-\infty}^{+\infty} \frac{e^{ix}}{x} dx = \pi i \ \text{Res}(\frac{e^{iz}}{z}, 0)) = \pi i,$$

hence

$$int_{-\infty}^{+\infty} \frac{\sin x}{x} dx = \pi. \tag{37}$$

Similar arguments give us the following "half-residue" formula for an integral along a smooth closed path γ in the complex plane:

$$\text{v.p.} \oint_\gamma f(z)dz = 2\pi i \left[\sum_{z_k} \text{Res}(f, z_k) + \frac{1}{2} \sum_{z'_k} \text{Res}(f, z'_k) \right]$$

where f is a single-valued function, analytic function everywhere inside γ except for the isolated singularities z_k and z'_k, where z_k lie strictly inside of γ and z'_k are are simple poles and lie on γ.
It is important in this formula that γ is smooth, so it has a tangent at the points z'_k. If γ is piece-wise smooth, then

$$\text{v.p.} \oint_\gamma f(z)dz = 2\pi i \left[\sum_{z_k} \text{Res}(f, z_k) + \sum_{z'_k} \nu_k \text{Res}(f, z'_k) \right]$$

where the numbers ν_k are such that the arc of γ that enters the pole at z'_k makes the angle $2\pi\nu_k$ with the arc leaving z'_k (so this formula contains the previous one, as when γ is smooth, this angle equals π).

8 Lecture 8: Introduction to Fourier Transform

We start a new topic: *Fourier Transform.* We are not dealing here with functions of complex variable and the functions of a real variable we consider here do not need to be analytic. However, the tools for computing oscillating integrals studied in the previous lectures will be utilised.

Definition 8.1. Let $f(x)$ be a function defined for real x. We define the *Fourier transform* $\hat{f}$ of f to be the function defined by:

$$\hat{f}(k) = \frac{1}{2\pi}\int_{-\infty}^{\infty} f(x)e^{-ikx}dx \tag{38}$$

If the integral eq. (38) exists for (almost) all real k then we may expand f as a *Fourier integral* by means of the inverse-Fourier transform formula:

$$f(x) = \int_{-\infty}^{\infty} \hat{f}(k)e^{ikx}dk \tag{39}$$

Remark

Different conventions have different factors in front of the integrals: some definitions have $\frac{1}{2\pi}$ only in front of the inverse Fourier transform, while others have a factor of $\frac{1}{\sqrt{2\pi}}$ in front of both integrals.

Remark

If x is a spatial variable, then we may think of a function $f(x)$ as a certain field, then the Fourier integral eq. (39) is a linear expansion of the field f over linear waves e^{ikx} with amplitudes $\hat{f}(k)$ and wave numbers k. If x is time, then $f(x)$ is a certain time-dependent signal, and the Fourier integral is the expansion of this signal over harmonic oscillations with frequencies k. Note that the linear wave $y(x) = e^{ikx}$ satisfies the equation $y'' + k^2y = 0$, i.e., it is a solution of one of the most basic differential equations in nature (the equation of a harmonic oscillator). It is the cornerstone of the modern scientific worldview that natural systems are described by differential equations (Newton's invention and Britain's most important contribution to humanity!). For different systems, the corresponding differential equations can be linear or nonlinear. However, even when we have a nonlinear system, if we consider it close to an equilibrium, the oscillations near an equilibrium are small, and, since a smooth function always admits a good linear approximation at a sufficiently small scale, we can replace the system by its linearisation. As a result, small oscillations are universally described by systems of harmonic oscillators, and the corresponding solutions are linear combinations of linear waves e^{ikx}, i.e., they are naturally represented by the Fourier integral. This explains the special role played by the Fourier transform and Fourier integral all over the physics, other sciences, and engineering. In general, as we will see below, Fourier transform is an effective tool for the study of *linear differential equations with constant coefficients.*

Let us list properties of the Fourier transform:

1. Multiplication by a constant:

$$g(x) = \alpha f(x) \quad \Longleftrightarrow \quad \hat{g}(k) = \alpha\hat{f}(k). \tag{40}$$

   This follows from the linearity of the integrals (38), (39).

2. Addition of functions:

$$g(x) = f_1(x) + f_2(x) \quad \iff \quad \hat{g}(k) = \hat{f}_1(k) + \hat{f}_2(k). \tag{41}$$

Again, this follows immediately from the linearity of the integral.

3. Scaling: if $\alpha \neq 0$ then

$$g(x) = f(\alpha x) \quad \iff \quad \hat{g}(k) = \frac{1}{|\alpha|}\hat{f}\left(k/\alpha\right). \tag{42}$$

We prove this by making a change of variables ($u = \alpha x$) in the integral:

$$\begin{aligned}
\hat{g}(k) =& \frac{1}{2\pi}\int_{-\infty}^{\infty} f(\alpha x)e^{-ikx}dx \\
=& \begin{cases} \frac{1}{2\pi}\int_{-\infty}^{\infty} f(u)e^{-iku/\alpha}du/\alpha & \text{if } \alpha > 0 \\ -\frac{1}{2\pi}\int_{-\infty}^{\infty} f(u)e^{-iku/\alpha}du/\alpha & \text{if } \alpha < 0 \end{cases} \\
=& \frac{1}{|\alpha|}\frac{1}{2\pi}\int_{-\infty}^{\infty} f(u)e^{-iku/\alpha}du \\
=& \frac{1}{|\alpha|}\hat{f}(k/\alpha).
\end{aligned}$$

Note that when $\alpha < 0$, while the variable x grows from $-\infty$ to $+\infty$, the variable u decreases from $+\infty$ to $-\infty$, so we need to swap the limits on the integral, which introduces a minus sign.

4. Parity:

$$f(x) \quad \text{is even} \quad \iff \quad \hat{f}(k) \quad \text{is even.} \tag{43}$$

$$f(x) \quad \text{is odd} \quad \iff \quad \hat{f}(k) \quad \text{is odd.} \tag{44}$$

This is an immediate consequence of the scaling property with $\alpha = -1$. Recall that by even we mean $f(-x) = f(x)$ and by odd we mean $f(-x) = -f(x)$.

5.

$$f(x) \quad \text{is real} \quad \iff \quad \hat{f}(k) = \hat{f}^*(-k) \tag{45}$$

To prove this we observe that:

$$\hat{f}^*(-k) = \frac{1}{2\pi}\int_{-\infty}^{\infty} f(x)e^{+i(-k)x}dx = \hat{f}(k). \tag{46}$$

From this and property 4 we see that f is real and even if and only if $\hat{f}$ is real and even. Similarly, f is real and odd if and only if $\hat{f}$ is purely imaginary and odd.

6. Translation of argument:

$$g(x) = f(x-a) \quad \iff \quad \hat{g}(k) = e^{-ika}\hat{f}(k). \tag{47}$$

This can be shown by making the integral substitution $u = x - a$:

$$\begin{aligned}\hat{g}(k) =& \frac{1}{2\pi}\int_{-\infty}^{\infty} f(x-a)e^{-ikx}dx \\ =& \frac{1}{2\pi}\int_{-\infty}^{\infty} f(u)e^{-ik(u+a)}du \\ =& \frac{e^{-ika}}{2\pi}\int_{-\infty}^{\infty} f(u)e^{-iku}du \\ =& e^{-ika}\hat{f}(k).\end{aligned}$$

7. Multiplication by e^{iax}:

$$g(x) = e^{iax}f(x) \qquad \Longleftrightarrow \qquad \hat{g}(k) = \hat{f}(k-a). \tag{48}$$

This follows straightforwardly:

$$\begin{aligned}\hat{g}(k) =& \frac{1}{2\pi}\int_{-\infty}^{\infty} f(x)e^{-i(k-a)x}dx \\ =& \hat{f}(k-a).\end{aligned}$$

8. Derivative:

$$g(x) = f'(x) \qquad \Longrightarrow \qquad \hat{g}(k) = ik\hat{f}(k). \tag{49}$$

Thus we see that Fourier transform changes differentiation into the much simpler operation of multiplication. This can be demonstrated by simply taking the derivative with respect to x in both sides of the Fourier integral (39):

$$\begin{aligned} & f(x) = \int_{-\infty}^{\infty} \hat{f}(k)e^{ikx}dk \\ \Longrightarrow \quad & f'(x) = \int_{-\infty}^{\infty} ik\hat{f}(k)e^{ikx}dk \\ \Longrightarrow \quad & \hat{g}(k) = ik\hat{f}(k).\end{aligned}$$

Note that we do not require the differentiability of $\hat{f}$: we differentiate the Fourier integral with respect to x which means taking the derivative of e^{ikx} inside the integral, and we can do it as many times as we want, as long the resulting integral (the Fourier integral for the derivative) converges.

9. Higher derivatives:

$$g(x) = f^{(n)}(x) \qquad \Longrightarrow \qquad \hat{g}(k) = (ik)^n\hat{f}(k). \tag{50}$$

This is obtained by iterating the previous rule n times.

Properties 9 and 10 (along with the linearity properties 1 and 2) are employed to the solution of linear differential equations with constant coefficients (not that it does not work when the coefficients are functions of x !). For example, consider the following linear n^{th}-order differential equation:

$$\frac{d^n y}{dx^n} + a_1\frac{d^{n-1}y}{dx^{n-1}} + ... + a_{n-1}\frac{dy}{dx} + a_n y = f(x), \tag{51}$$

where the coefficients a_r are constant. If the Fourier transform of f exists, then we can compute the Fourier transform of eq. (51) to find:

$$\left[(ik)^n + a_1 (ik)^{n-1} + ... + a_{n-1} (ik) + a_n\right] \hat{y}(k) = \hat{f}(k) \tag{52}$$

Then we can solve for $\hat{y}$ as:

$$\hat{y}(k) = \frac{1}{P(ik)} \hat{f}(k) \tag{53}$$

where:

$$P(z) = z^n + a_1 z^{n-1} + ... + a_{n-1} z + a_n \tag{54}$$

is the *characteristic polynomial* of our differential equation (51). Now, we can use the inverse Fourier transform to find a solution y_0:

$$y_0(x) = \int_{-\infty}^{\infty} \frac{1}{P(ik)} \hat{f}(k) e^{ikx} dk \tag{55}$$

Remark

When $\hat{f}(k)$ is bounded, the integrand of eq. (55) is guaranteed to decay as $|k| \to \infty$ since P is a polynomial so $|P(ik)| \to \infty$ as $|k| \to \infty$. Therefore if $\hat{f}$ is also analytic, then we may use Theorem 6.2 to compute the integral. Problems may appear if P has zeros along the imaginary axis; then $P(ik)$ vanishes at some real k and we may have to use the principal value of the integral.

A problem with formula (55) is that it provides only one solution of eq. (51). However, any n^{th}-order ordinary differential equation has many solutions (depending on initial conditions) and its general solution includes n independent constants of integration which can take arbitrary values (corresponding to arbitrary values of the initial conditions $y(0)$, $y'(0)$, ..., $y^{(n-1)}(0)$). Thus, we take y_0 as just a particular solution, from which we may construct the general solution:

$$y(x) = C_1 e^{\lambda_1 x} + ... + C_n e^{\lambda_n x} + y_0(x) \tag{56}$$

where $\lambda_1, ..., \lambda_n$ are the roots of the characteristic polynomial P given by (54), and $C_1, ..., C_n$ are arbitrary integration constants. We assume in this formula that the characteristic polynomial P has n *distinct* roots. When there are multiple roots, the formula for general solution may acquire additional polynomial factors in front of the exponents, but it does not change the conclusions much: the general solution of eq. (51) is the sum of y_0 given by eq. (55) and the general solution of the homogeneous equation

$$\frac{d^n y}{dx^n} + a_1 \frac{d^{n-1} y}{dx^{n-1}} + ... + a_{n-1} \frac{dy}{dx} + a_n y = 0. \tag{57}$$

It is easy to see that the exponents $e^{\lambda_j x}$ in eq. (56) satisfy eq. (57), because $P(\lambda_j) = 0$.
If $\text{Re}(\lambda_j) \neq 0$ for all $j = 1, ...n$, then the exponential terms in eq. (56) tend to infinity either as $x \to +\infty$ (if $\text{Re}(\lambda_j) > 0$) or as $x \to -\infty$ (if $\text{Re}(\lambda_j) < 0$). Therefore, if we wish to find a bounded solution we must set $C_1 = C_2 = ... = C_n = 0$ in this case, so the only solution which can be bounded is y_0. Moreover, in the case $\text{Re}(\lambda_j) \neq 0$ for all j, we have that $P(ik) \neq 0$ fore real k, and the integral in eq. (55) is well-defined.
We will return to this subject in the following lectures; in particular, we will show that y_0 is the only bounded solution if f is bounded and $\text{Re}(\lambda_j) \neq 0$ for all j. But if $\text{Re}(\lambda_j) = 0$ for some j, then the integrand of eq. (55) has singulaity at real $k = i\lambda_j$, and the integral may blow up. We will discuss bounded solutions of equation (51) in this case too - after we develop a more careful theory of Fourier transform in the next lectures.

9 Lecture 9: Fourier Transform for Absolutely Integrable Functions

In the previous lecture, we have introduced the Fourier integral

$$f(x) = \int_{-\infty}^{\infty} \hat{f}(k) e^{ikx} dk, \tag{58}$$

which means that we represent the function f as an integral of linear waves, with amplitudes $\hat{f}(k)$, where k is the wave number (or the frequency of the wave). To find $\hat{f}(k)$, we define the following Fourier transform

$$\hat{f}(k) = \frac{1}{2\pi} \int_{-\infty}^{\infty} f(x) e^{-ikx} dx, \tag{59}$$

and we claim that the above defined $\hat{f}(k)$ is indeed the Fourier coefficient in (58). As we mentioned, the Fourier transform is useful in the solution of linear differential equations with constant coefficients. To actually work with it, we need to answer 2 questions: for which functions $f(x)$ the integrals (59) and (58) are properly defined, and how to compute these integrals. The second question is resolved by going through examples, where we will use, e.g., the tools we learned from complex analysis. To answer the first question, we can, for instance, assume that f decays to zero sufficiently fast, like it is shown in Figure 27 and has no singularities (or some "tractable" singularities, so that we could use the principal value). Then the integral (59) converges for all k. However, this integral may also be defined when f is only a bounded function, or even for unbounded functions. We will discuss this later. talk about this in more details later, and now proceed to one of the most basic examples.

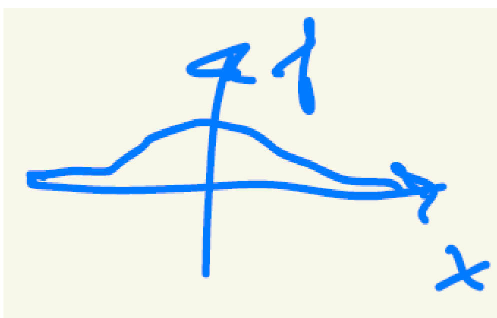


Figure 27: When $f(x)$ decays fast as $x \to \pm\infty$, the Fourier transform is well-defined.

Example Consider the *Gaussian function* defined as follows:

$$f(x) = e^{-\left(\frac{x}{a}\right)^2}. \tag{60}$$

It is easy to see that this function is even and rapidly decays to zero when $x \to \pm\infty$ (i.e., it practically vanishes for $|x| \gg a$), as shown in Figure 28. Therefore $\hat{F}(k)$ exists for any k and must be real and even.

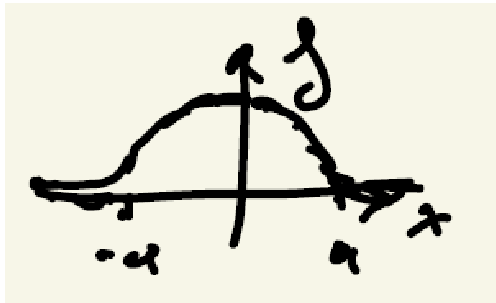


Figure 28: The Gaussian function decays to zero very fast for $|x| > a$.

Let us calculate the Fourier coefficients:

$$\hat{f}(k) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-(x/a)^2} e^{-ikx} dx = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-\frac{(x+ika^2/2)^2}{a^2} - \frac{k^2a^4}{4}} dx. \tag{61}$$

We claim that

$$\int_{-\infty}^{\infty} e^{-\frac{(x+ika^2/2)^2}{a^2}} dx = \int_{-\infty}^{\infty} e^{\frac{-x^2}{a^2}} dx.$$

Indeed, $\int_{-\infty}^{\infty} e^{-\frac{(x+ika^2/2)^2}{a^2}} dx$ is an integral of the analytic function $e^{-(z/a)^2}$ along the straight line Im $z = ka^2/2$ in the complex plane (see Fig. 29). Since there is no singularities between the lines Im $z = ka^2/2$ and Im $z = 0$, and $|e^{-(z/a)2}| = e^{-(x/a)^2}$ vanishes when $x \to \pm\infty$, using the residue formulas we can easily conclude that $\int e^{-(z/a)^2} dz$ along the path Im $z = ka^2/2$ and the path Im $z = 0$ is the same.

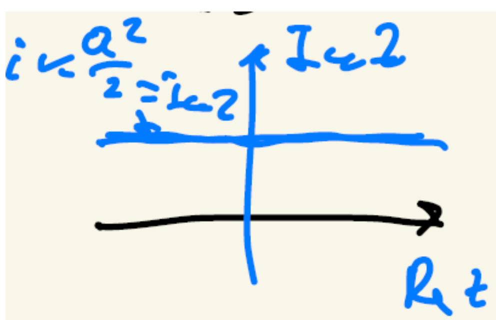


Figure 29: The integrals of $e^{-(z/a)^2}$ along the (blue) path Im $z = ka^2/2$ and along the real line (black) are the same, because e^{-z^2} has no singularities between these lines.

So, we rewrite the computation for $\hat{f}(k)$ in (61) as follows:

$$\hat{f}(k) = \frac{1}{2\pi} e^{-(ka/2)^2} \int_{-\infty}^{\infty} e^{\frac{-x^2}{a^2}} dx = \frac{1}{2\pi} e^{-(ka/2)^2} a \int_{-\infty}^{\infty} e^{-(x/a)^2} d(x/a) = \frac{a}{2\sqrt{\pi}} e^{-(ka/2)^2} \tag{62}$$

(we use here the fact that $\int_{-\infty}^{\infty} e^{-u^2} du = \sqrt{\pi}$.) We see that the Fourier transform of a Gaussian function is again a Gaussian function (of the wave number k). It practically vanishes at $|k| \gg \frac{2}{a}$, see Fig. 30. Note that when the "width" $2a$ of the Gaussian function $f(x)$ is small, the width of its Fourier transform is $\sim 4/a$ is large, and vice versa. This is an example to the uncertainty principle in quantum mechanics: We can think of the Gaussian $f(x)$ with a small a (with a proper normalisation factor) as a wave function of a particle localised in a small interval around zero. Then its Fourier transform $\hat{f}(k)$ will be the wave function in the momentum representation, and as we see, the better localisation in x is, the larger is the spread (the uncertainty) in the momenta k.

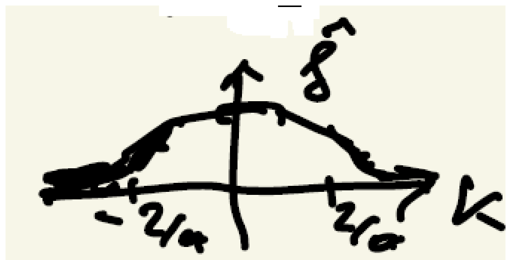


Figure 30: The Fourier transform of a Gaussian of the width $2a$ is again a Gaussian with the width $4/a$.

Let us check that we indeed recover $f(x)$ from the $\hat{f}(k)$ in (62) by means of the Fourier integral (58). Substitute (62) into (58) and get:

$$\frac{1}{2\sqrt{\pi}} a \int_{-\infty}^{\infty} e^{-(ka/2)^2+ikx} dk = \frac{1}{2\sqrt{\pi}} a \int_{-\infty}^{\infty} e^{-(ka/2-ix/a)^2-(x/a)^2} dk = \frac{1}{2\sqrt{\pi}} a e^{-(x/a)^2} \frac{2}{a} \int_{-\infty}^{\infty} e^{-(ka/2)^2} d(ka/2)$$
$$= e^{-(x/a)^2} \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} e^{-u^2} du = e^{-(x/a)^2},$$

as it should be.

In this particular example, the coefficients of the Fourier integral are indeed given by the Fourier transform. Out next goal is to show that the same is true for a sufficiently large class of functions f. A natural requirement for the function $f(x)$ is to be *absolutely integrable*, meaning that the integral of its absolute value converges:

$$\int_{-\infty}^{+\infty} |f(x)| dx < \infty.$$

It is said that f belongs to the space of absolutely integrable functions, denoted as L^1. For continuous functions, the condition $f \in L^1$ means, essentially, that f tends to zero sufficiently fast as $|x|$ grows (faster than $|x|^{-1}$, otherwise the integral would diverge).

Obviously, the condition $f \in L^1$ ensures that the integral (59) converges for every k (since $|f(x)e^{-ikx}| = |f(x)|$), so the Fourier transform $\hat{f}(k)$ is a well-defined function. In fact, we have a stronger statement:

Lemma 9.1. *If $f \in L^1$(i.e. $\int_{-\infty}^{\infty} |f(x)|dx < \infty$), then $\hat{f}(k)$ is a uniformly continuous function of k.*

Proof. We need to check whether $|\hat{f}(k_1) - \hat{f}(k_2)|$ tends to 0 as the distance between k_1 and k_2 goes to zero. the uniformity means that the rate of this convergence depends on $|k_1 - k_2|$ only. We have

$$|\hat{f}(k_1) - \hat{f}(k_2)| = \frac{1}{2\pi} \left| \int_{-\infty}^{\infty} f(x) \left(e^{-ik_1 x} - e^{-ik_2 x}\right) dx \right| \leq \frac{1}{2\pi} \int_{-\infty}^{\infty} |f(x)| \, |1 - e^{-i(k_1-k_2)x}| dx.$$

Thus, we need to show that

$$\int_{-\infty}^{\infty} |f(x)| \, |1 - e^{-i(k_1-k_2)x}| dx \to 0 \text{ as } k_1 - k_2 \to 0.$$

Figure 31: The contribution of large x to the integral of $|f|$ is small.

In order to estimate this integral, note that since $\int_{-\infty}^{+\infty} |f(x)|dx$ converges, and $|1 - e^{-i(k_1-k_2)x}| \leq 2$ is uniformly bounded for all x, and $k_{1,2}$, the contribution of large x is uniformly small. Hence, we can replace the integral over the whole real line by the integral over a sufficiently large interval and the error can be made as small as we need by taking the interval long enough. Now, on any fixed interval of integration, $e^{-i(k_1-k_2)x} \to 1$ as $k_1 - k_2 \to 0$, uniformly with respect to x. Thus, the factor $|1 - e^{-i(k_1-k_2)x}|$ goes uniformly to zero, and the integral over this interval becomes as small as we need when $k_1 - k_2$ gets small enough. So, by taking the interval of integration sufficiently long, and then taking $k_1 - k_2$ sufficiently small, we can indeed make the integral as small as we need. □

This lemma illustrates the general principle: *the large scale behaviour of the function translates into the small scale behaviour of its Fourier transform.* Here, the large scale feature of f is the existence of $\int_{-\infty}^{+\infty} |f(x)|dx$, and the small scale feature of $\hat{f}$ is its continuity. We have seen an example of this principle when analysed the Fourier transform of the Gaussian, and we will talk more about this in the next lecture.
Now we can prove the main result of this lecture.

Theorem 9.2. *Let*

$$\int_{-\infty}^{\infty} |f(x)|dx < \infty. \tag{63}$$

Compute the Fourier transform $$\hat{f}(k) = \frac{1}{2\pi} \int_{-\infty}^{\infty} f(y) e^{-iky} dy.$$

Let

$$\int_{-\infty}^{\infty} |\hat{f}(k)|dk < \infty. \tag{64}$$

Then

$$f(x) = \int_{-\infty}^{\infty} \hat{f}(k) e^{ikx} dk.$$

Proof. Since both integrals (63) and (64) are convergent, we can freely interchange integrals over x and over k in the double integrals below. So, we want to prove that

$$f(x) = \frac{1}{2\pi}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} e^{-iky} f(y) e^{ikx} dk dy = \frac{1}{2\pi}\int_{-\infty}^{+\infty} f(y)\left(\int_{-\infty}^{+\infty} e^{ik(x-y)} dk\right) dy.$$

(note that we replaced the variable of integration x by y in the formula for the Fourier transform). However, it is not clear how to interpret the integral $\displaystyle\int_{-\infty}^{+\infty} e^{ik(x-y)}dk$, so we use the following trick.
Since $\int_{-\infty}^{\infty} |\hat{f}(k)|dk < \infty$, we have that

$$\int_{-\infty}^{\infty} \hat{f}(k) e^{ikx} dk = \lim_{\varepsilon\to 0}\int_{-\infty}^{+\infty} e^{-(\varepsilon k)^2}\hat{f}(k) e^{ikx} dk.$$

Indeed, the function $e^{-\varepsilon k^2}$ is bounded (by 1), hence the integral is uniformly absolutely convergent for all ε, so the contribution of large $|k|$ to the integral is negligible for all ε, and on *any finite interval* of k the factor $e^{-\varepsilon k^2}$ tends to 1.
Thus, we need to prove that

$$f(x) = \lim_{\varepsilon\to 0}\int_{-\infty}^{+\infty} e^{-(\varepsilon k)^2}\hat{f}(k) e^{ikx} dk = \lim_{\varepsilon\to 0}\int_{-\infty}^{+\infty} f(y)\left(\frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{-(\varepsilon k)^2} e^{ik(x-y)} dk\right) dy. \qquad (65)$$

Here we apply

$$\int_{-\infty}^{+\infty} e^{-(\varepsilon k)^2} e^{iks} dk = \frac{\sqrt{\pi}}{\varepsilon} e^{-(\frac{s}{2\varepsilon})^2}$$

(by (62),(60) with $a = 2\varepsilon$, the function $e^{-(\varepsilon k)^2}$ is the Fourier transform of $\frac{\sqrt{\pi}}{\varepsilon} e^{-(\frac{s}{2\varepsilon})^2}$). We obtain

$$\lim_{\varepsilon\to 0}\int_{-\infty}^{+\infty} f(y)\left(\frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{-(\varepsilon k)^2} e^{ik(x-y)} dk\right) dy = \lim_{\varepsilon\to 0}\frac{1}{2\sqrt{\pi}\varepsilon}\int_{-\infty}^{+\infty} f(y) e^{-\frac{(x-y)^2}{4\varepsilon^2}} dy$$

Take any small $\delta > 0$ and write

$$\int_{-\infty}^{+\infty} f(y) e^{-\frac{(x-y)^2}{4\varepsilon^2}} dy = \int_{-\infty}^{x-\delta} f(y) e^{-\frac{(x-y)^2}{4\varepsilon^2}} dy + \int_{-\delta}^{+\delta} f(y) e^{-\frac{(x-y)^2}{4\varepsilon^2}} dy + \int_{x+\delta}^{+\infty} f(y) e^{-\frac{(x-y)^2}{4\varepsilon^2}} dy.$$

When $|y-x| > \delta$ we have $e^{-\frac{(x-y)^2}{4\varepsilon^2}} < e^{-\frac{\delta^2}{4\varepsilon^2}}$, which gets extremely small as $\varepsilon \to 0$. Therefore,

$$\left|\int_{-\infty}^{x-\delta} f(y) e^{-\frac{(x-y)^2}{4\varepsilon^2}} dy + \int_{x+\delta}^{+\infty} f(y) e^{-\frac{(x-y)^2}{4\varepsilon^2}} dy\right| \le e^{-\frac{\delta^2}{4\varepsilon^2}}\int_{-\infty}^{+\infty} |f(y)| dy \ll \varepsilon$$

(since $\int_{-\infty}^{+\infty} |f(y)|dy < \infty$ by the condition of the theorem), so

$$\lim_{\varepsilon\to 0}\frac{1}{2\sqrt{\pi}\varepsilon}\int_{-\infty}^{+\infty} f(y) e^{-\frac{(x-y)^2}{4\varepsilon^2}} dy = \lim_{\varepsilon\to 0}\frac{1}{2\sqrt{\pi}\varepsilon}\int_{x-\delta}^{x+\delta} f(y) e^{-\frac{(x-y)^2}{4\varepsilon^2}} dy$$

for any δ we choose.
In the same way we proved in Lemma 9.1 that the Fourier transform of an absolutely integrable function is a continuous function of k, one shows that the Fourier integral of an absolutely integrable function $\hat{f}(k)$ is continuous. Thus, condition (64) of the theorem implies that f is a uniformly continuous function. This

means that when δ is small, $f(y)$ is close to $f(x)$ for all y from the integration interval $[x-\delta, x+\delta]$, and we have

$$\lim_{\varepsilon\to 0} \frac{1}{2\sqrt{\pi}\varepsilon} \int_{x-\delta}^{x+\delta} f(y) e^{-\frac{(x-y)^2}{4\varepsilon^2}} dy \approx f(x) \lim_{\varepsilon\to 0} \frac{1}{2\sqrt{\pi}\varepsilon} \int_{x-\delta}^{x+\delta} e^{-\frac{(x-y)^2}{4\varepsilon^2}} dy,$$

where the accuracy of this approximation can be made as good as we want by taking δ small enough. Since

$$\lim_{\varepsilon\to 0} \frac{1}{2\sqrt{\pi}\varepsilon} \int_{x-\delta}^{x+\delta} e^{-\frac{(x-y)^2}{4\varepsilon^2}} dy = \lim_{\varepsilon\to 0} \frac{2\varepsilon}{2\sqrt{\pi}\varepsilon} \int_{-\delta/2\varepsilon}^{+\delta/2\varepsilon} e^{-\frac{(x-y)^2}{4\varepsilon^2}} d\frac{y-x}{2\varepsilon} = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{+\infty} e^{-s^2} ds = 1,$$

this gives us the sought formula (65).

Figure 32: The function $f(y)$ is continuous (blue graph), so in a small δ-neighbourhood of $y = x$ we can replace $f(y)$ by $f(x)$ in the integral; outside of this neighbourhood the factor $\frac{1}{\varepsilon} e^{-\frac{(x-y)^2}{4\varepsilon^2}}$ is so small (black graph), that the integration over these y values gives a negligible contribution to the total integral.

□

The theorem establishes the relation between the function and its Fourier coefficients for a sufficiently large class of functions f. However, it can still be restrictive: as we mentioned, conditions of the theorem imply, for example, that both $f(x)$ and its Fourier transform $\hat{f}(k)$ are continuous. In practice, it is quite often that we must transmit signals restricted to a certain bandwidth, i.e., $\hat{f}(k) = 0$ for all frequencies k outside a given finite interval. This is achieved by taking an arbitrary signal and filtering it by cutting off all the frequencies outside the bandwidth, i.e., by multiplying the original $\hat{f}(k)$ to a cut-off function, equal to 1 inside the allowed frequency interval and zero outside. The resulting Fourier spectrum will have discontinuities at the end points of the cut-off interval, hence the theory of such signals is not covered by the above theorem. We will return to this question and discuss how the Fourier transform theory is extended to this and other classes of functions.

10 Lecture 10: Fourier Transform: Smoothness and Decay Rate

In the previous lecture, we have already mentioned that the large scale behaviour of the function is reflected by the local behaviour of its Fourier transform and vice versa. For example, the convergence of an integral of $|f(x)|$ implies the continuity of $\hat{f}(k)$, and the convergence of an integral of $|\hat{f}(k)|$ implies the continuity of $f(x)$. This can be pushed further: as we show below, the Fourier coefficients of a sufficiently smooth L^1-function decay sufficiently fast as $|k| \to \infty$, and they depend sufficiently smoothly on k if the function decays to to zero sufficiently fast as $|x| \to \infty$.

Let $f(x)$ be a smooth function, let $\int_{-\infty}^{+\infty} |f(x)|dx < \infty$ and $\int_{-\infty}^{+\infty} |\hat{f}(k)|dk < \infty$. By Theorem 9.2,

$$f(x) = \int_{-\infty}^{+\infty} \hat{f}(k)e^{ikx}dk.$$

Since e^{ikx} is an analytic function of x, we can differentiate this integral with respect to x as long as the resulting integral (absolutely) converges. Thus, we have

$$f'(x) = \int_{-\infty}^{+\infty} \hat{f}(k)ike^{ikx}dk, \tag{66}$$

if

$$\int_{-\infty}^{+\infty} |k|\ |\hat{f}(k)|dk < \infty \tag{67}$$

(sometimes the result can hold true even without this condition, but this is condition is definitely sufficient). Since $\hat{f}$ is a continuous function of k, we find, in particular, that a sufficient condition for the existence and continuity of the derivative $f'(x)$ (defined by eq. (66))) is that the Fourier coefficients decay to zero sufficiently fast:

$$\left|\hat{f}(k)\right| = O\left(1/\left|k\right|^{2+\delta}\right), \qquad \delta > 0.$$

More generally, if

$$\int_{-\infty}^{+\infty} |k|^n|\hat{f}(k)|dk < \infty,$$

then f has n continuous derivatives and

$$f^{(n)}(x) = \int_{-\infty}^{+\infty} \hat{f}(k)(ik)^n e^{ikx}dk. \tag{68}$$

In particular, this holds if

$$\left|\hat{f}(k)\right| = O\left(1/\left|k\right|^{1+n+\delta}\right)$$

as $|k| \to \infty$, for some $\delta > 0$.

We can make the same argument with the formula

$$\hat{f}(k) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} f(x)e^{-ikx}dx$$

to obtain, for a continuous function $f(x)$, that:

$$\int_{-\infty}^{+\infty} |x|^n|f(x)|dx < \infty \quad \Longrightarrow \quad \hat{f}^{(n)}(k) \quad \text{exists and is continuous.} \tag{69}$$

This holds, for example, if

$$|f(x)| = O\left(\frac{1}{|x|^{n+1+\delta}}\right)$$

as $|x| \to \infty$, for some $\delta > 0$.

To summarise, we have shown that if $\hat{f}$ decays sufficiently fast as $|k| \to \infty$ then f has sufficiently many derivatives (and if f decays sufficiently fast as $|x| \to \infty$ then $\hat{f}$ has sufficiently many derivatives). Now we want to do the converse, assuming differentiability of f and observing the decay properties of $\hat{f}$ We start with the following lemma.

Lemma 10.1 (Riemann-Lebesgue)**.** *Suppose f satisfies*

$$\int_{-\infty}^{+\infty} |f(x)| dx < \infty.$$

Then $\hat{f}(k) \to 0$ as $|k| \to \infty$.

Proof. Here we give an outline of the proof of the lemma in the case where f is continuous. Recall the definition of the Fourier transform:

$$\hat{f}(k) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} f(x) e^{-ikx} dx$$

Since the $f \in L^1$, for sufficiently large R the tails of the integral (i.e. the contributions from regions where $|x| > R$) will be negligible. We may therefore approximate, for large R:

$$\hat{f}(k) \approx \frac{1}{2\pi} \int_{-R}^{R} f(x) e^{-ikx} dx. \tag{70}$$

Now, we may split up the remaining part of the integral into intervals of length $2\pi/k$, so:

$$\frac{1}{2\pi} \int_{-R}^{R} f(x) e^{-ikx} dx = \frac{1}{2\pi} \sum_{j=0}^{N-1} \int_{x_j}^{x_{j+1}} f(x) e^{-ikx} dx, \tag{71}$$

where $x_{j+1} = x_j + \frac{2\pi}{k}$. For large k these intervals will become very small, therefore on each interval we may use the fact that f is continuous to approximate $f(x)$ by its value at the end point of the interval:

$$\frac{1}{2\pi} \sum_{j=0}^{N-1} \int_{x_j}^{x_{j+1}} f(x) e^{-ikx} dx \approx \frac{1}{2\pi} \sum_{j=0}^{N-1} f(x_j) \int_{x_j}^{x_{j+1}} e^{-ikx} dx = 0 \tag{72}$$

(we use here that the integral of e^{-ikx} over a period is zero, and the length $2\pi/k$ of any interval $[x_j, x_{j+1}]$ is exactly the period).

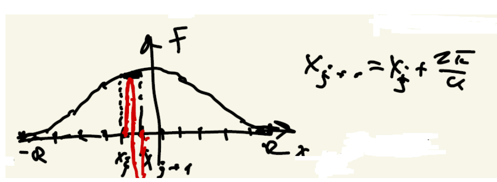


Figure 33: The integral of e^{ikx} (red graph) over a period is zero, so when f (black graph) is well-approximated by a piece-wise constant function, zero is a good approximation to the integral.

Thus, zero is a good approximation to our integral, and it gets better when the length of the intervals $[x_j, x_{j+1}]$ tends to zero, i.e., as $|k| \to \infty$. This proves the result when f is continuous, but the lemma still holds in the case where $f \in L^1$ is not continuous. The proof in the general case requires a theory of Lebesgue integral and is done by approximating f by piece-wise constant functions.

□

Now let us assume that $f(x)$ has derivative, and $f' \in L^1$, i.e., the integral of $|f'|$ exists:

$$\int_{-\infty}^{+\infty} |f'(x)| dx < \infty. \tag{73}$$

Then. we have by Lemma 10.1 that $\left|\widehat{f'}(k)\right| \to 0$ as $|k| \to \infty$, or equivalently that $\left|k\hat{f}(k)\right| \to 0$ as $|k| \to \infty$. We may also write this as

$$|\hat{f}'(k)| = o\left(\frac{1}{|k|}\right).$$

Applying the same argument to the n^{th} derivative gives:

$$\int_{-\infty}^{+\infty} |f^{(n)}(x)| dx < \infty \quad \Longrightarrow \quad |\hat{f}(k)| = o\left(\frac{1}{|k|^n}\right), \text{ as } |k| \to \infty. \tag{74}$$

Similarly, one can show that

$$\int_{-\infty}^{+\infty} |\hat{f}^{(n)}(k)| dk < \infty \quad \Longrightarrow \quad |f(x)| = o\left(\frac{1}{|x|^n}\right), \text{ as } |x| \to \infty. \tag{75}$$

Example Consider the following function:

$$f(x) = \frac{1}{1+x^2} \tag{76}$$

This function is smooth (infinitely differentiable) and all its derivatives are absolutely integrable; therefore we expect that the Fourier transform should decay very fast as $|k| \to \infty$. However, as $|x| \to \infty$ we have that $f(x) \sim 1/x^2$, i.e., it does not decay very fast and so we expect $\hat{f}$ to have only finitely many derivatives.

Figure 34: The Fourier transform $\hat{f}(k)$ is not differentiable at $k = 0$.

The Fourier transform may be calculated by Theorem 6.2 as:

$$\hat{f}(k) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \frac{e^{-ikx}}{1+x^2} dx = i \sum_{\mathrm{Im} z_j > 0} \mathrm{Res}\left(\frac{e^{-ikz}}{1+z^2}, z_j\right) \tag{77}$$

in the case $k < 0$. For $k > 0$ we have to use a different formula:

$$\hat{f}(k) = -i \sum_{\mathrm{Im} z_j < 0} \mathrm{Res}\left(\frac{e^{-ikz}}{1+z^2}, z_j\right). \tag{78}$$

The poles are given by $z_j = \pm i$. These are simple poles, so we may apply formula (12) which gives

$$\mathrm{Res}\left(\frac{e^{-ikz}}{1+z^2}, z_j\right) = \frac{e^{-ikz_j}}{2z_j}.$$

Thus:

$$\begin{aligned} \hat{f}(k) &= \begin{cases} \frac{e^{k}}{2} & \text{if} \quad k < 0 \\ \frac{e^{-k}}{2} & \text{if} \quad k > 0 \end{cases} \\ &= \frac{1}{2} e^{-|k|} \end{aligned} \tag{79}$$

Observe the following, in agreement with the theory:

- f is real and even and its Fourier transform is also real and even;
- f is smooth and its derivatives are absolutely-integrable; this should imply that $\hat{f}(k)$ decays faster than $\frac{1}{|k|^n}$ for any n, and indeed, it decays exponentially.
- $\hat{f}$ is not differentiable (the derivative does not exist at $k = 0$) which is consistent with the fact that $\int_{-\infty}^{\infty} |xf(x)|\, dx = \infty$.

11 Lecture 11: Delta function

Computations in Fourier analysis and its application become much more straightforward with the use of the so-called generalised functions. The most basic example is given by the Dirac delta-function. The idea of a delta-function is a density over a point mass or a point charge, or anything else localised at a single point. Since we have a finite mass squeezed at region of zero volume (a single point), the density at this point is infinite, and it is zero everywhere outside this point. This gives us the following formula for the delta-function (the density of a point mass at $x = 0$):

$$\delta(x) = \begin{cases} 0, & \text{if } x \neq 0, \\ +\infty, & \text{if } x = 0, \end{cases}$$

This is not specific enough, so we also add the normalization condition

$$\int_{-\infty}^{+\infty} \delta(x)dx = 1,$$

meaning that the total mass that is concentrated at $x = 0$ is equal to 1 (so the density of a point mass M is given by $M\delta(x)$).
As this formula by itself is non-sense, we need to give a meaning to it. A good way to do it is to regard the delta-function as a limit of an approximation by some functions, which decay fast to 0 as $|x|$ becomes non-zero, while the integral over the real line stays equal to 1. For instance, we can think of the delta-function as a limit of a step function supported in a small symmetric interval around zero and with height inverse proportional to the length of the interval:

$$\delta(x) = \lim_{\varepsilon \to 0} \begin{cases} \frac{1}{2\varepsilon}, & \text{if } |x| \neq \varepsilon, \\ 0, & \text{otherwise}, \end{cases}$$

see Figure 35.

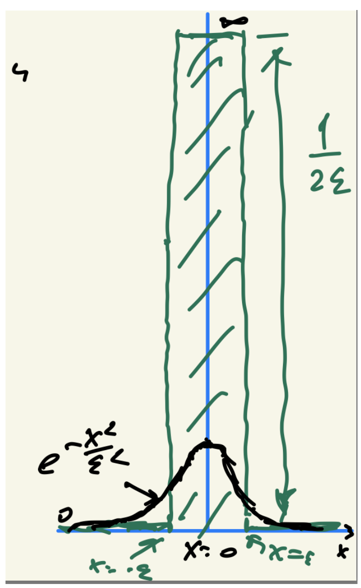


Figure 35: The image of the step function that converge to the delta-function (green). The Gaussian with a narrow support and height 1 (black) corresponds to the total mass $\varepsilon\sqrt{\pi}$, so it has to be multiplied to $(\varepsilon\sqrt{\pi})^{-1}$ to approximate the δ-function.

If we want to deal with a smooth approximating function, a popular choice is the Gaussian function with a narrow support, i.e., we may also define

$$\delta(x) = \lim_{\varepsilon \to 0} \frac{1}{\varepsilon\sqrt{\pi}} e^{\frac{-x^2}{\varepsilon^2}}. \tag{80}$$

Note that the factor in front of the Gaussian (the height) is chosen such that the integral over the real line is 1 for all ε, since

$$\int_{-\infty}^{\infty} e^{\frac{-x^2}{\varepsilon^2}} dx = \varepsilon \int_{-\infty}^{\infty} e^{-(x/\varepsilon)^2} d(x/\varepsilon) = \varepsilon\sqrt{\pi}.$$

Similarly, the delta-function concentrated on point a is an approximation of functions of total mass 1, whose height is very large and whose support is a very small interval around $x = a$:

$$\delta(x-a) = \begin{cases} 0, & \text{if } x \neq a, \\ \infty, & \text{if } x = a, \end{cases}$$

which satisfies that $\int_{-\infty}^{\infty} \delta(x)dx = 1$. Letting

$$\delta_\varepsilon(x) = \begin{cases} \frac{1}{2\varepsilon}, & \text{if } |x-a| \neq \varepsilon, \\ 0, & \text{otherwise,} \end{cases}$$

we obtain $\delta(x-a) = \lim_{\varepsilon \to 0} \delta_\varepsilon(x)$.
Now, for any continuous function f, we obtain

$$\int_{-\infty}^{\infty} f(x)\delta(x-a)dx = \lim_{\varepsilon \to 0} \int_{-\infty}^{\infty} f(x)\delta_\varepsilon(x)dx = \lim_{\varepsilon \to 0} \int_{-\varepsilon}^{\varepsilon} f(a+u)du.$$

Since f is continuous, its integral over the interval $[-\varepsilon, \varepsilon]$ can be approximated by $2\varepsilon f(a)$ when ε is small enough (see Fig. 36).

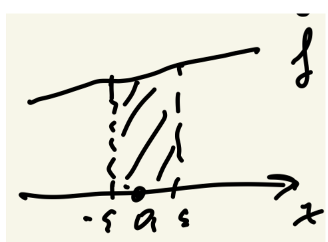


Figure 36: Integral of a continuous function f over $[-\varepsilon, \varepsilon]$ is approximately $2\varepsilon f(a)$.

This implies that

$$\int_{-\infty}^{\infty} f(x)\delta(x-a)dx = f(a).$$

Therefore, we can give an equivalent definition of the delta-function as follows: it is a linear operator, which, for every continuous functions f, returns a number $f(0)$ (hence $\delta(x-a)$ applied to a continuous function f returns $f(a)$).

Armed with this definition, we can further investigate properties of the delta-function.

1. For any $\lambda \neq 0$, we have
$$\delta(\lambda x) = \frac{1}{|\lambda|}\delta(x). \tag{81}$$
To prove this, consider $\int_{-\infty}^{\infty} f(x)\delta(\lambda x)dx$. Letting $y = \lambda x$, we obtain, for $\lambda > 0$,
$$\int_{-\infty}^{\infty} f(x)\delta(\lambda x)dx = \frac{1}{\lambda}\int_{-\infty}^{\infty} f(\frac{y}{\lambda})\delta(y)dy = \frac{1}{\lambda}f(0),$$
and, for $\lambda < 0$,
$$\int_{-\infty}^{\infty} f(x)\delta(\lambda x)dx = \frac{1}{\lambda}\int_{+\infty}^{-\infty} f(\frac{y}{\lambda})\delta(y)dy = -\frac{1}{\lambda}\int_{-\infty}^{+\infty} f(\frac{y}{\lambda})\delta(y) = -\frac{1}{\lambda}f(0)$$
(the "-" sign appears here because $y = \lambda x$ runs in the direction opposite to x when $\lambda < 0$, from $+\infty$ to $-\infty$, when x runs from $-\infty$ to $+\infty$, so we need to interchange the limits of integration). In any case, we see that the result of the integration of $f(x)$ with $\delta(\lambda x)$ is the same as $|\lambda|^{-1}$ time the result of the integration of $f(x)$ with $\delta(x)$, for any continuous function f, which gives (81).

One can also derive this formula by the comparison of the step-function approximations to $\delta(x)$ and $\delta(\lambda x)$, as shown in Figure 37. The support of the step function δ_ε which approximates $\delta(x)$ is $[-\varepsilon, \varepsilon]$, so the support of the step function $\delta_\varepsilon(\lambda x)$ is $[-\frac{\varepsilon}{|\lambda|}, \frac{\varepsilon}{|\lambda|}]$. Therefore, we need to multiply the height of the step function $\delta_\varepsilon(\lambda x)$ to $|\lambda|$ to make the integral equal to 1.

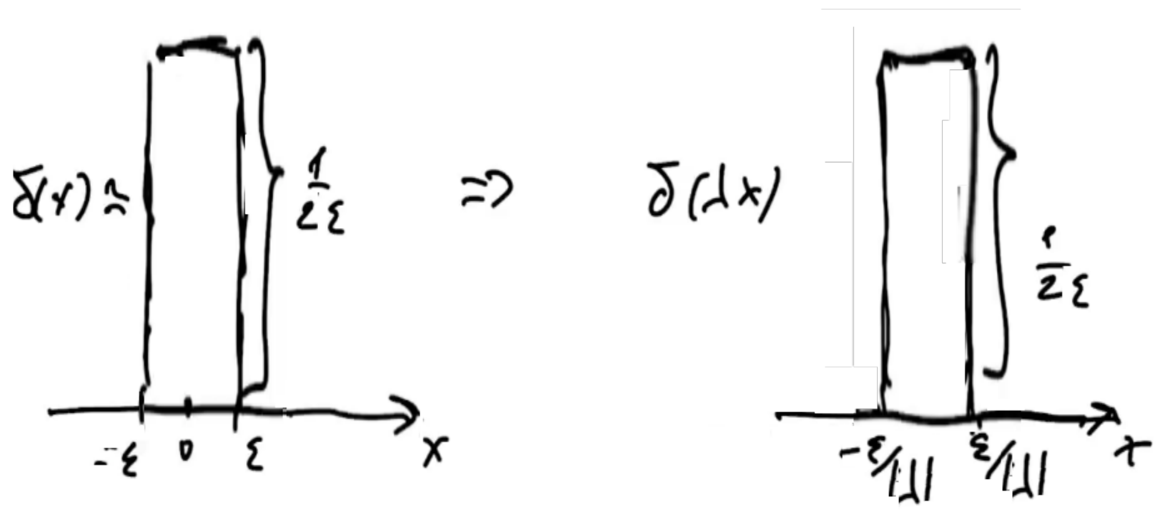


Figure 37: Step-function approximations to $\delta(x)$ (left) and to $\delta(\lambda x)$ (right).

2. Letting φ be a monotone smooth function, then
$$\delta(\varphi(x)) = \frac{1}{|\varphi'(x^*)|}\delta(x - x^*), \tag{82}$$
where x^* is the only root of φ (i.e., $\varphi(x^*) = 0$), see Fig. 38.

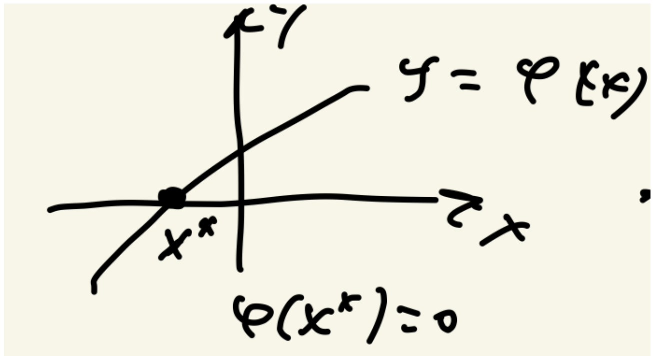


Figure 38: Image of a monotonically increasing smooth function with x^* as its root.

One simple example is when $\varphi(x) = \lambda x$, then formula (82) gives the result (81) of the first example. The proof of (82) is as follows: letting $y = \varphi(x)$, we have $dy = \varphi'(x)dx$. Therefore we get (for simplicity, suppose $\varphi' > 0$)

$$\int_{-\infty}^{\infty} f(x)\delta(\varphi(x))dx = \int_{-\infty}^{\infty} f(\varphi^{-1}(y))\delta(y)\frac{1}{\varphi'(\varphi^{-1}(y))}dy = \frac{1}{\varphi'(\varphi^{-1}(0))}f(\varphi^{-1}(0)) = \frac{1}{\varphi'(x^*)}f(x^*),$$

which means that Eq. (82) holds.

3. Letting φ be an arbitrary smooth functions with several simple roots $\{x_1, ..., x_n\}$, as shown in Figure 39,

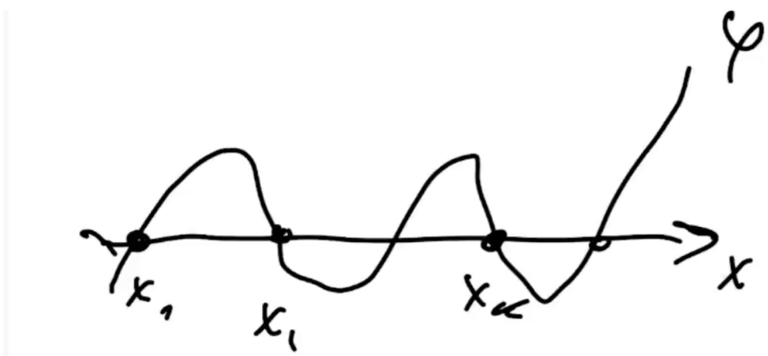


Figure 39: Image of a smooth function with several simple roots $x_1, \ldots, x_k$.

we find

$$\delta(\varphi(x)) = \sum_k \frac{1}{|\varphi'(x_k)|}\delta(x - x_k).$$

The roots x_k are assumed to be simple, meaning $\varphi'(x_k \neq 0$, so this is a well-defined formula. One justifies this formula as follows: it is clear that the result must be infinity at the points where $\varphi(x)$ is zero and zero everywhere else, so it is a linear combination of $\delta(x - x_k)$ with positive coefficients; once this is understood, the coefficient of $\delta(x - x_k)$ is found by integration against continuous functions f supported in a small neighbourhood of $x = x_k$, and since near $x = x_k$ the function φ is monotone, the resulting coefficient $|\phi'(x_k)|^{-1}$ is given by formula (82).

4. One can try to define the derivative of $\delta(x)$ as the limit of derivatives of smooth approximations to $\delta(x)$. The result is something like shown in Figure 40, which increases sharply when $x \to -0$, then rapidly decreases to become zero at $x = 0$ and further decreases to very large negative values for small $x > 0$, after which falls again very fast to zero. It is difficult to give a precise and rigorous description to such behaviour, so it is better to define $\delta'(x)$ formally.

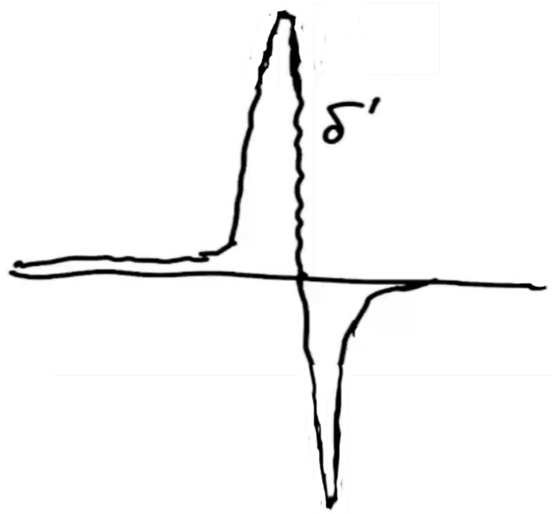


Figure 40: Image of the derivative of the delta function.

Namely, we define $\delta'(x)$ as such linear operator that for any continuously differentiable function f

$$\int_{\infty}^{\infty} f(x)\delta'(x)dx = -f'(0). \tag{83}$$

The rationale behind this definition is this. Notice that

$$\int_{\infty}^{\infty} (f(x)\delta(x))'dx = f(x)\delta(x)|_{x=\infty} - f(x)\delta(x)|_{x=-\infty} = 0$$

(because the delta-function vanishes at large positive and negative x). Thus we obtain

$$\int_{\infty}^{\infty} f(x)\delta'(x)dx = \int_{\infty}^{\infty} (f(x)\delta(x))'dx - \int_{\infty}^{\infty} f'(x)\delta(x)dx = -f'(0).$$

5. Continuing by induction, one defines the n-th derivative $\delta^{(n)}(x)$ by the rule

$$\int_{\infty}^{\infty} f(x)\delta^{(n)}(x)dx = (-1)^n f^{(n)}(0) \tag{84}$$

for any n-times continuously differentiable function f.

6. It follows immediately from the definition of the delta-function that its integral is

$$\int_{\infty}^{x} \delta(y-a)dy = \theta(x-a) = \begin{cases} 1 & \text{if } x > a, \\ 0 & \text{if } x < a. \end{cases}$$

The function θ is called the Heavyside step function; its shape is shown in Figure 41.

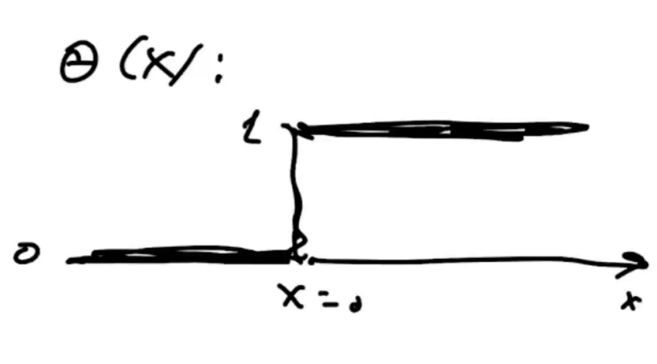


Figure 41: Image of the Heavyside step function θ.

In what follows, we consider several examples of the Fourier transform involving the delta-function.

1. Letting $F(x) = \delta(x)$, we have

$$\hat{\delta}(k) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \delta(x) e^{-ikx} dx = \frac{1}{2\pi}. \tag{85}$$

This means that if our signal $F(x)$ is a single pulse, then it's Fourier spectrum is constant over all frequencies. To demonstrate this properly, we cannot simply refer to the relation between the Fourier transform and Fourier integral established in the previous lectures, as this was proven (Theorem 9.2) only for a certain subclass of continuous functions, and the delta-function is not continuous. We, therefore, replace the delta function by its Gaussian approximation (80):

$$\delta(x) \approx \frac{1}{\varepsilon\sqrt{\pi}} e^{-\frac{x^2}{\varepsilon^2}}.$$

By formula (62) in lecture 9 (with $a = \varepsilon$), we get

$$\hat{\delta}(k) \approx G_\varepsilon(k) = \frac{1}{\varepsilon\sqrt{\pi}} \frac{\varepsilon}{2\sqrt{\pi}} e^{-\frac{(\varepsilon k)^2}{2^2}} = \frac{1}{2\pi} e^{-\frac{(\varepsilon k)^2}{4}}.$$

The result is flat on the very long interval of length $\sim \frac{2}{\varepsilon}$, as shown in Figure 42. When $\varepsilon \to 0$, the function G_ε tends to the constant value $\frac{1}{2\pi}$ on any interval of k.

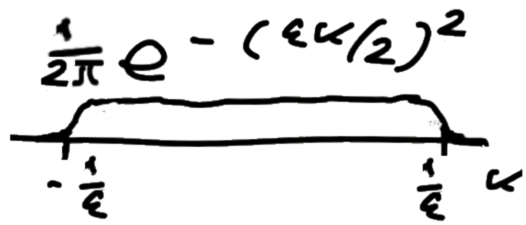


Figure 42: The approximation G_ε to the Fourier transform coefficient of δ is very close to the constant $\frac{1}{2\pi}$ on a very long interval of k.

Theorem 9.2 is applicable for the these Gaussian approximations, i.e., we have

$$\delta(x) \approx \int_{-\infty}^{\infty} G_\varepsilon(k) e^{ikx} dk.$$

Taking the limit $\varepsilon \to 0$, gives

$$\delta(x) = \int_{-\infty}^{\infty} \frac{1}{2\pi} e^{ikx} dk, \tag{86}$$

as was claimed.

2. Letting $\hat{F}(k) = \delta(k)$, we have

$$F(x) = \int_{-\infty}^{\infty} \delta(k) e^{ikx} dk = 1.$$

In other words, delta-function is the Fourier transform of $F(x) = 1$.

3. Similarly, letting $\hat{F}(k) = \delta(k-a)$, we have

$$F(x) = \int_{-\infty}^{\infty} \delta(k-a) e^{ikx} dk = e^{iax}.$$

4. Letting $F(x) = \cos ax = \frac{e^{iax}+e^{-iax}}{2}$, we obtain from the previous formula that

$$\hat{F}(k) = \frac{1}{2}\delta(k-a) + \frac{1}{2}\delta(k+a).$$

5. Letting $F(x) = \delta(x-a)$, we find

$$\hat{F}(k) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \delta(x-a) e^{-ikx} dx = \frac{1}{2\pi} e^{-iak}.$$

6. Letting $F(x) = \delta'(x)$, we obtain

$$\hat{F}(k) = ik\hat{\delta}(k) = \frac{ik}{2\pi}.$$

7. If, conversely, we let $\hat{F}(k) = \delta'(k)$, then

$$F(x) = \int_{-\infty}^{\infty} \delta'(k) e^{ikx} dk = -ix.$$

This implies that if $F(x) = x$, then $\hat{F}(k) = i\delta'(k)$.

12 Lecture 12: Fourier Transform beyond L^1.

We return to the analysis of the Fourier transform. Recall that if f is absolutely integrable, i.e. if $\int_{-\infty}^{\infty} |f(x)|\, dx < \infty$ then its Fourier transform $\hat{f}$ as given by eq. (38) exists and is continuous. If, also, $\hat{f}$ is absolutely integrable ($\int_{-\infty}^{\infty} \left|\hat{f}(k)\right| dk < \infty$) then, by Theorem 9.2, we may use the inverse Fourier transform eq. (39) to express f as a Fourier integral, and in this case f must be continuous.

However, there is a larger class of functions for which the Fourier transform makes sense. For example, as we have seen in the previous lecture, if $f(x) = $ constant, then its Fourier transform is a δ-function, but f is not absolutely integrable. We may also want to consider discontinuous functions. As we mentioned, the absolute integrability of the Fourier coefficients $\hat{f}(k)$ implies the continuity of $f(x)$, so if $f(x)$ is discontinuous, then the integral $\int_{-\infty}^{\infty} \left|\hat{f}(k)\right| dk$ must diverge. Similarly (see Lemma 9.1), when $\hat{f}(k)$ is discontinuous, then $\int_{-\infty}^{\infty} |f|(x) dx$ must diverge. Thus, in any case, we are interested in Fourier transform and Fourier integral for functions which do not decay fast at infinity and, as a result. are not absolutely integrable.

One of the approaches here is to consider the following generalisation of the Fourier integral and Fourier transform:

$$f(x) = \lim_{\varepsilon \to 0} \int_{-\infty}^{\infty} e^{-(\varepsilon k)^2} \hat{f}(k) e^{ikx} dk \tag{87a}$$

$$\hat{f}(k) = \lim_{\varepsilon \to 0} \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-(\varepsilon x)^2} \hat{f}(x) e^{-ikx} dx. \tag{87b}$$

This makes sense, as the Gaussian factor insides the integral tends to 1 as $\varepsilon \to 0$, on any finite interval of k or x. At the same time, since the Gaussian decays super-exponentially at infinity, these integrals exist for any finite ε for a very large class of functions (for example, for all bounded continuous functions), so whenever the limits at $\varepsilon \to 0$ exists in any reasonable sense, we can operate with in the same way (just with extra caution) as with the Fourier transform of absolutely integrable functions.
An important class of functions for which this definition of the Fourier transform works is the class L^2 of the *square-integrable functions*. We say that $f \in L^2$ if

$$\int_{-\infty}^{\infty} |f(x)|^2\, dx < \infty. \tag{88}$$

This is a weaker condition than $f \in L^1$ (the absolute integrability). For example, if $f(x) \sim 1/x$ at large x, it is in L^2 (since $f^2 \sim 1/x^2 \ll 1/|x|$ at large x), but f is not in L^1, since $\int \frac{dx}{x}$ is divergent as $x \to \infty$.

Theorem 12.1. *If $f \in L^2$ then $\hat{f}$ as given by eq.* (87b) *exists and is well-defined.*

We do not prove Theorem 12.1. However, we do prove the following formula, the *Parseval identity*:

$$\int_{-\infty}^{\infty} \left|\hat{f}(k)\right|^2 dk = \frac{1}{2\pi} \int_{-\infty}^{\infty} |f(x)|^2\, dx \tag{89}$$

Therefore if f is in L^2 then $\hat{f}$ is also in L^2. We show eq. (89) in the case in which f is continuous. Recall that we may deal with complex functions $\hat{f}(k)$, so $|\hat{f}(k)|^2 = \hat{f}(k)\hat{f}(k)^*$. We have

$$\int_{-\infty}^{\infty} \hat{f}(k)\hat{f}^*(k)dk = \int_{-\infty}^{\infty} \left(\frac{1}{2\pi}\int_{-\infty}^{\infty} f(y)e^{-iky}dy\right)\left(\frac{1}{2\pi}\int_{-\infty}^{\infty} f(x)^* e^{ikx}dx\right)dk =$$

$$= \frac{1}{(2\pi)^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(y)f(x)^* e^{ik(x-y)}dk\,dy\,dx = \frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(y)f(x)^*\delta(y-x)\,dy\,dx = \frac{1}{2\pi}\int_{-\infty}^{\infty} f(x)f(x)^*dx,$$

where we have used formula (86) with x replaced by $(x-y)$:

$$\int_{-\infty}^{\infty} e^{ik(x-y)}dk = 2\pi\delta(x-y). \tag{90}$$

Note that this proof works only for continuous functions $f(x)$, because the replacement of $\int_{-\infty}^{\infty} f(y)\delta(y-x)\,dy$ by $f(x)$ is valid only when f is continuous. Still the result stays true for general functions from L^2, but the argument would require more theory of function spaces (in particular, one uses the fact that every function from L^2 can be approximate by a continuous one).

The fact that the Fourier transform takes square-integrable functions of x to square-integrable functions of k is useful, e.g. in quantum mechanics. There, the Fourier transform of wave function is equivalent to the transition from coordinate to momentum representation, and it helps to know that the class of wave functions does not change when such transition is made.

The Parseval identity is a partial case of the following result: if $f, g \in L^2$, then

$$\int_{-\infty}^{\infty} \hat{f}(k)\hat{g}(k)^*dk = \frac{1}{2\pi}\int_{-\infty}^{\infty} f(x)g(x)^*dx. \tag{91}$$

If we take $g = f$, then formula (91) yields formula (89). In fact, formula (91) can be obtained from (89). Indeed, we have

$$\int_{-\infty}^{\infty} \hat{f}(k)\hat{f}(k)^*dk = \frac{1}{2\pi}\int_{-\infty}^{\infty} f(x)f(x)^*dx,$$

$$\int_{-\infty}^{\infty} \hat{g}(k)\hat{g}(k)^*dk = \frac{1}{2\pi}\int_{-\infty}^{\infty} g(x)g(x)^*dx,$$

and

$$\int_{-\infty}^{\infty} \widehat{f+g}(k)\widehat{f+g}(k)^*dk = \frac{1}{2\pi}\int_{-\infty}^{\infty} (f(x)+g(x))(f(x)+g(x))^*dx,$$

so by subtracting the first two lines from the third one, we obtain

$$\int_{-\infty}^{\infty} \hat{f}(k)\hat{g}(k)^*dk + \int_{-\infty}^{\infty} \hat{f}(k)^*\hat{g}(k)dk = \frac{1}{2\pi}\int_{-\infty}^{\infty} f(x)g(x)^*dx + \frac{1}{2\pi}\int_{-\infty}^{\infty} f(x)^*g(x)dx. \tag{92}$$

Similarly, from

$$\int_{-\infty}^{\infty} \widehat{f+ig}(k)\widehat{f+ig}(k)^*dk = \frac{1}{2\pi}\int_{-\infty}^{\infty} (f(x)+ig(x))(f(x)+ig(x))(x)^*dx,$$

we obtain

$$-\int_{-\infty}^{\infty} \hat{f}(k)\hat{g}(k)^*dk + \int_{-\infty}^{\infty} \hat{f}(k)^*\hat{g}(k)dk = -\frac{1}{2\pi}\int_{-\infty}^{\infty} f(x)g(x)^*dx + \frac{1}{2\pi}\int_{-\infty}^{\infty} f(x)^*g(x)dx, \tag{93}$$

and subtracting (93) from (92) gives (91).

In quantum mechanics one uses the bra-ket notation

$$\int_{-\infty}^{\infty} f(x)g^*(x)dx =: \langle g|f\rangle.$$

Then eq. (91) is written as

$$\langle g|f\rangle = 2\pi\langle \hat{g}|\hat{f}\rangle. \tag{94}$$

In other words, the inner product does not change when one goes from coordinate- to momentum- representation.

The L^2-theory of Fourier transform is also useful in signal transmission.

Example Suppose we know that the Fourier transform of f is given by:

$$\hat{f}(k) = \begin{cases} 1 & \text{if} \quad |k| < L, \\ 0 & \text{if} \quad |k| > L. \end{cases}$$

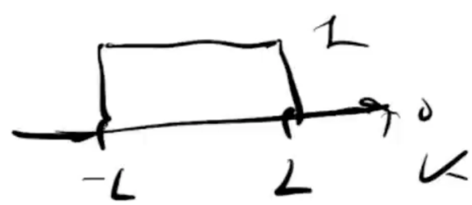


Figure 43: The discontinuous function $\hat{f}$ with finite support.

We know from the fact that $\hat{f}(k)$ is discontinuous that $f(x)$ cannot be in L^1. However, $\hat{f}(k)$ decays very fast as $|k| \to \infty$ (it is zero for $|k| > L$), therefore it is square-integrable, and, according the theory above, the corresponding function $f(x)$ is uniquely determined by the Fourier integral. We calculate:

$$\begin{aligned} f(x) &= \int_{\infty}^{\infty} \hat{f}(k)e^{ikx}dk \\ &= \int_{-L}^{L} e^{ikx}dk = \frac{e^{ikx}}{ix}\bigg|_{-L}^{L} = \frac{e^{iLx} - e^{-iLx}}{ix} \\ &= 2\frac{\sin xL}{x} \end{aligned}$$

As we see, f is not in L^1 indeed. However, it is in L^2, since $|f(x)|^2 = O(\frac{1}{x^2})$ as $|x| \to \infty$. Note that $f(x)$ is analytic (entire, in fact) function - this is in agreement with the theory discussed in Lecture 10, according to which the fast decay of the Fourier coefficients $\hat{f}(k)$ as $k \to \pm\infty$ ensures high regularity of $f(x)$.
This function is an example of a signal with a finite bandwidth - as we see, its Fourier spectrum is limited to the band $|k| < L$. It is a standard practice in radio transmission that each station has its own allocated bandwidth and should not translate signals outside of it, to avoid interference with others. This gives the finite bandwidth functions like $\frac{\sin x}{x}$ a special role.

The L^2 theory of Fourier transform is applicable to functions which decay slower than L^1 (roughly speaking, L^1 functions decay faster than $\frac{1}{|x|}$ as $|x| \to \infty$, while L^2 functions decay faster than $\frac{1}{\sqrt{|x|}}$). However, one can extend the theory even to functions which do not decay to zero at all. We will not pursue this in full generality, just consider several examples.

1. We have seen previously that if $f(x) = 1$, then it makes sense to put $\hat{f}(k) = \delta(k)$. Since $f(x)$ does not decay to zero, its Fourier transform is not a function, strictly speaking, but an operator which acts on continuous functions.

2. Take the following function

$$g(x) = \operatorname{sgn}(x) := \begin{cases} 1 & \text{if} \quad x > 0, \\ -1 & \text{if} \quad x < 0. \end{cases} \tag{95}$$

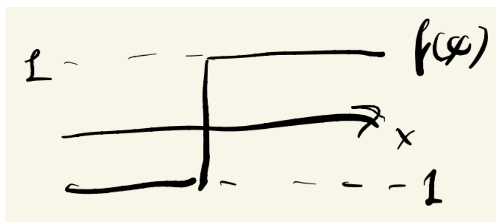


Figure 44: The sign function f is discontinuous and odd.

Can we define its Fourier transform? Before answering that, note that g is real and odd, therefore the Fourier transform must be purely imaginary and odd.

Let us discuss the Fourier integral for symmetric (even or odd) function more. Since we may expand $f(x) = \int_{-\infty}^{\infty} \hat{f}(k) e^{ikx} dk = \int_{-\infty}^{\infty} \hat{f}(k) \cos kx \, dk + i \int_{-\infty}^{\infty} \hat{f}(k) \sin kx \, dk$, we deduce the following (because the integral of any odd function over all real k is zero):

$$\begin{aligned} f \text{ even} \quad &\Longrightarrow \quad f(x) = \int_{-\infty}^{\infty} \hat{f}(k) \cos kx \, dk, \\ f \text{ odd} \quad &\Longrightarrow \quad f(x) = i \int_{-\infty}^{\infty} \hat{f}(k) \sin kx \, dk. \end{aligned}$$

For example, if we take $\hat{f}(k) = 1/k$, then it is an odd function and the corresponding Fourier integral can be found as:

$$f(x) = i\int_{-\infty}^{\infty} \frac{\sin kx}{k}\, dk \tag{96}$$

$$= \begin{cases} i\int_{-\infty}^{\infty} \frac{\sin kx}{kx}\, d(kx) & \text{if} \quad x > 0, \\ i\int_{+\infty}^{-\infty} \frac{\sin kx}{kx}\, d(kx) & \text{if} \quad x < 0, \end{cases} \tag{97}$$

$$= \begin{cases} i\int_{-\infty}^{\infty} \frac{\sin u}{u}\, du & \text{if} \quad x > 0, \\ -i\int_{-\infty}^{\infty} \frac{\sin u}{u}\, du & \text{if} \quad x < 0, \end{cases} \tag{98}$$

$$= \begin{cases} i\pi & \text{if} \quad x > 0, \\ -i\pi & \text{if} \quad x < 0, \end{cases} \tag{99}$$

$$= i\pi\,\mathrm{sgn}(x) \tag{100}$$

(we use here that $\int_{-\infty}^{+\infty} \frac{\sin u}{u} du = \pi$, see (37) in Lecture 7). So, up to a factor $i\pi$, we have found a representation of $\mathrm{sgn}(x)$ as a Fourier integral. Note that since $1/k$ has singularity at $k = 0$, we cannot immediately say that the function $\hat{f}(k) = 1/k$ is the Fourier transform of $f(x) = i\pi\,\mathrm{sgn}(x)$, because the Fourier integral $\int_{-\infty}^{\infty} \frac{1}{k} e^{ikx} dk$ is divergent at $k = 0$. It, however, converges in the principal-value sense, i.e., as the limit

$$\lim_{\varepsilon \to +0} \left(\int_{-\infty}^{-\varepsilon} + \int_{\varepsilon}^{+\infty} \right) \frac{1}{k} e^{ikx} dk,$$

and the result is exactly $i\int_{-\infty}^{\infty} \frac{\sin kx}{k}\, dk$. Therefore, we use the notation

$$\hat{f}(k) = (\text{v.p.})\frac{1}{k}$$

to stress that whenever we take an integral including $\hat{f}$, we should take only the principal value of it. Namely, we can think of $(\text{v.p.})\frac{1}{k}$ as a linear operator which, applied to a continuous function $h(k)$, returns the value of $\displaystyle\int_0^{\infty} \frac{h(k) - h(-k)}{k} dk$ (if this integral converges).
With keeping this in mind, we may now write:

$$g(x) = \mathrm{sgn}(x) \qquad \Longrightarrow \hat{g}(k) = (\text{v.p.})\frac{1}{i\pi k}. \tag{101}$$

3. Consider the Heaviside step function:

$$\theta(x) = \begin{cases} 1 & \text{if} \quad x > 0, \\ 0 & \text{if} \quad x < 0. \end{cases} \tag{102}$$

Since $\theta(x) = \frac{1}{2}(1 + \mathrm{sgn}(x))$, we immediately derive from the previous examples that the Fourier transform is

$$\hat{\theta}(k) = \frac{1}{2}\delta(k) + (\text{v.p.})\frac{1}{2\pi i k}. \tag{103}$$

We also know that $\theta'(x) = \delta(x)$, so if we take the Fourier transform of this we obtain $ik\hat{\theta}(k) = \hat{\delta}(k) = \frac{1}{2\pi}$. It is tempting then to write $\hat{\theta}(k) = \frac{1}{2\pi i k}$; however, we cannot simply multiply by $\frac{1}{ik}$ because this is not defined at $k = 0$. We must, therefore, use the principle value; there is also an additional $\delta(k)$ term in eq. (103). The message here is that we must be careful when dealing with functions which do not decay to zero and operators such as the delta function.

4. Suppose g is an L^1 function, i.e., $\int_{-\infty}^{\infty} |g(x)|\, dx < \infty$, and define f as:

$$f(x) = \int_{-\infty}^{x} g(y) dy. \tag{104}$$

Then we have $f'(x) = g(x)$ so taking the Fourier transform gives $ik\hat{f}(k) = \hat{g}(k)$. However, like in the previous example, we cannot simply divide $\hat{g}(k)$ by ik to get $\hat{f}(k)$. Note that since g is absolutely integrable, we would expect it to decay to zero as $|x| \to \infty$, but f, although bounded, need not decay. Therefore we must define properly what we mean by the Fourier transform of f. First, let us define $I := \int_{-\infty}^{\infty} g(x) dx = 2\pi\hat{g}(0)$ (the right-hand side follows from putting $k = 0$ in the Fourier transform formula $\hat{g}(k) = \frac{1}{2\pi}\int_{-\infty}^{\infty} g(x)e^{-ikx}dx$). We have $I = \lim_{x\to+\infty} f(x)$. Therefore f may be written as:

$$f(x) = I\theta(x) + h(x) \tag{105}$$

for some h such that $h(x) \to 0$ as $x \to \infty$. If the decay of $h(x)$ to zero is not anomalously slow, then we can define its Fourier transform $\hat{h}(k)$ and equation (105) gives

$$\hat{f}(k) = I\hat{\theta}(k) + \hat{h}(k), \tag{106}$$

where $\hat{\theta}(k)$ is given by (103).

Now, h satisfies $h'(x) = g(x) - I\delta(x)$ and therefore $ik\hat{h}(k) = \hat{g}(k) - \frac{I}{2\pi} = \hat{g}(k) - \hat{g}(0)$, giving

$$\hat{h}(k) = \frac{\hat{g}(k) - \hat{g}(0)}{ik}. \tag{107}$$

We know that since g is absolutely integrable, its Fourier transform $\hat{g}(k)$ is continuous (Lemma 9.1), hence $\hat{g}(k) - \hat{g}(0) \to 0$ as $k \to 0$. If the rate with which $\hat{g}(k) \to \hat{g}(0)$ as $k \to 0$ is not extremely slow, then $\hat{h}$ either is finite at $k = 0$, or even if it tends to infinity as $k \to 0$, this singularity is integrable, so formula (107) makes sense. We can now conclude that

$$\begin{aligned}\hat{f}(k) &= 2\pi\hat{g}(0)\left(\frac{1}{2}\delta(k) + (\text{v.p.})\frac{1}{2\pi ik}\right) + \frac{\hat{g}(k) - \hat{g}(0)}{ik} \\ &= \pi\hat{g}(0)\delta(k) + (\text{v.p.})\frac{\hat{g}(k)}{ik}\end{aligned} \tag{108}$$

In general, given a bounded function $f(x)$ such that $\lim_{x\to\infty} f(x) = I_+$ and $\lim_{x\to-\infty} f(x) = I_-$, we can consider the function $h(x) = f(x) - I_+\theta(x) - I_-\theta(-x)$. Then h decays as $x \to \pm\infty$, and if we can define its Fourier transform, then using formula (103) for the Fourier transform of θ, we obtain the Fourier transform of f.

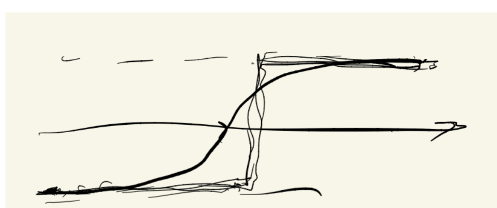


Figure 45: If f is bounded with well-defined limits at $\pm\infty$ then we may define $h = f(x) - I_+\theta(x) - I_-\theta(-x)$ to obtain the Fourier transform of f.

13 Lecture 13: Green Function

Let us define the *convolution* of a function f with a function g as:

$$(f * g)(x) := \int_{-\infty}^{\infty} f(y)g(x-y)dy = \int_{-\infty}^{\infty} f(x-y)g(y)dy \tag{109}$$

Note that this formula suggests that the involution is a commutative operation: $f * g = g * f$; we will indeed show the commutativity in a moment.
Let us prove the following:

$$\hat{h}(k) = \hat{f}(k)\hat{g}(k) \quad \Longleftrightarrow \quad h(x) = \frac{1}{2\pi}(f * g)(x), \tag{110}$$

i.e., the Fourier transform of a convolution of two functions is the product of their Fourier transforms, up to a constant factor.Indeed, if $\hat{h}(k) = \hat{f}(k)\hat{g}(k)$, then

$$\begin{aligned}
h(x) &= \int_{-\infty}^{\infty} \hat{f}(k)\hat{g}(k)e^{ikx}dk \\
&= \int_{-\infty}^{\infty} \left(\frac{1}{2\pi}\int_{-\infty}^{\infty} f(y)e^{-iky}dy\right)\hat{g}(k)e^{ikx}dk \\
&= \frac{1}{2\pi}\int_{-\infty}^{\infty} f(y)\left(\int_{-\infty}^{\infty} \hat{g}(k)e^{ik(x-y)}dk\right)dy \\
&= \frac{1}{2\pi}\int_{-\infty}^{\infty} f(y)g(x-y)dy \\
&= (f * g)(x).
\end{aligned}$$

Note that it follows that since the multiplication is a commutative operation, so the convolution must also be commutative.

Now we return to the solution of differential equations. Consider the linear ordinary differential equation with constant coefficients:

$$y^{(n)}(x) + a_1 y^{(n-1)}(x) + ... + a_{n-1}y'(x) + a_n y(x) = f(x), \tag{111}$$

We, on a formal level, solved this in Lecture 8 by taking the Fourier transform and using the property that Fourier transform exchanges derivatives with polynomial multiplication: $\widehat{y^{(n)}} = (ik)^n \hat{y}(k)$. We then obtained

$$P(ik)\hat{y}(k) = \hat{f}(k),$$

where P is the characteristic polynomial corresponding to equation (111), namely $P(\lambda) = \lambda^n + a_1\lambda^{n-1} + ... + a_n$. Then we may obtain a solution as

$$\hat{y}(k) = \frac{\hat{f}(k)}{P(ik)} \quad \Longrightarrow \quad y(x) = (f * G)(x), \tag{112}$$

where

$$\hat{G}(k) := \frac{1}{2\pi}\frac{1}{P(ik)}. \tag{113}$$

Note that solution (112) does not require the existence of the Fourier transform of f, so the right-hand side of (112) is applicable for a larger class of functions f: we only require that the integral $\int_{-\infty}^{\infty} f(s)G(x-s)ds$ exists.
The function $G(x)$ is referred to as the *Green function*. Since the Green function G satisfies

$$P(ik)\hat{G}(k) = \hat{\delta}(k), \tag{114}$$

it follows that $G(x)$ is a solution to the following ODE:

$$y^{(n)}(x) + a_1 y^{(n-1)}(x) + ... + a_{n-1}y'(x) + a_n y(x) = \delta(x). \tag{115}$$

One says that $G(x)$ is the *response to the point source* $\delta(x)$.
Once G that satisfies eq. (115) is found, one obtains the general solution to eq. (111) as

$$y(x) = \int_{-\infty}^{\infty} f(s)G(x-s)ds + y_0(x) \tag{116}$$

where $y_0(x)$ is the general solution to the homogeneous equation:

$$y^{(n)}(x) + a_1 y^{(n-1)}(x) + ... + a_{n-1}y'(x) + a_n y(x) = 0. \tag{117}$$

If λ is a root of P, i.e. $P(\lambda) = 0$, it is easy to see that $e^{\lambda x}$ is a solution to eq. (117) (if λ is a multiple root, i.e., $P(\lambda) = 0$ and $P'(\lambda) = 0$, then other solutions may arise, consisting of terms like $x^r e^{\lambda x}$). Thus, if the characteristic polynomial P has n distinct roots $\lambda_1, ..., \lambda_n$, then the general solution to the homogeneous equation is $y_0(x) = c_1 e^{\lambda_1 x} + ... + c_n e^{\lambda_n x}$ where $c_1, ..., c_n$ are arbitrary constants (and in the case of multiple roots polynomial factor may appear in front of the exponents here).

Formula (116) gives a general solution to eq. (111), where no reference to Fourier analysis is, in fact necessary. Note that there are infinitely many solution candidates for the Green function G because we may add to G any solution of eq. (117) and the resulting function will still satisfy eq. (111). A way to restrict the freedom in the choice of G is to require that it is uniformly bounded for all x. This can indeed be done when the characteristic polynomial $P(\lambda)$ has no purely imaginary roots, so $P(ik) \neq 0$ for all real values of k, and the integral in this formula

$$G(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{e^{ikx}}{P(ik)} dk \tag{118}$$

is well defined. If the equation is of order $n \geq 2$, then $\frac{1}{P(ik)} = O\left(\frac{1}{|k|^n}\right)$ is absolutely integrable, so $G(x)$ is uniformly bounded. The same is also true when $n = 1$, as we will see in the example below.
In fact, if $P(\lambda)$ has no purely imaginary roots, this bounded function $G(x)$ also decays exponentially as $x \to \pm\infty$. To see this, we, first, consider the case $n = 1$.

Example Consider the following ODE with $n = 1$:

$$y' + \alpha y = \delta(x). \tag{119}$$

The characteristic polynomial is $P(\lambda) = \lambda + \alpha$. Its only root is $\lambda = -\alpha$; it is real, so the requirement it has anon-zero real part means $\alpha \neq 0$. Let $\alpha > 0$ (the case $\alpha < 0$ is treated in the same way with the replacement $x \to -x$). If $x \neq 0$, then the delta-function in the right-hand side vanishes and y is a solution of the homogeneous equation

$$y' + \alpha y = 0. \tag{120}$$

So the solution for $x < 0$ is $y = Ae^{-\alpha x}$ and, for $x > 0$, the solution is $y = Be^{-\alpha x}$. However, we need not have $A = B$ because there is a discontinuity at $x = 0$. Indeed, if we integrate both sides of eq. (119) over the interval $(-\varepsilon, \varepsilon)$, we get

$$\begin{aligned} \int_{-\varepsilon}^{\varepsilon} y'(x)dx + \alpha \int_{-\varepsilon}^{\varepsilon} y(x)dx = 1 \\ \implies y(\varepsilon) - y(-\varepsilon) + \alpha \int_{-\varepsilon}^{\varepsilon} y(x)dx = 1 \\ \implies \quad Be^{-\alpha\varepsilon} - Ae^{\alpha\varepsilon} + O(\varepsilon) = 1 \\ \implies \quad B - A = 1, \end{aligned}$$

where in the final step we have taken the limit $\varepsilon \to 0$. However, we also want our solution to be bounded as $x \to \pm\infty$. In particular, for $x < 0$ we have $y = Ae^{-\alpha x}$ which does not decay as $x \to -\infty$; therefore we must set $A = 0$, from which we obtain $B = 1$. Thus we have the discontinuous solution as shown in Figure 46.

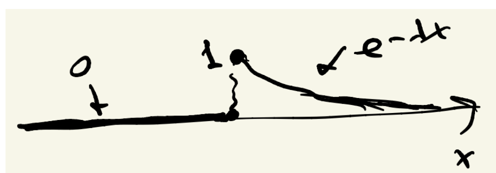


Figure 46: The bounded solution to eq. (119) is discontinuous: $y = 0$ for $x < 0$ and $y = e^{-\alpha x}$ for $x > 0$.

Like in this example, in the general case the Green function is obtained by gluing at $x = 0$ a pair of solutions to the homogeneous equation (117), one at $x > 0$, the other at $x < 0$. These solutions are of the form $c_1e^{\lambda_1 x} + ... + c_ne^{\lambda_n x}$, where $P(\lambda_j) = 0$. If $\text{Re}(\lambda_j) > 0$, then $e^{\lambda_j x} \to \infty$ as $x \to \infty$, but $e^{\lambda_j x} \to 0$ as $x \to -\infty$. And if $\text{Re}(\lambda_j) < 0$ then $e^{\lambda_j x} \to 0$ as $x \to \infty$, but $e^{\lambda_j x} \to \infty$ as $x \to -\infty$. Thus, when all roots λ_j have non-zero real parts, to have a bounded Green function, we must have

$$G(x) = \begin{cases} \sum_{\text{Re}(\lambda_j)>0} c_je^{\lambda_j x} & \text{if} \quad x < 0 \\ \sum_{\text{Re}(\lambda_j)<0} c_je^{\lambda_j x} & \text{if} \quad x > 0 \end{cases} \tag{121}$$

(in case some of λ_j is a multiple root, the coefficient c_j can to be replaced in this formula by a polynomial of degree equal to the multiplicity of the root, but the overall procedure remains the same). Coefficients c_j can be found by the jump condition: in equation eq. (115) we assume that $y, ..., y^{(n-2)}$ are continuous, so that the jump - equal to $1 = \int_{-\infty}^{\infty} \delta(x)dx$ - occurs in $y^{(n-1)}$. The condition of the continuity of $y, ..., y^{(n-2)}$ and the jump-1 condition on $y^{(n-1)}$ defines the coefficients c_j uniquely. Indeed, we saw that the bounded solution $G(x)$ always exists (defined by (118)), and it is also unique: if we have two different bounded solutions to eq. (115), then their difference must also be bounded, but this difference is a solution to the homogeneous equation (117), hence it is a sum of terms containing $e^{\lambda_j x}$ with $\text{Re}\lambda_j \neq 0$, hence it must go to infinity either at $x \to +\infty$ or $x \to -\infty$, a contradiction.

Example Consider the equation

$$y'' - y = \delta(x). \tag{122}$$

For this case the characteristic polynomial is $P(\lambda) = \lambda^2 - 1$ so the roots are $\lambda = \pm 1$. For $x \neq 0$ we have the solution $y = Ce^x + De^{-x}$. Therefore in order to obtain a bounded solution we take:

$$y = \begin{cases} C_1 e^x & \text{if} \quad x < 0 \\ C_2 e^{-x} & \text{if} \quad x > 0 \end{cases} \tag{123}$$

Then we have a jump condition on y' and a continuity condition on y:

$$y'(+0) - y'(-0) = 1,$$
$$y(+0) = y(-0),$$

where we have used the notation $y(\pm 0) := \lim_{\varepsilon \to +0} y(\pm\varepsilon)$. The second condition gives $C_1 = C_2$, and the first condition transforms to $C_1 + C_2 = 1$, so $C_1 = C_2 = \frac{1}{2}$, with the resulting solution being given as:

$$y = \frac{1}{2} e^{-|x|} \tag{124}$$

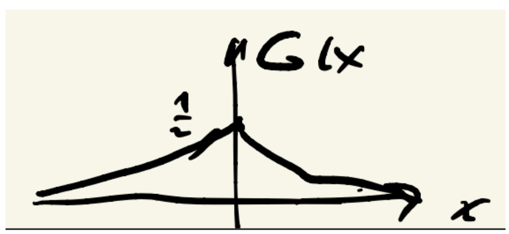


Figure 47: The bounded solution $G(x)$ to $y'' - y = \delta(x)$; the function $G(x)$ is continuous but the derivative $G'(x)$ is not.

In summary, if the characteristic polynomial P has no purely imaginary roots, the (bounded) Green function decays exponentially, hence it is absolutely integrable, which implies that the solution

$$y(x) = \int_{-\infty}^{\infty} f(s)G(x-s)ds = \int_{-\infty}^{\infty} f(x-s)G(s)ds \tag{125}$$

is uniformly bounded for all x whenever the function f (the right-hand side of eq. (111)) is uniformly bounded for all x. Moreover, it is the only bounded solution of eq. (111) in this case (since y_0 in eq. (116) - the general solution to the homogeneous equation (117) - is unbounded unless it is zero, as we explained).

But what if P has a purely imaginary root? This means $P(ik) = 0$ for some real k, leading to 2 problems: the integral in the formula (118) for the Green function becomes divergent, and the bounded solution to the corresponding equation (115) is not unique (or may even not exist).
Consider the following simple example:

$$y' = \delta(x) \tag{126}$$

The characteristic polynomial is $P(\lambda) = \lambda$, so $\lambda = 0$ is a root with the zero real part. Obviously, the solution to eq. (126) is $y(x) = \theta(x) + C$, i.e., all the solutions are bounded and we can take any of them as a Green function. After that one can apply formula (116), which returns

$$y(x) = \int f(x)dx + c_0$$

as a general solution to the equation

$$y' = f(x).$$

Note that this solution does not need to be bounded for a general bounded function $f(x)$ - one needs $\int_{-\infty}^{\infty} f(x)dx = 0$ for that. The same condition can be expressed as $\hat{f}(0) = 0$. Using formula (108) from the previous lecture, we can write the Fourier transform of the solution $y(x)$ as

$$\hat{y} = \underbrace{(\text{v.p})\frac{\hat{f}}{ik}}_{\text{partial solution}} + \underbrace{C\delta(k)}_{\substack{\text{general solution of} \\ \text{homogeneous system}}} \tag{127}$$

This picture easily extends to the general case of eq. (111) where we allow $P(\lambda)$ to have some purely imaginary roots $ik_1, ..., ik_m$ *which are all simple*. In this case,

$$\hat{G}(k) = \frac{1}{2\pi}(\text{v.p.})\frac{1}{P(ik)} \tag{128}$$

and the corresponding Green function

$$G(x) = \frac{1}{2\pi}(\text{v.p.})\int_{-\infty}^{\infty} \frac{1}{P(ik)} e^{ikx} dk$$

is bounded. Moreover, the bounded solutions of eq. (111) are given by the formula

$$y(x) = \int_{-\infty}^{\infty} f(s)G(x-s)ds + \sum_{j=1}^{m} C_j e^{ik_j x} \tag{129}$$

for arbitrary constants C_j, provided $\hat{f}(k_j) = 0$ for all j (meaning the external signal f should not be in resonance with the eigenfrequencies $k_1, \ldots, k_m$).

When the characteristic polynomial has *multiple roots on the imaginary axis*, the use of the Fourier transform method to find the Green function becomes more cumbersome, and it can be easier (and safer) to look for $G(x)$ by solving eq. (115) directly.

Example Consider the second-order equation:

$$y'' = f(x) \tag{130}$$

In this case we have $P(ik) = (ik)^2$ so we have a double root at $k = 0$. We find the Green function, i.e. the solution to $G''(x) = \delta(x)$ as:

$$G'(x) = \theta(x) + C \tag{131}$$

$$\implies \quad G(x) = Cx + B + x\theta(x) \tag{132}$$

Therefore there is no bounded solution. We may consider using eq. (112) to write the Fourier transform of the solution y as:

$$\hat{y}(k) = \frac{\hat{f}(k)}{-k^2} \tag{133}$$

However, the right-hand side of eq. (133) is not integrable in a principle value sense and therefore the use of the Fourier transform method becomes questionable. However, we can find the (unbounded) Green function, as we did to get eq. (132), without using the Fourier transform. Then suppose we choose $C = B = 0$ so that $G(x) = x\theta(x)$. The solution to eq. (130) is then

$$y(x) = \int_{-\infty}^{\infty} (x - s)\theta(x - s)f(s)ds + Cx + B \tag{134}$$

$$= \int_{0}^{\infty} sf(x - s)ds + Cx + B. \tag{135}$$

14 Lecture 14: Euler-Lagrange Equation

In this lecture we begin a new topic, the Calculus of Variations. Our main focus here is the Euler-Lagrange equation

$$\frac{d}{dx}\frac{\partial L}{\partial y'} = \frac{\partial L}{\partial y}.$$

It was discovered by Euler and Lagrange in the middle of eighteenth century as a solution to optimisation problems. Later, Lagrange showed that Newtonian mechanics can be formulated in similar terms. Remarkably, the two and half centuries of the development of science after that only confirmed that all fundamental physical theories can be expressed in terms of the Euler-Lagrange equations. As it stands now, the Lagrangian picture of the world is the central principle of Physics.

We start with the following optimisation problem: Given a function $L(y, v, x)$ (the Lagrangian) which depends on the three variables y, v, x, we substitute the first two arguments with the functions $y = y(x)$, and $v = y'(x)$, and ask which function $y(x)$ will give a minimal or maximal value to the functional

$$\int_a^b L(y(x), y'(x), x)dx, \tag{136}$$

subject to the condition that the values of y at the end points of the integration interval $[a, b]$ are fixed, $y(a) = y_a$, $y(b) = y_b$. It was discovered by Euler and Lagrange that this problem is reduced to the solution 0f a certain second-order differential equation.

Theorem 14.1. *If a twice continuously differentiable function $y(x)$ is a minimiser (or maximiser) of (136), then it must satisfy the equation*

$$\frac{d}{dx}\frac{\partial L}{\partial v}(y(x), y'(x), x) = \frac{\partial L}{\partial y}(y(x), y'(x), x). \tag{137}$$

Note that $\frac{d}{dx}$ here is not a partial derivative, i.e., by chain rule we have

$$\frac{d}{dx}\frac{\partial L}{\partial v}(y, y', x) = \frac{\partial^2 L}{\partial y \partial v}y' + \frac{\partial^2 L}{\partial v^2}y'' + \frac{\partial^2 L}{\partial v \partial x}.$$

Thus, the second derivative of y effectively enters equation (137), i.e., it is a second-order differential equation. The significance of this theorem is obvious. Instead of performing the monstrous task of computing the integral (136) for *all* legitimate functions $y(x)$ and then choosing the one that gives the optimal value to it, we only need to look among the solutions to the differential equations (137). It is a second-order differential equation; it is nonlinear, so no general formulas for its solution exist, however in many cases such formulas can be found (otherwise we can look for solutions numerically). A general solution to a second-order equation depends on two integration constants, whose values can be fixed to satisfy the boundary conditions $y(a) = y_a$, $y(b) = y_b$. This usually gives a small finite set of solutions - we then can compute the integral (136) on these functions only, after which we figure out which of them produces the sought optimum.

Let us prove Theorem 14.1.

Proof. If the function $y(x)$ optimises the integral (136), then for any perturbation function $\phi(x)$ the following function

$$F(\varepsilon) = \int_a^b L\left(y + \varepsilon\phi(x), y' + \varepsilon\phi'(x), x\right) dx$$

has an extremum at $\varepsilon = 0$. Thus, if y is a minimiser or maximiser, then $F'(\varepsilon) = 0$ at $\varepsilon = 0$.
We have

$$\frac{d}{d\varepsilon}F(\varepsilon)\bigg|_{\varepsilon=0} = \left(\int_a^b \frac{d}{d\varepsilon} L\left(y + \varepsilon\phi(x), y' + \varepsilon\phi'(x), x\right) dx\right)_{\varepsilon=0} = \int_a^b \left[\frac{\partial L}{\partial y}(y, y', x)\ \phi(x) + \frac{\partial L}{\partial v}(y, y', x)\ \phi'(x)\right] dx$$

$$= \int_{x=a}^{x=b} \frac{\partial L}{\partial y}\ \phi dx + \int_{x=b}^{x=a} \frac{\partial L}{\partial v} d\phi(x) = \int_a^b \frac{\partial L}{\partial y}\ \phi dx + \underbrace{\left(\overset{=0}{\frac{\partial L}{\partial v}\phi|_a^b} - \int_a^b \phi(x)\frac{d}{dx}\frac{\partial L}{\partial v}dx\right)}_{\text{integration by parts}}$$

$$= \int_a^b \left[\frac{\partial L}{\partial y}(y(x), y'(x), x) - \frac{d}{dx}\frac{\partial L}{\partial v}(y(x), y'(x), x)\right]\phi(x)dx$$

for any choice of ϕ such that $\phi(a) = \phi(b) = 0$ (this is a natural choice, as we keep the boundary conditions y_a at $x = a$ and y_b at $x = x_b$ the same for all ε).
Thus, the condition $F'(0) = 0$ recasts as

$$\int_a^b \left[\frac{\partial L}{\partial y}(y(x), y'(x), x) - \frac{d}{dx}\frac{\partial L}{\partial v}(y(x), y'(x), x)\right]\phi(x)dx = 0.$$

It is obvious that we can have this for all possible choices of the function ϕ only when the factor inside the integral vanishes identically, i.e., when the function $y(x)$ satisfies the Euler-Lagrange equation (137). In order to formally prove it (thus finishing the proof of the theorem), we assume that

$$\frac{\partial L}{\partial y}(y(x), y'(x), x) - \frac{d}{dx}\frac{\partial L}{\partial v}(y(x), y'(x), x) \neq 0$$

at some point $x = s \in (a, b)$. By continuity, it is non-zero for all x close to s, so if we take $\phi(x)$ to be close to the Dirac δ-function at $x = s$, i.e., $\phi(x) \approx \delta(x - s)$, then

$$\int_a^b \left[\frac{\partial L}{\partial y}(y(x), y'(x), x) - \frac{d}{dx}\frac{\partial L}{\partial v}(y(x), y'(x), x)\right]\phi(x)dx \approx \frac{\partial L}{\partial y}(y(s), y'(s), s) - \frac{d}{dx}\frac{\partial L}{\partial v}(y(s), y'(s), s) \neq 0.$$

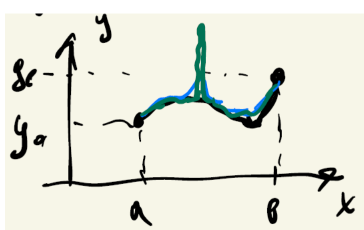


Figure 48: One can choose $\phi \approx \delta(x - s)$ for any $s \in (a, b)$.

□

Example Consider the minimisation problem

$$\int_0^1 \left((y')^2 + y^2\right) dx \to \min,$$

subject to boundary conditions $y(0) = 1$, $y(1) = 1$. We first differentiate $L = y'^2 + y^2$ and obtain

$$\frac{\partial L}{\partial y} = 2y, \qquad \frac{\partial L}{\partial y'} = 2y'.$$

Then, the Euler-Lagrange equation is

$$\frac{d}{dx}\frac{\partial L}{\partial y'} = 2y'' = \frac{\partial L}{\partial y} = 2y,$$

we obtain the equation

$$y'' = y.$$

The general solution is

$$y(x) = C_1 e^x + C_2 e^{-x}.$$

We may use the boundary conditions for $y(x)$ to conclude that $C_1 + C_2 = 1$ and $C_1 e + C_2 e^{-1} = 1$, hence

$$C_1 = \frac{1}{1+e}, \qquad C_2 = \frac{e}{1+e}.$$

Thus, the minimiser is

$$y(x) = \frac{1}{1+e}(e^x + e^{1-x}).$$

An important property of the Euler-Lagrange equation is the Law of Conservation of Energy.

Theorem 14.2. *If Lagrangian $L(y, y')$ does not depend on x, then the Hamiltonian function (or energy function)*

$$H(y, y') := -L(y, y') + \frac{\partial L}{\partial y'}(y, y')y'$$

stays constant on any solution of the Euler-Lagrange equation.

Proof. To show that $H(y(x), y'(x))$ is a constant, it is suffice to show that $\frac{d}{dx}H = 0$.

$$\begin{aligned}
\frac{d}{dx}H(y(x), y'(x)) &= -\frac{d}{dx}\left[L - \frac{\partial L}{\partial y'}y'\right] \\
&= -\frac{\partial L}{\partial y}y' - \frac{\partial L}{\partial y'}y'' + (\frac{d}{dx}\frac{\partial L}{\partial y'})y' + \frac{\partial L}{\partial y'}y'' \\
&= -\left(\frac{\partial L}{\partial y} - \frac{d}{dx}\frac{\partial L}{\partial y'}\right)y' \\
&= 0
\end{aligned}$$

by the Euler-Langrange equation. □

The law of conservation of energy allows to replace the second-order Euler-Lagrange equation by the first-order equation $H(y, y') = C$. It is a first order, as it involves only y and y', but not the second derivative y''. Usually, first-order differential equations are simpler than second-order ones.

Example We take the same Langrangian as before, i.e., $L(y, y') = y^2 + y'^2$ and we know from the previous example that the Euler-Langrange equation is the 2nd order differential equation $y'' = y$. Since $\frac{\partial L}{\partial y'} = 2y'$, we obtain the following Hamiltonian:

$$H = \frac{\partial L}{\partial y'} y' - L = 2y'^2 - L = (y')^2 - y^2.$$

By Theorem (14.2), we can replace the Euler-Lagrange equation by

$$(y')^2 = y^2 + C.$$

Even though it is a first-order equation, it does not look simpler than the Euler-Lagrange equation $y'' = y$, so maybe this is not a very good example. However, let us check directly that H is the conserved quantity here. Indeed,

$$\frac{d}{dx}\left((y')^2 - y^2\right) = 2y'y'' - 2yy' = 0$$

if $y'' = y$.

Let us look at a more interesting example. Consider the following Lagrangian

$$L(y, y') = y\sqrt{1 + y'^2},$$

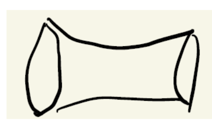


Figure 49: The surface of revolution of radius $r = y(x)$.

and let us minimise

$$\int_0^1 y\sqrt{1 + (y')^2} dx.$$

This corresponds to finding the surface of revolution with the minimal possible area, see Fig. 49.
Since the Lagrangian does not depend on x, the minimiser solves the equation $H(y, y') = C$ for some constant C. We have

$$\frac{\partial L}{\partial y'} = \frac{yy'}{\sqrt{1 + y'^2}},$$

hence

$$-H(y, y') = L - \frac{\partial L}{\partial y'} y' = y\sqrt{1 + y'^2} - \frac{yy'^2}{\sqrt{1 + y'^2}} = \frac{y(1 + y'^2)}{\sqrt{1 + y'^2}} - \frac{yy'^2}{\sqrt{1 + y'^2}}$$
$$= \frac{y}{\sqrt{1 + y'^2}} = C.$$

Squaring both sides of the above equality, we obtain

$$y^2 = C^2 + C^2(y')^2. \tag{138}$$

One of the solutions is given by $y(x) = C$ (hence $y'(x) = 0$). To find other solutions, we differentiate both sides with respect to x and find

$$2yy' = 2C^2 y' y'',$$

which gives

$$y'' = \frac{y}{C^2}$$

(we have already taken care about the solutions with $y'(x) = 0$). The general solution of the latter equation is

$$y = \tilde{A} e^{x/C} + \tilde{B} e^{-x/C}.$$

We can rewrite it as

$$y = A \cosh(x/C) + B \sinh(x/C)$$

with some different constants A and B (recall that $\cosh(x) = \frac{e^x + e^{-1}}{2}$, $\sinh(x) = \frac{e^x - e^{-1}}{2}$). Applying the identities $\cosh'(x) = \sinh(x)$ and $\sinh'(x) = \cosh$, we get

$$Cy' = A \sinh(x/C) + B \cosh(x/C),$$

$$y^2 = A^2 \cosh^2(x/C) + B^2 \sinh^2(x/C) + 2AB \sinh(x/C) \cosh(x/C),$$

$$Cy' = A^2 \sinh^2(x/C) + B^2 \cosh^{"}(x/C) + 2AB \sinh(x/C) \cosh(x/C).$$

Substituting this into (138) and using the identity $\cosh^2 + \sinh^2 = 1$, we find

$$C^2 = y^2 - (Cy')^2 = A^2 - B^2,$$

hence

$$y(x) = A \cosh(x/\sqrt{A^2 - B^2}) + B \sinh(x/\sqrt{A^2 - B^2}).$$

The constants A and B can be found once the boundary conditions have been specified.