Fractional calculus and generalised norms in condition monitoring of a load haul dumper
Contents

Introduction 4

1 Classical analysis 6
  1.1 Norms and function spaces ......................................... 6
  1.2 Convergence of function sequences ................................ 9
  1.3 Complex analysis ..................................................... 10
  1.4 Important theorems of analysis ..................................... 11
  1.5 Distributions ......................................................... 12
  1.6 Iterated differences and derivatives .............................. 16
  1.7 Iterated integrals .................................................... 18

2 Mathematical background for fractional calculus 20
  2.1 Gamma function ...................................................... 20
  2.2 Fourier series ....................................................... 23
    2.2.1 Series in $L_{1,\text{loc}}$ and $L_{2,\text{loc}}$ ....................... 23
    2.2.2 Pointwise convergence ....................................... 24
  2.3 Fourier transform .................................................. 26
    2.3.1 Transform in $S$ and $L_1$ ................................... 26
    2.3.2 Transform in $L_2$ ............................................. 28
    2.3.3 Transform in $S'$ ............................................... 29
    2.3.4 Properties ..................................................... 31
  2.4 Poisson summation formula and the sampling theorem ............ 34
  2.5 Discrete Fourier transform ....................................... 36
  2.6 Window functions .................................................. 37

3 Fractional calculus in time domain 39
  3.1 Riemann-Liouville fractional integral and derivative .......... 39
  3.2 Caputo fractional derivative ..................................... 45
  3.3 Grünwald-Letnikov fractional derivative and integral .......... 46
  3.4 Equivalences of definitions ..................................... 47

4 Fractional calculus in frequency domain 49
  4.1 Fourier differintegrals .......................................... 49
  4.2 Weyl differintegrals .............................................. 52
  4.3 Equivalences of definitions ..................................... 53
  4.4 Numerical algorithm .............................................. 56

5 Norms, means and other features of signals 60
  5.1 Generalised $l_p$ norms and Hölder means ....................... 60
  5.2 Measurement index ................................................. 64
6 Condition monitoring of a load haul dumper front axle 65
6.1 Data acquisition and acceleration signals ................. 65
6.2 Sensitivity surfaces and trend analysis ...................... 68
6.3 Complex differintegrals ..................................... 80
6.4 Conclusions ................................................. 84

References ....................................................... 85
Abstract

This thesis combines the concepts of fractional calculus, norms and means of vectors and machine condition monitoring. Fractional calculus is a branch of calculus that studies the concept of generalising differentiation and integration to arbitrary order. It is as old a discipline as classical calculus although not as widely known. Norm is a generalisation of the length or size of a mathematical object and it provides a single positive value to measure it. Machine condition monitoring is a discipline which focuses on extracting information mainly from industrial machines with noninvasive methods during their operation. This information is used to deduce the condition of the machine accurately.

In this thesis, the mathematical background for fractional calculus is built rigorously. Main focus is on application of Fourier analysis to fractional calculus. In addition to functions, distributions are utilised in the theory. Novel results are obtained on the continuity properties of some fractional derivatives and also on equivalences of different definitions. An effective numerical algorithm for calculating fractional derivatives and integrals in the frequency domain is presented.

The theory and numerical algorithm is utilised in calculating fractional derivatives and integrals of vibration signals. These signals are collected from the front axle of a load haul dumper working underground in the Pyhäsalmi mine. Generalised norms are calculated from the fractional derivatives and integrals and effective values for condition monitoring of the axle are found. Derivatives and integrals of complex order are also shown to change as the condition of the axle deteriorates during the 368 day long measurement period.
Introduction

This thesis is a treatment of fractional calculus and its application in machine condition monitoring. The thesis consists of the necessary mathematical background, fractional calculus in time and frequency domains, generalised norms and means of vectors and the practical application of the results in analysing vibration measurements from the front axle of a load haul dumper.

In section 1 important fundamental definitions and theorems of analysis are presented. In addition to typical functions, distributions are also utilised. Iterated differences and integrals are considered as those formulas will be used in defining certain fractional derivatives and integrals.

Section 2 deals with the concepts that are suitable as the generalisations needed to extend the normal derivatives and integrals to fractional ones. These concepts are the Gamma function, Fourier series and the Fourier transform. Fourier transform in this thesis is used to map a function from the time domain into frequency domain and it is the main tool that I will use in the later sections. An original proof about the speed of convergence of the Fourier transform is presented.

In the following two Sections 3 and 4 fractional calculus is considered in the time domain and frequency domain respectively. Original proofs about the equivalence of different definitions are presented and also about the continuity of the fractional derivatives of functions which have an ordinary derivative of bounded variation. An improved algorithm for the calculation of fractional derivatives and integrals in the frequency domain is presented shortly, since this algorithm is already published in a conference proceedings.

Section 5 deals with the norms and means of finite sequences and their generalisations. These are useful indices to calculate from measured displacement, velocity or acceleration signals.

Finally in Section 6 the theory is used in analysing acceleration signals measured from the front axle of a load haul dumper. Suitable orders of derivative and norm are searched which reveal the wear on the axle. Trend analysis of these features are plotted as well as the practicality of complex derivatives is visualised.

To emphasise the connections to physical quantities, the most frequently used variable in this thesis is \( t \) and it usually presents time in seconds (s). The second most important real variable is \( \nu \) which represents frequency (Hz = 1/s). The connection to angular frequency is \( \omega = 2\pi \nu \) (rad/s). The most frequent symbol for functions is \( f \) because of its popularity in mathematics. Functions are often referred to as signals as well, this again to emphasise the connection to measurements and the physical world.

I have adopted the term differintegral to describe operators which com-
bine both fractional integrals and derivatives. These operators are written $aD^z$, where $A$ is the definition in question, $a \in \mathbb{R}$ lower limit and $z \in \mathbb{C}$ the order of differintegration. Occasionally I will write fractional integrals $aI^z$ where $\text{Re}(z) > 0$ and the connection to differintegrals is $A^z = aI^{-z}$. Ordinary derivatives of order $n$ of function $f$ are written $\frac{d^n}{dt^n} f$, $f^{(n)}$ or $D^n f$. 
1 Classical analysis

1.1 Norms and function spaces

Functions or signals of interest for this thesis are of the type $f : \mathbb{R} \to \mathbb{C}$. These signals will be categorised by their integrability, continuity properties and periodicity. Many function spaces are also characterised by their norm that is a generalisation of length.

**Definition 1.1.** Let $V$ be a vector space. Norm is a function $\| \cdot \| : V \to \mathbb{R}$, which satisfies $\forall x, y \in V$ and $u \in \mathbb{C}$ the following three axioms

- $N1 \quad \|x\| \geq 0$, and $\|x\| = 0$ iff $x = 0$;
- $N2 \quad \|x + y\| \leq \|x\| + \|y\|;$
- $N3 \quad \|ux\| = |u| \|x\|$.

Axiom $N1$ is positive definiteness and $N2$ triangle inequality.

**Definition 1.2.** Function $f : \mathbb{R} \to \mathbb{C}$ is $T$-periodic, where $T > 0$, if $f(t) = f(t - T)$ for all $t \in \mathbb{R}$.

**Definition 1.3.** Function $f : U \to \mathbb{C}$, $U \subset \mathbb{R}$, is piecewise continuous if

1. function $f$ has only a finite number $N$ of discontinuities $t_1, t_2, \ldots, t_N$;
2. in each point of discontinuity $t_m$, $m = 1, 2, \ldots, N$, the left- and righthanded limits exist and are finite. Let us denote them $f(t_m-) = \lim_{h \to 0+} f(t_m - h)$ and $f(t_m+) = \lim_{h \to 0+} f(t_m + h)$.

If the endpoint is a point of discontinuity, only one of the above limits is required to exist.

If $f$ and $f'$ are piecewise continuous functions, then $f$ is piecewise smooth.

**Definition 1.4.** Subset $U \subset \mathbb{R}$ is a null set or has null Lebesgue measure if given any positive number $\epsilon$, there is a sequence $\{U_n\}$ of intervals in $\mathbb{R}$ such that $U \subset \bigcup_{n=1}^{\infty} U_n$ and the total length of the union is less than $\epsilon$.

Isolated points in $\mathbb{R}$ are clearly null sets but so are sets $\mathbb{N}, \mathbb{Z}$ and even $\mathbb{Q}$. Any interval of length $\lambda > 0$ is not a null set.

**Definition 1.5.** Functions $f$ and $g$, which are defined in $U \subset \mathbb{R}$, are equal almost everywhere, if the set $\{t \in U \mid f(t) \neq g(t)\}$ is a null set. This is shown

$$f(t) = g(t) \quad \text{a.e.}$$
Definition 1.6. Suppose that for the function \( f : [a, b] \to \mathbb{R} \) its absolute value \( |f| \) is integrable over the interval \([a, b]\) and

\[
\|f\|_1 = \int_a^b |f(t)| \, dt < \infty.
\]

Then one denotes \( f \in L_1[a, b] \). This is a special case of \( L_p[a, b] \)-spaces, where \( 1 \leq p < \infty \), \( |f|^p \) is integrable over \([a, b]\) and

\[
\|f\|_p = \left( \int_a^b |f(t)|^p \, dt \right)^{1/p} < \infty.
\]

If the function is defined on the whole real axis, one can also define function spaces \( f \in L_p(\mathbb{R}) = L_p \), for which

\[
\|f\|_p = \left( \int_{-\infty}^{\infty} |f(t)|^p \, dt \right)^{1/p} < \infty.
\]

The spaces \( L_\infty[a, b] \) and \( L_\infty(\mathbb{R}) = L_\infty \) consist of functions which are bounded almost everywhere with the norm

\[
\|f\|_\infty = \inf\{\lambda > 0 \mid |f(t)| \leq \lambda \text{ a.e. in } U\},
\]

where \( U \) is \([a, b]\) or \( \mathbb{R} \), respectively.

If \( f \cdot 1_{[a, b]} \in L_p[a, b] \) for all closed and bounded intervals \([a, b] \in \mathbb{R} \), one defines a function space \( f \in L_{p,\text{loc}}(\mathbb{R}) = L_{p,\text{loc}} \). Here \( 1_U \) is the characteristic function of the set \( U \)

\[
1_U(t) = \begin{cases} 1, & t \in U \\ 0, & t \notin U. \end{cases}
\]

Especially \( f \in L_{1,\text{loc}} \) is locally integrable.

Definition 1.7. Let \( 1 \leq p < \infty \). The space \( l_p(\mathbb{N}) \) consists of all sequences \( \{x_n\}_{n=1}^{\infty} \), that we may also write as vectors \( x = (x_1, x_2, \ldots) \), for which

\[
\|x\|_p = \left( \sum_{n=1}^{\infty} |x_n|^p \right)^{1/p} < \infty
\]

and the space \( l_\infty \) is the space of all bounded sequences with the norm

\[
\|x\|_\infty = \sup_{n=1,2,\ldots} |x_n| < \infty.
\]

Similarly we define \( l_p(\mathbb{Z}) \) when the sequence is defined on \( \mathbb{Z} \).
Definition 1.8. The support of function $f$, which is denoted $\text{supp}(f)$, is the closure of the set of points $t$ where $f(t) \neq 0$. Function $f$ has compact support if this set is bounded.

Definition 1.9. For function $f \in C^m(\mathbb{R}) = C^m$, $m \in \mathbb{N}_0 = \mathbb{N} \cup \{0\}$ its $m$th derivative $f^{(m)}$ is continuous. Especially function $f \in C^0 = C$ is a continuous function.

Definition 1.10. Function $f \in D$ is infinitely many times differentiable, i.e. $f \in C^\infty$ and has compact support, i.e. $f = f \cdot 1_{[a,b]}$ for some real numbers $a < b$.

Definition 1.11. Function $f$ in Schwartz space $f \in S(\mathbb{R}) = S$ has derivatives of all order, i.e. $f \in C^\infty$, and $f$ and all of its derivatives vanish rapidly

$$|f|_{j,m} = \sup_{t \in \mathbb{R}} \{|t|^j |f^{(m)}(t)|\} < \infty \quad \text{for all } j, m \in \mathbb{N}_0.$$  

The space $S$ is not a normed space, since $|f|_{j,m}$ is only a seminorm, for it may fail the axiom $\text{N1}$ of norms with some $j$ and $m$. The rapid descent of functions of Schwartz space means that they vanish at both infinities faster than reciprocal of any polynomial. The definition makes sense only on the whole interval $\mathbb{R}$. The Schwartz space is also very small compared to $L_p$ spaces, since we have the inclusion [1, p. 3]

$$S \subset L_p \text{ for all } 1 \leq p \leq \infty. \quad (1.3)$$

For spaces $L_p[a,b]$ we have the inclusion [2, p. 257]

$$L_p[a,b] \subset L_q[a,b] \text{ for all } 1 \leq q < p \leq \infty, \quad (1.4)$$

but the same is not true for spaces $L_p$, because with bigger $p$ the function $f \in L_p$ must behave better locally but may behave worse at the infinities.

The following two concepts only care about the behaviour of a function at the infinities.

Definition 1.12. Function $f : \mathbb{R} \to \mathbb{C}$ is big $O$ of $g$, i.e. $f \in O(g)$ if there exist $\lambda > 0$ and $t_0 \in \mathbb{R}$ such that

$$|f(t)| \leq \lambda |g(t)| \quad \text{for all } |t| \geq t_0.$$  

Function $f : \mathbb{R} \to \mathbb{C}$ is little $o$ of $g$, i.e. $f \in o(g)$ if for every $\epsilon > 0$ there exists $t_0 \in \mathbb{R}$ such that

$$|f(t)| \leq \epsilon |g(t)| \quad \text{for all } |t| \geq t_0.$$
If \( g(t) \neq 0 \) when \(|t| \geq \tau_0\) for some \( \tau_0 \in \mathbb{R} \), then \( f \in O(g) \) is equivalent to
\[
\lim_{|t| \to \infty} \frac{f(t)}{g(t)} = 0.
\]
Similarly a sequence \( \{x_n\}_{n=-\infty}^{\infty} \in O(y_n) \) if there exist \( \lambda > 0 \) and \( n_0 \in \mathbb{N} \) such that
\[
|x_n| \leq \lambda |y_n| \quad \text{for all } |n| \geq n_0,
\]
and \( \{x_n\}_{n=-\infty}^{\infty} \in o(y_n) \) if for every \( \epsilon > 0 \) there exists \( n_0 \in \mathbb{N} \) such that
\[
|x_n| \leq \epsilon |y_n| \quad \text{for all } |n| \geq n_0.
\]

**Definition 1.13.** Function \( f : [a,b] \to \mathbb{C} \) is of bounded variation, i.e. \( f \in BV[a,b] \) if its total variation is finite, i.e.
\[
\mathcal{T}(f,a,b) = \sup_{\mathcal{P}} \sum_{k=1}^{N} |f(t_k) - f(t_{k-1})| < \infty,
\]
where \( t_k, k = 0, 1, \ldots, N \) is a partition of the interval \([a,b]\) and the supremum is taken over all possible partitions \( \mathcal{P} \) of \([a,b]\). For the purposes of this thesis we extend this to \( f \in BV(\mathbb{R}) \) by extending the partition to the whole real line, i.e.
\[
\mathcal{T}(f, -\infty, \infty) = \sup_{\mathcal{P}} \sum_{k=-\infty}^{\infty} |f(t_{k+1}) - f(t_k)| < \infty.
\]

For a continuous function we can interpret its total variation as the length of the projection of its curve to the \( y \)-axis. Thus we see that a function of bounded variation is necessarily bounded and cannot oscillate infinitely on the interval in question. Also if \( a < d < b \), then
\[
\mathcal{T}(f, a, b) = \mathcal{T}(f, a, d) + \mathcal{T}(f, d, b). \tag{1.5}
\]
Function is of bounded variation if and only if it can be written as a difference of two increasing functions, which is called Jordan decomposition. It then follows from the properties of monotonic functions that a function of bounded variation is locally integrable and also differentiable a.e [3].

### 1.2 Convergence of function sequences

We will encounter sequences of functions which converge to another function in some sense. The type of convergence depends on the function space and the norm or seminorm in question.
Definition 1.14. Let $1 \leq p \leq \infty$. Sequence $\{f_n\}_{n=1}^{\infty}$, $f_n \in L_p$ converges in $L_p$ to $f \in L_p$, which we write $f_n \xrightarrow{L_p} f$ as $n \to \infty$, iff 
\[ \|f_n - f\|_p \to 0, \quad n \to \infty. \]

Definition 1.15. Sequence $\{f_n\}_{n=1}^{\infty}$, $f_n \in S$ converges in $S$ to $f \in S$, i.e. $f_n \xrightarrow{S} f$ as $n \to \infty$, iff 
\[ |f_n - f|_{j,m} \to 0, \quad n \to \infty \]
for all $j, m \in \mathbb{N}_0$.

Lemma 1.16. Convergence in $S$ implies convergence in $L_p$, $1 \leq p \leq \infty$.

Proof. First of all we know that $f \in S$ implies $f \in L_p$ (1.3). From the convergence $f_n \xrightarrow{S} f$ we know that especially 
\[ \sup_{t \in \mathbb{R}} |f_n(t) - f(t)| \to 0, \quad n \to \infty. \]

Since the values of $|f_n(t) - f(t)|$ go to zero as $n \to \infty$, it is clear that the integrals in the norms $\|f_n - f\|_p$ for $1 \leq p < \infty$ and the essential bound in $\|f_n - f\|_\infty$ go to zero as well.

\[ \square \]

Definition 1.17. Sequence $\{f_n\}_{n=1}^{\infty}$, $f_n \in D$ converges in $D$ to $f \in D$, i.e. $f_n \xrightarrow{D} f$ as $n \to \infty$, iff $f_n - f$ are all zero outside some fixed finite interval $U$ and for all $m \in \mathbb{N}_0$ we have 
\[ \sup_{t \in U} \left| \frac{d^m}{dt^m} (f_n(t) - f(t)) \right| \to 0, \quad n \to \infty. \]

1.3 Complex analysis

Definition 1.18. Simple curve in $\mathbb{C}$ is a curve that does not cross itself. Domain $G \subset \mathbb{C}$ is an open and connected set, i.e. between every point $s_1, s_2 \in G$ there exists a continuous function $f : [0,1] \to G$ for which $f(0) = s_1$ and $f(1) = s_2$. A domain $G$ is simply connected if any simple closed curve in $G$ can be continuously shrunk in $G$ to a point.

Intuitively we can say that there are no holes in a simply connected domain.
Definition 1.19. Let $\mathcal{L}$ be a continuous curve with equation
\[ s = \eta(t) = \mu(t) + i\rho(t), \quad a \leq t \leq b \]
and let $s_k = \eta(t_k)$, where values $t_k$, $k = 0, 1, \ldots, N$ are a partition of the interval $[a, b]$. The curve $\mathcal{L}$ is rectifiable if its length is finite, i.e.
\[ \sup_{\mathcal{P}} \sum_{k=1}^{N} |s_k - s_{k-1}| < \infty, \]
where the supremum is taken over all possible partitions $\mathcal{P}$ of $[a, b]$. This is equivalent to $\mu \in BV[a, b]$ and $\rho \in BV[a, b]$ [3, p. 116].

Definition 1.20. For a multivalued complex function $f$ its branch is a part of its range where $f$ is single-valued. The principal branch is typically chosen so that it coincides with the corresponding real-valued function for real arguments.

1.4 Important theorems of analysis

Theorem 1.21 (Fundamental theorem of calculus). Let $f \in L_1$ and define $F : \mathbb{R} \to \mathbb{C}$
\[ F(t) = \int_{-\infty}^{t} f(\tau) \, d\tau. \]
Then $F$ is differentiable a.e. and
\[ F'(t) = f(t) \quad a.e. \]

Proof. [4, p. 141] The lower limit of the integral can also be arbitrary $a$, in which case function $f$ can also be $f \in L_1[a, b]$. \hfill \Box

The fundamental theorem of calculus tells us that the differential operator is the left inverse of the integral operator.

Theorem 1.22 (Hölder inequality). Let $f \in L_p$ and $g \in L_q$, $1 \leq p \leq \infty$, and
\[ \frac{1}{p} + \frac{1}{q} = 1. \]
Then $fg \in L_1$ and
\[ \|fg\|_1 \leq \|f\|_p \|g\|_q. \] \hfill (1.6)

Proof. [5, p. 29] The case $p = 1$ and $q = \infty$ is not discussed in those references, but it follows with a simple estimate by taking out the norm $\|g\|_\infty$ from the integral. Theorem is also clearly valid for $f \in L_p[a, b]$ and $g \in L_q[a, b]$. \hfill \Box
Theorem 1.23 (Fubini theorem). Let $A, B \subset \mathbb{R}$ and $f : A \times B \to \mathbb{C}$. If at least one of the integrals

$$\int_B \int_A |f(t_1, t_2)| \, dt_1 \, dt_2, \quad \int_A \int_B |f(t_1, t_2)| \, dt_2 \, dt_1, \quad \int_{A \times B} |f(t_1, t_2)| \, dt$$

exists, then they all exist and are equal a.e.

Proof. [4, pp. 164-166], where the theorem is presented in a more general setting of measure spaces. \hfill \Box

Theorem 1.24 (Cauchy’s integral theorem). Let $G \subset \mathbb{C}$ be a simply connected domain and let $f$ be a single-valued analytic function on $G$. Then for any rectifiable closed curve $\mathcal{L}$ contained in $G$

$$\int_{\mathcal{L}} f(s) \, ds = 0.$$

Proof. [6, pp. 258-268] \hfill \Box

1.5 Distributions

A linear functional is a linear map from a vector space to real or complex numbers. Distributions are continuous linear functionals which map a set of smooth testing functions to $\mathbb{R}$ or $\mathbb{C}$. They generalise usual functions and are useful for example as initial values and solutions of differential equations and in the theory of Fourier transforms. The latter is of fundamental importance for this thesis. I will denote distributions with the same symbols as ordinary functions. When a distribution is generated with an ordinary function $f$, they are both referred to as $f$. The action of a distribution $f$ on a test function $\varphi$ is denoted as

$$\langle f, \varphi \rangle = \langle f, \varphi(t) \rangle,$$

(1.7)

where the variable $t$ is shown when necessary. For general distribution $f$ it is not meaningful to discuss anything pointwise and for this reason I will not write $f(t)$ unless $f$ is also a function.

Definition 1.25. Schwartz distribution $f \in D'$ is a linear continuous functional $f$ from the space $D$ to $\mathbb{C}$.

1. **Linearity** If $\varphi, \phi \in D$ and $u, v \in \mathbb{C}$, then

$$\langle f, u\varphi + v\phi \rangle = u\langle f, \varphi \rangle + v\langle f, \phi \rangle.$$
2. **Continuity** If a sequence of testing functions \( \{\varphi_n\}_{n=1}^{\infty} \) converges in \( D \) to zero, then

\[
\lim_{n \to \infty} \langle f, \varphi_n \rangle = 0.
\]

Every \( f \in L_{1,\text{loc}} \) defines a Schwartz distribution

\[
\langle f, \varphi \rangle = \int_{-\infty}^{\infty} f(t)\varphi(t) \, dt,
\]

since integral is a linear operator and we have the estimate

\[
\left| \langle f, \varphi_n \rangle \right| \leq \sup_{t \in U} |\varphi_n(t)| \int_U |f(t)| \, dt \to 0, \quad n \to \infty,
\]

whenever \( \varphi_n \xrightarrow{D} 0 \) and the support of every \( \varphi_n \) is contained in the finite subset \( U \) of \( \mathbb{R} \). If the action of a distribution \( f \) can written in the form (1.8) for some locally integrable function \( f \), it is a regular distribution. If this is not the case, the distribution is singular.

**Definition 1.26.** Tempered distribution \( f \in S' \) is a linear continuous functional \( f \) from the space \( S \) to \( \mathbb{C} \).

1. **Linearity** If \( \varphi, \phi \in S \) and \( u, v \in \mathbb{C} \), then

\[
\langle f, u\varphi + v\phi \rangle = u\langle f, \varphi \rangle + v\langle f, \phi \rangle.
\]

2. **Continuity** If a sequence of testing functions \( \{\varphi_n\}_{n=1}^{\infty} \) converges in \( S \) to zero, then

\[
\lim_{n \to \infty} \langle f, \varphi_n \rangle = 0.
\]

With a bigger family of test functions we get a smaller space of distributions. It means that \( S' \subset D' \) and it is actually a proper subspace of \( D' \) [7]. Every \( f \in L_p, 1 \leq p \leq \infty \) defines a tempered distribution

\[
\langle f, \varphi \rangle = \int_{-\infty}^{\infty} f(t)\varphi(t) \, dt,
\]

since from the inclusion (1.3) it follows that \( \varphi_n \in L_q, 1/p + 1/q = 1 \) and we may use H"older inequality Thm. 1.22

\[
\left| \int_{-\infty}^{\infty} f(t)\varphi_n(t) \, dt \right| \leq \int_{-\infty}^{\infty} |f(t)\varphi_n(t)| \, dt \leq \|f\|_p \|\varphi_n\|_q \to 0 \quad n \to \infty,
\]
since $\varphi_n \overset{S}{\to} 0$ which implies $\varphi_n \overset{L_q}{\to} 0$ according to Lemma 1.16.

For $f \in L_{1, \text{loc}}$ if we shift or translate the function by $\lambda > 0$, which we denote by $\sigma_\lambda f(t) = f(t - \lambda)$, then for any $\varphi \in D$

$$\int_{-\infty}^{\infty} f(t - \lambda)\varphi(t) \, dt = \int_{-\infty}^{\infty} f(\tau)\varphi(\tau + \lambda) \, d\tau.$$  

This motivates a definition for any distribution.

**Definition 1.27.** The shift by $\lambda > 0$ of distribution $f$ of space $D'$ or $S'$ is

$$\langle \sigma_\lambda f, \varphi \rangle = \langle f, \sigma_{-\lambda} \varphi \rangle.$$  

(1.9)

The Dirac delta distribution $\delta_\lambda$ is a tempered distribution which picks the value of a function at one point

$$\langle \delta_\lambda, \varphi \rangle = \varphi(\lambda).$$  

(1.10)

It represents a pulse, whose area is 1 and that is concentrated at the point $\lambda$. It is a singular distribution, although it is often written in the form (1.8). No function can satisfy such integral equation, so it is actually an abuse of notation. Dirac delta is also often referred to as a function for historical reasons.

**Definition 1.28.** The product $gf$ of distribution $f$ and function $g$ is defined as

$$\langle gf, \varphi \rangle = \langle f, g\varphi \rangle;$$  

(1.11)

and it is valid for example $f \in D'$ and $g \in C^\infty$ or $f \in S'$ and $g \in S$, because then $g\varphi \in D$ or $g\varphi \in S$, respectively.

It may be possible to multiply a distribution with a wider class of functions. For example for the Dirac delta distribution the product $g\delta_\lambda$ is well defined for any function $g$ that is defined at $\lambda$

$$\langle g\delta_\lambda, \varphi \rangle = \langle \delta_\lambda, g\varphi \rangle = g(\lambda)\varphi(\lambda).$$  

(1.12)

The product of two distributions is in general not possible to define.

**Definition 1.29.** Distribution $f$ is $T$-periodic, if

$$\langle \sigma_T f, \varphi \rangle = \langle f, \varphi \rangle.$$  

It is clear that $\langle f, \sigma_{-T} \varphi \rangle = \langle f, \varphi \rangle$ as well. From this definition and the definition of the shift of a distribution it follows another way to state the $T$-periodicity with the shifts of the test function

$$\langle f, \sigma_T \varphi \rangle = \langle \sigma_{-T} f, \varphi \rangle = \langle \sigma_T f, \varphi \rangle = \langle f, \varphi \rangle.$$  

14
Lemma 1.30. Every periodic distribution \( f \in D' \) with a period \( T > 0 \) is a tempered distribution, i.e. \( f \in S' \) as well.

**Proof.** [8, p. 229]

Let \( f \in L_{1, \text{loc}} \) define a regular distribution and assume that \( f' \in L_{1, \text{loc}} \) as well. Then it follows with integration by parts

\[
\langle f', \varphi \rangle = -\langle f, \varphi' \rangle, \quad \varphi \in D
\]

which suggests a definition of a distributional derivative.

**Definition 1.31.** For any distribution \( f \in D' \) or \( f \in S' \) its derivative of order \( m \in \mathbb{N} \) is

\[
\langle f^{(m)}, \varphi \rangle = (-1)^m \langle f, \varphi^{(m)} \rangle.
\]

(1.13)

With this definition every distribution is infinitely differentiable and the distributional derivative is also equal to the traditional derivative when function \( f \) is everywhere differentiable.

**Example 1.32.** The unit step function or Heaviside function is

\[
H(t) = \begin{cases} 
1, & t \geq 0 \\
0, & t < 0.
\end{cases}
\]

(1.14)

Clearly \( H \in L_{1, \text{loc}} \) and so \( H \in D' \). Thus we can calculate its distributional derivative

\[
\langle H', \varphi \rangle = -\langle H, \varphi' \rangle = -\int_0^\infty \varphi'(t) \, dt = \varphi(0) = \langle \delta_0, \varphi \rangle.
\]

It is obvious that any piecewise continuous function \( f \in L_1 \) can be written as a sum of some continuous and bounded function \( g \in L_\infty \cup C \) and shifted and scaled Heaviside functions

\[
f(t) = g(t) + \sum_{m=1}^N d_m H(t - t_m),
\]

(1.15)

where \( d_m \) is the size and direction of the jump discontinuity at the point \( t_m \). Now all the jump discontinuities can be differentiated as Dirac delta distributions

\[
f'(t) = g'(t) + \sum_{m=1}^N d_m \delta_{t_m},
\]

(1.16)

where \( f' \) and \( g' \) are the distributional derivatives of \( f \) and \( g \). Thus we have actually found a representation for the distributional derivative of a discontinuous function of the space \( L_1 \) in the sense of tempered distributions.
1.6 Iterated differences and derivatives

**Definition 1.33.** Backward difference operator with a step length of $h$ and centered at the point $t$ for function $f$ is

$$\Delta_h f(t) = f(t) - f(t - h), \quad (1.17)$$

and higher differences of order $m \in \mathbb{N}$ are defined by iterating the definition

$$\Delta_h^m f(t) = \Delta_h (\Delta_h^{m-1} f(t)), \quad (1.18)$$

where $\Delta_h^0 f(t) = f(t)$.

It follows from the definition that the $m$th backward difference can be expressed as a single summation. For this compact notation we will need binomial coefficient of $m, k \in \mathbb{N}_0$, $0 \leq k \leq m$

$$\binom{m}{k} \frac{m!}{k!(m - k)!}, \quad (1.19)$$

for which $\binom{m}{0} = \binom{m}{m} = 1$ for all $m \in \mathbb{N}_0$. The binomial coefficients satisfy the recursion [9, p. 11]

$$\binom{m+1}{k} = \binom{m}{k} + \binom{m}{k-1}. \quad (1.20)$$

**Lemma 1.34.**

$$\Delta_h^m f(t) = \sum_{k=0}^{m} (-1)^k \binom{m}{k} f(t - kh). \quad (1.21)$$

**Proof.** The Lemma is proved with induction. The basic step of $m = 1$ gives us the definition of $\Delta_h f(t)$. Let us then assume that the Lemma holds for some $m \in \mathbb{N}$. Then the backward difference of order $m + 1$ is

$$\Delta_h^{m+1} f(t) = \Delta_h (\Delta_h^m f(t))$$

$$= \Delta_h \left( \sum_{k=0}^{m} (-1)^k \binom{m}{k} f(t - kh) \right)$$

$$= \sum_{k=0}^{m} (-1)^k \binom{m}{k} \Delta_h f(t - kh)$$

$$= \sum_{k=0}^{m} (-1)^k \binom{m}{k} \left( f(t - kh) - f(t - kh - h) \right),$$
where the linearity of the difference operator was utilised. Next we break the summation into two parts to see its structure more clearly

\[
\Delta_{m+1}^{m+1} f(t) = \sum_{k=0}^{m} (-1)^k \binom{m}{k} f(t - kh) + \sum_{k=1}^{m+1} (-1)^k \left( \binom{m}{k-1} f(t - kh) \right)
\]

\[
= f(t) + f(t - (m + 1)h) + \sum_{k=1}^{m+1} (-1)^k \left[ \binom{m}{k} + \binom{m}{k-1} \right] f(t - kh)
\]

\[
= \sum_{k=0}^{m+1} (-1)^k \binom{m+1}{k} f(t - kh),
\]

and finally the recursion (1.20) and the equality \(1 = \binom{m}{m} = \binom{m+1}{m+1} = \binom{m}{0} = \binom{m+1}{0}\) were used to combine the summation again.

Since derivatives are defined as the limit of a difference quotient as \(h \to 0\), equation (1.21) divided by \(h\) should produce the \(m\)th derivative of a function in the limit.

**Lemma 1.35.** If \(f \in C^m[a, b]\), then for all \(n = 1, 2, \ldots, m\)

\[
D^n f(t) = \frac{d^n}{dt^n} f(t) = \lim_{h \to 0} \frac{\Delta^n h f(t)}{h^n} \quad a < t \leq b. \tag{1.22}
\]

**Proof.** The proof of this fundamental result is often omitted in the literature. This is the case for example in [10]. The proof is again based on induction and on the Lemma 1.34. The basic step of \(n = 1\) results in simply the definition of the derivative

\[
f'(t) = \lim_{h \to 0} \frac{1}{h} \left( f(t) - f(t - h) \right).
\]

Let us then assume that the Lemma holds for the \(n\)th derivative of the function. Investigate then the \(n + 1\)th derivative

\[
f^{(n+1)}(t) = \lim_{h \to 0} \frac{1}{h} \left( f^{(n)}(t) - f^{(n)}(t - h) \right)
\]

\[
= \lim_{h \to 0} \frac{1}{h} \left( \lim_{u \to 0} \frac{\Delta^n u f(t)}{u^n} - \lim_{u \to 0} \frac{\Delta^n u f(t - h)}{u^n} \right)
\]

\[
= \lim_{h \to 0} \lim_{u \to 0} \frac{1}{h u^n} \left( \Delta^n u f(t) - \Delta^n u f(t - h) \right)
\]

\[
= \lim_{h \to 0} \lim_{u \to 0} \frac{1}{h u^n} \left( \Delta^u f(t) \right)
\]

\[
= \lim_{h \to 0} \frac{\Delta^h_{n+1} f(t)}{h^{n+1}}.
\]
Since $f^{(n)}$ exists on $(a, b)$, we know that the limits of the backward difference operator exist for $f^{(n)}(t)$ and $f^{(n)}(t - h)$, and therefore, the summation of the limits could be expressed as a limit of summations in the third equality. Finally the limits could be combined since they now approach a common value simultaneously and the final step then followed from Lemma 1.34.

Clearly the Lemma is extended to the case $a = -\infty$, if $f \in C^m$.

1.7 Iterated integrals

Iterated integral of a function with the lower limit $a$ can be expressed as a single integral. This is again straightforwardly proved with induction.

**Definition 1.36.** Iterated integral operator with lower limit $a$ is

$$aI^n f(t) = \int_a^t \int_a^{t_1} \cdots \int_a^{t_{n-1}} f(t_n) \, dt_n \, dt_{n-1} \cdots dt_1. \quad (1.23)$$

**Lemma 1.37.** If $f \in L^1[a, b]$, then

$$aI^n f(t) = \frac{1}{(n-1)!} \int_a^t (t - \tau)^{n-1} f(\tau) \, d\tau. \quad (1.24)$$

**Proof.** The proof of this fundamental result is usually again omitted in the literature, for example in [10] and [11]. The basic case $n = 1$ is an identity.

Let us assume the case of $n$-fold iterated integral. Then we investigate the $n + 1$-fold iterated integral

$$aI^{n+1} f(t) = \int_a^t \int_a^{t_1} \cdots \int_a^{t_{n-1}} \int_a^{t_n} f(t_{n+1}) \, dt_{n+1} \, dt_n \, dt_{n-1} \cdots dt_1.$$ 

Using the assumption of $n$-fold iterated integral

$$aI^{n+1} f(t) = \int_a^t \frac{1}{(n-1)!} \int_a^{t_1} (t_1 - \tau)^{n-1} f(\tau) \, d\tau \, dt_1,$$

and changing the order of integration by Fubini theorem 1.23

$$aI^{n+1} f(t) = \frac{1}{(n-1)!} \int_a^t f(\tau) \int_\tau^t (t_1 - \tau)^{n-1} \, dt_1 \, d\tau$$

$$= \frac{1}{(n-1)!} \int_a^t f(\tau) \left( \frac{1}{n} (t_1 - \tau)^n \right) \bigg|_\tau^t \, d\tau$$

$$= \frac{1}{n!} \int_a^t (t - \tau)^n f(\tau) \, d\tau,$$
Figure 1: Integration limits in the iterated integral

which proves the Lemma through induction. The change of integration limits is visualised in Figure 1.

Clearly the Lemma is extended to the case $a = -\infty$, if $f \in L_1$. 

\[ \square \]
2 Mathematical background for fractional calculus

The two Lemmata 1.35 and 1.37 both contain factorials. It would seem that a generalisation of the factorial function is needed to generalise differentiation and integration to arbitrary order. Since the factorial function is defined only on the natural numbers, there are countless ways to draw a function between these points. But it turns out that only one function has some desirable smoothness properties to make it the best and only reasonable choice.

2.1 Gamma function

Definition 2.1. Euler’s definition of the Gamma function $\Gamma$ is the integral

$$\Gamma(z) = \int_0^\infty t^{z-1}e^{-t} \, dt, \quad z \in \mathbb{C}, \, \text{Re}(z) > 0.$$  \hspace{1cm} (2.1)

The integral converges absolutely, when $\text{Re}(z) > 0$. By integration by parts

$$\Gamma(z + 1) = \int_0^\infty t^ze^{-t} \, dt = -t^ze^{-t}\bigg|_0^\infty + z\int_0^\infty t^{z-1}e^{-t} \, dt = z\Gamma(z),$$  \hspace{1cm} (2.2)

and this may be used to construct analytic continuation of the Gamma function to the strip $-1 < \text{Re}(z) < 0$ by $\Gamma(z) = \Gamma(z + 1)/z$. Continuing this process reveals, that the Gamma function may be defined everywhere except at the points $\mathbb{Z}^- \cup \{0\}$ where it has simple poles. Formula

$$\Gamma(z)\Gamma(1-z) = \frac{\pi}{\sin(\pi z)}, \quad z \notin \mathbb{Z}$$ \hspace{1cm} (2.3)

provides an immediate analytic continuation from the right to the left half plane [12, pp. 65-67] From (2.2) and the fact $\Gamma(1) = \int_0^\infty e^{-t} \, dt = 1$ we find that the Gamma function truly generalises the factorial

$$\Gamma(n + 1) = n!, \quad n \in \mathbb{N}_0.$$ \hspace{1cm} (2.4)

The Gamma function is connected to the Beta function

$$B(\alpha, \gamma) = \int_0^1 (1-s)^{\alpha-1}s^{\gamma-1} \, ds, \quad \text{Re}(\alpha), \text{Re}(\gamma) > 0$$ \hspace{1cm} (2.5)

with the equality

$$B(\alpha, \gamma) = \frac{\Gamma(\alpha)\Gamma(\gamma)}{\Gamma(\alpha + \gamma)},$$ \hspace{1cm} (2.6)
which also continues the Beta function analytically [12, p. 76].

The following Lemma will be utilised in proving equalities of some definitions of fractional integrals and derivatives.

Lemma 2.2. \[ \int_{0}^{\infty} t^{z-1} e^{-\gamma t} \, dt = \frac{\Gamma(z)}{\gamma^z}, \quad \gamma \neq 0, \quad (2.7) \]

where \( \text{Re}(z) > 0 \) if \( \text{Re}(\gamma) > 0 \) and \( 0 < \text{Re}(z) < 1 \) if \( \text{Re}(\gamma) = 0 \) and the principal branch of \( \gamma^z \) is chosen so that \( \gamma^z > 0 \) if \( \gamma > 0 \) and \( z \) is real-valued.

Proof. I will follow [11, pp. 137-138]. Let us change variables in the integral definition of the Gamma function with \( \tau = \gamma t \), \( d\tau = \gamma \, dt \)

\[ \Gamma(z) = \int_{0}^{\infty} \tau^{z-1} e^{-\tau} \, d\tau = \gamma^z \int_{0}^{\infty} t^{z-1} e^{-\gamma t} \, dt, \]

so we get (2.7) in the case \( \text{Re}(z) > 0 \) and \( \gamma > 0 \). Both sides of the equation are analytic when \( \text{Re}(\gamma) > 0 \), so the equation is analytically continued to this case.

Then we only have to proof the boundary case \( \text{Re}(\gamma) = 0, \gamma \neq 0 \), i.e. \( \gamma = i\nu \)

\[ \int_{0}^{\infty} t^{z-1} e^{-i\nu t} \, dt = \frac{\Gamma(z)}{(i\nu)^z}, \quad 0 \leq \text{Re}(z) < 1. \quad (2.8) \]

The left side of (2.8) converges since we assumed in this boundary case that \( \text{Re}(z) < 1 \). The equality is proved with the substitution \( i\nu t = \theta \), \( d\theta = i\nu \, dt \)

\[ \int_{0}^{\infty} t^{z-1} e^{-i\nu t} \, dt = (i\nu)^{-z} \int_{L} \theta^{z-1} e^{-\theta} \, d\theta, \quad (2.9) \]

where \( L = (0, i \infty) \) when \( \nu > 0 \) and \( (-i \infty, 0) \) when \( \nu < 0 \). Now \( |e^{-\theta}| \) vanishes in the right half-plane as \( |\theta| \to \infty \), therefore, if we construct a path integral according to integration paths in Figure 2, then two of the sides contribute nothing to the integral and it follows by the Cauchy integral theorem 1.24

\[ (i\nu)^{-z} \int_{L} \theta^{z-1} e^{-\theta} \, d\theta = (i\nu)^{-z} \int_{0}^{\infty} \nu^{z-1} e^{-\nu} \, d\nu = (i\nu)^{-z} \Gamma(z). \quad (2.10) \]

The Cauchy integral theorem was stated for rectifiable closed curves, but the above stated equality follows with a limiting argument if we start by integrating over a small square with sides of length \( R \) and then let \( R \) tend to infinity.
Figure 2: Integration paths in Lemma 2.2. The upper one is for \( \nu > 0 \) and the lower one for \( \nu < 0 \)

**Definition 2.3.** Function \( f : U \to \mathbb{R}, U \subset \mathbb{R} \) is *convex* if for all \( \tau_1, \tau_2 \in U \) and \( t \in [0, 1] \)

\[
f(t\tau_1 + (1-t)\tau_2) \leq tf(\tau_1) + (1-t)f(\tau_2).
\]

Function \( f \) is *logarithmically convex* if \( \ln(f) \) is convex.

Geometrically convexity means that any line segment connecting two points of the convex function is above the function values. It is clear that a logarithmically convex function \( f \) is convex, because then it can be written as \( f(t) = e^{\ln(f(t))} \) and the exponential function is convex and increasing. Incredibly, logarithmic convexity is the feature that separates the Gamma function from every other generalisation of the factorial.

**Theorem 2.4 (Bohr-Mollerup).** Gamma function is the only generalisation of the factorial to positive real numbers \( t > 0 \) with the following three properties

1. \( f(1) = 1; \)
2. \( f(t+1) = tf(t); \)
3. \( f \) is logarithmically convex.

*Proof.* [13, p. 14]
2.2 Fourier series

Fourier’s pioneering ideas in representing a function as a series or integrals of sine and cosine functions \[14\] has had countless applications in physics and technology. The questions about the existence of such representations has also profoundly affected the development of analysis and even the definition of the integral. I will write Fourier series of functions as series of complex exponentials instead of sines and cosines because of it’s elegance and simple connections with the continuous and discrete Fourier transforms.

2.2.1 Series in \(L_{1,\text{loc}}\) and \(L_{2,\text{loc}}\)

**Definition 2.5.** Let \(f \in L_{1,\text{loc}}\) and \(T\)-periodic. Its Fourier series is

\[
f(t) \sim \sum_{k=-\infty}^{\infty} c_k e^{i2\pi kt/T}, \tag{2.11}
\]

where multipliers \(c_k\) are calculated as

\[
c_k = \frac{1}{T} \int_{0}^{T} f(t) e^{-i2\pi kt/T} \, dt, \quad k = 0, \pm 1, \pm 2, \ldots \tag{2.12}
\]

and we use the notation \(\sim\) since nothing is stated about the convergence of the series.

Because of the \(T\)-periodicity of \(f\) and complex exponentials, the integration in (2.12) can be calculated from any interval of length \(T\).

The following two convergence theorems will have analogous versions in the theory of Fourier transform. The calculation of Fourier coefficients for \(f \in L_{2,\text{loc}}\) is not straightforward, since they will converge only in the sense of norm in \(L_2[0, T]\) and the series will likewise converge in this norm. The method is similar to that of Fourier transform for \(L_2\) functions, which will be presented in the Section 2.3.2.

**Theorem 2.6.** If a \(T\)-periodic function \(f \in L_{1,\text{loc}}\) and \(\{c_k\} \in l_1(\mathbb{Z})\), then

\[
f(t) = \sum_{k=-\infty}^{\infty} c_k e^{i2\pi kt/T}, \quad \text{a.e.}
\]

If \(f\) is also continuous, we have equality for all \(t \in \mathbb{R}\).

**Proof.** [2, p. 27]
Theorem 2.7. If a $T$-periodic function $f \in L_{2,\text{loc}}$ then $\{c_k\} \in l_2(\mathbb{Z})$ and

$$f(t) = \sum_{k=-\infty}^{\infty} c_k e^{2\pi i k t/T}, \quad \text{a.e.}$$

Proof. [2, pp. 161-162]

2.2.2 Pointwise convergence

One of the first rigorous proof about the convergence of Fourier series is the following.

Theorem 2.8 (Dirichlet 1829). Let $f \in L_{1,\text{loc}}$ and $T$-periodic. If $f'(t-) \text{ ja } f'(t+)$ exist for all $t$ (for example when $f$ is piecewise smooth), the Fourier series of $f$ converges to

$$\frac{1}{2} (f(t-) + f(t+)) \quad \text{for every } t \in \mathbb{R}$$

Proof. This is often presented in the case of $2\pi$-periodic functions, as in the book [5, pp. 188-191]. Original proof by Dirichlet can be found in the paper [15].

The conditions of Theorem 2.8 can be weakened and still have the same convergence value.

Theorem 2.9 (Jordan 1881). Let $f$ be a function of bounded variation in $[0,T]$ (which implies that $f \in L_1[0,T]$) and $T$-periodic. Then the Fourier series of $f$ converges to

$$\frac{1}{2} (f(t-) + f(t+)) \quad \text{for every } t \in \mathbb{R}$$


For a piecewise smooth function $f$ or a function of bounded variation its Fourier series converges pointwise to $f(t)$ whenever $f$ is continuous at $t$ and to the average value of its limits in the points of discontinuity. For a continuous and piecewise smooth function a stronger result is valid.

Theorem 2.10. Let $f$ be a piecewise smooth function in $\mathbb{R}$ with a period $T$.

1. On every closed interval that does not include discontinuities of $f$, its Fourier series converges uniformly to $f$. 

2. If \( f \) is in addition continuous, its Fourier series converges absolutely to \( f \) for every \( t \in \mathbb{R} \).

**Proof.** Presented in the book [5, pp. 203-205] once again in the case of \( 2\pi \)-periodic functions. \( \square \)

A strong bound on the growth of the Fourier coefficients is achieved for functions of bounded variation. Before stating the result let us first calculate a simple orthogonality property of the complex exponentials that is needed in the proof. For all \( k, n \in \mathbb{Z} \), \( k \neq 0 \)

\[
\int_{Tn/k}^{T(n+1)/|k|} e^{-i2\pi kt/T} dt = -\frac{T}{i2\pi k} e^{-i2\pi(n+1)} - e^{-i2\pi n} = 0,
\]

when \( k > 0 \) and the signs in the exponents are reversed after the substitution if \( k < 0 \). It is also useful to note here that replacing \( k/T \) with a continuous real variable \( \nu \neq 0 \) changes nothing essential in the proof above, and therefore we have

\[
\int_{n/|\nu|}^{(n+1)/|\nu|} e^{-i2\pi\nu t} dt = 0,
\]

which will be utilised in the analogous treatment of Fourier transforms.

**Theorem 2.11.** If function \( f \) is of bounded variation on \([0, T]\) and \( T \)-periodic, then its Fourier coefficients satisfy \( c_k \in O(1/|k|) \). More precisely, we have the estimate

\[
|c_k| \leq T(f, 0, T) \frac{1}{|k|}.
\]

**Proof.** This proof is given as an assignment in [5, p. 167]. I will follow the original proof of [16] although it is written for \( 2\pi \)-periodic functions.

Let us suppose that \( k \neq 0 \) and fix \( k \) to some \( k \in \mathbb{Z} \). Then let us partition the interval \([0, T]\) with \( t_n = Tn/|k| \), \( n = 0, 1, \ldots, |k| \). Define a step function \( g \) for which \( g(t) = f(t_n) \) when \( t \in (t_{n-1}, t_n) \), \( n = 1, \ldots, |k| \). Then by (2.13) we have \( f_0^T g(t) e^{-i2\pi kt/T} dt = 0 \) and we can begin our estimation

\[
|c_k| = \left| \frac{1}{T} \int_0^T \left( f(t) - g(t) \right) e^{-i2\pi kt/T} dt \right| \leq \sum_{n=1}^{|k|} \frac{1}{T} \int_{t_{n-1}}^{t_n} \left| f(t) - f(t_n) \right| dt
\]

\[
\leq \sum_{n=1}^{|k|} \frac{T(f, t_{n-1}, t_n)}{|k|} \frac{T}{|k|} = T(f, 0, T) \frac{1}{|k|}.
\]
where in the second inequality the integral was estimated with its total variation times the length of the interval and finally property (1.5) was used.

2.3 Fourier transform

2.3.1 Transform in $S$ and $L_1$

For nonperiodic function it is possible to consider $T$ to be whole $\mathbb{R}$. Then the integration in the equation (2.12) is carried over $\mathbb{R}$ and the multiples of the fundamental frequency become a continuous frequency variable $\nu$.

**Definition 2.12.** Fourier transform of a function $f \in L_1$ is

$$\mathcal{F}\{f(t)\} = \hat{f}(\nu) = \int_{-\infty}^{\infty} f(t)e^{-i2\pi \nu t} \, dt \quad (2.16)$$

and its inverse transform is

$$\mathcal{F}^{-1}\{\hat{f}(\nu)\} = \int_{-\infty}^{\infty} \hat{f}(\nu)e^{i2\pi \nu t} \, d\nu. \quad (2.17)$$

Fourier transform (and its inverse) is well defined in $L_1$ since

$$\left| \int_{-\infty}^{\infty} f(t)e^{-i2\pi \nu t} \, dt \right| \leq \int_{-\infty}^{\infty} |f(t)||e^{-i2\pi \nu t}| \, dt = \int_{-\infty}^{\infty} |f(t)| \, dt = ||f||_1.$$  

The following is a fundamental mapping theorem of functions $f \in L_1$.

**Theorem 2.13** (Riemann-Lebesgue). *If* $f \in L_1$, *then* $\hat{f}$ is bounded, uniformly continuous and vanishes at both infinities, i.e.

$$\lim_{|\nu| \to \infty} \hat{f}(\nu) = 0. \quad (2.18)$$

*Proof.* For example in [5, pp. 170-171, 290-291].

Let’s present two theorems about the invertibility of Fourier transform in function spaces $L_1$ and $S$. Since $S \subset L_1$, Fourier transform is well defined for functions $f \in S$ as well.

**Theorem 2.14.** If $f \in S$, then also $\hat{f} \in S$ and

$$\mathcal{F}^{-1}\{\hat{f}(\nu)\} = f(t) \quad \text{for all} \ t \in \mathbb{R}$$

*Proof.* [17, pp. 141-142]

26
Theorem 2.15. If $f \in L_1$ and $\hat{f} \in L_1$, then

$$\mathcal{F}^{-1}\{\hat{f}(\nu)\} = f(t) \quad \text{a.e.}$$

If $f$ is also continuous, we have equality for all $t \in \mathbb{R}$.

Proof. [2, pp. 17-19], [3, pp. 87-89] □

Example 2.16. The rectangular function and triangular function are

$$\text{rect}(t) = H(t + 1/2)H(1/2 - t) = \begin{cases} 1, & |t| \leq \frac{1}{2} \\ 0, & |t| > \frac{1}{2} \end{cases} \quad (2.19)$$

$$\text{tri}(t) = \begin{cases} 1 - |t|, & |t| < 1 \\ 0, & |t| \geq 1 \end{cases} \quad (2.20)$$

For $T, \lambda > 0$ we have the Fourier transforms

$$\mathcal{F}\left\{\text{rect}\left(\frac{t}{T}\right)\right\} = T \frac{\sin(T\pi \nu)}{T\pi \nu}, \quad (2.21)$$

$$\mathcal{F}\{\text{tri}(t/T)\} = T \left(\frac{\sin(T\pi \nu)}{T\pi \nu}\right)^2, \quad (2.22)$$

$$\mathcal{F}\{e^{-\lambda|t|}H(t)\} = \frac{1}{\lambda + i2\pi \nu}, \quad (2.23)$$

$$\mathcal{F}\{e^{-\lambda|t|}\} = \frac{2\lambda}{\lambda^2 + (2\pi \nu)^2}, \quad (2.24)$$

$$\mathcal{F}\{e^{-\pi t^2}\} = e^{-\pi \nu^2}. \quad (2.25)$$

Proofs of these transforms are found for example in [2] or [18]. Most textbooks on Fourier analysis mention these transforms as well but the two examples mentioned use the same definition as this thesis. Transforms obtained using other definitions must be scaled accordingly. Different definitions often use angular frequency $\omega = 2\pi \nu$ and this choice makes it necessary to have powers of $2\pi$ in front of the transform or/and inverse transform. Also the sign of the exponent $-i2\pi \nu$ may be reversed for transform and its inverse.

It is important to notice, that $f \in L_1$ does not imply $\hat{f} \in L_1$. In fact, even for the simple box or rectangular function we get the sinc function that is not absolutely integrable.
2.3.2 Transform in $L_2$

"The idea of the Fourier transform for $L_2$ functions is classic analysis: We sneak up on it". [5, p. 356]

The symmetry which we saw in Theorem 2.14 is returned in the function space $L_2$. The downside is that the definition of Fourier transform in the space $L_2$ is a more delicate question than for spaces $L_1$ and $S$. The following method is presented for example in [5], wherefrom the above quote was taken. The basic method for this "sneaking up", is to define for function $f \in L_2$ the compactly supported function

$$f_n(t) = f(t) \cdot 1_{[-n,n]}(t), \quad n \in \mathbb{N}.$$  

Clearly $f_n \in L_1 \cap L_2$ for all $n \in \mathbb{N}$. Also $f_n - f \in L_2$ and therefore

$$\|f_n - f\|_2 \to 0, \quad \text{as } n \to \infty.$$

Now for all $n \in \mathbb{N}$ there exists Fourier transform $\hat{f}_n$. Because the function space $L_1 \cap L_2$ is dense in $L_2$ (i.e. every function from $L_2$ can be approximated with a function sequence from $L_1 \cap L_2$ in the $L_2$ norm), this suggests a definition of the Fourier transform in $L_2$ as a limit of transforms $\hat{f}_n$. But let us not jump into that conclusion before some preliminary results.

**Theorem 2.17** (Plancherel theorem). If $f \in L_1 \cap L_2$, then $\hat{f} \in L_2$ and

$$\|f\|_2^2 = \|\hat{f}\|_2^2. \quad (2.26)$$

**Proof.** [2, p. 156], [4, pp. 186-187], [5, pp. 352-353] \qed

**Lemma 2.18.** The sequence $\{\hat{f}_n\}_1^\infty$ constructed above converges to a function in $L_2$.

**Proof.** I will follow [5, p. 357]. Plancherel theorem 2.17 tells us that each $\hat{f}_n$ is in $L_2$. We also know that $\mathcal{F}\{f_n - f_m\} = \hat{f}_n - \hat{f}_m$ for each $n, m \in \mathbb{N}$ and by Plancherel theorem

$$\|\hat{f}_n - \hat{f}_m\|_2^2 = \|f_n - f_m\|_2^2. $$

Now for $m > n$

$$\|f_n - f_m\|_2^2 = \int_{-m}^{-n} |f(t)|^2 \, dt + \int_{n}^{m} |f(t)|^2 \, dt,$$

and both integrals approach 0 as $m, n \to \infty$. This proves that the sequence $\{\hat{f}_n\}_1^\infty$ is a Cauchy sequence in $L_2$ and thus the sequence converges in $L_2$. \qed
Now that we know that a limit of our function sequence exists, we may take this as a definition. Even more is true, since we now know that Fourier transform maps $L_2$ onto itself.

**Definition 2.19.** Fourier transform of a function $f \in L_2$ is the $L_2$ limit

$$\mathcal{F}\{f(t)\} = \hat{f}(\nu) = \lim_{n \to \infty} \hat{f}_n(\nu) \quad (2.27)$$

and its inverse transform is the $L_2$ limit

$$\mathcal{F}^{-1}\{\hat{f}(\nu)\} = \lim_{n \to \infty} \mathcal{F}^{-1}\{\hat{f}(\nu) \cdot 1_{[-n,n]}(\nu)\}. \quad (2.28)$$

Since $L_2$ limits are unique only almost everywhere, equations (2.27) and (2.28) define $\hat{f}$ and $\mathcal{F}^{-1}\{\hat{f}\}$ only almost everywhere.

**Theorem 2.20.** If $f \in L_2$ then also $\hat{f} \in L_2$ and

$$\mathcal{F}^{-1}\{\hat{f}(\nu)\} = f(t) \quad a.e.$$

*Proof.* [2, pp. 159-160], [5, p. 362] □

There is nothing special about the functions $f_n$ above. Indeed, any functions which are dense in $L_2$ could be used to construct a.e. equivalent definitions of Fourier transforms in $L_2$. For example in [3, pp. 209-210] it is shown that also $S$ is a dense subspace of $L_2$ and it is used to construct the transform in $L_2$.

When the Fourier transform of $f$ is invertible, function $f$ can be represented as a double integral. This representation is known as Fourier’s integral theorem

$$f(t) = \int_{-\infty}^{\infty} \left( \int_{-\infty}^{\infty} f(t)e^{-i2\pi \nu t} dt \right) e^{i2\pi \nu t} d\nu, \quad (2.29)$$

where the equality is either everywhere or almost everywhere depending on the function space in question.

### 2.3.3 Transform in $S'$

Theorem 2.14 makes it possible to define Fourier transform for the conjugate space $S'$. Like many other properties for distributions, this generalisation is again motivated by using an integration argument. Let $f \in L_1$ and $\varphi \in S$, then $\hat{f} \in L_\infty$ by Riemann-Lebesgue 2.13 and it defines a tempered distribution

$$\int_{-\infty}^{\infty} \hat{f}(\nu) \varphi(\nu) d\nu = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(t)e^{-i2\pi \nu t} \varphi(\nu) dtd\nu = \int_{-\infty}^{\infty} f(t)\hat{\varphi}(t) dt,$$

where we used Fubini theorem 1.23.
Definition 2.21. Fourier transform of a tempered distribution \( f \in S' \) is
\[
\langle \mathcal{F}\{f\}, \varphi \rangle = \langle f, \mathcal{F}\{\varphi\} \rangle,
\]
and its inverse transform is
\[
\langle \mathcal{F}^{-1}\{\hat{f}\}, \varphi \rangle = \langle \hat{f}, \mathcal{F}^{-1}\{\varphi\} \rangle.
\]

The power of distribution theory is illustrated in the elegant and simple inversion theorem for Fourier transform of the space in \( S' \).

Theorem 2.22. If \( f \in S' \) then also \( \hat{f} \in S' \) and
\[
\mathcal{F}^{-1}\{\hat{f}\} = f \quad \text{for every } \varphi \in S.
\]

Proof. It is evident from the definition that \( \hat{f} \in S' \) and its inverse is
\[
\langle \mathcal{F}^{-1}\{\hat{f}\}, \varphi \rangle = \langle \hat{f}, \mathcal{F}^{-1}\{\varphi\} \rangle = \langle f, \mathcal{F}\{\mathcal{F}^{-1}\{\varphi\}\} \rangle = \langle f, \varphi \rangle,
\]
where the last equality followed from the inversion of Schwartz functions Thm. 2.14. \( \square \)

The following example shows that the Dirac delta distribution (1.10) is mapped into a constant unit function.

Example 2.23.
\[
\langle \hat{\delta}_0, \varphi \rangle = \langle \delta_0, \hat{\varphi} \rangle = \hat{\varphi}(0) = \int_{-\infty}^{\infty} \varphi(t)e^{-i2\pi \cdot 0 \cdot t} dt = \langle 1, \varphi \rangle, \quad \varphi \in S.
\]
We will denote this simply \( \hat{\delta}_0 = 1 \). Similarly we will have \( \hat{1} = \delta_0 \) and since Fourier transforms of tempered distributions are invertible, these imply that \( \mathcal{F}^{-1}\{\delta_0\} = 1 \) and \( \mathcal{F}^{-1}\{1\} = \delta_0 \).

The following theorem connects Fourier series with the Fourier transform of periodic distributions. From Lemma 1.30 we already know that every periodic distribution \( f \in D' \) is also a tempered distribution and therefore has a Fourier transform in \( S' \).

Theorem 2.24. Let \( f \in D' \) be a periodic distribution (which by Lemma 1.30 is in \( S' \)) with a period \( T > 0 \). Then there exist constants \( c_k, k \in \mathbb{Z} \) such that
\[
\hat{f} = \sum_{k=-\infty}^{\infty} c_k \delta_{2\pi k/T}.
\]

Proof. [8, p. 231] \( \square \)
2.3.4 Properties

Shifting the function or its Fourier transform will mean multiplying with complex exponentials in the other domain.

**Lemma 2.25** (Shifting properties). *Let* \( \lambda \in \mathbb{R} \), *then it is valid for all the above definitions of Fourier transform that*

\[
F\{\sigma_\lambda f\} = e^{-i2\pi \lambda \nu} \hat{f},
\]

(2.33)

\[
F\{e^{i2\pi \lambda t} f\} = \sigma_\lambda \hat{f}.
\]

(2.34)

**Proof.** Let first \( f \in L_1 \). Then

\[
F\{f(t-\lambda)\} = \int_{-\infty}^{\infty} f(t-\lambda) e^{-i2\pi \nu t} \, dt = \int_{-\infty}^{\infty} f(\tau) e^{-i2\pi \nu (\tau+\lambda)} \, d\tau = e^{-i2\pi \lambda \nu} \hat{f}(\nu),
\]

and the second statement follows identically with a change of variables. Now the definition of the Fourier transform for \( L_2 \) functions brings this property to those functions as well. For \( f \in S' \) we have

\[
\langle F\{\sigma_\lambda f\}, \varphi \rangle = \langle \sigma_\lambda f, \hat{\varphi} \rangle = \langle f, \hat{\varphi}(\nu + \lambda) \rangle = \langle \hat{f}, e^{i2\pi (-\lambda) \nu} \hat{\varphi}(\nu) \rangle = \langle \hat{f}, e^{-i2\pi \lambda \nu} \varphi \rangle = \langle e^{-i2\pi \lambda \nu} \hat{f}, \varphi \rangle.
\]

\[\square\]

**Example 2.26.** Let us apply the shifting properties of Lemma 2.25 to Dirac delta distribution of Example 2.23

\[
F\{\delta_\lambda\} = e^{-i2\pi \lambda \nu},
\]

(2.35)

\[
\mathcal{F}^{-1}\{\delta_\lambda\} = e^{i2\pi \lambda t}.
\]

(2.36)

With integration by parts one can derive a very useful equation for the derivatives of a function.

**Theorem 2.27.** If \( f^{(m)} \in L_1 \) for all \( 0 \leq m \leq n \), then

\[
F\left\{ \frac{d^n}{dt^n} f(t) \right\} = (i2\pi \nu)^n F\{f(t)\}.
\]

(2.37)

**Proof.** [2, pp. 20-21] \[\square\]
Corollary 2.28. If $f^{(m)} \in L_1$ for all $0 \leq m \leq n$, then

$$\hat{f}(\nu) \in o\left(\frac{1}{|\nu|^n}\right)$$

(2.38)

Proof. Theorem 2.27 states that

$$\frac{\hat{f}(\nu)}{(i2\pi\nu)^{-n}} = \hat{f}^{(n)}(\nu),$$

where the right side goes to zero as $|\nu| \to \infty$ according to R-L lemma 2.13. Therefore, the result follows from the definition of the space little $o$.

\[ \square \]

Corollary 2.29. If in addition to assumptions of Theorem 2.27 we assume that $\hat{f}^{(m)} \in L_1$ for some $0 \leq m \leq n$, then

$$f^{(m)}(t) = \mathcal{F}^{-1}\left\{(i2\pi\nu)^m \mathcal{F}\{f(t)\}\right\} \quad \text{a.e.}$$

(2.39)

Proof. Follows directly from the inversion formula 2.15.

\[ \square \]

It follows from Thm. 2.14 that for functions $f \in S$ and for every $m \in \mathbb{N}$

$$\frac{d^m}{dt^m}f(t) = \mathcal{F}^{-1}\left\{(i2\pi\nu)^m \mathcal{F}\{f(t)\}\right\}.$$  

(2.40)

We may also differentiate Fourier transforms and the following frequency derivatives have similar forms. For example for $f \in S$ we have [2]

$$\frac{d^m}{d\nu^m} \hat{f}(\nu) = \mathcal{F}\{(-i2\pi t)^m f(t)\}. $$

(2.41)

Fourier transform of a derivative of a tempered distribution will have this familiar form also for every $m \in \mathbb{N}$

$$\langle \hat{f}^{(m)}, \varphi \rangle = \langle f^{(m)}, \varphi \rangle = \langle f, (-1)^m \hat{\varphi}^{(m)} \rangle = \langle f, (-1)^m \mathcal{F}\{(-i2\pi t)^m \varphi\} \rangle = \langle \hat{f}, \mathcal{F}\{(i2\pi t)^m \varphi\} \rangle = \langle \hat{f}, (-i2\pi t)^m \varphi \rangle = \langle \hat{f}, (i2\pi t)^m \varphi \rangle = \langle (i2\pi \nu)^m \hat{f}, \varphi \rangle$$

(2.42)

where we differentiated the distribution and used frequency derivatives (2.41).
Example 2.30. We can even calculate a formula for the Fourier transform of the derivatives of a periodic distribution of Theorem 2.24 by using the multiplication property of Dirac delta distribution (1.12) \( \langle (i2\pi \nu)^z \delta_{k/T}, \varphi \rangle = (i2\pi k/T)^z \varphi(k/T) \), for \( \varphi \in S \) to get

\[
\hat{f}^{(m)} = (i2\pi \nu)^m \sum_{k=-\infty}^{\infty} c_k \delta_{k/T} = \sum_{k=-\infty}^{\infty} \left( \frac{i2\pi k}{T} \right)^m c_k \delta_{k/T}.
\]

Now the inverse transform (2.36) gives us the derivatives of every periodic distribution \( f \in S' \)

\[
f^{(m)} = \sum_{k=-\infty}^{\infty} \left( \frac{i2\pi k}{T} \right)^m c_k F^{-1}\{\delta_{k/T}\} = \sum_{k=-\infty}^{\infty} \left( \frac{i2\pi k}{T} \right)^m c_k e^{i2\pi k/T}.
\] (2.43)

Convergence of the series is not guaranteed but nevertheless (2.43) always has meaning as a tempered distribution!

Dividing the Fourier transform with \( i2\pi \nu \) is related to integration with lower limit \(-\infty\).

Theorem 2.31. If \( f, tf \in L_1 \) and \( f \) has zero mean value, i.e. \( \int_{-\infty}^{\infty} f(t) \, dt = 0 \), then \( g(t) = \int_{-\infty}^{t} f(\tau) \, d\tau \in L_1 \) and we have the relation

\[
\hat{g}(\nu) = F\left\{ \int_{-\infty}^{t} f(\tau) \, d\tau \right\} = \frac{1}{i2\pi \nu} F\{f(t)\}, \quad \nu \neq 0,
\] (2.44)

and

\[
\hat{g}(0) = -\int_{-\infty}^{\infty} tf(t) \, dt.
\] (2.45)

Proof. [5, p. 336] \( \square \)

In Example 2.16 we saw that with discontinuous functions \( f \in L_1 \) we have a case of \( \hat{f}(\nu) \in O(1/|\nu|) \) and with continuous functions \( g \in L_1 \) we have \( \hat{g}(\nu) \in O(1/|\nu|^2) \), whereas from Corollary 2.28 we can only deduce that \( \hat{f}(\nu) \in o(1) \) and \( \hat{g}(\nu) \in o(1/|\nu|) \) respectively. Even though the little \( o \) is a stronger notation, it only implies the big \( O \) with the same function. This suggests that with some further assumptions a stronger result is valid. One feature that connects the examples above is that the discontinuous ones are also of bounded variation on \( \mathbb{R} \).

Theorem 2.32. If \( f \in L_1 \) is of bounded variation, then \( \hat{f} \in O(1/|\nu|) \). The actual estimate is

\[
|\hat{f}(\nu)| \leq T(f, -\infty, \infty) \frac{1}{|\nu|}.
\] (2.46)
Proof. I will prove this result in a similar fashion to Theorem 2.11.

Suppose that \( \nu \neq 0 \) and fix \( \nu \) to some \( \nu \in \mathbb{R} \). Then let us partition the real line with \( t_n = n/|\nu|, n \in \mathbb{Z} \). Define a step function \( g \) for which \( g(t) = f(t_n) \) when \( t \in (t_{n-1}, t_n) \). Then by (2.14) we have \( \int_{-\infty}^{\infty} g(t)e^{-i2\pi \nu t} \, dt = 0 \) and we can estimate

\[
|\hat{f}(\nu)| = \left| \int_{-\infty}^{\infty} (f(t) - g(t)) e^{-i2\pi \nu t} \, dt \right| \leq \sum_{n=-\infty}^{\infty} \int_{t_{n-1}}^{t_n} |f(t) - f(t_n)| \, dt
\leq \sum_{n=-\infty}^{\infty} T(f, t_{n-1}, t_n) \frac{1}{|\nu|}
= T(f, -\infty, \infty) \frac{1}{|\nu|}.
\]

For the rectangular function \( T(f, -\infty, \infty) = 1 \) and thus

\[
\mathcal{F}\left\{ \text{rect}\left(\frac{t}{T}\right) \right\} \leq \frac{1}{|\nu|}.
\]

Comparing this estimate to the actual transform (2.21) we see that the estimate can only be improved up to a constant.

**Corollary 2.33.** If \( f \in L_1 \) and \( f^{(n)} \in L_1 \cap BV(\mathbb{R}) \), then

\[
\hat{f}(\nu) \in O\left(\frac{1}{|\nu|^{n+1}}\right).
\]  

Proof. By Theorem 2.32 we have \( \hat{f}^{(n)} \in O(1/|\nu|) \) and since \( f \in L_1 \) is a tempered distribution, it has all distributional derivatives and their Fourier transforms (2.42). These Fourier transforms are functions, since \( \hat{f} \in L_\infty \) due to Theorem 2.13 and thus

\[
\hat{f}(\nu) = \frac{\hat{f}^{(n)}(\nu)}{(i2\pi \nu)^n},
\]

which proves that \( \hat{f}^{(n)} \in O(1/|\nu|^{n+1}) \).  \(\square\)

### 2.4 Poisson summation formula and the sampling theorem

**Lemma 2.34.** Let \( f \in L_{1,\text{loc}}(\mathbb{R}) \) and \( T \)-periodic. Define the function

\[
f_T = f \cdot 1_{[0,T]};
\]  

34
and its Fourier transform $\hat{f}_T$. Then the Fourier coefficients of $f$ are

$$c_k = \frac{1}{T} \hat{f}_T \left( \frac{k}{T} \right), \quad k = 0, \pm 1, \pm 2, \ldots,$$

(2.49)

Proof. This Lemma is an assignment in [2]. The Fourier coefficients of $f$ for all $k = 0, \pm 1, \pm 2, \ldots$ are

$$c_k = \frac{1}{T} \int_0^T f(t) e^{-i2\pi kt/T} \, dt = \frac{1}{T} \int_{-\infty}^{\infty} f_T(t) e^{-i2\pi kt/T} \, dt = \frac{1}{T} \hat{f}_T \left( \frac{k}{T} \right).$$

This statement is true even in a more generalised setting. When the periodic summation of a function equals a Fourier series where the coefficients are samples of the function’s Fourier transform, we have the Poisson summation formula after Siméon Denis Poisson

$$\sum_{m=-\infty}^{\infty} f(t + mT) = \frac{1}{T} \sum_{k=-\infty}^{\infty} \hat{f} \left( \frac{k}{T} \right) e^{i2\pi kt/T}.$$

(2.50)

A sufficient condition for the existence of such a formula is given in the next theorem.

**Theorem 2.35.** Let $f \in L_1(\mathbb{R})$ and $0 < T < \infty$. Poisson summation formula (2.50) holds for every $t \in \mathbb{R}$ if its left side converges for every $t \in \mathbb{R}$ to some continuous function and its right side converges for all $t \in \mathbb{R}$.

Proof. [2, p. 44]

The weak form of equation (2.50) is the case $t = 0$. It is enough to prove the sampling theorem that was independently discovered by several mathematicians in the beginning of 20th century. It is often named after Claude Shannon and Harry Nyquist.

**Theorem 2.36.** Let $f \in L_1(\mathbb{R})$ be continuous and its Fourier transform $\hat{f}$ zero outside the frequency interval $[-B, B]$ where, $0 < B < \infty$. Furthermore, let

$$\sum_{m=-\infty}^{\infty} \left| f \left( \frac{m}{2B} \right) \right| < \infty.$$

Then $f$ can be regained from its samples $f \left( m/(2B) \right)$, $m \in \mathbb{Z}$.

$$f(t) = \sum_{m=-\infty}^{\infty} f \left( \frac{m}{2B} \right) \frac{\sin(\pi(2Bt - m))}{\pi(2Bt - m)}.$$

(2.51)
The requirement that $\hat{f}$ is bandlimited, i.e. supp($\hat{f}$) = $[-B, B]$ is essential in sampling a signal. No information is lost from a bandlimited signal if it is sampled with a sampling frequency $\nu_s$ that is at least the Nyquist-frequency $2B$. If a signal contains higher frequencies than $B$ there will occur aliasing. This means that the copies of the spectrum, that are $\nu_s$ apart, overlap each other in the frequency domain. This can be prevented by using an anti-aliasing filter. No greater frequency than the Nyquist-frequency is required to reconstruct the signal if the algorithm for reconstruction models (2.51) accurately. In practise one often uses oversampling with $\nu_s >> 2B$ which makes it easier to approximate the signal even with a simple algorithm.

2.5 Discrete Fourier transform

In practise signals are sampled with sample time $\Delta t$ and length of the signal is finite. Let there be $N$ samples which are taken at times $t_n = \Delta t \cdot n$, $n = 0, 1, \ldots, N - 1$. The length of the measurement is $T = \Delta t \cdot N$, which means that after the last sample the measurement continues for time $\Delta t$. Let’s denote the signal values at the measurement points as $f(t_n) = f_n$. Now the complex exponent function in the formula (2.12) has at these discrete points the values

$e^{-i2\pi kt/T} = e^{-i2\pi kn\Delta t/(\Delta t \cdot N)} = e^{-i2\pi kn/N}$.

Integration is replaced with summation over $N$ points

$$c_k = \frac{1}{T} \int_0^T f(t)e^{-i2\pi kt/T}dt \approx \frac{1}{\Delta t \cdot N} \sum_{n=0}^{N-1} f_n e^{-i2\pi kn/N} \Delta t$$

$$= \frac{1}{N} \sum_{n=0}^{N-1} f_n e^{-i2\pi kn/N}.$$

**Definition 2.37.** Discrete Fourier transform (DFT) of sequence $f_n \in \mathbb{C}$ of length $N$ at the point $k$ is

$$\mathcal{F}\{f_n\} = F_k = \frac{1}{N} \sum_{n=0}^{N-1} f_n e^{-i2\pi kn/N} \quad (2.52)$$

and its inverse transform (IDFT) is

$$\mathcal{F}^{-1}\{F_k\} = f_n = \sum_{k=0}^{N-1} F_k e^{i2\pi kn/N}. \quad (2.53)$$
At first the DFT may look like just a crude approximation of the Fourier coefficients or the Fourier transform of a function, but it has better properties than a mere approximation typically has. Discrete Fourier transform and its inverse are complex-valued sequences of equal length. They are also $N$-periodic and the inverse returns the original sequence. DFT is therefore a discrete approximation of the signals spectrum but it also contains all the information needed to return the original sequence. For real-valued $f_n$ we have the symmetry

$$F_k = \overline{F_{N-k}}$$

and therefore the zero frequency $F_0 = F_N$ is real-valued. For real-valued signals one has to calculate only the terms $F_k$, for which $k \leq N/2$ [5, 19].

It is also beneficial to choose $N = 2^m$, $m \in \mathbb{N}$ if possible. Using brute force calculation of the DFT definition requires computational operations (complex multiplication and addition) of order $O(N^2)$. If $N$ is a power of two, an efficient algorithm called Fast Fourier Transform (FFT) calculates the same result in time $O(N \log_2(N))$. This is possible because the definition includes calculations of same terms multiple times [20].

The components of the DFT are intimately related to the Fourier coefficients $c_k$ of the original function. In fact, the following equality is valid.

**Theorem 2.38.** Let $\{c_k\}_{k=-\infty}^\infty$ be the Fourier coefficients of function $f$ and $\{F_k\}_{k=0}^{N-1}$ its DFT. Then

$$F_k = c_k, \quad 0 \leq k < N/2$$
$$F_{N+k} = c_k, \quad -N/2 < k < 0$$
$$F_{N/2} = c_{-N/2} + c_{N/2}.$$

**Proof.** [21, p. 183] \(\square\)

For bandlimited signals this result means, that the DFT can produce the correct Fourier coefficients, provided that the sampling frequency $\nu_s$ is high enough, that is actually again the Nyquist frequency $2B$, where $B$ is the highest frequency that the signal contains.

For the case of even $N$ the DFT cannot separate the coefficients $c_{-N/2}$ and $c_{N/2}$ but rather mixes them into one term. This is also an inconvenience for the numerical algorithm that is presented in section 4.4.

### 2.6 Window functions

Discrete Fourier analysis leads to periodic solutions both in time and frequency domain. If the original signal has significantly different values at
the endpoints, then its periodic continuation will have a jump discontinuity and hence slowly convergent Fourier series. This also means that its DFT is a poor approximation of the spectrum. Discontinuities can be removed by multiplying the signal with a suitable window function. When calculating differintegrals and not just the spectrum, it is desirable that the signal is distorted in the time domain as little as possible. Therefore, the window function should have quite rapid ascent and descent. We can also reject part of the signal from its beginning and end after differintegration.

A simple window function is constructed from the Hann window by using it in the ascent and descent parts of the window

\[
 w_1(t) = \begin{cases} 
 0, & \text{if } t \leq 0 \\
 0.5 \left(1 - \cos\left(\pi \frac{t}{T/\epsilon}\right)\right), & \text{if } 0 < t < T/\epsilon \\
 1, & \text{if } T/\epsilon \leq t \leq T/2 \\
 w_1(T - t), & \text{if } t > T/2,
\end{cases}
\]  

(2.56)

where \( \epsilon \) is the portion of \( T \) for ascent and descent. In the books [19, 22] examples are given of \( \epsilon = 10 \).

Lahdelma and Kotila introduced in the article [23] a window function

\[
 w_2(t) = \begin{cases} 
 0, & \text{if } t \leq 0 \\
 \left(\int_0^{T/\epsilon} e^{(\tau - T/\epsilon) \frac{1}{\epsilon}} d\tau\right)^{-1} \int_0^t e^{(\tau - T/\epsilon) \frac{1}{\epsilon}} d\tau, & \text{if } 0 < t < T/\epsilon \\
 1, & \text{if } T/\epsilon \leq t \leq T/2 \\
 w_2(T - t), & \text{if } t > T/2,
\end{cases}
\]  

(2.57)

and used the value \( \epsilon = 8 \) in their calculations. It is easy to see using 1.21 that all the derivatives of \( w_2 \) are continuous. Since the multiplication of two continuous functions is as well continuous, the window function \( w_2 \) preserves the continuity properties of the original signal.
3 Fractional calculus in time domain

3.1 Riemann-Liouville fractional integral and derivative

We have seen in Lemma 1.37 that an iterated integral with the lower limit \( a \) can be written as a single integral. Based on this observation and the property of Gamma function (2.4) \( \Gamma(n+1) = n! \), \( n \in \mathbb{N}_0 \) we define Riemann-Liouville fractional integral of order \( \alpha > 0 \).

**Definition 3.1.** The Riemann-Liouville fractional integral operator with lower limit \( a \) is

\[
_{a}^{RL}I_{a}^{\alpha}f(t) = \frac{1}{\Gamma(\alpha)} \int_{a}^{t} (t-\tau)^{\alpha-1} f(\tau) \, d\tau, \quad \alpha > 0.
\]  

(3.1)

The case \( \alpha = 0 \) is defined as the identity operator \( _{a}^{RL}I_{a}^{0}f = f \). The most used choices for \( a \) are \( 0 \) and \( -\infty \). Fractional integrals of any order \( \alpha > 0 \) exist a.e. for \( f \in L_{p}[a,b] \), where \( 1 \leq p \leq \infty \) and the space is kept invariant, i.e. \( _{a}^{RL}I_{a}^{\alpha}f \in L_{p}[a,b] \) with the same \( p \) [11, pp. 48-51]. This mapping doesn’t hold for the case \( a = -\infty \), but we have that \( _{-\infty}^{RL}I_{a}^{\alpha}f \) exists a.e. for \( f \in L_{p} \) if \( 0 < \alpha < 1 \) and \( 1 \leq p < 1/\alpha \) [11, pp. 94-95].

Fractional integrals satisfy the following semigroup property.

**Theorem 3.2.** Let \( f \in L_{p}[a,b] \), \( 1 \leq p \leq \infty \). Then it holds for \( \alpha, \gamma \geq 0 \) that

\[
_{a}^{RL}I_{a}^{\alpha}_{a}^{RL}I_{a}^{\gamma}f = _{a}^{RL}I_{a}^{\alpha+\gamma}f \quad \text{a.e.}
\]  

(3.2)

If \( f \) is in addition continuous or \( \alpha + \gamma \geq 1 \), then the equality holds for every \( t \). For \( a = -\infty \) and \( f \in L_{p} \) we have (3.2) if \( \alpha + \gamma < 1/p \).

**Proof.** This follows directly from the definition

\[
_{a}^{RL}I_{a}^{\alpha}f(t)
= \frac{1}{\Gamma(\alpha)\Gamma(\gamma)} \int_{a}^{t} (t-v)^{\alpha-1} \int_{a}^{v} (v-\tau)^{\gamma-1} f(\tau) \, d\tau \, dv
= \frac{1}{\Gamma(\alpha)\Gamma(\gamma)} \int_{a}^{t} f(\tau) \int_{\tau}^{t} (t-v)^{\alpha-1}(v-\tau)^{\gamma-1} \, dv \, d\tau
= \frac{1}{\Gamma(\alpha)\Gamma(\gamma)} \int_{a}^{t} f(\tau) \left( t - (\tau + s(t-\tau)) \right)^{\alpha-1} (s(t-\tau))^{\gamma-1}(t-\tau) \, ds \, d\tau
= \frac{1}{\Gamma(\alpha)\Gamma(\gamma)} \int_{a}^{t} f(\tau) (t-\tau)^{\alpha+\gamma-1} \int_{0}^{1} (1-s)^{\alpha-1}s^{\gamma-1} \, ds \, d\tau
= \frac{B(\alpha, \gamma)}{\Gamma(\alpha)\Gamma(\beta)} \int_{a}^{t} f(\tau)(t-\tau)^{\alpha+\gamma-1} \, d\tau
= _{a}^{RL}I_{a}^{\alpha+\gamma}f(t),
\]
where the order of integration was changed in the same way as in the proof of 1.37, substitution $v = \tau + s(t - \tau)$ was made and the property (2.6) of the Beta function was utilised.

Integral (3.1) diverges when $\alpha \leq 0$ and therefore it doesn’t immediately define fractional derivatives. A simple way to generalise Definition 3.1 into fractional differentiation is to calculate a suitable fractional integral and differentiate this result classically. These ideas made their first appearance in the papers of Abel [24, 25] where he solved the tautochrone problem of physics. He formulated the problem in to an integral equation which solution was essentially to invert a R-L type integral operator of order $1 - \alpha$, where $0 < \alpha < 1$. Abel mentioned nothing about fractional derivatives, but rather used the mathematical machinery and finally substituted $\alpha = 1/2$ to solve the problem.

The idea of generalising integration and differentiation through these methods is due to work of Riemann [26] and Liouville [27, 28]. Riemann dealt mainly with operators that had lower limit $a = 0$ and Liouville with the case $a = -\infty$. The original definition of Liouville was to expand a function in to series of exponentials and to differentiate it. Obviously the existence of such a series limits the functions to be differentiated, but this definition lead Liouville to find the integral representation.

**Definition 3.3.** The Riemann-Liouville fractional derivative operator with lower limit $a$ is

$$RL_a D^\alpha f(t) = \frac{d^n}{dt^n} RL_a I^{n-\alpha} f(t), \quad n = [\alpha].$$  \hspace{1cm} (3.3)

An immediate consequence of these definitions is that R-L differintegrals are linear operators, just as their ordinary counterparts. The semigroup property can be extended to fractional derivatives in simple form if we construct our $f$ suitably, that is, we assume that it is a fractional integral of some $\phi$. This guarantees enough smoothness for $f$ and its suitably rapid descent as $t \to a$.

**Theorem 3.4.** Let $\phi \in L_p[a, b]$, $1 \leq p \leq \infty$ and $\alpha, \gamma \geq 0$. Then for $f = RL_a I^{\alpha+\gamma}\phi$

$$RL_a D^\alpha RL_a D^\gamma f = RL_a D^{\alpha+\gamma} f.$$ \hspace{1cm} (3.4)

For $a = -\infty$ and $\phi \in L_p$ the Theorem holds if $\alpha + \gamma < 1/p$.

**Proof.** I will follow [10, pp. 29-30], although in that reference only the case $\phi \in L_1[a, b]$ is considered. By (3.3)

$$RL_a D^\alpha RL_a D^\gamma f = RL_a D^\alpha RL_a D^\gamma RL_a I^{\alpha+\gamma}\phi = D^{|\alpha|} RL_a I^{\alpha-\alpha} D^{|\gamma|} RL_a I^{\gamma-\gamma} RL_a I^{\alpha+\gamma}\phi.$$
The semigroup property of the fractional integrals (3.2) can be used

\[ RL_a D^\alpha RL_a D^\gamma f = D[^\alpha] RL_a I[^\alpha] - \alpha D[^\gamma] RL_a I[^\gamma] + \alpha \phi = D[^\alpha] RL_a I[^\alpha] - \alpha D[^\gamma] a I[^\gamma] RL_a I^\alpha \phi, \]

and now the fundamental theorem of calculus Thm. 1.21 cancels the operators of natural order

\[ RL_a D^\alpha RL_a D^\gamma f = D[^\alpha] RL_a I[^\alpha] - \alpha RL_a I^\alpha \phi = D[^\alpha] a I[^\alpha] \phi = \phi, \]

and finally we used Theorems 3.2 and 1.21 once more. The proof of \( D^{\alpha+\gamma} = \phi \) is proved with a similar method.

The concept of the fundamental theorem of calculus is also preserved.

**Theorem 3.5.** Let \( \alpha \geq 0 \). Then for all \( f \in L_p[a, b] \), \( 1 \leq p \leq \infty \)

\[ RL_a D^\alpha RL_a I^\alpha f = f \quad \text{a.e.} \tag{3.5} \]

For \( a = -\infty \) and \( f \in L_p \) the Theorem holds if \( \alpha < 1/p \).

**Proof.** This proof is also from [10, p. 30] but we have extended it as well to a wider class of functions. The case \( \alpha = 0 \) is clear since then the operators are identity operators. For \( \alpha > 0 \) (and \( \alpha < 1/p \) if \( f \in L_p \))

\[ RL_a D^\alpha RL_a I^\alpha f = D[^\alpha] RL_a I[^\alpha] - \alpha RL_a I^\alpha f = D[^\alpha] a I[^\alpha] \phi = \phi, \]

where the semigroup property (3.2) and Thm. 1.21 proved this result as well.

The generalisation to the complex order derivatives and integrals requires the manipulation of the multivalued power function \( (t - \tau)^{z-1} \), \( z = \alpha + i\beta \). We choose its principal branch, which means that

\[ (t - \tau)^{z-1} = (t - \tau)^{\alpha-1} e^{i\beta \ln(t - \tau)}, \quad \text{when } t - \tau > 0. \tag{3.6} \]

With this choice everything stated above holds if we replace \([\alpha] = [\text{Re}(z)]\).

**Definition 3.6.** The Riemann-Liouville fractional integral operator with lower limit \( a \) for complex \( z \in \mathbb{C} \) is

\[ RL_a I^z f(t) = \frac{1}{\Gamma(z)} \int_a^t (t - \tau)^{z-1} f(\tau) d\tau, \quad \text{Re}(z) > 0, \tag{3.7} \]

where \((t - \tau)^{z-1}\) is calculated as above.
Definition 3.7. The Riemann-Liouville fractional derivative operator with lower limit $a$ for complex $z \in \mathbb{C}$ is

$$RL_a D^z f(t) = \frac{d^n}{dt^n} RL^{n-z} f(t), \quad n = [\text{Re}(z)]. \quad (3.8)$$

Problematic values to define for the Riemann-Liouville operator are the fractional derivatives (or integrals) of purely imaginary order. These are discussed for example in [11, pp. 38-39].

The fractional derivative and integral are not local operators in the sense that values of the function are needed from the interval $(a, t]$ to calculate a differintegral at $t$. Nevertheless I will prove that for a compactly supported function its R-L differintegral is also compactly supported on the same interval with very minimal assumptions.

Theorem 3.8. Let $f(t) = 0$ when $t \notin [0, T]$. If $\lim_{t \to \infty} RL_a D^z f(t) = 0$, then $RL_a D^z f(t) = 0$ when $t \notin [0, T]$ for all $z \in \mathbb{C}$.

Proof. Let $n = [\text{Re}(z)]$ and first $t < 0$. Then the integrand in $RL_a I^z f(t)$ or in $RL_a I^{n-z} f(t)$ is always zero and therefore $RL_a D^z f(t) = 0$ is zero.

Then let $t > T$. Now we have

$$RL_a I^z f(t) = \frac{1}{\Gamma(z)} \int_a^T (t - \tau)^{z-1} f(\tau) \, d\tau = RL_a I^z f(T),$$

and the same is true for differentiation, when $RL_a I^{n-z} f(t) = RL_a I^{n-z} f(T)$ and therefore $RL_a D^z f(t) = RL_a D^z f(T)$.

So when $t > T$ the R-L differintegral gets the same value for every $t$. If $\text{Re}(z) > 0$ then this constant is differentiated to zero. For $\text{Re}(z) < 0$ we assumed that $\lim_{t \to \infty} RL_a D^z f(t) = 0$. Therefore this value must also be zero.

The assumption $\lim_{t \to \infty} RL_a D^z f(t) = 0$ is fulfilled for example if $RL_a D^z f \in L_p$, for $1 \leq p < \infty$. Actually for fractional derivatives this assumption is not even needed, since the ordinary derivative of a constant is zero. This condition is however necessary for fractional integrals. Consider for example $f = \text{rect}(t)$, $z = -1$ and $a < -1/2$. Then we have a simple integral

$$RL_a I^1 f(t) = \int_a^t \text{rect}(\tau) \, d\tau = \begin{cases} 0 & \text{if } t \leq -1/2 \\ t + 1/2 & \text{if } -1/2 < t < 1/2 \\ 1 & \text{if } t \geq 1/2, \end{cases}$$

which obviously doesn’t have compact support.
Example 3.9. Let us calculate the R-L differintegrals of two familiar functions

\[ f(t) = (t - a)^{y-1}, \quad \Re(y) > 0, \]
\[ g(t) = e^{\lambda t}, \quad \lambda > 0, \]

Now we can guess that \( a > -\infty \) is a good choice to differintegrate \( f \)

\[
\begin{align*}
R_L^a I^z f(t) &= \frac{1}{\Gamma(z)} \int_a^t (t - \tau)^{z-1}(\tau - a)^{y-1} d\tau \\
&= \frac{1}{\Gamma(z)} \int_0^1 (t - a - s(t - a))^{z-1}(s(t - a))^{y-1}(t - a) \, ds \\
&= \frac{1}{\Gamma(z)} (t - a)^{z+y-1} \int_0^1 (1 - s)^{s-1}(t - a)^{y-1} \, ds \\
&= \frac{\Gamma(y)}{\Gamma(z + y)} (t - a)^{z+y-1},
\end{align*}
\]

where substitution \( \tau = a + s(t - a) \), \( d\tau = (t - a)ds \) was done (the same as in the proof of Thm. 3.2) and the property (2.6) of the Beta function utilised.

Let \( n = \lceil \Re(z) \rceil \), then it follows with differentiation

\[
\begin{align*}
R_L^a D^z(t - a)^{y-1} &= \frac{d^n}{dt^n} R_L^a I^{n-z}(t - a)^{y-1} \\
&= \frac{d^n}{dt^n} \frac{\Gamma(y)}{\Gamma(n - z + y)} (t - a)^{n-z+y-1} \\
&= \frac{\Gamma(y)}{\Gamma(n - z + y)} \frac{\Gamma(n - z + y)}{\Gamma(-z + y)} (t - a)^{-z+y-1} \\
&= \frac{\Gamma(y)}{\Gamma(-z + y)} (t - a)^{-z+y-1}.
\end{align*}
\]

Next we choose \( a = -\infty \)

\[
\begin{align*}
R_L^a I^z g(t) &= \frac{1}{\Gamma(z)} \int_a^t (t - \tau)^{z-1}e^{\lambda \tau} d\tau \\
&= \frac{1}{\Gamma(z)} \int_0^1 (s/\lambda)^{z-1}e^{-s+\lambda t}(-\lambda)^{-1} \, ds \\
&= \frac{e^{\lambda t}}{\Gamma(z)\lambda^z} \int_0^\infty s^{z-1}e^{-s} \, ds \\
&= \frac{e^{\lambda t}}{\lambda^z},
\end{align*}
\]
where we substituted $t - \tau = s/\lambda$, $d\tau = -ds/\lambda$ and recognised the definition of the Gamma function (2.1). Now it follows with $n = \lceil \text{Re}(z) \rceil$

$$RL_a D^n e^\lambda t = \frac{d^n}{dt^n} \int_{-\infty}^{\infty} e^{\lambda t} \lambda^{n-z} f(t) dt = \lambda^n e^\lambda t.$$

Thus with a clever choice of the lower limits the familiar formulas for derivatives and integrals of polynomials and exponentials are recovered. It is quite obvious from the derivations above that different choices of $a$ would produce different functions, in fact, in most cases awfully complicated functions.

Next theorem from [11] is the key between R-L and Fourier differintegrals, which are introduced in the next chapter.

**Theorem 3.10.** If $f \in L_1$, then for $0 < \text{Re}(z) < 1$,

$$\mathcal{F}\{RL_{-\infty}^z f(t)\} = (i2\pi \nu)^{-z} \mathcal{F}\{f(t)\}, \quad \text{a.e.} \quad (3.9)$$

where $(i2\pi \nu)^z = e^{z \ln(2\pi |\nu|)+iz\pi^2 \text{sign}(\nu)}$.

**Proof.** The proof in [11, p. 138] is overly complicated and thus I have simplified it here. Let us find the Fourier transform

$$\mathcal{F}\{RL_{-\infty}^z f(t)\} = \frac{1}{\Gamma(z)} \int_{-\infty}^{\infty} e^{-i2\pi \nu t} \int_{-\infty}^{t} (t - \tau)^{z-1} f(\tau) d \tau dt$$

$$= \frac{1}{\Gamma(z)} \int_{-\infty}^{\infty} f(\tau) \int_{\tau}^{\infty} (t - \tau)^{z-1} e^{-i2\pi \nu t} dt d\tau$$

$$= \frac{1}{\Gamma(z)} \int_{-\infty}^{\infty} f(\tau) e^{-i2\pi \nu \tau} \int_{0}^{\infty} s^{z-1} e^{-i2\pi \nu s} ds d\tau$$

$$= \hat{f}(\nu) \frac{1}{\Gamma(z)} \int_{0}^{\infty} s^{z-1} e^{-i2\pi \nu s} ds$$

$$= (i2\pi \nu)^{-z} \hat{f}(\nu)$$

where in the second equality Fubini theorem was used and the integration limits changed in the same way as in the proof of 1.37. Then we substituted $t - \tau = s$, $dt = ds$ and finally used the result (2.8).

To extend the result to $\text{Re}(z) \leq 1$, the function and its integrals need to have zero mean value. To this end we introduce a special function space according to Lizorkin.
Definition 3.11. Lizorkin space $\Phi$ is defined as
\[ \Phi = \{ \phi \in S \mid \hat{\phi}^{(m)}(0) = 0, \ m = N_0 \}. \] (3.10)

The purpose of Lizorkin space is to make it invariant to fractional integration, that is, $\hat{\phi}$ does not become worse after division by $(i2\pi\nu)^z$.

Theorem 3.12. For $f \in \Phi$ equation (3.9) is valid for $\text{Re}(z) > 0$.

Proof. [11, p. 148-150]

The Fourier transforms of R-L derivatives is only briefly mentioned in [11] where the sequence of operators is also changed to justify the formula. In the next chapter we shall see that changing the order of differentiation and fractional integration might lead to a different operator, which the monumental book [11] dismisses entirely. So I will shortly elaborate the matter of Fourier transform of $\text{RL}_{-\infty} D^z f$ here in to a theorem.

Theorem 3.13. Let $m = \lfloor \text{Re}(z) \rfloor$ and $f \in L_1$. If $\text{RL}_{-\infty} D^{z-n} f \in L_1$ for $0 \leq n \leq m$, then
\[ \mathcal{F}\{ \text{RL}_{-\infty} D^z f(t) \} = (i2\pi\nu)^z \mathcal{F}\{ f(t) \}. \]

Proof. The assumptions on $f$ and its fractional derivatives enables us to use Theorems 2.27 and 3.10
\[ \mathcal{F}\{ \text{RL}_{-\infty} D^z f(t) \} = \mathcal{F}\{ D^m \text{RL}_{-\infty} I^{m-z} f(t) \} \]
\[ = (i2\pi\nu)^m \mathcal{F}\{ \text{RL}_{-\infty} I^{m-z} f(t) \} \]
\[ = (i2\pi\nu)^m (i2\pi\nu)^{z-m} \mathcal{F}\{ f(t) \} \]
\[ = (i2\pi\nu)^z \mathcal{F}\{ f(t) \}. \]

3.2 Caputo fractional derivative

The inverted order of fractional integration and ordinary differentiation results in a somewhat different fractional derivative operator. This was introduced by Caputo in geophysical studies [29, 30], although this order of operators has appeared in many other publications as well (even in [24]).

Definition 3.14. The Caputo fractional derivative operator is
\[ C^\alpha_a D^z f(t) = \text{RL}_{a} I^{n-z} \frac{d^n}{dt^n} f(t), \quad n = \lfloor \text{Re}(z) \rfloor. \] (3.11)
The Caputo fractional derivative is more useful in the theory of fractional differential equations due to it having initial values in terms of ordinary derivatives rather than fractional ones. Ordinary derivatives are the quantities that we can measure physically easily, and thus Caputo fractional derivative is more practical in physical situations. Even the Fourier transform of a Caputo derivative exists with simple assumptions on the function and its derivatives as opposed to R-L operator in Thm. 3.13.

**Theorem 3.15.** If $f^{(m)} \in L_1$ for $0 \leq m \leq n$, then for $m = \lfloor \text{Re}(z) \rfloor$

$$
\mathcal{F}\{ C_\infty D^z f(t) \} = (i2\pi \nu)^z \mathcal{F}\{ f(t) \}.
$$

**Proof.** Again we use Theorems 2.27 and 3.10 to prove the theorem

$$
\mathcal{F}\{ C_\infty D^z f(t) \} = \mathcal{F}\{ \text{RL}_\infty I^{m-z} f^{(m)}(t) \}
= (i2\pi \nu)^z m \mathcal{F}\{ f(t) \}
= (i2\pi \nu)^z (i2\pi \nu)^m \mathcal{F}\{ f(t) \}
= (i2\pi \nu)^z \mathcal{F}\{ f(t) \}.
$$

\[\square\]

### 3.3 Grünwald-Letnikov fractional derivative and integral

One form of fractional derivatives and integrals is especially important because of its easy applicability to numerical algorithms. It is based on differentials (1.21) and was developed independently by Anton Karl Grünwald in 1867 [31] and Aleksey Vasilievich Letnikov a year later [32].

In the equation (1.22) one can write the sum with the upper limit $\infty$, for \( \binom{n}{k} = 0 \), when $n \in \mathbb{N}$ and $n < k$. Now by replacing the factorials in the binomial coefficient (1.19) with the Gamma function

$$
\binom{z}{k} = \frac{\Gamma(z+1)}{k! \Gamma(z+1-k)},
$$
we arrive at a fractional derivative

$$
\lim_{h \to 0} \frac{1}{h^z} \sum_{k=0}^{\infty} (-1)^k \binom{z}{k} f(t - kh),
$$
if the series converges and the limit exists. Lower limit of fractional differentiation is now $-\infty$, since the series requires arbitrarily small values of $t$. By restricting the values of $h$ as it closes on zero, one has a definition with the lower limit $a$. 
Definition 3.16. The Grünwald-Letnikov fractional derivative operator is

$$GL_a D^z f(t) = \lim_{N \to \infty} \frac{\Delta^z h_N f(t)}{h_N^z}, \quad 0 < \text{Re}(z) \leq n,$$

(3.12)

where $h_N = (t - a)/N$ and $\Delta^z h_N f$ is the generalised difference

$$\Delta^z h_N f(t) = \sum_{k=0}^{N} (-1)^k \binom{z}{k} f(t - kh_N).$$

3.4 Equivalences of definitions

The calculation of Fourier transforms of different fractional operators with lower limit $-\infty$ has already implied their equivalence. We shall present here a few equivalence results without proofs for the interval $[a, b]$.

Theorem 3.17. Let $\alpha > 0$, $n = \lceil \alpha \rceil$ and $f \in C^n[a, b]$. Then

$$GL_a D^\alpha f(t) = RL_a D^\alpha f(t), \quad \text{for all } t \in (a, b).$$

Proof. [10, pp. 43-45]

Theorem 3.18. Let $\alpha > 0$ and $f \in C[a, b]$. Then

$$GL_a D^{-\alpha} f(t) = RL_a I^\alpha f(t), \quad \text{for all } t \in [a, b].$$

Proof. [10, p. 45]

Thus we see that the G-L definition defines fractional integration as well. In fact in [33] this fractional integral representation is derived by using the Riemann sums which approximate its Riemann integral.

The peculiar characteristic of the Caputo fractional derivative is revealed in the next theorem. It needs a definition of yet another function space.

Definition 3.19. Function $f : [a, b] \to \mathbb{C}$ has absolutely continuous $(n - 1)$st derivative, which we denote $f \in AC^n[a, b]$, if $f^{(n-1)}$ is differentiable a.e. and this derivative is $f^{(n)} \in L_1[a, b]$ and

$$f^{(n-1)}(t) = f^{(n-1)}(a) + \int_a^t f^{(n)}(\tau) \, d\tau.$$

(3.13)

We see that absolutely continuous functions are exactly those functions that have integration as the left inverse of their derivatives (except for the constant $f^{(n-1)}(a)$).
Theorem 3.20. Let $\alpha \geq 0$ and $m = \lceil \alpha \rceil$. If $f \in AC^m[a,b]$, then

$$C^a D^\alpha f(t) = RL D^\alpha \left( f(t) - \sum_{n=0}^{m-1} \frac{f^{(n)}(a)}{n!} (t-a)^n \right), \quad a.e.,$$

where the Taylor polynomial is 0 for the case $m = 0$.

Proof. [10, pp. 50-52]

Thus the inverted operator sequence causes exactly a difference of Taylor polynomial centered at $a$ in the two definitions.
4 Fractional calculus in frequency domain

4.1 Fourier differintegrals

Historically Fourier analysis was one of the first methods used to derive a formula for fractional derivatives of a function. In 1822 Fourier presented in his book *Théorie analytique de la chaleur* [14] a form of integral theorem (2.29) (here shown with modern notation)

\[
f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} f(\nu) \int_{-\infty}^{\infty} \cos(pt - p\nu) \, dp \, d\nu,
\]

and with it he presented the real order differintegrals of function \( f \)

\[
\frac{d^\alpha}{dt^\alpha} f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} f(\nu) \int_{-\infty}^{\infty} p^\alpha \cos \left( pt - p\nu + \frac{\alpha \pi}{2} \right) \, dp \, d\nu.
\]

I will use the modern Fourier integral theorem (2.29) instead and the result (2.40) to define differintegrals. This definition will of course apply to functions \( f \) whose Fourier transforms \( \hat{f} \) and inverse transforms of \( \hat{f} \) multiplied by \((i2\pi\nu)^z\) exist. For the complex exponentiation this means that we again choose only the principal branch of the function.

**Definition 4.1.** Let \( z \in \mathbb{C} \) and Fourier transform of \( f \) is \( \hat{f}(\nu) = \mathcal{F}\{f(t)\} \).

*The Fourier differintegral of \( f \) is*

\[
^{F}D^z f(t) = \mathcal{F}^{-1}\{(i2\pi\nu)^z \hat{f}(\nu)\} = \int_{-\infty}^{\infty} (i2\pi\nu)^z \hat{f}(\nu) e^{i2\pi\nu t} \, d\nu \tag{4.1}
\]

where \((i2\pi\nu)^z = e^{z \ln(2\pi|\nu|) + iz \frac{\pi}{2} \text{sign}(\nu)}\).

In modern fractional analysis this is not a frequently used definition of a differintegral. The reason is that the Fourier differintegral is in many cases equal to definitions in the time domain with the lower limit \(-\infty\) which shall also be demonstrated shortly. The author has found a similar definition for example in [34, 35]. For \( \text{Re}(z) < 0 \) it is convenient to assume that the signal has zero mean value, i.e. \( \int_{-\infty}^{\infty} f(t) \, dt = \hat{f}(0) = 0 \), or otherwise the zero frequency could be amplified in to a Dirac delta distribution (although we can manage this with our knowledge on distribution theory!).

In references [36, 37, 38, 39] Lahdelma et.al have defined real order derivatives and integrals of function \( X e^{i\omega t} \) as

\[
(i\omega)^\alpha X e^{i\omega t},
\]
and used this definition for analysing vibration measurements. When a signal has multiple frequencies, one can see $X$ as its Fourier transforms amplitude for each frequency $X = \hat{f}(\nu)$ and the differintegral of the signal is then according to Definition 4.1 the "sum", i.e. an integral, of all these differintegrals at different frequencies. The usefulness of Definition 4.1 in vibration mechanics and other physical applications is that the object may vibrate with any angular frequency $\omega = 2\pi \nu \in \mathbb{R}$ and the effect of these frequencies into the fractional derivatives and integrals is plainly seen. Even more importantly, it does not include a built-in ansatz, as our definitions in the time domain do. This is simply the fact that necessarily differintegrals are zero at $a$. In the case of vibration signals this means that with the Fourier definition one takes into account the whole past history of the vibration and we don’t assume that vibration amplitude was zero at some exact time in the past. In the article [36] Lahdelma has studied the transition of the real order derivatives and integrals $(i\omega)^\alpha X e^{i\omega t}$ in the complex plane as the order of differintegration changes and in the article [40] visualised the complex differintegrals $(i\omega)^z X e^{i\omega t}$.

An immediate consequence of the definition is the linearity of the Fourier differintegral. Furthermore, the semigroup property is a breeze to prove.

**Theorem 4.2.** If $F^D z + y f(t)$ exists for some $\Re(z), \Re(y) > 0$ (or $\Re(z), \Re(y) < 0$), then

$$F^D y F^D z f(t) = F^D z + y f(t) \quad a.e.$$ 

**Proof.** We know that the inverse transform of $(i2\pi \nu)^{z+y} \hat{f}$ exists, and therefore $F^{-1}\{(i2\pi \nu)^z \hat{f}(\nu)\}$ exists as well and

$$F^D y F^D z f(t) = F^{-1}\{F^{-1}\{(i2\pi \nu)^y F\{F^{-1}\{(i2\pi \nu)^z \hat{f}(\nu)\}\}\}\}$$

$$= F^{-1}\{(i2\pi \nu)^y (i2\pi \nu)^z \hat{f}(\nu)\}$$

$$= F^D z + y f(t).$$

\[\square\]

In similar fashion we see also that for $\Re(z) > 0$ the operator $F^D z f$ is the left inverse of $F^D z f$, but the converse might not hold since $F^D z f$ attenuates the zero frequency.
Example 4.3. Let $\lambda \in \mathbb{R}$ and $z \in \mathbb{C}$, then
\[
F D^z e^{it\lambda} = F^{-1} \left\{ (i2\pi\nu)^z \delta_{\lambda/2\pi} \right\} \\
= F^{-1} \left\{ (i2\pi \frac{\lambda}{2\pi})^z \delta_{\lambda/2\pi} \right\} \\
= (i\lambda)^z F^{-1} \{ \delta_{\lambda/2\pi} \} \\
= (i\lambda)^z e^{it\lambda},
\]
where we used the multiplication property of Dirac delta distribution (1.12)
\[
\langle (i2\pi\nu)^z \delta_{\lambda/2\pi}, \varphi \rangle = (i2\pi \frac{\lambda}{2\pi})^z \varphi(\lambda/2\pi), \text{ for } \varphi \in S.
\]
Now we are able to calculate the differintegrals of cosine and sine functions as well
\[
F D^z \cos(t\lambda) = F D^z \frac{1}{2} (e^{it\lambda} + e^{-it\lambda}) \\
= \frac{1}{2} ((i\lambda)^z e^{it\lambda} + (-i\lambda)^z e^{-it\lambda}) \\
= \frac{1}{2} \lambda^z (i^z e^{it\lambda} + (-i)^z e^{-it\lambda}) \\
= \frac{1}{2} \lambda^z (e^{i(t\lambda + z\pi/2)} + e^{-i(t\lambda + z\pi/2)}) \\
= \lambda^z \cos \left( t\lambda + z \frac{\pi}{2} \right),
\]
\[
F D^z \sin(t\lambda) = F D^z \frac{1}{i2} (e^{it\lambda} - e^{-it\lambda}) \\
= \lambda^z \sin \left( t\lambda + z \frac{\pi}{2} \right).
\]

If function $f \in C^n$ then how smooth are its real order derivatives? It should not be surprising that derivatives $F D^\alpha f$, $0 < \alpha < n$ are continuous, but more interesting is, that derivatives $n < \alpha < n + 1$ are in most cases continuous as well.

**Theorem 4.4.** If for the signal $f$ its $n + 1$th derivative is $f^{(n+1)} \in L_1$ and $f^{(n+1)} \in BV(\mathbb{R})$, then functions $F D^\alpha f \in C$ for $n < \alpha < n + 1$.

**Proof.** Let us write $\alpha = n + \gamma$, where $0 < \gamma < 1$ and consider the Fourier derivative
\[
F D^\alpha f(t) = F^{-1} \{(i2\pi\nu)^\alpha \hat{f}(\nu)\} = F^{-1} \{(i2\pi\nu)^{n+\gamma} \hat{f}(\nu)\}.
\]
Now according to Corollary 2.33 we have $\hat{f} \in O(1/|\nu|^{n+2})$ and therefore
\[
(i2\pi\nu)^{n+\gamma} \hat{f} \in O(1/|\nu|^{2-\gamma}). \tag{4.2}
\]
Since \( f \in L_1 \) it follows from R-L lemma 2.13 that \( \hat{f} \in C \) and since \((i2\pi \nu)^{n+\gamma} \in C\) as well, we have \((i2\pi \nu)^{n+\gamma} \hat{f} \in C\). This means that \((i2\pi \nu)^{n+\gamma} \hat{f} \) cannot have local singularities. This together with (4.2) means that it is integrable, i.e. \((i2\pi \nu)^{n+\gamma} \hat{f} \in L_1 \) and therefore it again follows from R-L lemma (applied to inverse Fourier transform) that \( F^D f \in C \).

\[\square\]

### 4.2 Weyl differintegrals

In practical calculations one only has a discretised sample of a signal and it is also of finite length. In this case it is often plausible to extend the signal periodically to utilise the theory of Fourier series.

One can calculate derivatives of a function term by term from their series representations for smooth enough functions. For Fourier series the following theorem is valid.

**Theorem 4.5.** Let \( f \) be continuous \( T \)-periodic function and \( f' \) piecewise smooth. Then \( f' \) has a pointwise converging series representation.

\[
\frac{1}{2}(f'(t-) + f'(t+)) = \sum_{k=-\infty}^{\infty} \left( \frac{i2\pi k}{T} \right) c_k e^{i2\pi kt/T}.
\]

**Proof.** Presented in the book [5, p. 221] again in \( 2\pi \)-periodic case. \[\square\]

One can therefore often differentiate functions term by term using their Fourier series. For smooth enough functions, i.e. \( f^{(n)} \) piecewise smooth, one can even calculate the \( n \)th derivative

\[
f^{(n)}(t) = \frac{d^n}{dt^n} \sum_{k=-\infty}^{\infty} c_k e^{i2\pi kT/T} = \sum_{k=-\infty}^{\infty} \left( \frac{i2\pi k}{T} \right)^n c_k e^{i2\pi kT/T} \quad \text{a.e.} \tag{4.3}
\]

Equation (4.3) motivates a definition of fractional derivative for periodic functions, which I call by the name of Weyl fractional derivative. Hermann Weyl studied in an article in 1917 [41] the use of Fourier series especially for fractional integration of 1-periodic functions.

**Definition 4.6.** Weyl differintegral for \( T \)-periodic functions is

\[
WD^z f(t) = \sum_{k=-\infty}^{\infty} \left( \frac{i2\pi k}{T} \right)^z c_k e^{i2\pi kT/T}, \tag{4.4}
\]

where \((i2\pi k/T)^z = e^{z\ln(2\pi|k|/T)+iz\pi/2\text{sign}(k)}\).
Again for fractional integration, i.e. the case $\text{Re}(z) < 0$ it is useful to consider that the signal has zero mean value over one period
\[
c_0 = \frac{1}{T} \int_0^T f(t) \, dt = 0.
\]
This means that we remove constants from the signals to make the definition meaningful. There is no such restriction on derivatives $W^D\alpha$ since differentiated signals will have $c_0$ multiplied by zero. Weyl actually studied the fractional integrals of Fourier series in convolution form and this presentation is also used in [11].

Theorem 2.11 now enables us to prove the analogous version of Theorem 4.4.

**Theorem 4.7.** Let $f \in L_{1,\text{loc}}$ be a $T$-periodic function and $f^{(n+1)} \in BV[0,T]$. Then functions $W^D\alpha f \in C$ for $n < \alpha < n + 1$.

**Proof.** Since $f \in L_{1,\text{loc}}$ defines a tempered distribution, we know from (2.43) that the Fourier coefficients for the distributional derivative $f^{(n+1)}$ exist. Furthermore since $f^{(n+1)} \in BV[0,T]$ we know by Theorem 2.11 that those coefficients satisfy
\[
\left( \frac{i2\pi k}{T} \right)^{n+1} c_k \in O(1/|k|),
\]
and therefore $c_k \in O(1/|k|^{n+2})$. Let us write $\alpha = n + \gamma$, where $0 < \gamma < 1$ and consider the Weyl derivative
\[
W^D\alpha f(t) = \sum_{k=-\infty}^{\infty} \left( \frac{i2\pi k}{T} \right)^{n+\gamma} c_k e^{i2\pi kt/T},
\]
where the coefficients are now
\[
\left( \frac{i2\pi k}{T} \right)^{n+\gamma} c_k \in O \left( \frac{1}{|k|^{2-\gamma}} \right),
\]
and thus the series $W^D\alpha f$ converges absolutely. Now it follows from Theorem 2.6 that this series defines a continuous function. \qed

### 4.3 Equivalences of definitions

Based on our calculations of Fourier transforms for different definitions in Chapter 3 and Theorem 2.22 we can state that if the Fourier transform of R-L or Caputo differintegral equals $(i2\pi \nu)^z \hat{f}$ and this is in $S'$, then it is equal to
\( F^\alpha D^\alpha f \) in the sense of tempered distributions. If \( F^\alpha D^\alpha f \in L_1 \), then the equality is a.e.

How are \( F^\alpha D^\alpha \) and \( W^\alpha D^\alpha \) related? If the signal \( f \) is periodic, it usually means that \( f \notin S(\mathbb{R}) \) and \( f \notin L_1(\mathbb{R}) \), but \( f \) might still define a distribution.

**Theorem 4.8.** If \( f \in L_{1,\text{loc}} \) is \( T \)-periodic and either \( F^\alpha D^\alpha f \) or \( W^\alpha D^\alpha f \) exists (as a function or distribution) for some \( z \in \mathbb{C} \), then they both exist and are equal

\[
F^\alpha D^\alpha f = W^\alpha D^\alpha f, \quad \text{in } S'.
\]

**Proof.** This proof goes along similar lines as the calculations in Example 2.30. According to Lemma 1.30 function \( f \in L_{1,\text{loc}} \) defines a tempered distribution and therefore it has a Fourier transform. It also has a Fourier series representation which we use to calculate its Fourier transform

\[
\mathcal{F}\{f(t)\} = \mathcal{F}\left\{\sum_{k=-\infty}^{\infty} c_k e^{i2\pi k t / T}\right\} = \sum_{k=-\infty}^{\infty} c_k \mathcal{F}\{e^{i2\pi k t / T}\} = \sum_{k=-\infty}^{\infty} c_k \delta_{k/T},
\]

which is a series of pulses \( 1/T \) apart and we have found the representation of Theorem 2.24. Furthermore

\[
F^\alpha D^\alpha f(t) = \mathcal{F}^{-1}\left\{(i2\pi \nu)^\alpha \sum_{k=-\infty}^{\infty} c_k \delta_{k/T}\right\}
\]

\[
= \sum_{k=-\infty}^{\infty}\left(\frac{i2\pi k}{T}\right)^\alpha c_k \mathcal{F}^{-1}\{\delta_{k/T}\}
\]

\[
= \sum_{k=-\infty}^{\infty}\left(\frac{i2\pi k}{T}\right)^\alpha c_k e^{i2\pi k t / T} = W^\alpha D^\alpha f(t),
\]

where we used the multiplication property of Dirac delta distribution (1.12)

\[
\langle (i2\pi \nu)^\alpha \delta_{k/T}, \varphi \rangle = (i2\pi k/T)^\alpha \varphi(k/T), \text{ for } \varphi \in S.
\]

The attained equality is in the sense of tempered distributions. When the differintegrals are in the space \( L_{1,\text{loc}} \) we have equality a.e. in \( \mathbb{R} \).

**Theorem 4.9.** Let \( f \in L_{1,\text{loc}} \) be a \( T \)-periodic function and \( f_T = f \cdot 1_{[0,T]} \) and \( z \in \mathbb{C} \). If \( \text{Re}(z) < 0 \) let us also assume that \( f \) has zero mean value, i.e. \( c_0 = 0 \). If \( W^\alpha D^\alpha f \) exists for all \( t \in \mathbb{R} \) and \( F^\alpha D^\alpha f_T \) and \( RL^\alpha D^\alpha f_T \) are equal and continuous, we have

\[
F^\alpha D^\alpha f_T = RL^\alpha D^\alpha f_T = (W^\alpha D^\alpha f) \cdot 1_{[0,T]}, \quad \text{for all } t \in \mathbb{R}.
\]
Proof. Let us write the Fourier coefficients of $W D^z f$
\[ c_{k,z} = \left( \frac{i2\pi k}{T} \right)^z \]
\[ c_k = \left( \frac{i2\pi k}{T} \right)^z \frac{1}{T} \int_0^T f(t)e^{-i2\pi kt/T} \, dt, \quad k = 0, \pm 1, \pm 2, \ldots \]

Let us denote the Fourier transform of $F D^z f_T$
\[ F_{T,z}(\nu) = (i2\pi \nu)^z \widehat{f}_T(\nu) = (i2\pi \nu)^z \int_0^T f(t)e^{-i2\pi \nu t} \, dt. \]

So we have
\[ c_{k,z} = \frac{1}{T} F_{T,z} \left( \frac{k}{T} \right), \quad k = 0, \pm 1, \pm 2, \ldots \]
and we can write the periodic summation of $F D^z f_T$ according to Poisson summation formula (2.50)
\[ \sum_{m=-\infty}^{\infty} F D^z f_T(t + mT) = \frac{1}{T} \sum_{k=-\infty}^{\infty} F_{T,z} \left( \frac{k}{T} \right) e^{i2\pi kT/T} \]
\[ = \sum_{k=-\infty}^{\infty} c_{k,z} e^{i2\pi kT} \]
\[ = W D^z f(t). \]

Let us check the conditions of Theorem 2.35. The left side of the Poisson summation formula is a continuous function because $F D^z f_T$ is continuous and has bounded support on $[0, T]$ according to Thm. 3.8. The convergence of the right side for all $t \in \mathbb{R}$ is presumed.

We could arrive at the equivalence 4.9 by multiplying a $C^\infty$ function with $w_2$ (2.57). A situation where there clearly can’t be an equivalence is the case when $f$ is not periodic or with bounded support. Then one can’t form the Fourier series of $f$ and thus Weyl differintegrals are not defined.

I present one more equivalence from [11]. In this case we interpret the R-L integral as a special limit.

**Theorem 4.10.** Let $f \in L_1(0, T)$ and $T$-periodic and it has zero mean value, i.e. $c_k = 0$. Then
\[ W D^\alpha f(t) = \frac{R_L}{-\infty} I^\alpha f(t), \quad 0 < \alpha < 1, \]
when $R_L I^\alpha f$ is understood as
\[ \frac{1}{\Gamma(\alpha)} \int_{-\infty}^t (t - \tau)^{\alpha-1} f(\tau) \, d\tau = \frac{1}{\Gamma(\alpha)} \lim_{n \to \infty} \int_{t-2\pi n}^t (t - \tau)^{\alpha-1} f(\tau) \, d\tau. \]

4.4 Numerical algorithm

The numerical work in this thesis favours the frequency domain methods for calculating differintegrals. The reasons are the vibration measurements in question, accuracy and speed. Vibration measurements from a rotating machine are well suitable for Fourier analysis because of their natural periodic nature. The use of DFT is a spectral method which is both accurate (if we know that there is no aliasing in the signal) and fast because of the FFT. To generate such a spectral method in the time domain typically means that if the differintegral is approximated as a matrix-vector multiplication, this matrix is full of values because differintegrals are not local operators. This typically means a calculation cost of $O(n^2)$ operations. Moreover there are multiple sources of error. Besides discretising, there is also error from cutting the calculation in the definitions which have lower limit $a = -\infty$.

The following algorithm by the author has appeared in [42] and therefore is presented only shortly here. Some points which were not discussed in that paper are more thoroughly studied in this thesis.

The starting point for numerical calculations are the measured discrete values \( \{f_n\}_{n=0}^{N-1} \) and calculated values \( \{F_k\}_{k=0}^{N-1} \) of the DFT. The IDFT formula (2.53) then defines a trigonometric interpolation of the original function if we write it with a continuous time variable \( t \). The problem is with the values \( k \) because the \( N \)-periodicity of IDFT means that replacing \( k \) with \( k + m_kN \), where \( m_k \in \mathbb{Z} \), does not change the values of the interpolant at the original discrete points. The trigonometric interpolation is thus

\[
f(t) = \sum_{k=0}^{N-1} F_k e^{2\pi i (k + m_k N)t/T},
\]

which is equal to the values \( f_n \) for all \( m_k \in \mathbb{Z} \) at the points \( t_n = \Delta t \cdot n \), \( n = 0, 1, \ldots, N-1 \) and \( T = \Delta t \cdot N \). Bigger \( k \) will mean more oscillation between sample points. This would distort the differintegrals undesirably. The smoothness requirement that \( \frac{1}{T} \int_0^T |f'(t)|^2 \, dt \) is minimised or the frequencies are limited to the interval

\[
|k + m_k N| \leq \frac{N}{2},
\]

will result in the same unique trigonometric interpolation of minimal oscil-
\[ f(t) = F_0 + \sum_{0 < k < N/2} (F_k e^{i2\pi kt/T} + F_{N-k} e^{-i2\pi kt/T}) + F_{N/2} \cos \left( \frac{\pi N t}{T} \right) \]
\[ = \sum_{0 \leq k < N/2} F_k e^{i2\pi kt/T} + \sum_{-N/2 < k < 0} F_{N+k} e^{i2\pi kt/T} + F_{N/2} \cos \left( \frac{\pi N t}{T} \right). \]

These ideas were presented in the reference [43] where they are utilised in the calculation of first and second derivatives of the function and solving differential equations numerically. In the article [42] the author differintegrated the interpolant (4.7) \( z \in \mathbb{C} \) times with the Fourier definition to get
\[ f^{(z)}(t) = \sum_{0 \leq k < N/2} \left( \frac{2\pi k i}{T} \right)^z F_k e^{i2\pi k t/T} + \sum_{-N/2 < k < 0} \left( \frac{2\pi k i}{T} \right)^z F_{N+k} e^{i2\pi k t/T} + \left( \frac{\pi N}{T} \right)^z F_{N/2} \cos \left( \frac{\pi N t}{T} + \frac{z \pi}{2} \right). \]

In this thesis the author has shown rigorously how the complex exponential and cosine functions are differintegrated in the Example 4.3.

The algorithm then follows by computing the interpolation at the sample points \( t_n = \frac{T}{N} n \)
\[ f^{(z)}_n = \sum_{0 \leq k < N/2} \left( \frac{2\pi k i}{T} \right)^z F_k e^{i2\pi k n/N} + \sum_{-N/2 < k < 0} \left( \frac{2\pi k i}{T} \right)^z F_{N+k} e^{i2\pi k n/N} + \left( \frac{\pi N}{T} \right)^z F_{N/2} \cos \left( \pi n + \frac{z \pi}{2} \right), \]
where the last term is
\[ \left( \frac{\pi N}{T} \right)^z F_{N/2} \frac{1}{2} (e^{izn + iz\pi} + e^{-izn - iz\pi}) \]
\[ = \left( \frac{\pi N}{T} \right)^z \frac{1}{2} (i^z + (-i)^z) F_{N/2}(e^{iz\pi})^n \]
\[ = \left( \frac{\pi N}{T} \right)^z \cos \left( \frac{z \pi}{2} \right) F_{N/2}(-1)^n. \]

After these manipulations the formula is again in the form of IDFT and thus the discrete differintegrated values may be calculated. The multiplier for the last term is zero only when cosine is zero, which happens at \( z = 2n - 1, n \in \mathbb{N} \).
These calculations were performed just because of a single problematic term in the DFT. If $N$ is odd, nothing special is required. However, to utilise the speed of the FFT algorithm we need $N$ that is a power of two [20].

Algorithm 1. Differintegration in frequency space

1. Calculate the DFT $F_k$, $0 \leq k \leq N-1$ of the sequence $f_n$, $0 \leq n \leq N-1$ with the FFT algorithm

   $$F_k = \mathcal{F}\{f_n\}.$$  

2. Calculate a new sequence $G_k$, $0 \leq k \leq N-1$ and $G_0 = 0$

   $$G_k = \left(\frac{2\pi ki}{T}\right)^z F_k, \quad 0 < k < N/2$$

   $$G_{N+k} = \left(\frac{2\pi ki}{T}\right)^z F_{N+k}, \quad -N/2 < k < 0$$

   $$G_{N/2} = \left(\frac{\pi N}{T}\right)^z \cos\left(\frac{\pi}{2}\right) F_{N/2} \quad \text{(if } N \text{ is even)}.$$  

3. Calculate the differintegrated sequence $x_n^{(z)}$ with the IFFT

   $$f_n^{(z)} = \mathcal{F}^{-1}\{G_k\}.$$  

Suppose that the signal $f \in L_1[0,T]$ is real-valued and bandlimited to the frequency interval $[-M/2T, M/2T]$. It may have been for example filtered with a lower pass filter before sampling so that it is actually representable with a finite number of frequencies (and there is no aliasing) and can be written as a finite sum

$$f(t) = \sum_{k=-M/2}^{M/2} c_k e^{i2\pi kt/T},$$  

for which the Weyl differintegral exists for all $z \in \mathbb{C}$

$$W^D_z f(t) = \sum_{k=-M/2}^{M/2} \left(\frac{i2\pi k}{T}\right)^z c_k e^{i2\pi kt/T}.$$  

58
But we also have the differintegral of the trigonometric interpolation of minimal oscillation based on the DFT

\[ f^{(z)}(t) = \sum_{0 \leq k < N/2} \left( \frac{i2\pi k}{T} \right)^z F_k e^{i2\pi kt/T} + \sum_{-N/2 < k < 0} \left( \frac{i2\pi k}{T} \right)^z F_{N+k} e^{i2\pi kt/T} \]

\[ + \left( \frac{\pi N}{T} \right)^z F_{N/2} \cos \left( \frac{\pi N t}{T} + \frac{\pi}{2} \right), \]

and the result from Thm. 2.38 shows that if \( N > M \), Algorithm 1 will produce exactly the Weyl differintegral of the function. The needed sampling frequency (Nyquist-frequency) is therefore higher than \( 2 \cdot M/(2T) = M/T \).
5 Norms, means and other features of signals

5.1 Generalised $l_p$ norms and Hölder means

We will now look into norms and means of discrete vectors in more detail. Let $x \in l_p(\mathbb{Z})$ for $1 \leq p \leq \infty$. Then the $l_p$ norms of Definition 1.7 decrease if $p$ increases

$$\|x\|_p \geq \|x\|_q,$$  \hspace{1cm} \text{when } p < q, \hspace{1cm} (5.1)

and equality is achieved iff $x$ is a constant vector $x = (x, x, \ldots)$. Notice also that this inequality holds only for $l_p$ spaces which have null measure and the inequality is reversed in the case of $L_p[a, b]$ spaces which have finite measure. [44]

From now on we will focus on $l_p$ spaces with finite number $N$ of dimensions. Then we immediately notice that the $l_\infty$ norm (1.2) is simplified to

$$\|x\|_\infty = \max_{n=1,...,N} |x_n|.$$

and the $l_2$ norm is the Euclidean length

$$\|x\|_2 = \left( \sum_{n=1}^{N} |x_n|^2 \right)^{1/2}.$$

One can generalise $l_p$ norms with weight factors $w_n$

$$\|x\|_{p,w} = \left( \sum_{n=1}^{N} w_n |x_n|^p \right)^{1/p},$$

for which $\sum_{n=1}^{N} w_n = 1$. Weight factors $w_n = 1/N$ are especially interesting, and Lahdelma et.al have used them to define a useful norm [38, 39].

**Definition 5.1.** Let $1 \leq p < \infty$. The $\overline{l}_p$ norm is

$$\|x\|_{p,\overline{w}} = \left( \sum_{n=1}^{N} \frac{1}{N} |x_n|^p \right)^{1/p} = \frac{1}{\sqrt{N}} \|x\|_p. \hspace{1cm} (5.2)$$

Norm $\overline{l}_p$ is therefore a $1/N$ weighted $l_p$ norm and it is also called power mean or Hölder mean after Otto Hölder. The $1/N$ weights makes it more reasonable to compare vectors of different length with this norm, whereas increasing $N$ will always increase $l_p$ norm. The order of growth is reversed when compared to $l_p$ norm, i.e. $\overline{l}_p$ norm grows with $p$

$$\|x\|_{p,\overline{w}} \leq \|x\|_{q,\overline{w}}, \hspace{1cm} \text{when } p < q, \hspace{1cm} (5.3)$$
but the limit \( p \to \infty \) is the same

\[
\|x\|_\infty = \lim_{p \to \infty} \|x\|_{p, \frac{1}{p}} = \max_{n=1, \ldots, N} |x_n|,
\]

(5.4)

which is the peak value [45]. Notice that in the book [45] the Hölder mean is defined with arbitrary weights and all the theorems are also proved with that definition. Only equal weights will be utilised in this thesis. Norm \( p = 1 \) is the absolute mean or arithmetic mean

\[
\|x\|_{1, \frac{1}{N}} = \frac{1}{N} \sum_{n=1}^{N} |x_n|,
\]

(5.5)

and norm \( p = 2 \) is the root mean square (rms)

\[
\|x\|_{2, \frac{1}{N}} = \left( \frac{1}{N} \sum_{n=1}^{N} |x_n|^2 \right)^{1/2},
\]

(5.6)

which is used to describe for example the energy content of a signal. Thus such traditional features such as peaks and rms values accentuate certain features of vibration signals.

There should be no reason to limit our attention to 'so few' features, when \( l_p \) norms can be calculated for any \( p \geq 1 \). If we don’t mind violating the axioms of norms, we can extend these calculations even further. If \( p < 1 \), the triangle inequality is no longer satisfied and the following features are called quasinorms.

**Definition 5.2.** Quasinorm satisfies axioms N1 and N3 of norms 1.1, but axiom N2 is replaced with

\[
\|x + y\| \leq k(\|x\| + \|y\|)
\]

with some constant \( k > 1 \).

Inequality (5.3) is valid for all \( p \in \mathbb{R} = -\infty \cup \mathbb{R} \cup \infty \), if we define all the limiting values correctly. Quasinorm \( p = 0 \) is meaningful as a limit \( p \to 0 \) and turns out to be the geometric mean

\[
\|x\|_{0, \frac{1}{N}} = \lim_{p \to 0} \|x\|_{p, \frac{1}{p}} = \left( \prod_{n=1}^{N} |x_n| \right)^{1/N}.
\]

(5.7)

Recognising quasinorm \( p = -1 \) as the harmonic mean

\[
\|x\|_{-1, \frac{1}{N}} = \frac{N}{\sum_{n=1}^{N} \frac{1}{|x_n|}},
\]

(5.8)
we see that the result (5.3) implies the inequality of arithmetic, geometric and harmonic means. Limit $p \to -\infty$ also exists and is naturally the minimum value $\parallel x \parallel_{-\infty} = \lim_{p \to -\infty} \parallel x \parallel_{\frac{1}{p}} = \min_{n=1,...,N} |x_n|$. (5.9)

Lahdelma et.al [46] have utilised the cases $0 \leq p < 1$ to good effect. With $p < 0$ none of the vector values can be 0. Let’s collect these generalisations under one definition, which is the Hölder mean or generalised $l_p$ norm. The word generalised then not only refers to weights $1/N$ but also to inclusion of quasinorms.

Definition 5.3. For all $p \in \mathbb{R}$ we define Hölder mean or $l_p$ norm of vector $x \in \mathbb{R}^N$, $x_n \neq 0 \forall n$ as

$$
\parallel x \parallel_{p, \frac{1}{N}} = \begin{cases} 
\left( \sum_{n=1}^{N} \frac{1}{N} |x_n|^p \right)^{1/p} & \text{if } p \in \mathbb{R} \setminus \{0\} \\
\left( \prod_{n=1}^{N} |x_n|^p \right)^{1/N} & \text{if } p = 0 \\
\max_{n=1,...,N} |x_n| & \text{if } p = \infty \\
\min_{n=1,...,N} |x_n| & \text{if } p = -\infty
\end{cases}
$$

(5.10)

Some basic properties of Hölder means are collected in the next theorem.

Theorem 5.4. For vectors $x, y \in \mathbb{R}^N$ the i.e. $x \leq y$ states that $x_n \leq y_n \forall 1 \leq n \leq N$. Then the Hölder mean $p \in \mathbb{R}$ satisfies the following properties

- (Continuity) $\lim_{y \to 0} \parallel x + y \parallel_{p, \frac{1}{N}} = \parallel x \parallel_{p, \frac{1}{N}}$;
- (Homogeneity) $\parallel \lambda x \parallel_{p, \frac{1}{N}} = \lambda \parallel x \parallel_{p, \frac{1}{N}}, \lambda > 0$;
- (Internality) $\min_{n=1,...,N} |x_n| \leq \parallel x \parallel_{p, \frac{1}{N}} \leq \max_{n=1,...,N} |x_n|$, with equality iff $x$ is a constant vector $x = (x, x, \ldots, x)$;
- (Monotonicity) $x \leq y \implies \parallel x \parallel_{p, \frac{1}{N}} \leq \parallel y \parallel_{p, \frac{1}{N}}$, with equality iff $x = y$;
- (Reflexivity) If $x$ is a constant vector $x = (x, x, \ldots, x)$, then $\parallel x \parallel_{p, \frac{1}{N}} = x$;
• (Symmetry) $\|x\|_{p, \frac{1}{N}}$ is not changed if the elements of $x$ are permuted.

**Proof.** Stated as immediate in [45, p. 175]. \square

A generalisation of the triangle inequality is also valid. This is also sometimes referred to as Minkowski’s inequality.

**Theorem 5.5.** Let $x, y \in \mathbb{R}^N$ and $1 \leq p \leq \infty$. Then

$$\|x + y\|_{p, \frac{1}{N}} \leq \|x\|_{p, \frac{1}{N}} + \|y\|_{p, \frac{1}{N}},$$

with equality iff either

1. $p = 1$,
2. $1 < p < \infty$ and $x$ and $y$ are linearly dependent, or
3. $p = \infty$ and for some $m$, $1 \leq m \leq N$, $\max_{n=1,\ldots,N} |x_n| = x_m$ and $\max_{n=1,\ldots,N} |y_n| = y_m$.

If $p < 1$ then reverse inequality holds.

**Proof.** [45, p. 213] \square

The reversed triangle inequality for the case $p < 1$ is more informative than classification of these cases as quasinorms, where constant $k > 1$ can be as big as necessary.

From the results above we can see that Hölder means or $\ell_p$ norms form a continuous measurement scale for the magnitude of vector signals. Large values of $p$ amplify the biggest elements of vectors and small values of $p$ the smallest ones. Compare this with the Fourier and Weyl differintegrals, where the bigger degrees of differintegral amplify high frequencies and smaller degrees amplify lower frequencies. We will see in Chapter 6 that vibration signals respond quite nicely to these calculations and that we can search optimal degrees of differintegral and $\ell_p$ norms for early detection of different faults.
5.2 Measurement index

Lahdelma presented in 1992 [47] the measurement index, or MIT index, utilising rms values of displacement and its derivatives and integrals of order \( n \in \mathbb{N} \). Later it has been generalised to \( l_p \) norms and real order derivatives and integrals [38, 39]. Thus it is formulated as

\[
\tau^{MIT}_{\alpha_1, \alpha_2, \ldots, \alpha_n} = \frac{1}{\tau} \sum_{k=1}^{n} b_{\alpha_k} \frac{\|x^{(\alpha_k)}\|_{p_k}}{\|r^{(\alpha_k)}\|_{p_k}},
\]

(5.11)

where \( \sum_{k=1}^{n} b_{\alpha_k} = 1 \), \( \tau \) is signal length and vector \( r \) is a reference signal of the machine in good condition. MIT index can also be combined with other quantities that are related to its condition, such as temperature, pressure or some statistical features of different signals. The weight factors \( b_{\alpha_k} \) are used to take into account the severity of different faults occurring at different \( \alpha_k \) and \( p_k \). The inverse of MIT index is defined as condition index SOL

\[
SOL = \frac{1}{MIT}.
\]

(5.12)

When the machine is in good condition, the two indices are equal, i.e. \( MIT = SOL = 1 \). As its condition becomes weaker, growth of the MIT index and decrease of the SOL index usually occur. The word 'usually' is necessary, since, for example, minor wear on new parts may cause smoother operation at the beginning of their usage and hence temporarily reduce MIT. The downside of MIT is that in some cases the base signals for normal condition might be hard to evaluate, although this is the case with almost every condition monitoring method. One may also need to monitor the device for long periods to determine good (preferably optimal) values for \( \alpha, p, N, \tau \) and \( b \).
6 Condition monitoring of a load haul dumper front axle

6.1 Data acquisition and acceleration signals

The experimental part consists of analysing vibration measurements from the front axle of a load haul dumper (LHD) working underground in the Pyhäsalmi mine. The measurement has been running since 22.4.2013 and a shorter period of 271 days (22.4.2013 - 18.1.2014) of these measurements have already been analysed in the paper [42]. Initial problems with the measurement rig, which caused some short stoppages to data acquisition, are mentioned in that paper as well. There have been no major problems after 18.1.2014. In this thesis I analyse these measurements from the period 22.4.2013 - 13.5.2014, which covers 368 days.

The measurement rig consists of four SKF CMPT 2310 accelerometers which were mounted externally onto the LHD’s front axle housing. They measure horizontal and vertical vibrations near the planetary gearboxes on either side. The four vibration measurements plus a tachometer pulse from the drive shaft are recorded with a National Instruments CompactRIO 9024 data logger into a solid-state drive (SSD) as files of one minute length. Sampling frequency is 12800 Hz, and a built-in antialiasing filter guarantees that there are no aliases at frequencies that are less than \(0.45 \cdot 12800 \text{ Hz} = 5760 \text{ Hz}\). The measurement points are named right vertical (RV), left vertical (LV), right horizontal (RH) and left horizontal (LH). From now on \(x^{(2)}\) stands for the measured accelerometer signals and \(\alpha = 2\) in graphs also means that we are operating on the original signal.

The LHD moves to its operating area in the beginning of most measurement days. During this transition its speed is relatively constant and vibration levels quite steady. For calculations 151 signals were visually selected from these transitions. The rotational frequency of the drive shaft was approximately 13.5 Hz. From each signal a 4 second sample was multiplied with the window function (2.57) using \(\epsilon = 10\) and then differintegrated with Algorithm 1. High-pass filtering with an ideal filter was performed at cut-off frequency 3 Hz to remove unreliable low-frequency components, and low-pass filterings at cut-off frequencies 2000 Hz, 3000 Hz and 5000 Hz were also performed with an ideal filter. From both ends of the signal 20% was rejected and thus a 2.4 second signal remained to be used in calculation of \(l_p\) norms. All the calculations were performed with Matlab.

Figure 3 shows the 4s samples from the first and last days of the measurements and their amplitude spectra from the point RV. The vibration level
has pretty much doubled and the biggest components in the spectra are located at higher frequencies and are bigger. Figure 4 shows the corresponding signals from the point LV. Here the vibration level hasn’t risen that much but almost doubled nevertheless. Similar changes in the spectra has occurred as in the measurements from the point RV.

Figures 5 and 6 show the corresponding comparisons of the first and last days horizontal measurements. Again a doubling in vibration level has occurred and a shift in frequency content towards higher frequencies. We can immediately deduce just by looking at the signals and their spectra that an increase in high frequency vibration has occurred everywhere. We also note that there is a lot of variation between measurement points.

![Figure 3: Signals and their amplitude spectra from the point RV at the beginning and end of the measurement period](image)
Figure 4: Signals and their amplitude spectra from the point LV at the beginning and end of the measurement period

Figure 5: Signals and their amplitude spectra from the point RH at the beginning and end of the measurement period
Figure 6: Signals and their amplitude spectra from the point LH at the beginning and end of the measurement period

6.2 Sensitivity surfaces and trend analysis

Figure 7 shows a surface that has $2^4 MIT_p^p$ values from the point RV with $0 \leq p \leq 10$ and $-2 \leq \alpha \leq 6$ using a step of 0.1. The values for $2^4 MIT_p^p$ are calculated using the last (13.5.2014) and first day (22.4.2013) measurements. The measurements from the first day are used as reference measurements in all the calculations to follow. If calculated from some other signal from a nearby day, the surface would have some variation but still the local maxima containing ridge between $\alpha = 4$ and $\alpha = 5$ has constantly increased with time. The value $\alpha = 4.4$ was chosen in [42] so we shall continue to see how those derivatives develop with time. It is interesting that both differentiation and integration seem to increase sensitivity to change in condition.

Figure 8 shows the trend of the measurements from RV using the values $2^4 MIT_2^2$ for the whole measurement period. Only after the month long stoppage in the measurements, which occurred after 200 days, do we see any significant increase. There seems to be two separate "ladders" at 1.5 and finally around 3 where the values stay for a while.

Figure 9 shows the trend of the same signals using derivative $x^{(4,4)}$ and $p = 10$. Now we can see an increase to nearly 2 in just 100 days. After 200 days the $2^4 MIT_{4,4}^{10}$ values rise to over 2 and then in 300 days to 5 and then
bounced between 4 and 6. Very interestingly, we see three distinct phases in this increase and thus the MIT index seems to have a ladder-like behaviour with time in this case. The variation seen in these trends is most probably due to changing terrain and minor fluctuations in the speed of the vehicle.
Figure 8: Trend of $^{2.4}MIT_2$ from the point RV in the frequency range 3 - 5000 Hz

Figure 9: Trend of $^{2.4}MIT_{4.4}$ from the point RV in the frequency range 3 - 5000 Hz
Figure 10: $^{2.4}MIT^p_\alpha$ surface from the point LV in the frequency range 3 - 5000 Hz

Figure 10 shows a surface of $^{2.4}MIT^p_\alpha$ values from the point LV with $0 \leq p \leq 10$ and $-2 \leq \alpha \leq 6$ using a step of 0.1. This time a prominent ridge is present around $\alpha = 4.2$. In [42] value $\alpha = 3.6$ was chosen for trend analysis and I will continue using it. Actually any value near 4 seems to perform almost identically in trend analysis. Figures 11 and 12 show trends from the point LV of rms of the acceleration signal and $x^{(3,6)}$ respectively. Only after 200 days some changes are recognisable and then it seems that the $MIT$ indices increase linearly with time and quicker with $\alpha = 3.6$. 
Figure 11: Trend of $^{2.4}MIT_2^2$ from the point LV in the frequency range 3 - 5000 Hz

Figure 12: Trend of $^{2.4}MIT_{3.6}^2$ from the point LV in the frequency range 3 - 5000 Hz
In [42] it was noticed that low-pass filtering effectively increased sensitivity and so the calculations were repeated in the frequency ranges 3-3000 Hz and 3-2000 Hz. Figure 13 shows a surface of $2.4 \text{MIT}_p^\alpha$ values from the point RV with $0 \leq p \leq 10$ and $-2 \leq \alpha \leq 18$ using a step of 0.2 in the frequency range 3-3000 Hz and Figure 14 shows the corresponding values calculated from the point LV in the frequency range 3-2000 Hz. In [42] the optimal values were located at the ridge $\alpha = 10$, but now we see that with time the differentiation increases sensitivity even more and we could use derivatives of order 16 for example. Nevertheless, these very high derivatives actually give very similar results to the value $\alpha = 10$ used in [42] in trend analysis and thus we will continue using it and the easy to calculate norm $p = 1$. The trends of these $2.4 \text{MIT}_1^{10}$ values are plotted in Figures 15 and 16. The $2.4 \text{MIT}_1^{10}$ values from the point RV are doubled after only 50 days and the values finally stay around 15. For the point LV the values still won’t increase before 200 days, but after that, they increase linearly up to 9.
Figure 13: $^{2.4}MIT^{\rho}_{\alpha}$ surface from the point RV in the frequency range 3 - 3000 Hz

Figure 14: $^{2.4}MIT^{\rho}_{\alpha}$ surface from the point LV in the frequency range 3 - 2000 Hz
Figure 15: Trend of $^{2,4}MIT_{10}^1$ from the point RV in the frequency range 3 - 3000 Hz

Figure 16: Trend of $^{2,4}MIT_{10}^1$ from the point LV in the frequency range 3 - 2000 Hz
In the earlier analysis in [42] it was stated that horizontal measurements have changed very little and no clear trend is imminent. Well the newest measurements have revealed changes in horizontal vibrations as well. Figures 17 and 18 show the trends of rms values of these acceleration signals. Values \(2.4 \text{MIT}_2^2\) from the point RH have a curious drop between 200 and 280. The step like development in time (though one negative step this time) has much resemblance to Figure 9. Notice also this time correspondence of the trends from LV and LH.

Figure 19 shows a surface of \(2.4 \text{MIT}_a^p\) values from the point RH with \(0 \leq p \leq 10\) and \(-2 \leq \alpha \leq 8\) using a step of 0.1 and 20 shows similar values from the point LH. Once again the order of norm has little effect, but the maximum values are reached around \(\alpha = 5.8\) and \(\alpha = 3.5\). Using these derivative values and \(p = 1\) we obtain trends in Figures 21 and 22 where sensitivities are almost 5 and 4 respectively.
Figure 17: Trend of $^{2.4}MIT_2^2$ from the point RH in the frequency range 3 - 5000 Hz

Figure 18: Trend of $^{2.4}MIT_2^2$ from the point LH in the frequency range 3 - 5000 Hz
Figure 19: $^{2.4}MIT_\alpha^p$ surface from the point RH in the frequency range 3 - 5000 Hz

Figure 20: $^{2.4}MIT_\alpha^p$ surface from the point LH in the frequency range 3 - 5000 Hz
Figure 21: Trend of $^{2-4}MIT_{5.8}^1$ from the point RH in the frequency range 3 - 5000 Hz

Figure 22: Trend of $^{2-4}MIT_{3.5}^1$ from the point LH in the frequency range 3 - 5000 Hz
6.3 Complex differintegrals

Complex differintegrals is a signal processing method with tremendous potential in fault diagnosis. It has been proven to be useful in diagnosing simulated faults in test rigs [23, 48].

In Figures 23 and 24 we can see complex integrals of 1s samples from the signals from RV and RH at the beginning of the measurement period. These signals can be found from Figures 3 and 5. The effect of window function (2.57) is left visible at the ends of the signals. In these calculations $\epsilon = 10$. Two complex derivatives are visible in Figure 25 from the point RV. As integration makes signals smoother, we see that the complex integrals become simpler helix like objects. Differentiation on the other hand makes the helix more tightly wrapped.

![Complex integrals](image)

(a) $x^{(0+i)}$

(b) $x^{(1+i)}$

Figure 23: Complex integrals of $x^{(2)}$ from the point RV on 22.4.2013
The added third dimension in the complex differintegrals enables us also to look into its projections. Especially interesting are the projections to the complex plane along the time axis. These projections are shown from the integrals of Figure 24 in Figure 26 and from the derivatives of Figure 25 in Figure 27 along with the same calculations from the last measurement day. The diameter of the projections have increased and this is very evident in the differentiated signals. Different faults may reveal themselves in different patterns, as is studied in detail in [48]. These projections could be very suitable for pattern recognition software or even for maintenance personnel to monitor.

The unit of the complex derivative $x^{(z)}$ is m/s$^2$. This is uncommon in physics, but an unavoidable consequence of the Definition 4.1.
Figure 25: Complex derivatives of $x^{(2)}$ from the point RV on 22.4.2013

Figure 26: Projections of complex integrals of $x^{(2)}$ from the point RH
Figure 27: Projections of complex derivatives of $x^{(2)}$ from the point RV

(a) $x^{(2+i)}$, 22.4.2013
(b) $x^{(3+i)}$, 22.4.2013
(c) $x^{(2+i)}$, 13.5.2014
(d) $x^{(3+i)}$, 13.5.2014
6.4 Conclusions

Fractional calculus is a natural extension to integer order derivatives and integrals, although not without its difficulties. One must be familiar both with the mathematical theory and the application in hand to utilise fractional calculus effectively. Especially in numerical fractional calculus there is always the possibility that the method chosen will provide you with another fractional operator than the one you wished for. For vibration measurements from rotating machines the theory and methods presented in this thesis provide a useful toolkit to acquire new signals and analyse them. Frequency domain calculations are fast to perform and also very reliable when the signal is free from aliases.

The condition monitoring of a mobile vehicle is a challenge. In this thesis I utilised only a small part of the collected signals when the LHD was moving at relatively constant speed and without load. The signals from the working stages are more difficult to utilise in condition monitoring, although they could tell us something about the stress that the machine goes through in action.

Vertical vibrations on the right side of the axle have risen during the whole measurement period of 386 days. Interestingly, the trend analysis indicated that the $MIT$ indices have risen in steps. Similar three distinct steps are also seen on the right horizontal vibrations, although more weakly. This leads me to suggest that possibly three distinct faults have occurred on the right side planetary gearbox components. This means most probably faults on bearings or gear teeth.

On the left side it seems that the gearbox components have started to deteriorate only after 200 days. After that, the increase in $MIT$ indices has been linear. Overall, differentiation increases sensitivity a lot. For vertical vibrations $2.4 MIT_{10}^{1}$ values after low-pass filtering at 3000 Hz or 2000 Hz are very effective, giving sensitivities of about 15 and 9 for the right and left vertical measurements respectively. Derivatives of about $\alpha = 4$ or higher provide good sensitivity from all measurement points when filtered at 5000 Hz. Order of norm has very little effect in these signals. Complex derivatives are also sensitive to the changes in the condition of the axle and could possible be tuned to detect certain faults if more information was available.

Measurements are still running and with good luck they will be until the axle needs repair. Then it should be interesting to see the damage in the axle inspected and the data compared to vibration measurements.
References


