5
Auxiliary tools from analysis and measure theory

For our investigations on radial basis functions in the following chapters it is crucial to collect several results from different branches of mathematics. Hence, this chapter is a repository of such results. In particular, we are concerned with special functions such as Bessel functions and the Γ-function. We discuss the features of Fourier transforms and give an introduction to the aspects of measure theory that are relevant for our purposes. The reader who is not interested in these technical matters could skip this chapter and come back to it whenever necessary. Nonetheless, it is strongly recommended that one should at least have a look at the definition of Fourier transforms to become familiar with the notation we use. Because of the diversity of results presented here, we cannot give proofs in every case.

5.1 Bessel functions

Bessel functions will play an important role in what follows. Most of what we discuss here can be found in the fundamental book [187] by Watson.

The starting point for introducing Bessel functions is to remind the reader of the classical Γ-function and some of its features.

Definition 5.1 *The Γ-function is defined by*

$$\Gamma(z) := \lim_{n\to\infty} \frac{n!\, n^z}{z(z+1)\dots(z+n)}$$

for $z \in \mathbb{C}$.

It is a meromorphic function, well investigated in classical analysis. Some of its relevant properties are collected in the next proposition:

Proposition 5.2 *The Γ-function has the following properties:*

(1) $1/\Gamma(z)$ *is an entire function;*
(2) $\Gamma(1) = 1$, $\Gamma(1/2) = \sqrt{\pi}$*;*
(3) $\Gamma(z) = \int_0^\infty e^{-t} t^{z-1} dt$ *for* $\Re(z) > 0$ *(Euler's representation);*
(4) $\Gamma(z+1) = z\Gamma(z)$ *(recurrence relation);*

(5) $\Gamma(z)\Gamma(1-z) = \pi/\sin(\pi z)$ *(reflection formula);*

(6) $1 \leq \dfrac{\Gamma(x+1)}{\sqrt{2\pi x}\, x^x e^{-x}} \leq e^{1/(12x)}$, $x > 0$ *(Stirling's formula);*

(7) $\Gamma(2z) = \dfrac{2^{2z-1}}{\sqrt{\pi}}\Gamma(z)\Gamma(z+1/2)$ *(Legendre's duplication formula).*

The Γ-function and its properties are well known. Proofs of the formulas just stated can be found in any book on special functions. A particular choice would be the book [102] by Lebedev. Now, we continue by introducing Bessel functions.

Definition 5.3 *The Bessel function of the first kind of order $\nu \in \mathbb{C}$ is defined by*

$$J_\nu(z) := \sum_{m=0}^{\infty} \frac{(-1)^m (z/2)^{2m+\nu}}{m!\,\Gamma(\nu+m+1)}$$

for $z \in \mathbb{C} \setminus \{0\}$.

The power z^ν in this definition is defined by $\exp[\nu \log(z)]$, where log is the principal branch of the logarithm, i.e. $-\pi < \arg(z) \leq \pi$.

The Bessel function can be seen as a function of z and also as a function of ν and the following remarks are easily verified. Obviously, $J_\nu(z)$ is holomorphic in $\mathbb{C} \setminus [0, \infty)$ as a function of z for every $\nu \in \mathbb{C}$. Moreover, the expansion converges pointwise also for $z < 0$. If $\nu \in \mathbb{N}$ then J_ν has an analytic extension to $\mathbb{C}$. If $\Re(\nu) \geq 0$ then we have a continuous extension of $J_\nu(z)$ to $z = 0$. Finally, if $z \in \mathbb{C} \setminus \{0\}$ is fixed then $J_\nu(z)$ is a holomorphic function in $\mathbb{C}$ as a function of ν. We state further elementary properties in the next proposition.

Proposition 5.4 *The Bessel function of the first kind has the following properties:*

(1) $J_{-n} = (-1)^n J_n$ *if* $n \in \mathbb{N}$*;*

(2) $\dfrac{d}{dz}\{z^\nu J_\nu(z)\} = z^\nu J_{\nu-1}(z)$*;*

(3) $\dfrac{d}{dz}\{z^{-\nu} J_\nu(z)\} = -z^{-\nu} J_{\nu-1}(z)$*;*

(4) $J_{1/2}(z) = \sqrt{\dfrac{2}{\pi z}}\sin z, \qquad J_{-1/2}(z) = \sqrt{\dfrac{2}{\pi z}}\cos z.$

Proof For the first property we simply use that $1/\Gamma(n) = 0$ for $n = 0, -1, -2, \ldots$. The second and the third property follow by differentiation under the sum using also the recurrence formula for the Γ-function. The latter together with $\Gamma(1/2) = \sqrt{\pi}$ demonstrates the last item. $\square$

There exist many integral representations for Bessel functions. The one that matters here is given in the next proposition.

Proposition 5.5 *If we denote for $d \geq 2$ the unit sphere by $S^{d-1} = \{x \in \mathbb{R}^d : \|x\|_2 = 1\}$ then we have, for $x \in \mathbb{R}^d$,*

$$\int_{S^{d-1}} e^{ix^T y} dS(y) = (2\pi)^{d/2} \|x\|_2^{-(d-2)/2} J_{(d-2)/2}(\|x\|_2). \tag{5.1}$$

Proof Obviously, both sides of (5.1) are radially symmetric. Hence with spherical coordinates and $r = \|x\|_2$ we can derive

$$\int_{S^{d-1}} e^{ix^T y} dS(y) = \int_{S^{d-1}} e^{iry_1} dS(y) = \frac{2\pi^{(d-1)/2}}{\Gamma((d-1)/2)} \int_0^\pi e^{ir\cos\theta} \sin^{d-2}\theta \, d\theta$$

using that the surface area of S^{d-2} is given by

$$\omega_{d-2} = \frac{2\pi^{(d-1)/2}}{\Gamma((d-1)/2)}.$$

The initial and the last integral are obviously restrictions of entire functions to $r > 0$. Hence, we can calculate the last integral by expanding the exponent in the integral and integrating term by term, which gives

$$\int_0^\pi e^{ir\cos\theta} \sin^{d-2}\theta \, d\theta = \sum_{k=0}^\infty \frac{i^k r^k}{k!} a_k$$

with $a_k := \int_0^\pi \cos^k\theta \sin^{d-2}\theta \, d\theta$. By induction it is possible to show $a_{2k+1} = 0$ and

$$a_{2k} = \frac{(2k)!\,\Gamma((d-1)/2)\,\Gamma(1/2)}{2^{2k} k!\,\Gamma((k+d)/2)}.$$

Collecting everything together gives the stated representation. The exchange of integration and summation can easily be justified. Since we will give similar arguments in later proofs, this time we will leave the details to the reader. □

Our next result is concerned with the asymptotic behavior of the Bessel functions of the first kind.

Proposition 5.6 *The Bessel function has the following asymptotic behavior:*

(1) $J_\nu^{(\ell)}(r) = \mathcal{O}(1/\sqrt{r})$ *for* $r \to \infty$ *and* $\nu \in \mathbb{R}$, $\ell \in \mathbb{N}_0$*;*

(2) $J_\nu(r) = \sqrt{\frac{2}{\pi r}} \cos\left(r - \frac{\nu\pi}{2} - \frac{\pi}{4}\right) + \mathcal{O}(r^{-3/2})$
for $r \to \infty$ *and* $\nu \in \mathbb{R}$*;*

(3) $J_{d/2}^2(r) \leq \frac{2^{d+2}}{r\pi}$ *for* $r > 0$ *and* $d \in \mathbb{N}$*;*

(4) $\lim_{r\to 0} r^{-d} J_{d/2}^2(r) = \frac{1}{2^d \Gamma^2(d/2+1)}$ *for* $d \in \mathbb{N}$.

Proof The last property is an immediate consequence of the definition of the Bessel functions. The penultimate property is obviously true in the case $d = 1$, since in this case $J_{1/2}(r) = \sqrt{2/(\pi r)} \sin r$. The case $d \geq 2$ is more complicated. It is based mainly on Weber's "crude" inequality for Hankel functions (see Watson [187], p. 211). The complete proof needs too many details on Bessel functions to be presented here; it can be found in the paper [145] by Narcowich and Ward. To give the proof for the second property would go beyond the scope of this book; hence, we refer the reader to Watson [187], p. 199 or, alternatively, to Lebedev [102], p. 122. In the case $\ell = 0$, the first property is a weaker version of the second property. For higher derivatives we use repeatedly the recurrence relation $2J_\nu'(z) = J_{\nu-1}(z) - J_{\nu+1}(z)$, which is a consequence of the formulas given in Proposition 5.4, to derive the desired asymptotic behavior. □

We now turn to the Laplace transforms of some specific functions involving Bessel functions.

Lemma 5.7 *For $\nu > -1$ and every $r > 0$ it is true that*

$$\int_0^\infty J_\nu(t)t^{\nu+1}e^{-rt}dt = \frac{2^{\nu+1}\Gamma(\nu+3/2)r}{\sqrt{\pi}(r^2+1)^{\nu+3/2}}.$$

Proof Let us start by looking at the binomial series. For $0 \le r < 1$ and $\mu > 0$ we have

$$(1+r)^{-\mu} = \sum_{m=0}^\infty \frac{(-1)^m\Gamma(\mu+m)}{m!\,\Gamma(\mu)} r^m.$$

Hence, if we replace r by $1/r^2$ this gives

$$r^{2\mu}(1+r^2)^{-\mu} = \sum_{m=0}^\infty \frac{(-1)^m\Gamma(\mu+m)}{m!\,\Gamma(\mu)} r^{-2m}$$

for $r > 1$. Moreover, Legendre's duplication formula with $z = \nu + m + 1 > 0$ yields

$$\Gamma(\nu+m+3/2) = \frac{\sqrt{\pi}\,\Gamma(2\nu+2m+2)}{2^{2\nu+2m+1}\Gamma(\nu+m+1)}.$$

Thus, setting $\mu = \nu + 3/2 > 1/2$ shows that

$$\frac{2^{\nu+1}\Gamma(\nu+3/2)r}{\sqrt{\pi}(r^2+1)^{\nu+3/2}} = \sum_{m=0}^\infty \frac{(-1)^m\Gamma(2\nu+2m+2)}{2^{2m+\nu}m!\,\Gamma(\nu+m+1)} r^{-2m-2\nu-2}.$$

Now we will have a look at the integral. Using the definition of the Bessel function, and interchanging summation and integration, allows us to make the following derivation:

$$\begin{aligned}\int_0^\infty J_\nu(t)t^{\nu+1}e^{-rt}dt &= \sum_{m=0}^\infty \frac{(-1)^m}{2^{2m+\nu}m!\,\Gamma(\nu+m+1)}\int_0^\infty t^{2m+2\nu+1}e^{-rt}dt\\ &= \sum_{m=0}^\infty \frac{(-1)^m\Gamma(2m+2\nu+2)}{2^{2m+\nu}m!\,\Gamma(m+\nu+1)} r^{-2m-2\nu-2}.\end{aligned}$$

In the last step we used that the integral representation of the Γ-function can be expressed as

$$\int_0^\infty t^{2m+2\nu+1}e^{-rt}dt = r^{-2m-2\nu-2}\Gamma(2m+2\nu+2).$$

Moreover, the interchange of summation and integration can be justified as follows. First note that Stirling's formula allows the bound $\Gamma(\nu+m+1) \ge (1/c_\nu)m!$, where the constant

c_ν depends only on $\nu > -1$. Then

$$\begin{aligned}\sum_{m=0}^{\infty} \frac{(t/2)^{\nu+2m}}{m!\,\Gamma(\nu+m+1)} t^{\nu+1} e^{-rt} &\le c_\nu t^{2\nu+1} \sum_{m=0}^{\infty} \frac{t^{2m}}{2^{2m}(m!)^2}\, e^{-rt} \\ &\le c_\nu t^{2\nu+1} e^{-rt} \sum_{m=0}^{\infty} \frac{t^{2m}}{(2m)!} \\ &\le c_\nu t^{2\nu+1} e^{-rt} e^{t},\end{aligned}$$

which is clearly in $L_1[0,\infty)$ provided that $r > 1$. Hence, Lebesgue's convergence theorem justifies the interchange.

Up to now we have shown that the stated equality holds for all $r > 1$. But, since both sides are analytic functions in $\Re(r) > 0$ and $|\Im(r)| < 1$, the equality extends to all $r > 0$ by analytic continuation. □

The following result is in the same spirit.

Lemma 5.8 *For $r > 0$ it is true that*

$$\int_0^\infty J_0(t) e^{-rt} dt = \frac{1}{(1+r^2)^{1/2}}.$$

Proof From the duplication formula of the Γ-function it follows that

$$\Gamma(m+1/2) = \frac{(2m)!\sqrt{\pi}}{2^{2m} m!}.$$

Hence, as in the proof of Lemma 5.7 we get the representation

$$(1+r^2)^{-1/2} = \sum_{m=0}^{\infty} \frac{(-1)^m \Gamma(m+1/2)}{m!\,\Gamma(1/2)} r^{-2m-1} = \sum_{m=0}^{\infty} \frac{(-1)^m (2m)!}{2^{2m}(m!)^2} r^{-2m-1}.$$

Thus, by interchanging summation and integration we derive

$$\begin{aligned}\int_0^\infty J_0(t) e^{-rt} dt &= \sum_{m=0}^{\infty} \frac{(-1)^m}{2^{2m}(m!)^2} \int_0^\infty t^{2m} e^{-rt} dt \\ &= \sum_{m=0}^{\infty} \frac{(-1)^m \Gamma(2m+1)}{2^{2m}(m!)^2} r^{-2m-1} \\ &= (1+r^2)^{-1/2}\end{aligned}$$

for $r > 1$. The interchange of summation and integration can be justified as in Lemma 5.7. The equality extends to $r > 0$ by analytic continuation as before. □

Our final result on Bessel functions of the first kind deals with J_0 again.

Lemma 5.9 *For the Bessel function J_0 the following two properties are satisfied:*

(1) $\int_0^r J_0(t)dt > 0$ *for all* $r > 0$*;*

(2) $\int_0^\infty J_0(t)dt = 1.$

The second integral is intended as an improper Riemann integral.

Proof The first result is known as Cooke's inequality. The easiest way to prove it is to use another representation of the Bessel function, namely

$$J_0(t) = \frac{2}{\pi} \int_1^\infty \frac{\sin(ut)}{(u^2-1)^{1/2}}\, du;$$

see Watson [187], p. 170. This shows that

$$\int_0^r J_0(t)dt = \frac{2}{\pi} \int_1^\infty \frac{1-\cos(ru)}{u(u^2-1)^{1/2}}\, du > 0$$

for all $r > 0$.

Finally, let us discuss the second result. We know by Lemma 5.8 that $\int_0^\infty J_0(t)e^{-rt}dt = (1+r^2)^{-1/2}$ for all $r > 0$. Hence we want to let $r \to 0$. Unfortunately, since J_0 is not in $L_1(\mathbb{R})$ we cannot use classical convergence arguments and so we have to be more precise. The idea is to use the triangle inequality twice to get the bound

$$\begin{aligned} \left|1 - \int_0^R J_0(t)dt\right| &\le \left|1-(1+r^2)^{-1/2}\right| + \left|\int_0^R J_0(t)(e^{-rt}-1)dt\right| \\ &\quad + \left|\int_R^\infty J_0(t)e^{-rt}dt\right| \end{aligned}$$

for an arbitrary $r \in (0, 1]$. Next suppose that for every $\epsilon > 0$ we can find an $R_0 > 0$ such that the last integral becomes uniformly less than $\epsilon/3$ for all $r \in (0, 1]$ provided $R > R_0$. If this is true then we can fix such an $R > R_0$ and then choose $0 < r \le 1$ such that the first two terms also become small (using that J_0 is bounded), which establishes the second result. Hence it remains to prove this uniform bound on the last integral. To this end we use the asymptotic expansion of J_0 from Proposition 5.6, i.e.

$$J_0(t) = \sqrt{\frac{2}{\pi}} \frac{\cos(t-\pi/4)}{\sqrt{t}} + S(t),$$

with a remainder $S(t)$ satisfying $|S(t)| \le Ct^{-3/2}$ for $t \ge 1$. The remainder part of the integral in question can easily be bounded since $|\int_R^\infty S(t)e^{-rt}dt| \le C\int_R^\infty t^{-3/2}dt = CR^{-1/2}$ with a generic constant $C > 0$ that is independent of $r > 0$. The main part is bounded by integration by parts:

$$\begin{aligned} \int_R^\infty \frac{e^{-rt}}{\sqrt{t}} \cos\left(t - \frac{\pi}{4}\right) dt = &-\frac{1}{\sqrt{2t}} e^{-rt} \frac{(1+r)\cos t + (r-1)\sin t}{1+r^2}\Bigg|_R^\infty \\ &- \int_R^\infty \frac{e^{-rt}}{\sqrt{8}\, t^{3/2}} \frac{(1+r)\cos t + (r-1)\sin t}{1+r^2}\, dt. \end{aligned}$$

This expansion shows that the integral on the left-hand side can also be bounded uniformly by a constant times $R^{-1/2}$ for all $r \in (0, 1]$. □

After discussing Bessel functions of the first kind we come to another family of Bessel functions, called *Bessel functions of imaginary argument* or *modified Bessel functions of the third kind*, sometimes also Mcdonald's functions.

Definition 5.10 *The modified Bessel function of the third kind of order $\nu \in \mathbb{C}$ is defined by*

$$K_\nu(z) := \int_0^\infty e^{-z\cosh t} \cosh(\nu t)\, dt$$

for $z \in \mathbb{C}$ with $|\arg(z)| < \pi/2$; $\cosh t = (e^t + e^{-t})/2$.

It follows immediately from the definition that $K_\nu(x) > 0$ for $x > 0$ and $\nu \in \mathbb{R}$. The modified Bessel functions satisfy recurrence relations similar to the Bessel functions of the first kind (see Watson [187], p. 79). The only one that matters here is

$$\frac{d}{dz}\left[z^\nu K_\nu(z)\right] = -z^\nu K_{\nu-1}(z). \tag{5.2}$$

Moreover, there exist several other representation formulas. The following one is of particular interest for us. Again, its proof goes beyond the scope of this book, so for this we refer the reader again to Watson [187], p. 206.

Proposition 5.11 *For $\nu \geq 0$ and $x > 0$ the modified Bessel function of the third kind has the representation*

$$K_\nu(x) = \left(\frac{\pi}{2x}\right)^{1/2} \frac{e^{-x}}{\Gamma(\nu + 1/2)} \int_0^\infty e^{-u} u^{\nu-1/2} \left(1 + \frac{u}{2x}\right)^{\nu-1/2} du.$$

This representation gives some insight into lower bounds on the decay of the modified Bessel functions.

Corollary 5.12 *For every $\nu \in \mathbb{R}$ the function $x \mapsto x^\nu K_{-\nu}(x)$ is nonincreasing on $(0, \infty)$. Moreover, it has the lower bound*

$$K_\nu(x) \geq \sqrt{\frac{\pi}{2}} \frac{e^{-x}}{\sqrt{x}}, \qquad x > 0,$$

if $|\nu| \geq 1/2$. In the case $|\nu| < 1/2$ the lower bound is given by

$$K_\nu(x) \geq \frac{\sqrt{\pi}\, 3^{|\nu|-1/2}}{2^{|\nu|+1}\Gamma(|\nu| + 1/2)} \frac{e^{-x}}{\sqrt{x}}, \qquad x \geq 1.$$

Proof The recurrence relation (5.2) together with $K_{-\nu} = K_\nu$ and the fact that $K_{\nu-1}$ is positive on $(0, \infty)$ gives the monotonic property of $x^\nu K_{-\nu}(x)$. To prove the lower bound we can restrict ourselves to $\nu \geq 0$ because of $K_\nu = K_{-\nu}$. If $\nu \geq 1/2$ then we get from Proposition 5.11

$$K_\nu(x) \geq \left(\frac{\pi}{2x}\right)^{1/2} \frac{e^{-x}}{\Gamma(\nu + 1/2)} \int_0^\infty e^{-u} u^{\nu-1/2} du = \left(\frac{\pi}{2x}\right)^{1/2} e^{-x}.$$

However, if $0 \leq \nu < 1/2$ then we have to be more careful. We use that for $u \in [0, 1]$ and

$x \geq 1$ it is obviously true that $u \leq 1$ and $1 + u/(2x) \leq 1 + 1/(2x) \leq 3/2$, so that

$$\int_0^\infty e^{-u} u^{\nu-1/2} \left(1 + \frac{u}{2x}\right)^{\nu-1/2} du \geq \int_0^1 e^{-u} u^{\nu-1/2} \left(1 + \frac{u}{2x}\right)^{\nu-1/2} du$$

$$\geq \left(\frac{3}{2}\right)^{\nu-1/2} \int_0^1 e^{-u} du \geq \frac{1}{2}\left(\frac{3}{2}\right)^{\nu-1/2},$$

which finishes the proof. □

Finally, we derive upper bounds on K_ν. We will need these bounds for complex-valued ν. We start with their behavior if the argument tends to infinity.

Lemma 5.13 *The modified Bessel function K_ν, $\nu \in \mathbb{C}$, has the asymptotic behavior*

$$|K_\nu(r)| \leq \sqrt{\frac{2\pi}{r}} e^{-r} e^{|\Re(\nu)|^2/(2r)}, \qquad r > 0. \tag{5.3}$$

Proof With $b = |\Re(\nu)|$ we have

$$\begin{aligned} |K_\nu(r)| &\leq \frac{1}{2}\int_0^\infty e^{-r\cosh t} |e^{\nu t} + e^{-\nu t}| dt \\ &\leq \frac{1}{2}\int_0^\infty e^{-r\cosh t} [e^{bt} + e^{-bt}] dt \\ &= K_b(r). \end{aligned}$$

Furthermore, from $e^t \geq \cosh t \geq 1 + t^2/2$ for $t \geq 0$ we can conclude that

$$\begin{aligned} K_b(r) &\leq \int_0^\infty e^{-r(1+t^2/2)} e^{bt} dt \\ &= e^{-r} e^{b^2/(2r)} \frac{1}{\sqrt{r}} \int_{-b/\sqrt{r}}^\infty e^{-s^2/2} ds \\ &\leq \sqrt{2\pi} e^{-r} e^{b^2/(2r)} \sqrt{\frac{1}{r}}. \end{aligned}$$

□

While the last lemma describes the asymptotic behavior of the modified Bessel functions for large arguments, the next lemma describes the behavior in a neighborhood of the origin. Nonetheless, in the case $\Re(\nu) \neq 0$ it holds for all $r > 0$.

Lemma 5.14 *For $\nu \in \mathbb{C}$ the modified Bessel functions satisfy, for $r > 0$,*

$$|K_\nu(r)| \leq \begin{cases} 2^{|\Re(\nu)|-1}\Gamma(|\Re(\nu)|) r^{-|\Re(\nu)|}, & \Re(\nu) \neq 0, \\ \dfrac{1}{e} - \log\dfrac{r}{2}, & r < 2, \quad \Re(\nu) = 0. \end{cases} \tag{5.4}$$

Proof Let us first consider the case $\Re(\nu) \neq 0$. Again, we set $b = |\Re(\nu)|$ and know already that $|K_\nu(r)| \leq K_b(r)$ from the proof of the last lemma.

From the definition of the modified Bessel function, however, we can conclude for every $a > 0$ that

$$\begin{aligned}K_b(r) &= \frac{1}{2}\int_{-\infty}^{\infty} e^{-r\cosh t} e^{bt} dt \\ &= \frac{1}{2}\int_{-\infty}^{\infty} e^{-r(e^t+e^{-t})/2} e^{bt} dt \\ &= a^{-b}\frac{1}{2}\int_{0}^{\infty} e^{(-r/2)(u/a+a/u)} u^{b-1} du\end{aligned}$$

by substituting $u = ae^t$. By setting $a = r/2$ we obtain

$$K_b(r) = 2^{b-1} r^{-b} \int_0^{\infty} e^{-u} e^{-r^2/(4u)} u^{b-1} du \le 2^{b-1}\Gamma(b) r^{-b}.$$

For $\Re(\nu) = 0$ we use $\cosh t \ge e^t/2$ to derive

$$\begin{aligned}K_0(r) &= \int_0^{\infty} e^{-r\cosh t} dt \le \int_0^{\infty} e^{-\frac{r}{2}e^t} dt \\ &= \int_{r/2}^{\infty} e^{-u}\frac{1}{u} du \quad \le \int_1^{\infty} e^{-u} du + \int_{r/2}^{1} \frac{1}{u} du \\ &= \frac{1}{e} - \log\frac{r}{2}.\end{aligned}$$

□

5.2 Fourier transform and approximation by convolution

One of the most powerful tools in analysis is the Fourier transform. Not only will it help us to characterize positive definite functions, it will also be necessary in several other places. Hence, we will dwell on this subject maybe at first sight longer than necessary. We start with the classical L_1 theory.

Definition 5.15 *For $f \in L_1(\mathbb{R}^d)$ we define its Fourier transform by*

$$\widehat{f}(x) := (2\pi)^{-d/2} \int_{\mathbb{R}^d} f(\omega) e^{-ix^T\omega} d\omega$$

and its inverse Fourier transform by

$$f^{\vee}(x) := (2\pi)^{-d/2} \int_{\mathbb{R}^d} f(\omega) e^{ix^T\omega} d\omega.$$

We will always use this symmetric definition. But the reader should note that there are other definitions on the market, which differ from each other and this definition only by the way in which the 2π terms are distributed.

For a function $f \in L_1(\mathbb{R}^d)$ the Fourier transform is continuous. Moreover, the following rules are easily established. The overstrained reader might have a look at the book [180] by Stein and Weiss.

Theorem 5.16 *Suppose $f, g \in L_1(\mathbb{R}^d)$; then the following is true.*

(1) $\int_{\mathbb{R}^d} \widehat{f}(x)g(x)dx = \int_{\mathbb{R}^d} f(x)\widehat{g}(x)dx$.

(2) The Fourier transform of the convolution

$$f * g(x) := \int_{\mathbb{R}^d} f(y)g(x-y)\,dy$$

is given by $\widehat{f * g} = (2\pi)^{d/2}\widehat{f}\ \widehat{g}$.

(3) With $\widetilde{f}(x) := \overline{f(-x)}$ we find that $\widehat{f * \widetilde{f}} = (2\pi)^{d/2}|\widehat{f}|^2$.

(4) For $T_a f(x) := f(x-a)$, $a \in \mathbb{R}^d$, we have $\widehat{T_a f}(x) = e^{-ix^T a}\widehat{f}(x)$.

(5) For $S_\alpha f(x) := f(x/\alpha)$, $\alpha > 0$, we have $\widehat{S_\alpha f} = \alpha^d S_{1/\alpha}\widehat{f}$.

(6) If, in addition, $x_j f(x) \in L_1(\mathbb{R}^d)$ then $\widehat{f}$ is differentiable with respect to x_j and

$$\frac{\partial \widehat{f}}{\partial x_j}(x) = (-iy_j f(y))^\wedge(x).$$

If $\partial f/\partial x_j$ is also in $L_1(\mathbb{R}^d)$ then

$$\widehat{\frac{\partial f}{\partial x_j}}(x) = ix_j\widehat{f}(x).$$

Obviously the last item extends to higher-order derivatives in a natural way. The following space will turn out to be the natural playground for Fourier transforms.

Definition 5.17 *The Schwartz space $\mathcal{S}$ consists of all functions $\gamma \in C^\infty(\mathbb{R}^d)$ that satisfy*

$$|x^\alpha D^\beta \gamma(x)| \le C_{\alpha,\beta,\gamma}, \qquad x \in \mathbb{R}^d,$$

for all multi-indices $\alpha, \beta \in \mathbb{N}_0^d$ with a constant $C_{\alpha,\beta,\gamma}$ that is independent of $x \in \mathbb{R}^d$. The functions of $\mathcal{S}$ are called test functions or good functions.

In other words, a good function and all of its derivatives decay faster than any polynomial. Of course all functions from $C_0^\infty(\mathbb{R}^d)$ are contained in $\mathcal{S}$ but so also are the functions

$$\gamma(x) = e^{-\alpha\|x\|_2^2}, \qquad x \in \mathbb{R}^d,$$

for all $\alpha > 0$, and we are going to compute their Fourier transforms now. Because of the scaling property of the Fourier transform it suffices to pick one specific $\alpha > 0$.

Theorem 5.18 *The function $G(x) := e^{-\|x\|_2^2/2}$ satisfies $\widehat{G} = G$.*

Proof First note that, because

$$\begin{aligned}\widehat{G}(x) &= (2\pi)^{-d/2}\int_{\mathbb{R}^d} e^{-\|y\|_2^2/2}e^{-ix^T y}dy\\ &= \prod_{j=1}^d \left((2\pi)^{-1/2}\int_{-\infty}^{\infty} e^{-y_j^2/2}e^{-ix_j y_j}dy_j\right),\end{aligned}$$

the Fourier transform of the d-variate Gaussian G is the product of the univariate Fourier transforms of the univariate Gaussian $g(t) = e^{-t^2/2}$, and it suffices to compute this univariate

Fourier transform. Cauchy's integral theorem yields

$$\begin{aligned}\widehat{g}(r) &= (2\pi)^{-1/2}\int_{-\infty}^{\infty} e^{-(t^2/2+irt)}dt \\ &= (2\pi)^{-1/2}e^{-r^2/2}\int_{-\infty}^{\infty} e^{-(t+ir)^2/2}dt \\ &= (2\pi)^{-1/2}e^{-r^2/2}\int_{-\infty}^{\infty} e^{-t^2/2}dt \\ &= e^{-r^2/2}.\end{aligned}$$

□

We need another class of functions. In a certain way they are the opposite of good functions. Actually, they define continuous linear functionals on $\mathcal{S}$, but this will not really matter for us.

Definition 5.19 *We say that a function f is slowly increasing if there exists a constant $m \in \mathbb{N}_0$ such that $f(x) = \mathcal{O}(\|x\|_2^m)$ for $\|x\|_2 \to \infty$.*

Our next result concerns the approximation of functions by convolution.

Theorem 5.20 *Define $g_m(x) = (m/\pi)^{d/2}e^{-m\|x\|_2^2}$, $m \in \mathbb{N}$, $x \in \mathbb{R}^d$. Then the following hold true:*

(1) $\int_{\mathbb{R}^d} g_m(x)dx = 1$;
(2) $\widehat{g}_m(x) = (2\pi)^{-d/2}e^{-\|x\|_2^2/(4m)}$,
(3) $\widehat{\widehat{g}}_m(x) = g_m(x)$,
(4) $\Phi(x) = \lim_{m\to\infty}\int_{\mathbb{R}^d}\Phi(\omega)g_m(\omega - x)d\omega$, provided that $\Phi \in C(\mathbb{R}^d)$ is slowly increasing.

Proof (1) follows from

$$\int_{\mathbb{R}^d} g_m(x)dx = (2\pi)^{-d/2}\int_{\mathbb{R}^d} e^{-\|y\|_2^2/2}dy = 1.$$

To prove (2) we remark that $g_m = (m/\pi)^{d/2}S_{1/\sqrt{2m}}G$. Thus Theorems 5.16 and 5.18 lead to

$$\widehat{g}_m = (m/\pi)^{d/2}(2m)^{-d/2}S_{\sqrt{2m}}\widehat{G} = (2\pi)^{-d/2}S_{\sqrt{2m}}G.$$

For (3) note that

$$\widehat{\widehat{g}}_m = (2\pi)^{-d/2}(S_{\sqrt{2m}}G)^{\wedge} = (2\pi)^{-d/2}(\sqrt{2m})^d S_{1/\sqrt{2m}}\widehat{G} = g_m.$$

For (4) we first restrict ourselves to the case $x = 0$. From (1) we see that

$$\int_{\mathbb{R}^d}\Phi(\omega)g_m(\omega)d\omega - \Phi(0) = \int_{\mathbb{R}^d}[\Phi(\omega) - \Phi(0)]g_m(\omega)d\omega.$$

Now choose an arbitrary $\epsilon > 0$. Then there exists a $\delta > 0$ such that $|\Phi(\omega) - \Phi(0)| < \epsilon/2$ for $\|\omega\|_2 \le \delta$. Furthermore, since Φ is slowly increasing, there exists an $\ell \in \mathbb{N}_0$ and an

$M > 0$ such that $|\Phi(\omega)| \le M(1 + \|\omega\|_2)^\ell$, $\omega \in \mathbb{R}^d$. This means that we can find a generic constant C_δ such that we can bound the integral as follows:

$$\begin{aligned}\left|\int_{\mathbb{R}^d} [\Phi(\omega) - \Phi(0)] g_m(\omega) d\omega\right| &\le \int_{\|\omega\|_2 \le \delta} |\Phi(\omega) - \Phi(0)| g_m(\omega) d\omega \\ &\quad + C_\delta \int_{\|\omega\|_2 > \delta} g_m(\omega) \|\omega\|_2^\ell d\omega \\ &\le \frac{\epsilon}{2} + C_\delta m^{-\ell/2} \int_{\|\omega\|_2/\sqrt{2m} > \delta} e^{-\|\omega\|_2^2/2} \|\omega\|_2^\ell d\omega \\ &\le \epsilon\end{aligned}$$

for sufficiently large m. The case $x \ne 0$ follows immediately by replacing Φ by $\Phi(\cdot + x)$ in the previous case. □

The approximation process described in Theorem 5.20, item (4), is a well-known method of approximating a function. It is sometimes also called *approximation by mollification or regularization*. Let us stay a little longer with this process. In particular we are now interested in replacing the Gaussians by an arbitrary compactly supported C^∞-function. Moreover, we are interested in weaker forms of convergence.

Lemma 5.21 *Suppose that* $f \in L_p(\mathbb{R}^d)$, $1 \le p < \infty$, *is given. Then we have* $\lim_{x \to 0} \|f - f(\cdot + x)\|_{L_p(\mathbb{R}^d)} = 0$.

Proof Let us denote $f(\cdot + x)$ by f_x. We start by showing the result for a continuous function g with compact support. Choose a compact set $K \subseteq \mathbb{R}^d$ such that the support of g_x is contained in K for all $x \in \mathbb{R}^d$ with $\|x\|_2 \le 1$. Since g is continuous, it is uniformly continuous on K. Hence for a given $\epsilon > 0$ we find a $\delta > 0$ such that $|g(y) - g(x + y)| < \epsilon$ for all $y \in K$ and all $\|x\|_2 < \delta$. This finishes the proof in this case because

$$\|g - g_x\|_{L_p(\mathbb{R}^d)} = \|g - g_x\|_{L_p(K)} \le \epsilon [\mathrm{vol}(K)]^{1/p}.$$

Now, for an arbitrary $f \in L_p(\mathbb{R}^d)$ and $\epsilon > 0$ we choose a function $g \in C_0(\mathbb{R}^d)$ with $\|f - g\|_{L_p(\mathbb{R}^d)} < \epsilon/3$. By substitution this also means that $\|f_x - g_x\|_{L_p(\mathbb{R}^d)} < \epsilon/3$, giving

$$\begin{aligned}\|f - f_x\|_{L_p(\mathbb{R}^d)} &\le \|f - g\|_{L_p(\mathbb{R}^d)} + \|g - g_x\|_{L_p(\mathbb{R}^d)} + \|g_x - f_x\|_{L_p(\mathbb{R}^d)} \\ &< \frac{2\epsilon}{3} + \|g - g_x\|_{L_p(\mathbb{R}^d)},\end{aligned}$$

and this becomes smaller than ϵ for sufficiently small $\|x\|_2$. □

Note that the result is wrong in the case $p = \infty$. In fact, $\|f - f(\cdot - x)\|_{L_\infty(\mathbb{R}^d)} \to 0$ as $x \to 0$ implies that f is almost everywhere uniformly continuous.

The previous result allows us to establish the convergence of approximation by convolution, not only pointwise but also in L_p. The following results are formulated for $g \in C_0^\infty(\mathbb{R}^d)$, but the proofs show that for example the second and third items hold even if we have only $g \in L_1(\mathbb{R}^d)$.

Theorem 5.22 *Suppose an even and nonnegative $g \in C_0^\infty(\mathbb{R}^d)$ is given, normed by $\int g(x)dx = 1$. Define $g_m(x) = m^d g(mx)$. Then the following are true.*

(1) If $f \in L_1^{\text{loc}}(\mathbb{R}^d)$ then $f * g \in C^\infty(\mathbb{R}^d)$ and $D^\alpha(f * g) = f * (D^\alpha g)$.
(2) If $f \in L_p(\mathbb{R}^d)$ with $1 \le p \le \infty$ then $f * g \in L_p(\mathbb{R}^d)$ and $\|f * g\|_{L_p(\mathbb{R}^d)} \le \|f\|_{L_p(\mathbb{R}^d)}\|g\|_{L_1(\mathbb{R}^d)}$.
(3) If $f \in L_p(\mathbb{R}^d)$ with $1 \le p < \infty$ then $\|f - f * g_n\|_{L_p(\mathbb{R}^d)} \to 0$ for $n \to \infty$.
(4) If $f \in C(\mathbb{R}^d)$ then $f * g_n \to f$ uniformly on every compact subset of $\mathbb{R}^d$.

Proof The first property is an immediate consequence of the theory of integrals depending on an additional parameter. The second property is obviously true for $p = 1, \infty$. Moreover, if $1 < p < \infty$ we obtain, with $q = p/(p-1)$,

$$\begin{aligned}|f * g(x)| &\le \int_{\mathbb{R}^d} |f(y)||g(x-y)|^{1/p}|g(x-y)|^{1/q}dy \\ &\le \left(\int_{\mathbb{R}^d} |f(y)|^p|g(x-y)|dy\right)^{1/p}\left(\int_{\mathbb{R}^d}|g(x-y)|dy\right)^{1/q}\end{aligned}$$

or in other words

$$\begin{aligned}\|f * g\|^p_{L_p(\mathbb{R}^d)} &\le \left(\int_{\mathbb{R}^d}|g(y)|dy\right)^{p/q}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}|f(y)|^p|g(x-y)|dydx \\ &= \|g\|^p_{L_1(\mathbb{R}^d)}\|f\|^p_{L_p(\mathbb{R}^d)}.\end{aligned}$$

Using Minkowski's inequality again, we can derive

$$\|f - f * g_n\|^p_{L_p(\mathbb{R}^d)} \le \int_{\mathbb{R}^d}\|f - f(\cdot + y/n)\|^p_{L_p(\mathbb{R}^d)}g(y)dy$$

in the same fashion as before. But since $\|f - f(\cdot + y/n)\|_{L_p(\mathbb{R}^d)} \to 0$ for $n \to \infty$ by Lemma 5.21 and since $\|f - f(\cdot + y/n)\|_{L_p(\mathbb{R}^d)} \le 2\|f\|_{L_p(\mathbb{R}^d)}$, Lebesgue's dominated convergence theorem yields the third property.

Finally, let f be a continuous function and $\epsilon > 0$ be given. If $K \subseteq \mathbb{R}^d$ is a compact set then we have $K \subseteq B(0, R)$ for $R > 0$. Choose $\delta > 0$ such that $|f(x) - f(y)| < \epsilon$ for all $x, y \in B(0, R+1)$ with $\|x - y\|_2 < \delta$. Without restriction we can assume that g is supported in $B(0, 1)$. This allows us to conclude, for $x \in K$, that

$$|f(x) - f * g_n(x)| \le \int_{B(x,1/n)}|f(x) - f(y)|g_n(x-y)dy \le \epsilon$$

whenever $1/n < \delta$, which establishes uniform convergence on K. □

We now come back to the initial form of approximation by convolution, as given in Theorem 5.20, to prove one of the fundamental theorems in Fourier analysis.

Theorem 5.23 *The Fourier transform defines an automorphism on $\mathcal{S}$. The inverse mapping is given by the inverse Fourier transformation. Furthermore, the $L_2(\mathbb{R}^d)$-norms of a function and its transform coincide: $\|\widehat{f}\|_{L_2(\mathbb{R}^d)} = \|f\|_{L_2(\mathbb{R}^d)}$.*

Proof From Theorem 5.16 we can conclude that Fourier transformation maps $\mathcal{S}$ back into $\mathcal{S}$ and that the same is valid for the inverse transformation. Using Theorem 5.16 and Theorem 5.20 we obtain, again using Lebesgue's convergence theorem,

$$\begin{aligned} f(x) &= \lim_{m\to\infty} \int_{\mathbb{R}^d} f(\omega) g_m(\omega - x) d\omega \\ &= \lim_{m\to\infty} \int_{\mathbb{R}^d} f(\omega + x) \widehat{\widehat{g}}_m(\omega) d\omega \\ &= \lim_{m\to\infty} \int_{\mathbb{R}^d} \widehat{f}(\omega) e^{ix^T\omega} \widehat{g}_m(\omega) d\omega \\ &= (2\pi)^{-d/2} \int_{\mathbb{R}^d} \widehat{f}(\omega) e^{ix^T\omega} d\omega. \end{aligned}$$

Finally, we now have for an arbitrary $g \in \mathcal{S}$,

$$\widehat{\widehat{\overline{g}}}(x) = (2\pi)^{-d/2} \int_{\mathbb{R}^d} \widehat{\overline{g}}(\omega) e^{-ix^T\omega} d\omega = (2\pi)^{-d/2} \overline{\int_{\mathbb{R}^d} \widehat{g}(\omega) e^{ix^T\omega} d\omega} = \overline{g}(x),$$

which, together with Theorem 5.16, leads to

$$\int_{\mathbb{R}^d} f(x)\overline{g}(x) dx = \int_{\mathbb{R}^d} f(x) \widehat{\widehat{\overline{g}}}(x) dx = \int_{\mathbb{R}^d} \widehat{f}(x) \widehat{\overline{g}}(x) dx,$$

so that not only are the norms equal but also the inner products. □

Obviously, the proof can be extended to the following situation.

Corollary 5.24 *If $f \in L_1(\mathbb{R}^d)$ is continuous and has a Fourier transform $\widehat{f} \in L_1(\mathbb{R}^d)$ then f can be recovered from its Fourier transform:*

$$f(x) = (2\pi)^{-d/2} \int_{\mathbb{R}^d} \widehat{f}(\omega) e^{ix^T\omega} d\omega, \qquad x \in \mathbb{R}^d.$$

Another consequence of Theorem 5.23 is that it allows us to extend the idea of the Fourier transform to $L_2(\mathbb{R}^d)$, even for functions that are not integrable and hence do not possess a classical Fourier transform. Theorem 5.23 asserts that Fourier transformation constitutes a bounded linear operator defined on the dense subset $\mathcal{S}$ of $L_2(\mathbb{R}^d)$. Therefore, there exists a unique bounded extension T of this operator to all $L_2(\mathbb{R}^d)$, which we will call Fourier transformation on $L_2(\mathbb{R}^d)$. We will also use the notation $\widehat{f} = T(f)$ for $f \in L_2(\mathbb{R}^d)$. In general, $\widehat{f}$ for $f \in L_2(\mathbb{R}^d)$ is given as the $L_2(\mathbb{R}^d)$-limit of $\{\widehat{f_n}\}$ if $f_n \in S$ converges to f in $L_2(\mathbb{R}^d)$.

Corollary 5.25 *(***Plancherel***) There exists an isomorphic mapping $T : L_2(\mathbb{R}^d) \to L_2(\mathbb{R}^d)$ such that:*

(1) $\|Tf\|_{L_2(\mathbb{R}^d)} = \|f\|_{L_2(\mathbb{R}^d)}$ for all $f \in L_2(\mathbb{R}^d)$;
(2) $Tf = \widehat{f}$ for all $f \in L_2(\mathbb{R}^d) \cap L_1(\mathbb{R}^d)$;
(3) $T^{-1}g = g^\vee$ for all $g \in L_2(\mathbb{R}^d) \cap L_1(\mathbb{R}^d)$.

The isomorphism is uniquely determined by these properties.

Proof The proof follows from the explanation given in the paragraph above this corollary and the fact that $\mathcal{S} \subseteq L_1(\mathbb{R}^d) \cap L_2(\mathbb{R}^d)$, so that $L_1(\mathbb{R}^d) \cap L_2(\mathbb{R}^d)$ is also dense in $L_2(\mathbb{R}^d)$ with respect to the $L_2(\mathbb{R}^d)$-norm. $\square$

Finally, we will take a look at the Fourier transform of a radial function. Surprisingly, it turns out to be radial as well. This is of enormous importance in the theory to come.

Theorem 5.26 *Suppose $\Phi \in L_1(\mathbb{R}^d) \cap C(\mathbb{R}^d)$ is radial, i.e. $\Phi(x) = \phi(\|x\|_2)$, $x \in \mathbb{R}^d$. Then its Fourier transform $\widehat{\Phi}$ is also radial, i.e. $\widehat{\Phi}(\omega) = \mathcal{F}_d\phi(\|\omega\|_2)$ with*

$$\mathcal{F}_d\phi(r) = r^{-(d-2)/2} \int_0^\infty \phi(t)t^{d/2} J_{(d-2)/2}(rt)dt.$$

Proof The case $d = 1$ follows immediately from

$$J_{-1/2}(t) = \left(\frac{2}{\pi t}\right)^{1/2} \cos t.$$

In the case $d \geq 2$ we set $r = \|x\|_2$. Splitting the Fourier integral and using (5.1) yields

$$\begin{aligned}
\widehat{\Phi}(x) &= (2\pi)^{-d/2} \int_{\mathbb{R}^d} \Phi(\omega) e^{-ix^T\omega} d\omega \\
&= (2\pi)^{-d/2} \int_0^\infty t^{d-1} \int_{S^{d-1}} \Phi(t\|\omega\|_2) e^{-itx^T\omega} dS(\omega)dt \\
&= (2\pi)^{-d/2} \int_0^\infty \phi(t)t^{d-1} \int_{S^{d-1}} e^{-itx^T\omega} dS(\omega)dt \\
&= r^{-(d-2)/2} \int_0^\infty \phi(t)t^{d/2} J_{(d-2)/2}(rt)dt.
\end{aligned}$$

$\square$

5.3 Measure theory

We assume the reader to have some familiarity with measure and integration theory. The convergence results of Fatou, Beppo Levi, and Lebesgue for integrals defined by general measures should be known. Other results like Riesz' representation theorem and Helly's theorem are perhaps not standard knowledge. Hence, we now review the material relevant to us. The reader should be aware of the fact that terms like Borel measure and Radon measure have different meanings throughout the literature. Hence, when using results from measure theory it is crucial to first have a look at the definitions. Here, we will mainly use the definitions and results of Bauer [9] since his definition of a measure is to a certain extent constructive. Another good source for the results here is Halmos [78]. We will not give proofs in this short section. Instead, we refer the reader to the books just mentioned. Moreover, Helly's theorem can be found in the book [46] by Donoghue.

Let Ω be an arbitrary set. We denote the set of all subsets of Ω by $\mathcal{P}(\Omega)$. We have to introduce several names and concepts now.

Definition 5.27 *A subset $\mathcal{R}$ of $\mathcal{P}(\Omega)$ is called a ring on Ω if*
(1) $\emptyset \in \mathcal{R}$,
(2) $A, B \in \mathcal{R}$ *implies* $A \setminus B \in \mathcal{R}$,
(3) $A, B \in \mathcal{R}$ *implies* $A \cup B \in \mathcal{R}$.

The name is motivated by the fact that a ring $\mathcal{R}$ is indeed a ring in the algebraical sense if one takes the intersection $\cap$ as multiplication and the symmetric difference $\triangle$, defined by $A \triangle B := (A \setminus B) \cup (B \setminus A)$, as addition.

We are concerned with certain functions defined on a ring.

Definition 5.28 *Let $\mathcal{R}$ be a ring on a set Ω. A function $\mu : \mathcal{R} \to [0, \infty]$ is called a pre-measure if*
(1) $\mu(\emptyset) = 0$,
(2) for disjoint $A_j \in \mathcal{R}$, $j \in \mathbb{N}$, *with* $\cup A_j \in \mathcal{R}$ *we have* $\mu(\cup A_j) = \sum \mu(A_j)$.

The last property is called σ-additivity.

Note that for the second property we consider only those sequences $\{A_j\}$ of disjoint sets whose union is also contained in $\mathcal{R}$. This property is not automatically satisfied for a ring. It is different in the situation of a σ-algebra.

Definition 5.29 *A subset $\mathcal{A}$ of $\mathcal{P}(\Omega)$ is called a σ-algebra on Ω if*
(1) $\Omega \in \mathcal{A}$,
(2) $A \in \mathcal{A}$ *implies* $\Omega \setminus A \in \mathcal{A}$,
(3) $A_j \in \mathcal{A}$, $j \in \mathbb{N}$, *implies* $\cup_{j \in \mathbb{N}} A_j \in \mathcal{A}$.

Obviously, each σ-algebra is also a ring. Moreover, each ring $\mathcal{R}$, or more generally each subset $\mathcal{R}$ of $\mathcal{P}(\Omega)$, defines a smallest σ-algebra that contains $\mathcal{R}$. This σ-algebra is denoted by $\sigma(\mathcal{R})$ and is obviously given by

$$\sigma(\mathcal{R}) = \cap\{\mathcal{A} : \mathcal{A} \text{ is a } \sigma\text{-algebra and } \mathcal{R} \subseteq \mathcal{A}\}.$$

We also say that $\mathcal{R}$ generates the σ-algebra $\sigma(\mathcal{R})$.

Definition 5.30 *A pre-measure defined on a σ-algebra is called a measure. The sets in the σ-algebra are called measurable with respect to this measure.*

Obviously measurability depends actually more on the σ-algebra than on the actual measure.

Since any ring $\mathcal{R}$ is contained in $\sigma(\mathcal{R})$ it is natural to ask whether a pre-measure μ on $\mathcal{R}$ has an extension $\widetilde{\mu}$ to $\sigma(\mathcal{R})$, meaning that $\widetilde{\mu}(A) = \mu(A)$ for all $A \in \mathcal{R}$. The answer is affirmative.

Proposition 5.31 *Each pre-measure μ on a ring $\mathcal{R}$ on Ω has an extension $\widetilde{\mu}$ to $\sigma(\mathcal{R})$.*

The measures introduced so far will also be called *nonnegative measures* in contrast with *signed measures*. A signed measure is a function $\mu : \mathcal{A} \to \overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, \infty\}$ defined on a σ-algebra $\mathcal{A}$, which is σ-additive but not necessarily nonnegative.

A measure is called finite if $\mu(\Omega) < \infty$. The total mass of a nonnegative measure is given by $\|\mu\| := \mu(\Omega)$. A signed measure μ can be decomposed into $\mu = \mu_+ - \mu_-$ with two nonnegative measures μ_+, μ_-. In this case the total mass is defined by $\|\mu\| = \|\mu_+\| + \|\mu_-\|$. We will also use the notation $\|\mu\| = \int_\Omega |d\mu|$.

In the case where Ω is a topological space, further concepts are usually needed.

Definition 5.32 *Let Ω be a topological space and $\mathcal{O}$ denote its collection of open sets. The σ-algebra generated by $\mathcal{O}$ is called the Borel σ-algebra and denoted by $\mathcal{B}(\Omega)$. If Ω is a Hausdorff space then a measure μ defined on $\mathcal{B}(\Omega)$ that satisfies $\mu(K) < \infty$ for all compact sets $K \subseteq \Omega$ is called a Borel measure. The carrier of a Borel measure μ is the set* $\Omega \setminus \{U : U \text{ is open and } \mu(U) = 0\}$.

Note that a Borel measure is more than just a measure defined on Borel sets. The assumption that Ω is a Hausdorff space ensures that compact sets are closed and therefore measurable. A finite measure defined on Borel sets is automatically a Borel measure.

If Q is a subset of a Hausdorff space Ω then $\mathcal{B}(Q)$ is given by $\mathcal{B}(Q) = Q \cap \mathcal{B}(\Omega)$, using the induced topology on Q.

In case of $\mathbb{R}^d$, it is well known that $\mathcal{B}(\mathbb{R}^d)$ is also generated by the set of all semi-open cubes $[a, b) := \{x \in \mathbb{R}^d : a_j \le x_j < b_j\}$. To be more precise, it is known that the set $\mathcal{F}^d$, which contains all finite unions of such semi-open cubes, is a ring. Hence, any pre-measure defined on $\mathcal{F}^d$ has an extension to $\mathcal{B}(\mathbb{R}^d)$.

After introducing the notation for measures, the next step is to introduce measurable and integrable functions with respect to a certain measure. Since this is standard again, we omit the details here and proceed by stating those results that we will need later on.

Theorem 5.33 (Riesz) *Let Ω be a locally compact metric space. If λ is a linear and continuous functional on $C_0(\Omega)$, which is nonnegative, meaning that $\lambda(f) \ge 0$ for all $f \in C_0(\Omega)$ with $f \ge 0$, then there exists a nonnegative Borel measure μ on Ω such that*

$$\lambda(f) = \int_\Omega f(x) d\mu(x)$$

for all $f \in C_0(\Omega)$. If Ω possesses a countable basis then the measure μ is uniquely determined.

A metric space Ω possesses a countable basis if there exists a sequence $\{U_j\}_{j\in\mathbb{N}}$ of open sets such that each open set U is the union of some of these sets.

Theorem 5.34 (Helly) *Let $\{\nu_k\}$ be a sequence of (signed) Borel measures on the compact metric space Ω of uniformly bounded total mass. Then there exists a subsequence $\{\nu_{k_j}\}$ and a Borel measure ν on Ω such that*

$$\lim_{j\to\infty} \int_\Omega f(x) d\nu_{k_j}(x)$$

exists for all $f \in C(\Omega)$ and equals $\int_\Omega f(x) d\nu(x)$.

Finally, we need to know a result on the uniqueness of measures.

Theorem 5.35 (Uniqueness theorem) *Suppose that Ω is a metric space and μ and ν are two finite Borel measures on Ω. If*

$$\int f d\mu = \int f d\nu$$

for all continuous and bounded functions f, then $\mu \equiv \nu$.

In this book we are mainly confronted with situations where Ω is $\mathbb{R}^d$, [0, 1], or $[0, \infty)$ endowed with the induced topology. These sets are obviously metric (hence Hausdorff), locally compact, and possess a countable basis. Moreover, they are complete with respect to their metric. Such spaces are sometimes called *Polish spaces* and have some remarkable properties. For example, every finite Borel measure is regular, meaning inner and outer regular. This will reassure readers who might have wondered about regularity.

Finally, whenever we work on $\mathbb{R}^d$ and do not specify a σ-algebra or a measure, we will assume tacitly that we are employing Borel sets and Lebesgue measure.