1

Mathematical foundations

1.1 Set theory

The language of naïve set theory is ubiquitous in geometry and even more so in stochastic geometry. The reader will find a thorough introduction in specialised textbooks. The following briefly summarises notation and defines important sets and operations which will often be employed later.

A *set* is a collection of mathematical objects taken from a suitable domain of discourse. If x is an *element* of a set S this is written as $x \in S$. All sets appearing in this book are constructed from two fundamental sets, which are the set of the natural numbers $\{1, 2, \ldots\}$ and the set of the real numbers (the *real line*) $\mathbb{R} = (-\infty, \infty)$. All the constructions here are suitably regular, and the more profound aspects of mathematical logic and set theory are ignored.

The notation for the sets of natural and real numbers illustrates two useful conventions for the description of sets. The braces { } in the example above enclose a description of the set of natural numbers by (implicit, infinite) enumeration. The notation (u, v) for two real numbers, perhaps equal to $-\infty$ or $+\infty$, describes the set of all real numbers x such that $u < x < v$. This set (u, v) is an *open interval* of the real line $\mathbb{R}$. *Closed* and *half-open intervals* are given by

$$\begin{aligned} (u, v] &= \{x \in \mathbb{R} : u < x \leq v\} \quad \text{(half-open)}, \\ [u, v) &= \{x \in \mathbb{R} : u \leq x < v\} \quad \text{(half-open)}, \\ [u, v] &= \{x \in \mathbb{R} : u \leq x \leq v\} \quad \text{(closed)}. \end{aligned}$$

Here the braces { } enclose a description of a set as the collection of elements of another set satisfying some property. Contraction of this notation is often used, as for example:

$$\begin{aligned} &\{x \in \mathbb{R} : x = y + z \text{ with } 0 < y < 1 \text{ and } 0 < z < 1\} \\ &\quad = \{x : x = y + z \quad \text{with } 0 < y < 1 \text{ and } 0 < z < 1\} \\ &\quad = \{y + z : 0 < y < 1 \text{ and } 0 < z < 1\}. \end{aligned}$$

Stochastic Geometry and its Applications, Third Edition.
Sung Nok Chiu, Dietrich Stoyan, Wilfrid S. Kendall and Joseph Mecke.

Note that this set is actually the open interval (0, 2). Call A a *subset* of a set S (and S a *superset* of A), and write $A \subset S$, if all elements of A are also elements of S. (This book does not use the symbol $\subseteq$, thus $\subset$ includes also the case that $A = S$.) If $A, B \subset S$ for some set S then their *union, intersection*, and *difference* are

$$\begin{aligned} &\text{union} && A \cup B = \{x \in S : x \in A \text{ or } x \in B\}, \\ &\text{intersection} && A \cap B = \{x \in S : x \in A \text{ and } x \in B\}, \\ &\text{difference} && A \setminus B = \{x \in S : x \in A \text{ and } x \notin B\}. \end{aligned}$$

Also define the *complement* A^{c} of A in S as

$$\begin{aligned} A^{\mathrm{c}} &= \{x \in S : \ x \notin A\} \\ &= S \setminus A. \end{aligned}$$

Notice that the definition of A^{c} depends on the – usually implicit – choice of superset S. The empty set $\emptyset$ is the set that contains no elements. Formally, it is

$$\emptyset = S \setminus S = A \setminus A$$

for any A.

Special collections of sets (σ-*algebras*) are considered in Section 1.9.

Two sets A and B can be used to form the *Cartesian product* $A \times B$ given by the ordered pairs (a, b), that is,

$$A \times B = \{(a, b) : a \in A \text{ and } b \in B\}.$$

More generally, the Cartesian product of n sets $A_1, \ldots, A_n$ is

$$A_1 \times \cdots \times A_n = \{(a_1, \ldots, a_n) : a_1 \in A_1, \ldots, a_n \in A_n\}.$$

An important example is given by

$$\begin{aligned} \mathbb{R}^2 &= \mathbb{R} \times \mathbb{R} \\ &= \{(x_1, x_2) : x_1, x_2 \in \mathbb{R}\} \end{aligned}$$

which is the Cartesian plane. The higher-dimensional counterparts are

$$\mathbb{R}^3 = \mathbb{R} \times \mathbb{R} \times \mathbb{R}$$

and

$$\mathbb{R}^d = \{(x_1, \ldots, x_d) : x_1, \ldots, x_d \in \mathbb{R}\}.$$

The spaces $\mathbb{R}^2$ and $\mathbb{R}^3$ are often referred to as the *plane* and *space*, respectively, and $\mathbb{R}^d$ as the *d-dimensional space*. Because of additional structures such as topology (see Section 1.2) and linearity (Section 1.3), the term *Euclidean space* is used. An element $x = (x_1, x_2, \ldots, x_d) \in \mathbb{R}^d$ is usually referred to as a *point* in geometry. However, in stochastic geometry a distinction must be drawn. The study of stochastic geometry frequently concerns random collections of points, referred to as *point processes*. It is convenient to refer to members of such processes as *points of the process*, or simply *points*. Therefore points that are merely locations

in $\mathbb{R}^d$ with no membership presumed of some special collection of points are referred to as *location points* or *locations*. A particular case of a location point is the *origin*

$$o = (0, \ldots, 0).$$

Here this book makes a notational difference between a real number (0) and a point of $\mathbb{R}^d$, while otherwise normal italic letters are used both for points and numbers, with one exception:

r is always a real number and $\mathbf{r}$ a point of $\mathbb{R}^d$ of distance r from the origin o.

1.2 Topology in Euclidean spaces

A concept of distance is associated with Euclidean spaces. The *Euclidean metric* measures the distance between two location points $x = (x_1, x_2, \ldots, x_d)$ and $y = (y_1, y_2, \ldots, y_d)$ as

$$\|x - y\| = \sqrt{(x_1 - y_1)^2 + \cdots + (x_d - y_d)^2}. \tag{1.1}$$

This distance is used to define the *closed ball* $B(a, r)$ of *centre* a and of *radius* r

$$B(a, r) = \{x \in \mathbb{R}^d : \|x - a\| \leq r\}. \tag{1.2}$$

Here r is a positive real number and a is a location point. The *open ball* $B^{\text{int}}(a, r)$ is defined similarly but with strict rather than weak inequality, that is,

$$B^{\text{int}}(a, r) = \{x \in \mathbb{R}^d : \|x - a\| < r\}.$$

This means that the boundary of the ball, that is, the *sphere*, does not belong to the set.

The metric (or equivalently the closed and open balls) can be used to define special properties that might be possessed by subsets A of Euclidean space. The set A is said to be *bounded* if there is a ball $B(a, r)$ such that

$$A \subset B(a, r).$$

A sequence $x_1, x_2, \ldots$ is said to *converge* to x if

$$\lim_{n \to \infty} \|x_n - x\| = 0.$$

A set A is said to be *open* if for each $x \in A$ a positive number ε can be found (depending on x) such that $B(x, \varepsilon) \subset A$. The system of open sets of $\mathbb{R}^d$ is denoted by $\mathbb{O}$. (In this and other similar notations the exponent 'd' is omitted whenever the actual dimension is clear in the context.) Examples of open sets in the one-dimensional case of $d = 1$ are the open intervals (u, v). In the general case of $\mathbb{R}^d$ examples include the open balls $B^{\text{int}}(a, r)$ described above, and the open hypercubes

$$(u_1, v_1) \times \cdots \times (u_d, v_d).$$

Finite intersections and arbitrary unions of open sets are again open.

A set A is said to be *closed* if its complement A^{c} in $\mathbb{R}^d$ is open. The system of closed sets of $\mathbb{R}^d$ is denoted by $\mathbb{F}$. Examples of closed sets in the case $d = 1$ are the closed intervals $[a, b]$. In the general case of $\mathbb{R}^d$ examples include the closed balls $B(a, r)$ and the closed hypercubes

$$[u_1, v_1] \times \cdots \times [u_d, v_d]$$

and the hyperplanes

$$\left\{ x = (x_1, \ldots, x_d) \in \mathbb{R}^d : \sum_{i=1}^{d} a_i x_i = b \right\}$$

for some constants b and $a_1, \ldots, a_d$ with $a_1, \ldots, a_d$ not all zero. Finite unions of closed sets and arbitrary intersections of closed sets are again closed.

The *interior* A^{int} of a general set A is the union of all open sets contained in A. The *closure* A^{cl} of A is the intersection of all closed sets containing A. Thus A^{int} is the largest open set contained in A, while A^{cl} is the smallest closed set containing A. Hence

$$A^{\mathrm{int}} \subset A \subset A^{\mathrm{cl}}.$$

Moreover $A^{\mathrm{int}} = \left((A^{\mathrm{c}})^{\mathrm{cl}}\right)^{\mathrm{c}}$. A set A is open precisely when $A^{\mathrm{int}} = A$, and closed precisely when $A^{\mathrm{cl}} = A$. If $A = (A^{\mathrm{int}})^{\mathrm{cl}}$ then A is said to be *regular closed*.

The difference $\partial A = A^{\mathrm{cl}} \setminus A^{\mathrm{int}}$ is the *boundary* of A. An important example is the *sphere* of centre a and radius r, which is the boundary of $B(a, r)$ and is given by

$$\partial B(a, r) = \{x \in \mathbb{R}^d : \|a - x\| = r\}.$$

The particular case $\partial B(o, 1)$ is the *unit sphere* of $\mathbb{R}^d$, and is denoted by S^{d-1}.

A set $K \subset \mathbb{R}^d$ is said to be *compact* if it is both closed and bounded. The system of all compact subsets of $\mathbb{R}^d$ is denoted by $\mathbb{K}$. Moreover denote the system of non-empty compact sets by

$$\mathbb{K}' = \mathbb{K} \setminus \{\emptyset\}.$$

Examples of compact sets include the closed balls and the closed hypercubes.

The *distance between a point x and a closed set F* is denoted by $d(x, F)$ and defined as

$$d(x, F) = \inf\{\|x - y\| : y \in F\}.$$

In image analysis one speaks of the *Euclidean distance function* (or *transform*) which assigns to each $x \in F^{\mathrm{c}}$ the value $d(x, F)$. The *signed* Euclidean distance function is defined for all $x \in \mathbb{R}^d$ and takes for $x \in F$ the value $-d(x, \partial F)$. Ohser and Schladitz (2009) discuss algorithms for efficient calculations of the Euclidean distance function and give examples for its use.

1.3 Operations on subsets of Euclidean space

Euclidean space is a *vector space* since it allows the vector space operations of

$$\begin{aligned} &\text{addition} && x + y = (x_1 + y_1, \ldots, x_d + y_d), \\ &\text{scalar multiplication} && c \cdot x = cx = (cx_1, \ldots, cx_d) \end{aligned}$$

for location points $x = (x_1, \ldots, x_d)$ and $y = (y_1, \ldots, y_d)$ in $\mathbb{R}^d$ and real numbers c. This vector structure allows for the definition of set operations special to Euclidean space as follows:

Multiplication by real numbers

$$cA = \{c \cdot x : x \in A\}$$

for real numbers c and $A \subset \mathbb{R}^d$. The case $c = -1$ leads to the particular case of *reflection*

$$\check{A} = -A = \{-x : x \in A\} \qquad \text{for } A \subset \mathbb{R}^d.$$

If $A = \check{A}$ then A is said to be *symmetric*.

Translation

$$A_x = A + x = \{y + x : y \in A\} \qquad \text{for } x \in \mathbb{R}^d \text{ and } A \subset \mathbb{R}^d.$$

Minkowski-addition

$$A \oplus B = \{x + y : x \in A,\ y \in B\} \qquad \text{for } A, B \subset \mathbb{R}^d. \tag{1.3}$$

Applying to a set A the operation of Minkowski-addition by a set B enlarges, translates, and deforms the set A; see Figure 1.1 on p. 9. Minkowski-addition is both associative and commutative. Other properties are summarised in the following formulae:

$$A_x = A \oplus \{x\}, \tag{1.4}$$

$$A \oplus B = \bigcup_{y \in B} A_y = \bigcup_{x \in A} B_x, \tag{1.5}$$

$$A \oplus B = \{x : A \cap (\check{B})_x \neq \emptyset\}, \tag{1.6}$$

$$A \oplus (B_1 \cup B_2) = (A \oplus B_1) \cup (A \oplus B_2). \tag{1.7}$$

If $A_1 \subset A_2$ then $A_1 \oplus B \subset A_2 \oplus B$. Formula (1.5) shows that Minkowski-addition can be represented as the union of the translates B_x as x runs through A. In the special case of $B = B^{\text{int}}(o, r)$ the Minkowski-sum $A \oplus B$ is the union of all location points that are of distance smaller than r from A.

Minkowski-subtraction

$$A \ominus B = \bigcap_{y \in B} A_y. \tag{1.8}$$

Equivalent forms are

$$A \ominus B = (A^{\mathrm{c}} \oplus B)^{\mathrm{c}} \tag{1.9}$$

and

$$A \ominus B = \{x : (\check{B})_x \subset A\}, \tag{1.10}$$

where the complement is with respect to the superset $\mathbb{R}^d$. From the definition it is easy to show that

$$A_x \ominus B = A \ominus B_x \qquad \text{for all } x \in \mathbb{R}^d. \tag{1.11}$$

In general, the Minkowski-subtraction is not an inverse for Minkowski-addition, but the following relationship holds

$$(A \ominus \check{B}) \oplus B \subset A \subset (A \oplus \check{B}) \ominus B.$$

Note that $\{o\} \subset B \ominus \check{B}$, and equality holds if B is bounded.

If B is the ball $B(o, r)$ then $A \ominus B$ is the union of all location points lying within A and such that balls centred at them and of radius r are completely contained in A.

Sometimes the notation

$$A_{\oplus r} = A \oplus B(o, r)$$

and

$$A_{\ominus r} = A \ominus B(o, r)$$

is used, where r is a positive number.

The family $\mathbb{K}'$ of non-empty compact subsets of $\mathbb{R}^d$ can be made into a metric space by using the *Hausdorff metric*:

$$\delta(K_1, K_2) = \inf\{r : K_1 \subset K_2 \oplus B(o, r) \text{ and } K_2 \subset K_1 \oplus B(o, r)\} \qquad \text{for } K_1, K_2 \in \mathbb{K}'.$$

This δ defines distances between compact sets which are positive even when the sets have a non-empty intersection. It is easy to see that δ is compatible with the Euclidean metric in the sense that

$$\delta(\{x_1\}, \{x_2\}) = \|x_1 - x_2\| \qquad \text{for } x_1, x_2 \in \mathbb{R}^d.$$

And the distance between a point x and a compact K as defined in Section 1.2 satisfies

$$d(x, K) = \delta(\{x\}, K).$$

Furthermore, the family $\mathbb{K}'$ equipped with the metric δ is *complete*: if $K_1, K_2, \ldots \in \mathbb{K}'$ with $\delta(K_m, K_n) \to 0$ as $m, n \to \infty$ then there is a $K_\infty \in \mathbb{K}'$ with $\delta(K_m, K_\infty) \to 0$ as $m \to \infty$. The metric is also *countably separated*: there is a sequence $K_1, K_2, \ldots \in \mathbb{K}'$ such that for any $K \in \mathbb{K}'$ and any $\varepsilon > 0$ there is a K_n with $\delta(K_n, K) < \varepsilon$. (A possible choice of such a sequence $K_1, K_2, \ldots$ are the finite subsets of $\mathbb{Q}^d$, where $\mathbb{Q}$ denotes the set of rational numbers. These finite subsets can be enumerated, as every mathematician knows.) So $\mathbb{K}'$ is a *complete separable metric space* when endowed with the Hausdorff metric: that is to say, it is a *Polish space*.

1.4 Mathematical morphology and image analysis

In many fields of technology and science there is a great need for methods of analysis for large quantities of data in the form of *images*. Examples are satellite photographs, geological maps, microscope images of sections of metals, minerals, cellular tissue, and data coming from computerised tomography. The sheer quantity of data requires the use of automatic and quantitative methods. This section describes briefly some ideas of mathematical morphology applied in this context. In particular, the reader will get an idea how statistical procedures on random sets (such as described throughout the rest of this book) can be performed automatically.

Technical equipment yields image data which usually have the form of two- or three-dimensional arrays of *pixels* with grey values. Such a greyscale image can be reduced to a binary image by operations like thresholding, in which grey-tone values lower than a chosen threshold are set to white, and the others to black. The following discussion considers only such binary images, which correspond to random sets, where for example the black pixels stand for some set and the white for its complement.

Once the image is reduced to such an array of pixels then it is possible (with suitable equipment and software) to determine important image characteristics automatically. These quantities are represented by numbers of pixels in specific subsets of the image. Examples are the areas or volumes of the white and black parts of the image, the length or area content of the boundary between these parts, and the number of components of the black part. If one conceives of the image as a realisation of a random structure then these measurements lead to statistical estimates of various characteristics of the random structure. For example, the area fraction p is estimated by the proportion of area covered by black pixels (see Sections 3.4 and 6.4.2), while the specific length of boundary is estimated by the proportion of length of boundary between black and white to unit area in the image (see Sections 7.3 and 8.3).

However, more is possible. A powerful idea is to subject a given image to repeated transformations and to perform measurements on the transformed images (Serra, 1982). On the one hand, automatic operations can be applied to the image to free it from the inevitable image defects and artifacts of the image-processing procedure. This is often done as a preliminary step before visual inspection. On the other hand, important functions can be measured quickly and easily. Examples of these functions are the set covariance, chord length distribution function and contact distribution functions, which will be explained in Section 6.3.

Important image transformations have their origin in *mathematical morphology*, as introduced by Matheron and Serra. Many of its transformations employ 'structuring elements'. This theory is a direct application of the Euclidean set operations discussed in the previous section and it is a useful technical aid in the analysis of images. For thorough discussions, see Heijmans (1994), Goutsias and Heijmans (2000), Serra (1982) and Soille (1999), and for the three-dimensional case, see Ohser and Schladitz (2009).

In the following A will denote the set which is of interest, also called the *image*: in practical applications this is often the set of black pixels. The structuring element B will often be a ball, a disc, a line segment, or a two-point set. In practice, when images are pixel sets, balls or discs are approximated by also pixel sets and the set operations are approximated by operations on lattices, see for example Goutsias and Heijmans (2000). However, the following description follows Euclidean geometry.

The free `R` package `spatstat` can perform these operations when B is a disc, whilst the commercial `Image Processing Toolbox` of `MATLAB` allows users to create and manipulate the structuring element.

Dilation

This is the operation

$$A \mapsto A \oplus \check{B}. \tag{1.12}$$

The set A is enlarged (but not scaled) and, at least in the case where the structuring element B is a ball, smoothed. In particular the action of dilation fills in cavities, repairs fissures, and joins together a fragmented image. If A is a realisation of a random closed set, then dilation by rB followed by measurement of the area of $A \oplus r\check{B}$ allows estimation of the contact distribution function $H_B(r)$ of the random set (see Sections 6.3.3 and 6.4.5).

Erosion

This is in some sense dual to the operation of dilation, and is given by

$$A \mapsto A \ominus \check{B}. \tag{1.13}$$

Erosion shrinks the set A, tending to produce smaller fragments, even separating connected sets into several subsets. This can be helpful in the estimation of the number of particles composing an image, where some particles are in contact. The erosion operation is of great importance for the quantitative estimation of various summary characteristics of random sets. For example, if one takes for the structuring element B a two-point set $\{x, y\}$ with $\|x - y\| = r$, then the normalised area of the eroded set $A \ominus \check{B}$ is an estimate for the set covariance $C(r)$ (see Section 6.3.2).

If A is a union of non-overlapping discs, then an estimate of the diameter distribution can be obtained by means of successive erosions of A by discs $B(o, r_1)$, $B(o, r_2)$, ... with $r_1 < r_2 < \cdots$, and counting the numbers of components of the eroded A.

Ohser and Schladitz (2009, pp. 86–7) show how the coordination number of sinter particles can be determined by erosion.

Opening

The opening of A by B can be viewed as an attempt to reverse an erosion by a dilation. It is given by

$$A \mapsto A \circ B = (A \ominus \check{B}) \oplus B. \tag{1.14}$$

The opening $A \circ B$ of a set A by B has an appearance similar to that of the original set A, but is built only on the portions of the image that survive the initial erosion. Thus $A \circ B \subset A$; small disconnected fragments of the image disappear under opening and this is useful in systematically eliminating possible image defects or noise.

A set A is *(morphologically) B-open* if $A = A \circ B$. For example, in the plane a union of discs with radii larger than or equal to r is $B(o, r)$-open.

Closing

This is dual to opening and can be viewed as an attempt to reverse a dilation by an erosion. It is given by

$$A \mapsto A \bullet B = (A \oplus \check{B}) \ominus B. \tag{1.15}$$

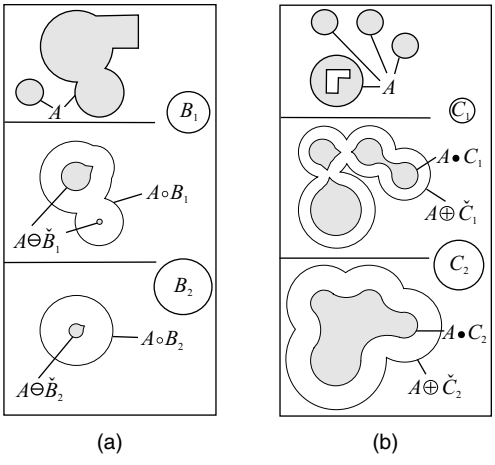

Figure 1.1 (a) The operations of erosion and opening with discs applied to a planar set A. Components that overlap are separated while small components and roughnesses vanish or are reduced. (b) The operations of dilation and closing applied to a set. Gaps are closed up, concavities vanish or are reduced, and clusters of small particles are merged.

The same as the opening $A \circ B$, the closing $A \bullet B$ bears an approximate resemblance to A, but now $A \subset A \bullet B$. As in the case of opening, the closing operation is useful in cleaning up an image. The action of closing tends to close up small holes, to join up close but separated subsets, and to smooth out the boundaries of an image. Opening and closing are simple examples of morphological filtration operations.

A set A is *(morphologically) B-closed* if $A = A \bullet B$. For example, in the plane a union of non-intersecting discs with radii smaller than r and with inter-centre distance larger than $4r$ is $B(o, r)$-open.

Figure 1.1 displays typical results of the application of these transformations. The example discussed there shows how repeated application of the transformations of mathematical morphology with perhaps different structuring elements leads to composite image transformations; see Ohser and Schladitz (2009, p. 88) for a three-dimensional example.

1.5 Euclidean isometries

A transformation $x \mapsto \boldsymbol{m}x$ is said to be a Euclidean *isometry* if it leaves invariant the distance between points x and y for all x and y. That is to say,

$$\|x - y\| = \|\boldsymbol{m}x - \boldsymbol{m}y\| \qquad \text{for all } x, y \in \mathbb{R}^d.$$

It can be shown that every isometry of Euclidean space can be represented in the form

$$\boldsymbol{m}x_k = b_k + \sum_{l=1}^{d} a_{kl} \cdot x_l \tag{1.16}$$

for $k = 1, \ldots, d$, $b = (b_1, \ldots, b_d) \in \mathbb{R}^d$, and $\mathbf{A} = (a_{kl})$ an orthogonal matrix. Such a matrix has the determinant $\det \mathbf{A} = \pm 1$. The isometry is said to be a *proper isometry* if $\det \mathbf{A} = +1$. So reflections are *not* proper isometries.

If $\mathbf{A}$ is the unit matrix then the isometry is said to be a *translation* and sometimes denoted by

$$T_b x = x - b. \tag{1.17}$$

The set A_x in (1.4) is the *translated set* of A and can be written as

$$T_{-x} A = A_x.$$

Note the composition formula

$$T_x T_y = T_{x+y}. \tag{1.18}$$

Rotations about the origin are proper isometries given by (1.16) for which $b = o$ and are denoted by $\boldsymbol{r}x$ or $\mathbf{A}x$, where $\mathbf{A}$ is an orthogonal matrix with $\det A = 1$. A further composition formula is

$$\boldsymbol{r} T_x y = T_{\boldsymbol{r}x} \boldsymbol{r} y. \tag{1.19}$$

Formula (1.16) makes it plain that every proper isometry is the composition of a rotation about the origin with a translation.

An alternative term for a proper isometry is *(rigid) motion*.

1.6 Convex sets in Euclidean spaces

A subset K of $\mathbb{R}^d$ is called *convex* if for every pair of points x, y in K the intervening line segment also lies in K. That is to say

$$cx + (1 - c)y \in K \qquad \text{whenever } x, y \in K \text{ and } 0 \le c \le 1.$$

Important examples of convex sets include the *affine linear subspaces*, which are subsets L of $\mathbb{R}^d$ with the property that for every x, y in L the whole line $\{cx + (1 - c)y : c \in \mathbb{R}\}$ through x and y also lies in L. The origin o does not necessarily belong to such an L.

The *dimension* of an affine linear subspace is the smallest integer k such that L can be given by the formula

$$L = \left\{ c_1 z_1 + \cdots + c_{k+1} z_{k+1} : c_1, \ldots, c_{k+1} \in \mathbb{R}, \sum_{i=1}^{k+1} c_i = 1 \right\}$$

for some points $z_1, \ldots, z_{k+1}$ in $\mathbb{R}^d$. Affine linear subspaces of dimension k are called *k-flats* or *k-planes*; if a *k-flat* contains o then it is called a k-subspace. The $(d-1)$-flats are called *hyperflats* or *hyperplanes*. The 1-flats are called *lines*. In the spatial case $d = 3$ the term *2-flat* is abbreviated to *flat* and the term *2-plane* to *plane*. Note that the hyperflats of the plane $\mathbb{R}^2$ are its lines.

The k-flats are closed sets but are not bounded and therefore not compact. Convex sets which are also compact are sometimes called *convex bodies*. Examples are the closed balls, the closed and bounded hypercubes, and the closed discs (intersections of closed balls with 2-flats). The system of all convex bodies is denoted by $C(\mathbb{K})$. In the case $d = 1$ the system $C(\mathbb{K})$ coincides with the system of all closed and bounded intervals. Open balls and open hypercubes are not convex bodies, though they are convex. The sphere

$$\partial B(o, r) = \{x \in \mathbb{R}^d : \|x - a\| = r\}$$

and the torus are of course not convex.

The smallest convex set which contains a given set A is called the *convex hull* of A and denoted by $\operatorname{conv} A$. For example, the convex hull of a sphere is the corresponding closed ball.

An important functional characteristic for a convex body K is the *support function* defined by

$$s(K, u) = \max_{x \in K} \langle u, x \rangle \qquad \text{for } u \in \mathbb{R}^d, \tag{1.20}$$

where $\langle u, x \rangle = u_1 x_1 + \cdots + u_d x_d$ is the scalar product of $u = (u_1, \ldots, u_d)$ and $x = (x_1, \ldots, x_d)$. The support function is convex and positively homogeneous (i.e., $s(K, \alpha u) = \alpha \cdot s(K, u)$ for all $\alpha > 0$). It determines K uniquely.

For $u \in S^{d-1}$, $s(K, u)$ is the signed distance from the origin of the support hyperplane to K with exterior normal vector u; the distance is negative if and only if u points into the open half space containing the origin o. The function $s(K, \cdot)$ is completely determined by its values on S^{d-1} because of positive homogeneity. Therefore in this book sometimes $s(K, u)$ is considered as a function on S^{d-1}, given by (1.20) with $\mathbb{R}^d$ replaced by S^{d-1}. See Schneider (1993, Section 1.7) for more information on the support function.

If K is symmetric (so $\check{K} = K$), then the support function is uniquely determined by its values on one hemisphere of the sphere S^{d-1}. In this case the *modified support function* $s_m(K, \cdot)$ is defined on the set L_1 of all lines through the origin. For $\ell \in L_1$ let $e(\ell)$ be the point on $\ell \cap S^{d-1}$ in the upper hemisphere ($x_d \geq 0$). Then

$$s_m(K, \ell) = s\big(K, e(\ell)\big) \qquad \text{for } \ell \in L_1. \tag{1.21}$$

In the planar case ($d = 2$) the line ℓ is uniquely given by the angle α formed by ℓ and the x_1-axis in the upper half-plane, and hence $s_m(K, \cdot)$ becomes a function defined on $(0, \pi]$.

Some set operations preserve the class $C(\mathbb{K})$. In particular if K_1, K_2 belong to $C(\mathbb{K})$ then so do the sets $c \cdot K_1$ for real c, $\check{K}_1$, $K_1 \cap K_2$ and $K_1 \oplus K_2$.

A *convex body functional* $h(K)$, defined on $C(\mathbb{K})$, assigns a real value $h(K)$ to each $K \in C(\mathbb{K})$. Of particular interest are those nonnegative convex body functionals which possess the following properties:

isometry-invariance	$h(\boldsymbol{m}K) = h(K)$ if $K \in C(\mathbb{K})$ and $\boldsymbol{m}$ is an isometry,
monotonicity	if $K_1 \subset K_2$ then $h(K_1) \leq h(K_2)$,
C-additivity	$h(K_1) + h(K_2) = h(K_1 \cup K_2) + h(K_1 \cap K_2)$ for $K_1, K_2 \in C(\mathbb{K})$, if $K_1 \cup K_2 \in C(\mathbb{K})$.

Important examples of convex body functionals for the cases $d = 1, 2, 3$ are

the *length* $l(K)$	if	$d = 1$,
the *area* $A(K)$	if	$d = 2$,
the *volume* $V(K)$	if	$d = 3$,
the *boundary length* $L(K)$	if	$d = 2$,
the *surface area* $S(K)$	if	$d = 3$.

In the case $d = 2$, if K is a line segment then $L(K)$ is defined as *twice* the length of K. Likewise in the case $d = 3$ if K is actually a subset of a flat then $S(K)$ is *twice* the area of K.

The *parallel set* of distance r of a set $A \subset \mathbb{R}^d$ is the set $A_{\oplus r} = A \oplus B(o, r)$. The operation of taking a parallel set preserves the properties of convexity, of compactness, and of being a ball.

Expressing the length ($d = 1$), area ($d = 2$) and volume ($d = 3$) of the parallel set as functions of the distance r is of particular interest. For the case $d = 1$ this is given simply by

$$l\big(K \oplus B(o, r)\big) = l(K) + 2r. \tag{1.22}$$

In the cases $d = 2$ and $d = 3$ the *Steiner formula* holds

$$A\big(K \oplus B(o, r)\big) = A(K) + L(K)\, r + \pi r^2, \tag{1.23}$$

$$V\big(K \oplus B(o, r)\big) = V(K) + S(K)\, r + 2\pi \bar{b}(K) r^2 + \frac{4\pi r^3}{3}. \tag{1.24}$$

Here $\bar{b}(K)$ is yet another convex body functional, called the *average breadth* or *average width*.

The average breadth can be defined as follows. For each line ℓ through the origin let $b_\ell(K)$ be the least distance between two parallel hyperplanes perpendicular to ℓ and enclosing K entirely between them. Then $\bar{b}(K)$ is defined to be the mean value of $b_\ell(K)$ averaging over all lines ℓ through the origin using the uniform direction distribution. This can be given in an explicit formula as

$$\bar{b}(K) = \frac{1}{2\pi} \int_0^{\pi/2} \int_0^{2\pi} l(K|_{S_{\beta,\lambda}}) \sin\beta \, \mathrm{d}\lambda \, \mathrm{d}\beta, \tag{1.25}$$

where $l(K|_{S_{\beta,\lambda}})$ is the length of the orthogonal projection of K on $S_{\beta,\lambda}$, the line which passes through the origin and through the point $(\sin\beta\cos\lambda, \sin\beta\sin\lambda, \cos\beta)$.

In the special case of polyhedra the following formula holds (see Santaló, 1976, p. 226):

$$\overline{b}(K) = \frac{1}{4\pi} \sum_i l_i \alpha_i, \tag{1.26}$$

where l_i is the length of the i^{th} edge and α_i is the angle between the normals of the faces which meet at the i^{th} side, where $0 < \alpha_i \leq \pi$.

Table 1.1 displays average breadths, together with volumes and surface areas, for various convex bodies $K \subset \mathbb{R}^3$. Other formulae for average breadths can be found in Hadwiger (1957, p. 215) and Santaló (1976, pp. 226, 229, 230).

Table 1.1 Volumes, surface areas, and average breadths for convex bodies K in $\mathbb{R}^3$.

K	V	S	$\overline{b}$
Ball of radius r	$\frac{4}{3}\pi r^3$	$4\pi r^2$	$2r$
Cylinder, radius r, height h	$\pi r^2 h$	$2\pi r(r+h)$	$\frac{h+\pi r}{2}$
Disc, radius r	0	$2\pi r^2$	$\frac{\pi r}{2}$
Square plate, side a	0	$2a^2$	a
Convex flat (a planar convex subset in $\mathbb{R}^3$), area A and perimeter L	0	$2A$	$\frac{L}{4}$
Segment, length l	0	0	$\frac{l}{2}$
Spheroid, equator radius a, half axes of meridian ellipse a and λa, $\alpha = \sqrt{1-\lambda^2}$, $\beta = \sqrt{\lambda^2 - 1}$	$\frac{4}{3}\pi\lambda a^3$	$2\pi\lambda a^2(\frac{1}{\lambda} + \frac{\lambda}{\beta}\sin^{-1}\frac{\beta}{\lambda})$, if $\lambda > 1$ $2\pi\lambda a^2(\frac{1}{\lambda} - \frac{\lambda}{\alpha}\ln\frac{1-\alpha}{\lambda})$, if $\lambda < 1$	$a\{\lambda + \frac{1}{\beta}\ln(\beta+\lambda)\}$, if $\lambda > 1$ $a(\lambda + \frac{1}{\alpha}\sin^{-1}\alpha)$, if $\lambda < 1$
Cube, side a	a^3	$6a^2$	$\frac{3}{2}a$
Rectangular parallelepiped, sides a, b, c	abc	$2(ab+bc+ca)$	$\frac{a+b+c}{2}$
Tetrahedron, side a	$\frac{\sqrt{2}}{12}a^3$	$\sqrt{3}\,a^2$	$\frac{3a}{2\pi}\cos^{-1}(-\frac{1}{3})$
Octahedron, side a	$\frac{\sqrt{2}}{3}a^3$	$2\sqrt{3}\,a^2$	$\frac{3a}{\pi}\cos^{-1}\frac{1}{3}$

The average breadth $\overline{b}$ is closely related to M, the integral of mean curvature, by

$$\overline{b} = \frac{M}{2\pi}. \tag{1.27}$$

The *integral of mean curvature* M is the surface integral over ∂K, the boundary of K, of the mean of the two principal curvatures $1/r_1(x)$ and $1/r_2(x)$ of the surface

$$M = \int_{\partial K} m(x)\,\mathrm{d}S,$$

where

$$m(x) = \frac{1}{2}\left(\frac{1}{r_1(x)} + \frac{1}{r_2(x)}\right)$$

is the mean curvature of ∂K in the surface point x and $\mathrm{d}S$ is the surface element. For M to be well-defined by this integral the surface ∂K must satisfy suitable regularity conditions, though of course the average breadth makes sense for any convex body.

Formulae (1.23) and (1.24) given above can be generalised to the case of dimension d. Suppose K is in $C(\mathbb{K})$. Then the volume $v_d\big(K \oplus B(o,r)\big)$ of a parallel body for K is given by the *Steiner formula*:

$$v_d\big(K \oplus B(o,r)\big) = \sum_{k=0}^{d} \binom{d}{k} W_k(K) r^k. \tag{1.28}$$

This formula introduces the important *quermassintegrals* or *Minkowski functionals* $W_k(K)$. They are isometry-invariant, monotone, C-additive convex body functionals, defined directly by the formula

$$W_k(K) = \frac{b_d}{b_{d-k}} \int_{\mathbb{L}_k} v_{d-k}(K|_{E^\perp}) U_k(\mathrm{d}E), \tag{1.29}$$

in which

v_k	is the k-dimensional Lebesgue measure (see p. 30),
$\mathbb{L}_k$	is the set of all k-subspaces,
$K\|_{E^\perp}$	is the orthogonal projection of K on $E^\perp$,
$E^\perp$	is the $(d-k)$-subspace orthogonal to $E \in \mathbb{L}_k$,
U_k	is the uniform probability distribution on $\mathbb{L}_k$.

Furthermore

$$b_k = \frac{\sqrt{\pi^k}}{\Gamma(1 + k/2)} \tag{1.30}$$

is the volume of the unit ball in $\mathbb{R}^k$. Important special cases are

$$b_0 = 1, \qquad b_1 = 2, \qquad b_2 = \pi, \qquad b_3 = 4\pi/3.$$

If K is the unit ball, then

$$W_k\big(B(o,1)\big) = b_d \qquad \text{for } k = 0, 1, \ldots, d. \tag{1.31}$$

For a general convex body K, the quantity $W_0(K)$ is equal to the volume $v_d(K)$; the quantity $dW_1(K)$ is the $(d-1)$-*dimensional measure of area of the boundary* ∂K; $W_d(K)$ is constant (independent of K) and equal to b_d, and $(2/b_d)W_{d-1}(K)$ is the average breadth. In the particular cases of $d = 1, 2, 3$ this gives

$$d = 1: \quad W_0(K) = l(K), \quad W_1(K) = 2, \tag{1.32}$$

$$d = 2: \quad W_0(K) = A(K), \quad W_1(K) = \frac{L(K)}{2}, \quad W_2(K) = \pi, \tag{1.33}$$

$$\begin{aligned} d = 3: \quad & W_0(K) = V(K), \quad W_1(K) = \frac{S(K)}{3}, \\ & W_2(K) = \frac{2\pi}{3}\bar{b}(K) = \frac{M(K)}{3}, \quad W_3(K) = \frac{4\pi}{3}. \end{aligned} \tag{1.34}$$

The Minkowski functionals W_k are closely related to the so-called *intrinsic volumes* V_k:

$$b_{d-k}V_k(K) = \binom{d}{k} W_{d-k}(K) \qquad \text{for } k = 0, 1, \ldots, d. \tag{1.35}$$

Which functional should be used is merely a matter of mathematical convention. The V_k can be considered more natural and depend only on K but not the dimension of its surrounding space (Schneider and Weil, 2008, p. 600). Therefore this book prefers the V_k. Substituting W_0, W_1, W_d and W_{d-1} into Formula (1.35) gives, for example,

$$V_0(K) = 1, \tag{1.36}$$

$$V_{d-1}(K) = \frac{1}{2}S_d(K), \tag{1.37}$$

$$V_d(K) = v_d(K), \tag{1.38}$$

where $S_d(K)$ is the $(d-1)$-dimensional area of ∂K. The relationships corresponding to Formulae (1.28) and (1.31)–(1.34), respectively, are

$$v_d\big(K \oplus B(o,r)\big) = \sum_{k=0}^{d} b_{d-k}V_k(K)r^{d-k}, \qquad \text{(Steiner formula)} \tag{1.39}$$

$$V_k\big(B(o,1)\big) = \binom{d}{k}\frac{b_d}{b_{d-k}} \qquad \text{for } k = 0, 1, \ldots, d, \tag{1.40}$$

$$d = 1: \quad V_0(K) = 1, \quad V_1(K) = l(K), \tag{1.41}$$

$$d = 2: \quad V_0(K) = 1, \quad V_1(K) = \frac{L(K)}{2}, \quad V_2(K) = A(K), \tag{1.42}$$

$$d = 3: \quad V_0(K) = 1, \quad V_1(K) = 2\bar{b}(K) = \frac{M(K)}{\pi},$$
$$V_2(K) = \frac{S(K)}{2}, \quad V_3(K) = V(K). \tag{1.43}$$

Hadwiger's characterisation theorem states that every nonnegative, motion-invariant, monotone, C-additive convex body functional h can be written in the form

$$h(K) = \sum_{k=0}^{d} a_k V_k(K), \tag{1.44}$$

where the a_k are nonnegative constants depending on h. Here 'nonnegative monotone' can be replaced by 'continuous' (using Hausdorff metric) if the a_k are allowed to be general real constants. For a modern proof see Schneider and Weil (2008, pp. 628–30), the idea of which can be traced back to Klain (1995).

The volume itself is an intrinsic volume as noted above. Formula (1.39) can be generalised to apply not only to the volume $v_d = V_d$ but to the other intrinsic volumes. The generalised Steiner formula is

$$V_k\big(K \oplus B(o,r)\big) = \sum_{j=0}^{k} \binom{d-j}{d-k} \frac{b_{d-j}}{b_{d-k}} V_j(K) r^{k-j} \qquad \text{for } k = 0, 1, \ldots, d. \tag{1.45}$$

In particular, Formula (1.45) gives

$$d = 2: \qquad L\big(K \oplus B(o,r)\big) = L(K) + 2\pi r \tag{1.46}$$

$$d = 3: \qquad S\big(K \oplus B(o,r)\big) = S(K) + 4\pi\bar{b}(K)\, r + 4\pi r^2, \tag{1.47}$$

$$\bar{b}\big(K \oplus B(o,r)\big) = \bar{b}(K) + 2r. \tag{1.48}$$

Matheron (1978), Miles (1974b) and Weil (1982b) give similar formulae for $V\big(K \ominus B(o,r)\big)$ under suitable smoothness conditions on ∂K.

The intrinsic volumes for intersections of convex bodies with flats satisfy the *Crofton formula*

$$\int_{\mathbb{L}_k} \int_{E^\perp} V_j(K \cap E_x) v_{d-k}(\mathrm{d}x) U_k(\mathrm{d}E) = \frac{\binom{k}{k-j} \frac{b_k}{b_j}}{\binom{d}{k-j} \frac{b_d}{b_{d-k+j}}} V_{d-k+j}(K) \tag{1.49}$$

for $0 \le j \le k \le d-1$ and $K \in C(\mathbb{K})$. (Recall that E_x is the translate of E by x.) The Lebesgue measure v_{d-k} on $E^\perp$ and the intrinsic volume V_j on E_x are defined by identifying $E^\perp$ and E_x with $\mathbb{R}^{d-k}$ and $\mathbb{R}^k$, respectively.

If $j = 0$ then (1.49) gives

$$V_k(K) = \frac{\binom{d}{d-k} \frac{b_d}{b_k}}{b_{d-k}} \int_{\mathbb{L}_{d-k}} v_k(K|_{E^\perp}) U_{d-k}(\mathrm{d}E), \tag{1.50}$$

which, after substituting into (1.35), is equivalent to Formula (1.29). The intrinsic volumes of intersections of a given convex body with another moved convex body are given by the so-called *principal kinematic formula* (see Schneider and Weil, 2008, Theorems 5.1.3 and 5.1.5).

The results of this and Section 1.8 belong to the fields of *convex geometry* and *integral geometry*. References to these branches of mathematics are Gruber and Wills (1993), Schneider (1993), Klain and Rota (1997) and Schneider and Weil (2008). Modern integral geometry studies generalisations of intrinsic volumes known as *curvature measures*; see Schneider (1993), Schneider and Weil (2008) and Section 7.3.4.

1.7 Functions describing convex sets

1.7.1 General

This section presents some functions that describe size and shape of deterministic convex bodies in $\mathbb{R}^d$, additionally to the support function already mentioned in Section 1.6. This includes the set covariance, the chord length probability density function and the erosion–dilation functions. All these functions serve as descriptors of sets, for example in statistical analysis, where functional data are simpler to handle than set-valued data. They also appear in formulae for other characteristics or can be directly obtained by statistical analysis.

All these functions are independent of the positions and orientations of the sets in space, thus they coincide for congruent sets; it is not necessary to define 'centres' in the sets. That is why these functions are also well suited to shape statistics; see Stoyan and Stoyan (1994). However, they do not uniquely characterise the corresponding convex bodies, that is, there may be different sets with the same function.

In the planar case, both the erosion–dilation function and the chord length probability density function can be easily determined by image analysis.

1.7.2 Set covariance

The *set covariance* $\gamma_K(r)$, *geometric covariogram* or *distance probability*, introduced by Porod (1951), of a convex body K is defined by

$$\gamma_K(\mathbf{r}) = v_d\big(K \cap (K - \mathbf{r})\big), \tag{1.51}$$

where $\mathbf{r}$ is a vector of $\mathbb{R}^d$ with $\|\mathbf{r}\| = r$.

In integral form, the set covariance function $\gamma_K(\mathbf{r})$ can be written as

$$\gamma_K(\mathbf{r}) = \int_{\mathbb{R}^d} \mathbf{1}_K(x)\mathbf{1}_K(x - \mathbf{r})\mathrm{d}x \qquad \text{for } \mathbf{r} \in \mathbb{R}^d, \tag{1.52}$$

where $\mathbf{1}_K(x)$ is the indicator function of K, defined by

$$\mathbf{1}_K(x) = \begin{cases} 1 & \text{for } x \in K \\ 0 & \text{otherwise.} \end{cases} \tag{1.53}$$

A physicist would perhaps write the integral as

$$\int_K \mathrm{d}v_1 \int_K \mathrm{d}v_2 \delta(\mathbf{r}_1 + \mathbf{r} - \mathbf{r}_2),$$

using the Dirac delta function δ (see Section 1.9).

It is natural to conjecture that the set covariance $\gamma_K(\mathbf{r})$ is able to characterise the set K uniquely, up to translations and reflections. This has been proved for planar convex K (Averkov and Bianchi, 2009) and for convex polyhedra in $\mathbb{R}^3$ (Bianchi, 2009), but in general it is not true for convex polytopes in $\mathbb{R}^d$, $d \geq 4$ (Bianchi, 2005).

Frequently, the *isotropised* set covariance $\overline{\gamma}_K(r)$ is used, the average of $\gamma_K(\mathbf{r})$ over all possible directions of $\mathbf{r}$, assuming uniform directions:

$$\overline{\gamma}_K(r) = \int_{S^{d-1}} \gamma_K(r\boldsymbol{u}) U_1(\mathrm{d}\boldsymbol{u}). \tag{1.54}$$

Here r is a positive number, and $\boldsymbol{u}$ is a unit vector representing a point on the unit sphere S^{d-1} and is identified with the line containing the vector $\boldsymbol{u}$, so that with slight abuse of notation, the distribution U_1, defined on p. 14, also denotes the uniform distribution on S^{d-1}.

In physics $\overline{\gamma}_K(r)$ is usually normalised such that it has the value 1 for $r = 0$. This function can be measured directly by scattering methods.

Important analytical properties of the function γ_K are known; see Matheron (1975). In particular, the derivative can be found: if $r > 0$ and $\boldsymbol{u}$ is a unit vector then $\gamma_K(r\boldsymbol{u})$ has

$$\frac{\mathrm{d}}{\mathrm{d}r}\gamma_K(r\boldsymbol{u}) = -\Big(v_{d-1}\left(\left(K \cap (K + r\boldsymbol{u})\right)|_{\boldsymbol{u}^\perp}\right)\Big).$$

Here $A|_{\boldsymbol{u}^\perp}$ is the orthogonal projection of A on the hyperplane that has $\boldsymbol{u}$ as normal vector. The geometric configuration is illustrated in Figure 1.2.

It is also the case that $\gamma_K(r\boldsymbol{u})$ is convex in r. This follows from the observation that $-\mathrm{d}\gamma_K(r\boldsymbol{u})/\mathrm{d}r$ is decreasing in r. Thus also $\overline{\gamma}_K(r)$ is decreasing and convex in r.

The isotropised set covariance $\overline{\gamma}_K(r)$ is closely related to the so-called *distance distribution* $P(r)$, which is the probability density function of the random distance between two

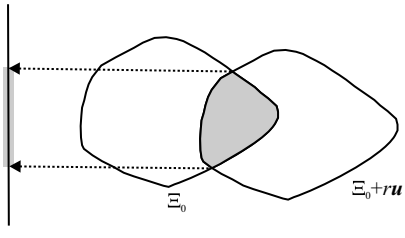

Figure 1.2 The orthogonal projection $\left(K \cap (K + r\boldsymbol{u})|_{\boldsymbol{u}^\perp}\right)$ in the case $d = 2$. It is the shaded interval on the hyperplane (which in this two-dimensional case is a line) perpendicular to the vector $\boldsymbol{u}$.

independent random points that are uniformly distributed in K. In the spatial case ($d = 3$)

$$P(r) = \frac{4\pi r^2 \overline{\gamma}_K(r)}{\big(V(K)\big)^2}, \tag{1.55}$$

and for general dimension d

$$P(r) = \frac{2\pi^{d/2}}{\Gamma(d/2)} \frac{r^{d-1}\overline{\gamma}_K(r)}{v_d(K)^2} \qquad \text{for } r \geq 0. \tag{1.56}$$

There is also a relation to the chord length distribution function; see below. The function $P(r)$ can be easily determined based on physical data, but it is not very sensitive with respect to shape variation of K; see Glatter (1979).

In general, $\overline{\gamma}_K(r)$ is not an easy function to calculate. The following are the formulae of the positive part of $\overline{\gamma}_K(r)$ for some particular bodies K:

(a) $K =$ a ball of radius R in $\mathbb{R}^3$,

$$\overline{\gamma}_K(r) = \frac{4}{3}\pi R^3 \left(1 - \frac{3r}{4R} + \frac{r^3}{16R^3}\right) \qquad \text{for } 0 \leq r \leq 2R; \tag{1.57}$$

(b) $K =$ a disc of radius R in $\mathbb{R}^2$,

$$\overline{\gamma}_K(r) = 2R^2 \cos^{-1}\left(\frac{r}{2R}\right) - \frac{r\sqrt{4R^2 - r^2}}{2} \qquad \text{for } 0 \leq r \leq 2R; \tag{1.58}$$

(c) $K =$ a rectangle of area A and side-length ratio β (≥ 1) in $\mathbb{R}^2$,

$$\overline{\gamma}_K(r) = \frac{A}{\pi} \cdot \begin{cases} \pi - 2x - \dfrac{2x}{\beta} + \dfrac{x^2}{\beta} & \text{for } 0 \leq x \leq 1, \\ 2\sin^{-1}\left(\dfrac{1}{x}\right) - \dfrac{1}{\beta} - 2(x - u) & \text{for } 1 < x \leq \beta, \\ 2\sin^{-1}\left(\dfrac{\beta - uv}{x^2}\right) + 2u + \dfrac{2v}{\beta} - \beta - \dfrac{1 + x^2}{\beta} & \text{for } \beta < x < \sqrt{\beta^2 + 1}, \\ 0 & \text{for } x \geq \sqrt{\beta^2 + 1}, \end{cases} \tag{1.59}$$

where

$$x = \frac{r}{\sqrt{A/\beta}}, \quad u = \sqrt{x^2 - 1}, \quad v = \sqrt{x^2 - \beta^2}. \tag{1.60}$$

The formula for a parallelepiped and formulae for other bodies can be found in Gille (1988) and Gille (2014).

A good approximation for $d = 3$ and for small r is

$$\overline{\gamma}_K(r) \approx V(K) - \frac{S(K)}{4} r, \tag{1.61}$$

and for $d = 2$

$$\overline{\gamma}_K(r) \approx A(K) - \frac{L(K)}{\pi} r. \tag{1.62}$$

Better but more complicated approximations are given in Ciccariello (1995).

The set covariance plays an important rôle in small-angle scattering experiments, since the intensity collected there is the Fourier transform of the autocorrelation function of the scattering density in the sample, see Guinier and Fournet (1995). A particular case is that of a monodisperse, isotropic and dilute 'particulate sample', where the sample can be understood as a system of homogeneous particles of the same shape and size, random orientation and small number density. In this case the sample autocorrelation is proportional to $\overline{\gamma}_K(r)$.

1.7.3 Chord length distribution

Random lines generate chords of random length in intersected convex sets. The standard case is that where the lines are uniform random in the sense of the motion-invariant measure discussed in Section 8.2.2, restricted to the set of lines which actually hit the convex body K. (Note that there are many other ways of defining random chords; see for example Solomon, 1978; Coleman, 1989; Chiu and Larson, 2009, and other literature on the so-called Bertrand paradox (e.g. M. G. Kendall and Moran, 1963, pp. 9–10).) To each random chord there corresponds its length, and its distribution function is called the *chord length distribution function*. For any K with inner points there exists the corresponding probability density function, which is denoted here by $f_K(l)$. Many formulae for $f_K(l)$ and its moments can be found in the literature, in particular in Gille (2014). Some of them will be given here.

Planar case

The mean chord length $\overline{\ell}_K$ is given by

$$\overline{\ell}_K = \frac{\pi A(K)}{L(K)}. \tag{1.63}$$

The third moment of the chord length is

$$\overline{\ell^3}_K = \int_0^\infty l^3 f_K(l) \mathrm{d}l = \frac{3\big(A(K)\big)^2}{L(K)} \tag{1.64}$$

(Santaló, 1976).

For the second moment of the chord length there exists no general formula comparable to (1.63) or (1.64). Of course, for many examples of planar convex sets formulae for $\overline{\ell^2}$ are known; see for example Voss (1982).

For a disc of radius R it holds

$$\overline{\ell^n} = \begin{cases} \dfrac{1 \cdot 3 \cdot \ldots \cdot n}{2 \cdot 4 \cdot \ldots \cdot (n+1)} \dfrac{\pi}{2} (2R)^n & \text{for odd } n, \\ \\ \dfrac{2 \cdot 4 \cdot \ldots \cdot n}{3 \cdot 5 \cdot \ldots \cdot (n+1)} (2R)^n & \text{for even } n. \end{cases} \tag{1.65}$$

The chord length density function $f_K(l)$ is known for many sets K. The following formulae give it for some K:

(a) $K =$ a disc of radius R,

$$f_K(l) = \frac{l}{2R\sqrt{4R^2 - l^2}} \qquad \text{for } 0 \le l \le 2R; \tag{1.66}$$

(b) $K =$ an ellipse with semiaxis lengths a and b $(a \ge b)$,

$$f_K(l) = 3abl\overline{\ell} \int_{\max\{\ell,\, 2b\}}^{2a} \frac{1}{x^3 \sqrt{x^2 - 4b^2} \sqrt{x^2 - l^2} \sqrt{4a^2 - x^2}} \, \mathrm{d}x; \tag{1.67}$$

the integral can be evaluated numerically;

(c) $K =$ a rectangle with side lengths a and b $(a \ge b)$ (Gille, 1988, 2014),

$$f_K(l) = \begin{cases} \dfrac{1}{a+b} & \text{for } 0 \le l \le b, \\ \\ \dfrac{ab^2}{l^2(a+b)\sqrt{l^2 - b^2}} & \text{for } b < l \le a, \\ \\ \dfrac{ab}{l^2(a+b)} \left(\dfrac{a}{\sqrt{l^2 - a^2}} + \dfrac{b}{\sqrt{l^2 - b^2}} \right) - \dfrac{1}{a+b} & \text{for } a < l \le \sqrt{a^2 + b^2}; \end{cases} \tag{1.68}$$

(d) $K =$ an equilateral triangle of side length a (Sulanke, 1961, Gille, 2014),

$$f_K(l) = \begin{cases} \dfrac{1}{a} \left(\dfrac{1}{2} + \dfrac{\pi}{3} \sqrt{\dfrac{1}{3}} \right) & \text{for } 0 \le l \le \dfrac{\sqrt{3}a}{2}, \\ \\ -\dfrac{1}{l} \sqrt{1 - \dfrac{3a^2}{4l^2}} + \dfrac{1}{a} \left(\dfrac{1}{2} - \dfrac{2\pi}{3\sqrt{3}} + \dfrac{2}{\sqrt{3}} \sin^{-1} \left(\dfrac{\sqrt{3}a}{2l} \right) \right) & \text{for } \dfrac{\sqrt{3}a}{2} < l \le a. \end{cases} \tag{1.69}$$

The case of a general triangle is considered in Ciccariello (2010).

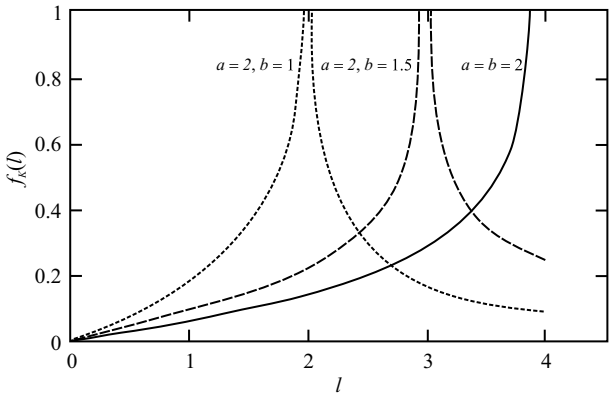

Figure 1.3 Chord length density functions for ellipses with $a = 2$ and $b = 1$, 1.5 and 2.

Figure 1.3 shows chord length density functions for two ellipses and a disc, and Figure 1.4 those for a square and a rectangle.

The form of the chord length density function is related to certain features of the corresponding body K. For example, algebraic singularities of this function correspond to parallel pieces of the contour and the form of $f_K(l)$ for l close to its maximum is essentially related to smaller details of the contour.

Spatial case

The mean chord length $\bar{\ell}_K$ is given by

$$\bar{\ell}_K = \frac{4V(K)}{S(K)}. \tag{1.70}$$

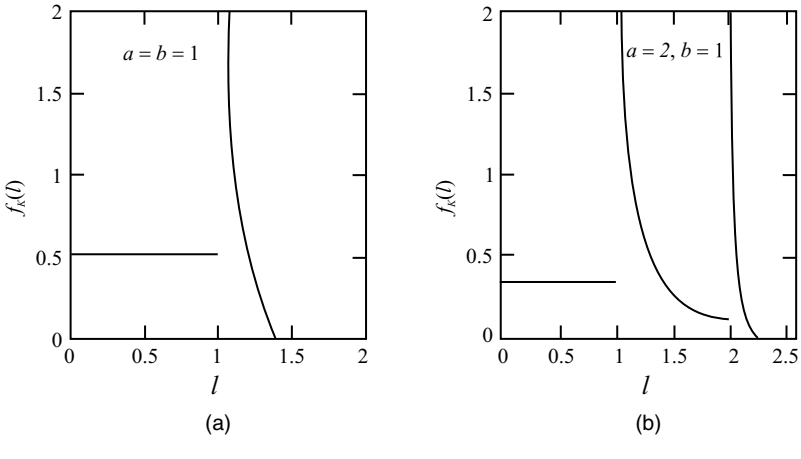

Figure 1.4 Chord length density functions for (a) a square with side length 1 and (b) a rectangle with side lengths 2 and 1.

The fourth moment is

$$\overline{\ell^4}_K = \frac{12\big(V(K)\big)^2}{\pi S(K)}. \tag{1.71}$$

Clearly, for particular bodies also the second and third moment of the chord length are known. For a ball of radius R it holds

$$\overline{\ell^n}_K = \frac{2^{n+1}}{n+2} R^n. \tag{1.72}$$

Note that there is a close relationship between the isotropised set covariance $\overline{\gamma}_K(r)$ and the chord length density function $f_K(l)$: for $d = 2$

$$f_K(l) = \frac{\pi}{L(K)} \frac{\mathrm{d}^2 \overline{\gamma}_K(l)}{\mathrm{d}l^2}, \tag{1.73}$$

and for $d = 3$

$$f_K(l) = \frac{4}{S(K)} \frac{\mathrm{d}^2 \overline{\gamma}_K(l)}{\mathrm{d}l^2} \qquad \text{for } l \geq 0 \tag{1.74}$$

(see Kingman, 1969; Gille, 2014).

These formulae are frequently used to determine chord length density functions. The case of K being a ball is treated in Section 10.4 in the context of stereology and Figure 10.4 shows the chord length density function for a parallelepiped.

In the literature further formulae can be found for $\overline{\gamma}_K(r)$ and $f_K(l)$ for various bodies K. Among the bodies for which calculations have been made are: ellipsoids, cylinders (finite and infinite), hemispheres, tetrahedra, prisms, parallelepipeds and for various isotropic random sets such as the typical cells of the Voronoi, Poisson plane and dead leaves tessellation. The results are scattered in many different journals and collected in Gille (2014). Gille asserts that there are around ten different mathematical techniques that have been applied to determine chord length distributions, which depend on the geometry of the bodies of interest.

Chords also play a rôle for nonconvex sets. There an intersecting line may produce a sequence of segments, which are best analysed together as a totality. Miles (1972a, 1985) shows how the segments can be related together so as to obtain useful formulae, for example, for moments of length. Ciccariello (2009) determines chord length distributions for planar butterfly-shaped objects, considering all section segments as single chords. There $\mathrm{d}^2\overline{\gamma}_K(r)/\mathrm{d}r^2$ can be negative.

Many physical experimental techniques indirectly measure chord length distributions. One of the physical principles is the interference of two monochromatic waves, originating at the endpoints of a chord: scattering investigations of electromagnetic waves (light, X-rays) or neutrons yield a scattering intensity curve, also called ‘diffraction pattern’. The scattering intensity I is recorded as a function of the scattering vector q, $I(q)$. Such patterns describe order ranges from nanometers to micrometers. Finally, chord length distributions result from $I(q)$ by integral transformations; see Wu and Schmidt (1973) and Burger and Ruland (2001). The series of formulae for chord length distribution functions above and in Gille (2014) may

be of value in finding suitable models when empirical distribution functions are given. Another example of the application of chord length distribution is in the classification of microscopic images of for example starch grains (Tong *et al.*, 2008).

While these applications are related to small-scale objects, chords also play a rôle in the geometrical investigation of very large objects, in astronomy in the context of occultations, for example of asteroids, which yield chord length information, see Hestroffer *et al.* (2002).

1.7.4 Erosion–dilation functions

The *erosion–dilation function* $E_K(r)$ of the convex body K is defined as

$$E_K(r) = \begin{cases} \nu_d\big(K \oplus B(o,r)\big) & \text{for } r > 0, \\ \nu_d\big(K \ominus B(o,|r|)\big) & \text{for } r \le 0. \end{cases} \tag{1.75}$$

That is, $E_K(r)$ is the volume of the inner ($r \le 0$) and outer ($r > 0$) parallel set of K. While for $r > 0$ the value of $E_K(r)$ is simply given by the Steiner formula (1.28) or (1.39), the function is more complicated for negative r. The erosion–dilation function can be easily generalised for the case of nonconvex sets — nothing in Formula (1.75) need change. In Section 6.3.3 an analogous function for unbounded sets will be used.

A normalised version of the erosion function is the *spherical erosion function* $Q_K(r)$, which has the nature of a distribution function:

$$Q_K(r) = 1 - \frac{\nu_d\big(K \ominus B(o,r)\big)}{\nu_d(K)} \qquad \text{for } r \ge 0. \tag{1.76}$$

It is the distribution function of the distance of a uniform random point in K to the boundary of K.

Like the erosion-based function, also functions that include morphologically 'opened' or 'closed' sets can be used, for example

$$G_K(r) = 1 - \frac{\nu_d\big(K \circ B(o,r)\big)}{\nu_d(K)} \qquad \text{for } r \ge 0. \tag{1.77}$$

See Serra (1982, p. 333ff) and Ripley (1988, Chapter 6).

Note that other test sets, for example point pairs or segments, can be used instead of balls $B(o,r)$.

1.8 Polyconvex sets

Some of the results concerning convex sets can be generalised for sets in a wider class known as the *convex ring* $\mathcal{R}$; see Klain and Rota (1997) and Schneider and Weil (2008). The convex ring $\mathcal{R}$ is the system of all subsets A of $\mathbb{R}^d$ which can be expressed as finite unions of convex bodies:

$$A = \bigcup_{i=1}^{n} K_i \qquad \text{for } K_i \in C(\mathbb{K}).$$

If A_1 and A_2 both belong to $\mathcal{R}$ then so do $A_1 \cup A_2$ and $A_1 \cap A_2$. The elements of the convex ring are called *polyconvex* sets.

The *extended convex ring* is the system

$$\mathbb{S} = \{K : K \cap K' \in \mathcal{R} \text{ for all convex bodies } K'\}.$$

Each element of $\mathbb{S}$, which may be a union of infinitely many convex sets, yields a polyconvex set when intersected with a convex body. The elements of $\mathbb{S}$ are called *locally polyconvex*.

An *additive functional* h on $\mathcal{R}$ is a map $h : \mathcal{R} \to \mathbb{R}$ with the properties

$$h(\emptyset) = 0$$

and

$$h(A_1 \cup A_2) + h(A_1 \cap A_2) = h(A_1) + h(A_2)$$

for $A_1, A_2 \in \mathcal{R}$.

Isometry-invariance of functionals on $\mathcal{R}$ is defined as for convex body functionals.

A very important example of an additive and isometry-invariant functional on $\mathcal{R}$ is the *connectivity number* or *Euler–Poincaré characteristic* χ. It can be understood as a generalisation of convex object count. Indeed, for a convex non-empty K it takes the value

$$\chi(K) = 1$$

and

$$\chi(A) = n,$$

if A is the union of n disjoint convex bodies. And it holds $\chi(\emptyset) = 0$. This together with the additivity property defines $\chi(A)$ for general A in $\mathcal{R}$. If A is a union of convex bodies

$$A = \bigcup_{i=1}^{n} K_i \qquad \text{for } K_i \in C(\mathbb{K}),$$

then the additivity property gives the 'inclusion–exclusion' formula

$$\chi(A) = \sum_i \chi(K_i) - \sum_{1 \le i_1 < i_2 \le n} \chi(K_{i_1} \cap K_{i_2}) + \cdots + (-1)^{n-1} \chi(K_1 \cap \cdots \cap K_n). \quad (1.78)$$

It can be shown that $\chi(A)$ is independent of the representation of A as a finite union of convex bodies. Figure 1.5 shows four planar sets of connectivity numbers $+1$, 0 and -1 respectively.

The connectivity number can, the same as volume and surface content, be defined also for closed sets outside the extended convex ring; see for example Wilder (1963) and Richeson (2008). For such cases the following informal rules hold:

- In the two-dimensional case $\chi(A)$ equals the number of outer boundaries minus the number of inner boundaries. An outer boundary is a boundary such that an observer moving along it anticlockwise sees the set on his left hand.
- In the three-dimensional case $\chi(A)$ equals the number of connected components of A, minus the number of independent two-dimensional holes or tunnels, and plus the number of three-dimensional holes or voids.

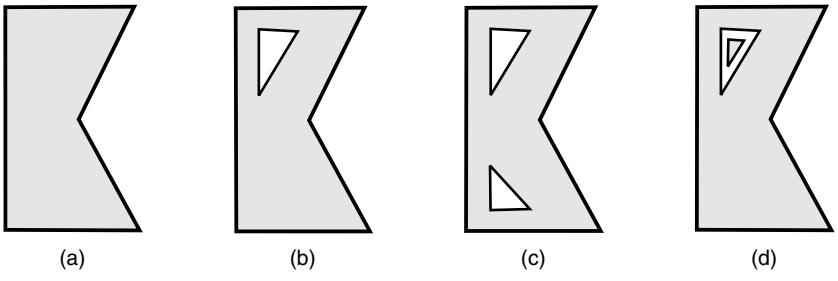

Figure 1.5 Four planar sets with various connectivity numbers: (a) $+1$, (b) 0, (c) -1, (d) $+1$.

A reader who wants to get some feeling of the three-dimensional connectivity number should note that the connectivity number of a solid torus as well as a tea cup is 0, that of a closed ball is 1 and that of a sphere as well as the surface of any convex body is 2. The set A formed by the edges of a cube has $\chi(A) = -4$.

The connectivity number makes it possible to generalise the intrinsic volumes, still denoted by V_k, to act on $\mathcal{R}$ by means of additivity, as shown by Schneider (1980) and Schneider and Weil (2008, p. 190). The term $\nu_k(K|_{E^\perp})$ in Formula (1.50) comes from the inner integral of the left-hand side of Formula (1.49) and must be replaced by

$$\int_{E^\perp} \chi(K \cap E_x)\nu_k(\mathrm{d}x)$$

so that V_0 coincides with χ on $\mathcal{R}$, and hence for a polyconvex set $A \in \mathcal{R}$, the (generalised) intrinsic volumes are defined as

$$V_k(A) = \frac{\binom{d}{d-k}\frac{b_d}{b_k}}{b_{d-k}} \int_{\mathbb{L}_{d-k}} \int_{E^\perp} V_0(A \cap E_x)\nu_k(\mathrm{d}x)U_{d-k}(\mathrm{d}E) \tag{1.79}$$

for $k = 1, \ldots, d$. These functionals are additive and isometry-invariant, but not continuous and may be negative for $k \leq d-1$. The quantity $V_d(A)$ is the d-dimensional volume of A, $V_{d-1}(A)$ is half of the area of ∂A, and $V_0(A)$ is the connectivity number of A.

Minkowski functionals of polyconvex sets, still denoted by $W_k(A)$ and connected with the $V_k(A)$ by Formula (1.35), may also be used.

The intrinsic volumes can be generalised also in another way, to the *positive extension* introduced by Matheron (1975) and Schneider (1980) and referred to in Section 7.3.4. This generalisation yields nonnegative functionals and loses additivity. A quantity that arises by this route of generalisation is $M^+(A)$, the *integral of mean positive curvature*; see also p. 294. A further positive characteristic which is a counterpart to the connectivity number is the *convexity number* χ^+. The following gives an account of its definition after Matheron (1975, pp. 122–3), but using the notation developed in this chapter.

First define the *convexity number of* $A \in \mathcal{R}$ *with respect to the unit vector* $\boldsymbol{u}$. Let $S_{\boldsymbol{u},r}$ for real r be the hyperplane through the point $r\boldsymbol{u}$ and perpendicular to $\boldsymbol{u}$. For A in $\mathcal{R}$ let $K_1, K_2, \ldots, K_n$ be the connected components of $A \cap S_{\boldsymbol{u},r}$. The connected component K_i is said to be an *entering set* for A (with respect to $\boldsymbol{u}$ and $S_{\boldsymbol{u},r}$) if there is an open and connected set G containing K_i such that $G \cap K_j$ is empty for $j \neq i$ and $A \cap G \cap S_{\boldsymbol{u},r-\varepsilon}$ is empty for all

sufficiently small positive ε. Let $n(r)$ denote the number of entering sets associated with $S_{\boldsymbol{u},r}$. If $\boldsymbol{u}$ is a fixed unit vector then this number $n(r)$ is nonzero for only a finite number of real numbers r, so that it is possible to put

$$\chi^+(A, \boldsymbol{u}) = \sum_r n(r).$$

This is the convexity number of A with respect to the unit vector $\boldsymbol{u}$. Figure 7.2 on p. 295 demonstrates an application of the convexity number in the planar case with $\boldsymbol{u} = (0, 1)$ and $\boldsymbol{u} = (0, -1)$.

The (mean) convexity number χ^+ is then simply the rotation average of $\chi^+(A, \boldsymbol{u})$ with respect to the invariant probability measure U_1 on the unit sphere S^{d-1}:

$$\chi^+(A) = \int_{S^{d-1}} \chi^+(A, \boldsymbol{u})\, U_1(\mathrm{d}\boldsymbol{u}).$$

One might ask whether the various integral-geometric formulae (the Steiner formula, Hadwiger's characterisation theorem, the principal kinematic formula, and the Crofton formula) hold in the convex ring if the standard intrinsic volumes are systematically replaced by their additive generalisations. This is indeed the case; see Schneider and Weil (2008). In particular,

$$v_d\big(A \oplus B(o, r)\big) = v_d(A) + S(A)r + o(r), \tag{1.80}$$

for regular closed A; see Schneider (1993, Theorem 4.4.1). More general sets are considered in Kiderlen and Rataj (2006).

Still more is possible: the theory can be extended to sets of positive reach and sets which are finite unions of such sets; see Zähle (1984a, 1987a) and Rother and Zähle (1990). For $A \subset \mathbb{R}^d$, reach(A) is defined to be

$$\operatorname{reach}(A) = \sup\{r\colon \text{ for all } x \in A \oplus B(o, r) \text{ there exists a unique point of } A \text{ nearest to } x\}.$$

If A is of positive reach, then A is closed. If A is convex, then reach$(A) = \infty$. An example of a set which is of positive reach but not in the convex ring is the sphere $\partial B(o, R)$. Its reach is R.

1.9 Measure and integration theory

In modern probability theory and geometry the concept of measure plays a central rôle. Consequently these ideas play a large part in the present book. However, the authors expect that many of the readers of this book are not familiar with this theory, and have therefore written the book so that most of it can be understood without a deeper knowledge of these mathematical theories.

The following introduction is of course too short for a thorough understanding of measure and integration theory. The aim is to remind mathematicians of what they have learned before and to help non-mathematicians understand the notation.

Measures are real-valued functions, defined on families of sets and enjoying properties of additivity and positivity.

To introduce the idea of a measure, the following three simple and linked examples may be helpful. Consider the contents of a body K in space that is made up of some material. Suppose that the mass of this body is not necessarily homogeneously distributed in it. Associated with each portion A of K ($A \subset K$) are the quantities $\alpha_1(A)$, the volume of A, and $\alpha_2(A)$, the mass of K that is contained in A. Suppose that in the body K there are also small grains of another material. Let the number of these grains lying in A be given by $\alpha_3(A)$. Thus there are three set functions α_1, α_2 and α_3 defined on the portions of K. If a portion A is divided into disjoint parts A_1 and A_2 then

$$\alpha_k(A) = \alpha_k(A_1 \cup A_2) = \alpha_k(A_1) + \alpha_k(A_2) \qquad \text{for } k = 1, 2, 3.$$

Set functions with this property are called *additive*.

For mathematical reasons, it is useful to stipulate that such functions should have the stronger property of *σ-additivity*. That is to say, if a set can be divided into a countable disjoint union of subsets then the value of the set function on the whole set should equal the sum of the values of the set function on the subsets. Together with various mathematical technicalities, this consideration leads to the idea that set functions of the kind given above are naturally defined on systems of sets that are closed under the basic set operations of union, intersection, and complementation, and also under the operations of taking countably infinite unions and intersections.

Such systems of sets are called *σ-algebras*. Technically a σ-algebra is a system $\mathcal{X}$ of subsets of some ground set X satisfying three conditions:

(S1) $X \in \mathcal{X}$;

(S2) if $A \in \mathcal{X}$, then $A^{\mathrm{c}} \in \mathcal{X}$;

(S3) if $A_1, A_2, \ldots \in \mathcal{X}$, then $\bigcup_{k=1}^{\infty} A_k \in \mathcal{X}$.

From these properties others follow immediately:

$\emptyset \in \mathcal{X}$;

if $A, B \in \mathcal{X}$, then $A \setminus B \in \mathcal{X}$;

if $A_1, A_2, \ldots \in \mathcal{X}$, then $\bigcap_{k=1}^{\infty} A_k \in \mathcal{X}$.

Trivial examples of σ-algebras are given by the 'minimal' σ-algebra $\{\emptyset, X\}$ and the 'maximal' σ-algebra, the *power set* $\mathcal{P}(X)$, that is, the system of all subsets of X.

Often one has to consider the *trace* $\mathcal{X}_A$ of the σ-algebra $\mathcal{X}$ on $A \subset X$, defined by

$$\mathcal{X}_A = \{B \cap A : B \in \mathcal{X}\},$$

which is also a σ-algebra.

A very important example of a σ-algebra is given by the family $\mathcal{B}^d$ of *Borel sets of* $\mathbb{R}^d$. This is the smallest σ-algebra on $\mathbb{R}^d$ that contains all the open subsets of $\mathbb{R}^d$. It can be shown that $\mathcal{B}^d$ is strictly smaller than the power set $\mathcal{P}(\mathbb{R}^d)$, that is, there are subsets of $\mathcal{B}^d$ which are not Borel.

Loosely speaking all 'reasonable' sets are Borel sets; it is difficult to give examples of non-Borel sets. The σ-algebra $\mathcal{B}^d$ contains all the subsets of $\mathbb{R}^d$ that can be constructed from the open subsets by the basic set operations and by limits. Since these operations can be iterated this is an exceedingly large class of sets. It certainly includes all closed sets, and thus all compact sets, and thus all elementary geometrical bodies discussed in this book. Denote the subclass of all bounded Borel sets of $\mathbb{R}^d$ by $\mathcal{B}_o^d$.

The pair $[X, \mathcal{X}]$, formed by a set X and a σ-algebra $\mathcal{X}$ of subsets of X, is called a *measurable space* and the A in $\mathcal{X}$ are called *measurable sets*.

A function $f : X \to \mathbb{R}$ is said to be *$\mathcal{X}$-measurable* if for each Borel set $B \in \mathcal{B}^1$ the inverse image $f^{-1}(B) = \{x \in X : f(x) \in B\}$ belongs to the σ-algebra $\mathcal{X}$ associated with X. Elementary algebraic operations and operations such as the taking of absolute values, and the taking of limits when applied to measurable functions all yield measurable functions. A particular example of a measurable function is the *indicator function* $\mathbf{1}_A(x)$ of a measurable set A as defined in (1.53) on p. 17.

The coordinate functions $x = (x_1, \ldots, x_d) \mapsto x_i$ are special examples of Borel-measurable functions from $\mathbb{R}^d$ to $\mathbb{R}$.

All continuous functions and hence all differentiable functions on $\mathbb{R}^d$ are Borel-measurable.

If $\mathbb{R}$ is replaced by a general set Y with associated σ-algebra $\mathcal{Y}$ in the definition of a measurable function then a transformation

$$f : X \to Y$$

is said to be $(\mathcal{X}, \mathcal{Y})$-measurable if $f^{-1}(B) \in \mathcal{X}$ for each $B \in \mathcal{Y}$.

Suppose that $[X, \mathcal{X}]$ is a measurable space. A *measure* on $[X, \mathcal{X}]$ is a function $\mu : \mathcal{X} \to [0, \infty]$ with the following two properties:

$$\text{(M1)} \qquad \mu(\emptyset) = 0,$$

$$\text{(M2)} \qquad \mu\left(\bigcup_{k=1}^{\infty} A_k\right) = \sum_{k=1}^{\infty} \mu(A_k)$$

for all $A_1, A_2, \ldots \in \mathcal{X}$ with $A_i \cap A_j = \emptyset$ whenever $i \neq j$. Note that $\mu(A)$ may equal ∞. Property (M2) is referred to as the property of *σ-additivity*. A particular case of this property is finite additivity:

$$\mu(A_1 \cup \cdots \cup A_n) = \mu(A_1) + \cdots + \mu(A_n)$$

if $A_1, \ldots, A_n$ are pairwise disjoint. Another important consequence is that if $A, B \in \mathcal{X}$ and $B \subset A$ then

$$\mu(A \setminus B) = \mu(A) - \mu(B).$$

The measures which occur in this book generally have two other important properties. They are defined on the Borel σ-algebra of $\mathbb{R}^d$ for some d, and they are *locally finite*, that is, finite on bounded sets; such measures are referred to as *Radon measures*. (If $\mu(X) < \infty$ then μ is called *finite*. The set X can be nevertheless an unbounded subset of $\mathbb{R}^d$.) The relationship between measure theory and the topology of the ground space is an important and extensive theory; however, the two definitions above will suffice for the discussions below.

Simple examples of measures on measurable spaces are given by the *Dirac measures* δ_x for x in X. They are defined by

$$\delta_x(A) = \begin{cases} 1 & \text{if } x \in A, \\ 0 & \text{otherwise.} \end{cases} \tag{1.81}$$

On $[\mathbb{R}^d, \mathcal{B}^d]$ the Dirac measures are Radon measures.

A further and most important example is that of *Lebesgue measure* ν_d on $[\mathbb{R}^d, \mathcal{B}^d]$. A fundamental theorem of measure theory has as a special case the result that locally finite measures on $[\mathbb{R}^d, \mathcal{B}^d]$ are unambiguously characterised by their values on hypercubes. Lebesgue measure is characterised by

$$\nu_d(Q) = (v_1 - u_1) \cdot \ldots \cdot (v_d - u_d)$$

if $Q = [u_1, v_1] \times \ldots \times [u_d, v_d]$. This implies that the values of the ν_d-measure of geometrical objects such as balls, cylinders, solid toruses, and so forth, coincide with the volumes in elementary geometry. Indeed, in the case $d = 3$ the Lebesgue measure ν_3 is equal to the volume measure V. Some of such values are given in Table 1.1 on p. 13. In the cases $d = 1$ and $d = 2$, it is $\nu_1 = l$ the length measure and $\nu_2 = A$ the area measure.

If B is a bounded Borel set then $\nu_d(B) < \infty$. This follows from the positivity of ν_d and its finiteness on all hypercubes including those which contain the bounded set B. Thus ν_d is indeed a Radon measure.

Lebesgue measure has the property of being *isometry-invariant*:

$$\nu_d(\boldsymbol{m}A) = \nu_d(A) \tag{1.82}$$

for all isometries $\boldsymbol{m}$ and Borel sets A. Another important theorem states that every isometry-invariant Radon measure μ (indeed every translation-invariant Radon measure) on $[\mathbb{R}^d, \mathcal{B}^d]$ is a constant multiple of the Lebesgue measure,

$$\mu = c\nu_d \qquad \text{for some } c \geq 0. \tag{1.83}$$

Further measures of geometrical interest are the *k-dimensional Hausdorff measures* h_k,

$$h_k(B) = 2^{-k} b_k \lim_{\delta \to 0+} \inf \left\{ \sum_{i=1}^{\infty} \big(\operatorname{diam}(M_j)\big)^k : B \subset \bigcup_{i=1}^{\infty} M_j,\ \operatorname{diam}(M_j) \leq \delta \right\}$$
$$\text{for } B \in \mathcal{B}^d,\ k = 0, 1, \ldots, d, \tag{1.84}$$

where diam denotes the diameter, the M_i are compact subsets of $\mathbb{R}^d$, and b_k, defined in Formula (1.30), denotes the volume of the k-dimensional unit ball. For $k = d$ the measure coincides with the Lebesgue measure, for $k = 0$ it is the counting measure assigning to a set the number of its elements. Further, $h_k(B)$ for a k-dimensional submanifold B of $\mathbb{R}^d$ coincides with its usual k-dimensional Lebesgue measure: for example, if B is a curve in $\mathbb{R}^d$, then $h_1(B)$ is its length; see Morgan (2009) for more details.

Dirac measures and Lebesgue measure are quite different in that Dirac measures are concentrated on single points while Lebesgue measure gives zero mass to every point. Measures μ on $[\mathbb{R}^d, \mathcal{B}^d]$ that are concentrated on a countably infinite collection of points $x_1, x_2, \ldots$ (so that $\mu(\mathbb{R}^d \setminus \{x_1, x_2, \ldots\}) = 0$) are called *atomic measures*, and are called *counting measures*

if in addition they give each point a mass of one or zero. A measure is called a *diffuse measure* if, like Lebesgue measure, it gives zero mass to each single point. Property (M2) then ensures that every countable set has also Lebesgue measure zero. Since the set of all rational numbers in $\mathbb{R}$ is countable, it has Lebesgue measure zero, while the set I of all irrational numbers in $[0, 1]$ has Lebesgue measure 1.

A measure μ on a measurable space $[X, \mathcal{X}]$ defines a triplet $[X, \mathcal{X}, \mu]$, which is called a *measure space*. Associated with each measure space is the *integral with respect to* μ. For a measurable real-valued function $f(x)$, $x \in X$, the integral is written as

$$\int f(x)\,\mu(\mathrm{d}x).$$

Such integrals appear at many places in this book, where the variable x is often not a real number or a point of $\mathbb{R}^d$ but, for example, a compact set or a point pattern.

The construction of this general integral starts with the particular case of an indicator function $f(x) = \mathbf{1}_A(x)$ where

$$\int \mathbf{1}_A(x)\,\mu(\mathrm{d}x) = \mu(A) \qquad \text{for } A \in \mathcal{X}. \tag{1.85}$$

Remember that for the classical Riemann integral and $A = [a, b]$

$$\int_{-\infty}^{+\infty} \mathbf{1}_A(x)\,\mathrm{d}x = \int_a^b \mathrm{d}x = b - a,$$

which can be seen as a particular case of (1.85) since $\nu_1(A) = b - a$.

Further steps of the construction are

$$\int \big(c_1 f_1(x) + c_2 f_2(x)\big)\,\mu(\mathrm{d}x) = c_1 \int f_1(x)\,\mu(\mathrm{d}x) + c_2 \int f_2(x)\,\mu(\mathrm{d}x) \qquad \text{for } c_1, c_2 \geq 0 \tag{1.86}$$

and

$$\int \sum_{k=1}^{\infty} f_k(x)\,\mu(\mathrm{d}x) = \sum_{k=1}^{\infty} \int f_k(x)\,\mu(\mathrm{d}x) \tag{1.87}$$

for nonnegative functions $f_1(x), f_2(x), \ldots$. For such functions an integral value of $+\infty$ is an accepted convention.

For an arbitrary measurable function $f(x)$ the positive and negative part are considered separately to obtain

$$\int f(x)\,\mu(dx) = \int f(x)\mathbf{1}_B(x)\,\mu(\mathrm{d}x) - \int \big(-f(x)\big)\mathbf{1}_{B^c}(x)\,\mu(\mathrm{d}x), \tag{1.88}$$

where

$$B = \{x \in X : f(x) > 0\}.$$

Here one says that the integral exists if

$$\int |f(x)|\,\mu(dx) = \int f(x)\mathbf{1}_B(x)\,\mu(\mathrm{d}x) + \int \big(-f(x)\big)\mathbf{1}_{B^c}(x)\,\mu(\mathrm{d}x) < \infty.$$

Occasionally such integrals are written as

$$\int f \, \mathrm{d}\mu = \int f(x) \, \mu(\mathrm{d}x) = \int f(t) \, \mu(\mathrm{d}t),$$

where the *dummy variables* x and t of integration can be chosen arbitrarily from previously undefined symbols.

The integral derived from the measure space $[\mathbb{R}^d, \mathcal{B}^d, \nu_d]$ is called the *Lebesgue integral*. In the case $d = 2$ this integral can be interpreted graphically. A nonnegative function $f : \mathbb{R}^2 \to [0, \infty)$ can be considered as giving the height $f(x)$ at each point x of a surface lying over the plane $(x_3 = 0)$ in $\mathbb{R}^3$. Then the integral $\int f(x)\nu_2(\mathrm{d}x)$ can be interpreted as the volume lying between the surface and the plane $(x_3 = 0)$. A similar interpretation is possible for the case $d = 1$.

The Lebesgue integral can be applied to functions where the Riemann integral does not exist. Consider the case $d = 1$ and $f(x)$ the Dirichlet function, which is the indicator function of the set I of all irrational numbers in $[0, 1]$:

$$\mathbf{1}_I(x) = \begin{cases} 1 & \text{for } x \text{ irrational and } 0 \le x \le 1, \\ 0 & \text{otherwise.} \end{cases} \tag{1.89}$$

Here the Riemann integral does not exist, since the upper and lower Riemann sums are always 1 and 0, respectively. In contrast, Formula (1.85) yields the Lebesgue integral value 1 since $\nu_1(I) = 1$.

A frequent contraction of notation involves removing indicator functions from the integrand and inserting the corresponding sets as *domain of integration* or *range of integration*. In the case $d = 1$

$$\int \mathbf{1}_{[a,b]}(x) f(x) \, \nu_1(\mathrm{d}x) = \int_{[a,b]} f(x) \, \nu_1(\mathrm{d}x),$$

and in the general case

$$\int \mathbf{1}_A(x) f(x) \, \mu(\mathrm{d}x) = \int_A f(x) \, \mu(\mathrm{d}x).$$

Where no domain of integration is given for the integral it is to be understood that the domain of integration is the whole of the appropriate space.

For elementary functions $f(x)$ on $\mathbb{R}$ the expression $\int_{[a,b]} f(x)\nu_1(\mathrm{d}x)$ equals the well-known elementary Riemann integral

$$\int_a^b f(x) \, \mathrm{d}x.$$

The Lebesgue integral simply fits better to many theoretical calculations in stochastic geometry.

If two measures μ_1 and μ_2 are given on the same measurable space $[X, \mathcal{X}]$ then sometimes one can be given in terms of the other by an integral formula involving a *density* $g : X \to [0, \infty)$:

$$\mu_2(A) = \int_A g(x) \, \mu_1(\mathrm{d}x) \qquad \text{for all } A \in \mathcal{X}.$$

In this case μ_2 is said to be *absolutely continuous with respect to* μ_1, and written as

$$\mu_2 \ll \mu_1.$$

The Radon–Nikodym theorem gives a necessary and sufficient condition for one Radon measure μ_2 to be absolutely continuous with respect to another Radon measure μ_1:

$$\text{if } \mu_1(A) = 0 \text{ implies } \mu_2(A) = 0 \text{ for all } A \text{ in } \mathcal{X}, \text{ then } \mu_2 \ll \mu_1.$$

In practice the actual calculation of the density function $g(x)$ for given μ_1 and μ_2 may be difficult; see Rao (1993).

Probability theory makes much use of the ideas of measure theory. If a measure space $[\Omega, \mathcal{A}, \mathbf{P}]$ is such that $\mathbf{P}(\Omega) = 1$ then it is called a *probability (measure) space* and the measure $\mathbf{P}$ is called a *probability measure*. Then Ω is called the *sample space* and $\mathcal{A}$ is the *σ-algebra of events*; the elements of Ω are called *sample points*; the subsets of Ω that belong to $\mathcal{A}$ are called *events*. Real-valued $\mathcal{A}$-measurable functions defined on Ω are called *random variables*. By tradition they are usually not denoted by f, g or h but by X, Y or Z. They assign real numbers to sample points. The measurability condition ensures that for a random variable X it is possible to define probabilities such as

$$\mathbf{P}(X \leq x) = \mathbf{P}(\{\omega \in \Omega : X(\omega) \leq x\}) = F(x). \tag{1.90}$$

The function $F(x)$ is called the *distribution function of* X.

The definition of more general random variables is analogous. They are also measurable mappings, but then the image space is not $\mathbb{R}$ but a suitable other space. In the case of random sets, for example, the image space is $\mathbb{F}$, the space of all closed subsets of $\mathbb{R}^d$; and for defining measurability, it is necessary to introduce a σ-algebra of subsets of $\mathbb{F}$.

If X is a real-valued random variable with $\int |X(\omega)|\mathbf{P}(\mathrm{d}\omega) < \infty$ then $\mathbf{E}(X)$ denotes its expectation or mean,

$$\mathbf{E}(X) = \int X(\omega)\mathbf{P}(\mathrm{d}\omega) = \int_{-\infty}^{\infty} x \,\mathrm{d}F(x). \tag{1.91}$$

where the last expression is a so-called *Riemann–Stieltjes integral*.

If $\mathbf{P}$ is a probability measure on $[\mathbb{R}^1, \mathcal{B}^1]$ and absolutely continuous with respect to ν_1, then X is said to be an absolutely continuous random variable and there exists a nonnegative function $f(x)$, the *probability density function*, satisfying

$$f(x) = F'(x)$$

and the integral given in (1.91) is rewritten as

$$\int_{-\infty}^{\infty} x\, f(x)\mathrm{d}x.$$

If $\mathbf{P}$ is atomic, then X is a discrete random variable and the integral in (1.91) is equal to

$$\sum_i x_i p_i,$$

where x_i are the atoms with weights p_i. In physics literature, the latter would be written as

$$\int_{-\infty}^{\infty} x\, f(x)\,\mathrm{d}x$$

with

$$f(x) = \sum_i \delta_{x_i}(x) = \sum_i \delta(x - x_i)$$

using Dirac delta functions.

Analogously, the notation

$$\mathbf{E}\big(g(X)\big) = \int_{-\infty}^{\infty} g(x)\,\mathrm{d}F(x) \tag{1.92}$$

is used for a measurable function $g(x)$.

Furthermore, $\mathbf{cov}(X, Y)$ is the *covariance* of two random variables X and Y,

$$\mathbf{cov}(X, Y) = \mathbf{E}\Big(\big(X - \mathbf{E}(X)\big)\big(Y - \mathbf{E}(Y)\big)\Big), \tag{1.93}$$

and in particular

$$\mathbf{var}(X) = \mathbf{cov}(X, X) = \mathbf{E}\Big(\big(X - \mathbf{E}(X)\big)^2\Big) \tag{1.94}$$

is the *variance* of X.

Thorough introductions to the theory of measure, integration and probability can be found in many textbooks; for example, Billingsley (1995), Kallenberg (2002), Pollard (2002), Capiński and Kopp (2004), Rosenthal (2006), Yeh (2006), Bogachev (2007), Athreya and Lahiri (2010), Durrett (2010) and Gut (2013).