9

Random tessellations, geometrical networks and graphs

9.1 Introduction and definitions

A *tessellation* or *mosaic* is a division of the plane into polygons, or of space into polyhedra. Figures 9.1, 9.4 and 9.17 show patterns that can be interpreted as the result of *tessellating* a plane. Such geometrical patterns occur in many natural situations. Examples include crystalline structures, crack patterns, and foam structures. Various models are described in Section 9.2.

This chapter mainly considers the case of random tessellations composed of *convex* polygons or polyhedra, although some of the results given are also true in the nonconvex case; see the references on p. 370. Precise mathematical definitions are given here for the case of planar tessellations. The generalisation to the spatial case and the d-dimensional case is straightforward.

There is a vast mathematical literature on random tessellations. There are specialised books such as Møller (1994), Okabe *et al.* (2000) and van de Weygaert *et al.* (2014), while other books contain chapters on tessellations, for example Matheron (1975), Schneider and Weil (2008) and Ohser and Schladitz (2009).

The system of edges of a tessellation is an example of a network. Geometrical properties of such and other *random geometrical networks* are often studied without reference to tessellations. For random geometrical networks the edge lengths and positions of vertices play a rôle, whilst for *random graphs* only the topological properties such as connectivity are of interest. There is a rapidly growing literature on random graphs, which will be reviewed briefly in Section 9.12.

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

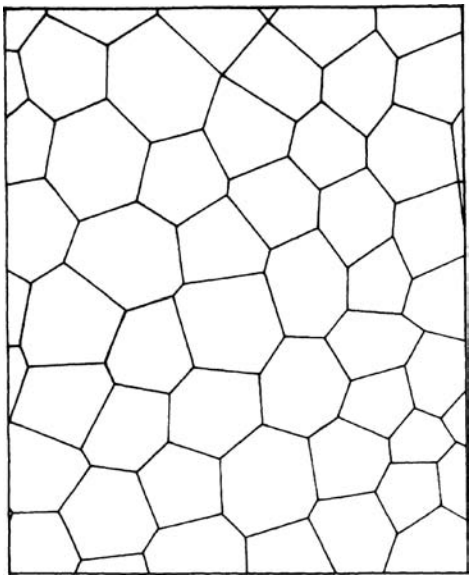

Figure 9.1 Plan view of basaltic columnar jointing at Burg Stolpen (Saxony) as measured by Professor R. A. Koch and his students (see Koch *et al.*, 1983). This pattern can be interpreted as a part of a realisation of a stationary isotropic tessellation. It is not a Dirichlet tessellation; see Stoyan and Stoyan (1980a).

Planar random tessellations

Let $\mathbb{C}$ be the set of all planar compact convex sets C with polygonal boundaries. A subset of polygons $\theta \subset \mathbb{C}$ is said to be a *tessellation* if the interiors of the constituent polygons are pairwise disjoint, if their union fills the plane, and if the family is locally finite. These three conditions can be expressed formally by:

(a) $C_1^{\text{int}} \cap C_2^{\text{int}}$ is empty if C_1 and $C_2 \in \theta$ and $C_1 \neq C_2$,

(b) $\bigcup_{C \in \theta} C = \mathbb{R}^2$,

(c) if B is a bounded planar set then the set $\{C \in \theta : C \cap B \neq \emptyset\}$ is finite.

The polygons $C \in \theta$ are the *cells* of the tessellation θ. The terminology suggested in Weiss and Cowan (2011) is adopted here. *Vertices* and *edges* are the primitive elements of the planar graph of a tessellation (also called, respectively, 0- and 1-faces of the tessellation), whilst *corners* and *sides* are 0-faces and 1-faces of the constituent polygons; see Figure 9.11 on p. 361. Because of convexity, the interior angle at a corner of a polygon is strictly less than π. If a vertex appears at an inner point of a polygonal side instead of at a corner, then it is called a π-*vertex*. An edge has a vertex at each of its two endpoints but no vertex in its interior; a side is the union of edges which it contains, and its endpoints are not π-vertices but polygonal corners. Tessellations without π-vertices are called *regular* or *face-to-face*.

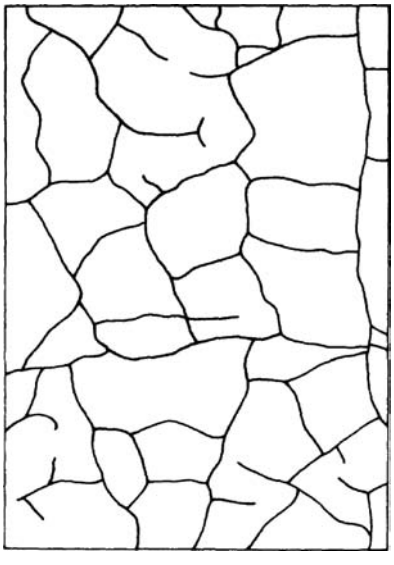

Figure 9.2 Mud-cracks in a ditch. The pattern shows some anisotropy caused by the geometry of the ditch. It can be interpreted as an intermediate stage in the formation of a tessellation with nonconvex cells. As given, the pattern forms a geometrical network.

Denote by E_θ the set-theoretic union of all edges of θ. Clearly, this *edge system* E_θ is a fibre system, more precisely a segment system, and thus a random closed set. Evidently, θ can be recovered from E_θ. Thus ideas of the theories of random closed sets and of fibre processes can be applied to characterise tessellations.

Let $\mathbb{T}$ denote the class of all tessellations and let $\mathscr{T}$ denote the σ-algebra on $\mathbb{T}$ generated by sets of the form

$$\{\theta \in \mathbb{T} : E_\theta \cap K \neq \emptyset\},$$

where K runs through all compact subsets of $\mathbb{R}^2$. If tessellations are identified with their edge systems then $\mathscr{T}$ is the trace of the hitting σ-algebra $\mathcal{F}$ on $\mathbb{T}$.

A *planar random tessellation* (or simply *tessellation* for brevity) is a random variable Θ with values in $[\mathbb{T}, \mathscr{T}]$. Its *distribution* is the induced probability measure on $[\mathbb{T}, \mathscr{T}]$. The definitions of stationarity and isotropy for random tessellation follow the usual form: A tessellation Θ and its distribution P are said to be *stationary* if for all x in $\mathbb{R}^2$ the translated tessellation

$$\Theta + x = \{C + x : C \in \Theta\} = \{C : C - x \in \Theta\}$$

has the same distribution as the original tessellation Θ, or

$$P(Y_x) = P(Y), \qquad \text{for all } x \in \mathbb{R}^2 \text{ and } Y \in \mathscr{T},$$

where $Y_x = \{\theta \in \mathbb{T} : \theta_{-x} \in Y\}$ and Y belongs to $\mathscr{T}$. If there is invariance with respect to rotations about the origin then Θ and P are called *isotropic*; isotropic and stationary tessellations are also called *motion-invariant*.

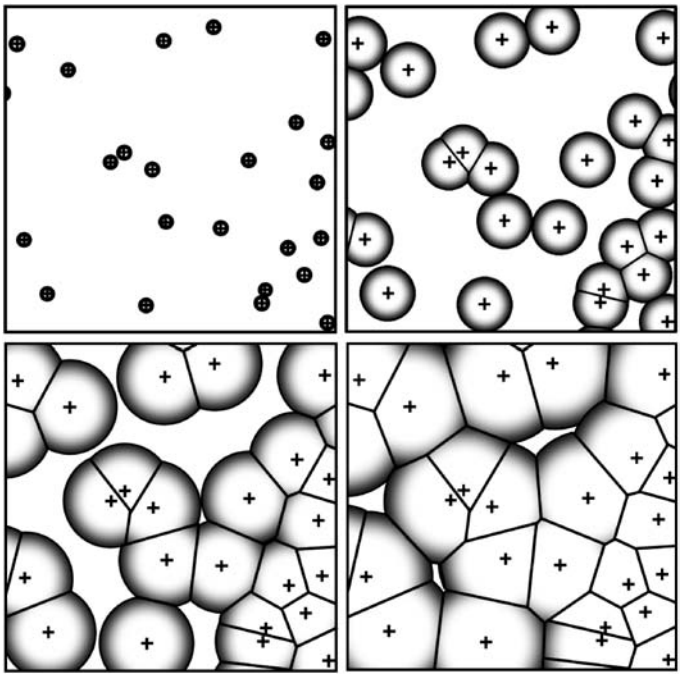

Figure 9.3 Four steps of the growth process which leads to a Dirichlet tessellation. Courtesy of L. Muche.

As noted above, the definition in the three-dimensional and d-dimensional cases is directly analogous; see Schneider and Weil (2008, Chapter 10).

Random tessellations of other spaces, for example the sphere, are also of interest. Miles (1971b) and Arbeiter and Zähle (1994) study such tessellations and give mean-value formulae as well as formulae for intersections with subspheres. Isokawa (2000) investigates tessellations in three-dimensional hyperbolic space.

Another avenue of generalisation is the study of random cell complexes (Zähle, 1987b, 1988; Leistritz and Zähle, 1992). An example is a system of groups of cells randomly scattered in the plane without overlappings and with gaps between them (as in Figure 9.3).

9.2 Mathematical models for random tessellations

In this section a representative selection of models is discussed and a partial summary of the literature is given. The purpose is to acquaint the reader with various types and methods of construction.

Line and plane tessellations

A nondegenerate line process Ψ in $\mathbb{R}^2$ generates a tessellation Θ, whose edge system E_Θ is simply the union of all lines of Ψ. Figure 8.7 on p. 309, showing part of a line process, can

therefore also be interpreted as part of a *line tessellation*. In an analogous fashion hyperplane processes provide *hyperplane tessellations* of $\mathbb{R}^d$.

The distributional properties of Θ depend on those of Ψ. In particular, if Ψ is stationary then so is Θ. In general little is known about general line tessellations, with the exception of some simple assertions derived from geometrical considerations, such as that in the planar case the mean number of sides of a randomly chosen polygon is 4 if no line intersections involve more than two lines.

Much work has been done for the special case of *Poisson* line and plane tessellations built on Poisson line and plane processes. Some important formulae and results are presented Section 9.5 below.

Voronoi tessellation and related models

Voronoi tessellations

Dirichlet (1850) and Voronoi (1908) consider regular tessellations of planes and higher-dimensional spaces, motivated by problems in number theory. The application of Dirichlet and Voronoi tessellations to irregular and random point patterns appear to have arisen independently in applications in meteorology (Thiessen and Alter, 1911), metallurgy and crystallography (Kolmogorov, 1937; Johnson and Mehl, 1939) and ecology (Pielou, 1977; Matérn, 1986); see Okabe *et al.* (2000) for a historical sketch of the development of ideas.

Let $\varphi = \{x_1, x_2, \ldots\}$ be a locally finite system of points in $\mathbb{R}^d$. Each location in $\mathbb{R}^d$ is associated to its *nearest point(s)* belonging to φ. The *neighbourhood* or *Voronoi cell* $C(x_i, \varphi)$ of a point x_i, called the *nucleus* or *generator*, of φ is defined by

$$C(x_i, \varphi) = \{y \in \mathbb{R}^d : \|y - x_i\| \leq \|y - x_j\| \text{ for all } j \neq i\}. \tag{9.1}$$

The $C(x_i, \varphi)$ are all convex polygons (with the x_i as inner points) but it is possible for some polygons to be unbounded. If this possible deficiency does not arise then the $T(y)$ constitute a tessellation of $\mathbb{R}^d$, the *Voronoi tessellation* $\mathcal{V}(\varphi)$ relative to (or generated by) φ. Some authors refer to the $d = 2$ case as the *Dirichlet* or *Thiessen* tessellation; here the first of these terms is used. When φ is a lattice (as it was in Dirichlet's and Voronoi's cases), physicists and metallurgists call the cells Wigner–Seitz cells or zones.

The Voronoi tessellation can be also interpreted as a result of a growth process. The points x_i of φ play the rôle of nuclei, in which growth begins at the same instant. The speed of growth is uniform in all directions, so that the nuclei grow into discs (or balls in the spatial case) while empty space is available. When two discs come in contact, growth stops at the contact point and a facet begins to develop, while in other places growth continues. Eventually, the whole space is divided into cells. Figure 9.3 shows some steps of the growth process in the planar case. Muche (1993) and Schulz *et al.* (1993) study characteristics of the 'incomplete' tessellation obtained after time t.

This simple growth model makes the Voronoi tessellation attractive as a model for microstructures which result from crystallisation. Surprisingly, it seems to be also a good model for the universe; see Zaninetti (2006), van de Weygaert (2007) and van de Weygaert *et al.* (2014). The idea is that the 'empty space' grows while the galaxies tend to lie on the cell facets, edges and vertices. The so-called Abell clusters of galaxies are believed to be positioned in vertices. This greatly refines the Neyman–Scott model mentioned in Section 5.3.

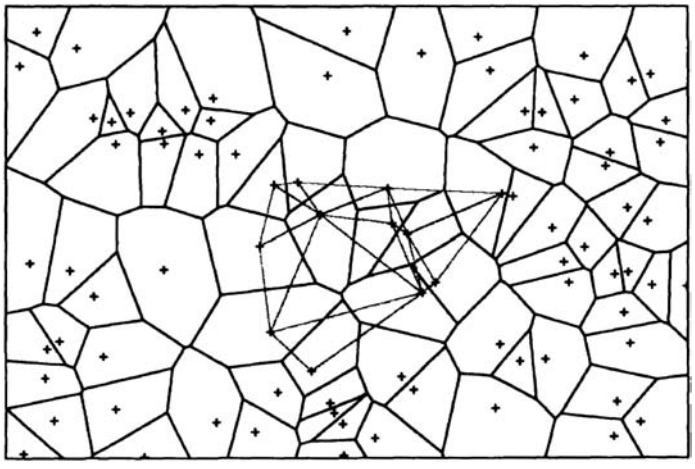

Figure 9.4 A typical realisation of the Dirichlet tessellation $\mathscr{V}$ relative to a sample of a homogeneous Poisson point process and some triangles of the corresponding Delaunay tessellation.

The regular square tessellation is a particular example of a Voronoi tessellation. It arises from a regular square grid of nuclei. This tessellation is degenerate; if by contrast a set of nuclei is in general quadratic position, meaning that no three nuclei lie on the same line and no four nuclei on the same circle, then this set must give rise to a tessellation with vertices at each of which only three cells meet. In the d-dimensional case 'general quadratic position' means that almost surely never $k+1$ points lie in a $(k-1)$-dimensional affine subspace of $\mathbb{R}^d$ for $k = 2, \ldots, d$ and never $d+2$ points lie on the surface of a sphere. Then at each vertex $d+1$ Voronoi cells meet.

If the nuclei form a stationary point process Φ of finite positive intensity ϱ then almost surely all the $C(x_i, \Phi)$ are bounded so $\mathscr{V}(\Phi)$ is indeed a random tessellation. If Φ is stationary then $\mathscr{V}(\Phi)$ inherits this property.

Known formulae for $\mathscr{V}(\Phi)$ are almost entirely confined to the case when Φ is a homogeneous Poisson process, and are presented in Section 9.7. Figure 9.4 shows a realisation of $\mathscr{V}(\Phi)$ relative to a sample of a homogeneous Poisson process Φ in the planar case. Furthermore, Figure 9.5 shows a simulated cell of a spatial Voronoi tessellation.

Voronoi tessellations relative to non-Poisson point processes are studied by simulation in Hermann *et al.* (1989), Lorz (1990), Kohutek and Saxl (1993), Lorz and Hahn (1993), van de Weygaert (1994), Saxl and Ponížil (2001, 2002) and Dereudre and Lavancier (2011). Heinrich (1998) considers contact and chord length distribution for the case of Poisson cluster and Gibbs processes. See Okabe *et al.* (2000, Section 5.12) for a survey.

The Voronoi tessellation has been often used as a model for natural phenomena in a variety of fields. Here are some now classical papers:

- *agriculture and forestry*: Klier (1969), Fischer and Miles (1973), Gavrikov *et al.* (1993), Kessler and Werner (2003);
- *astrophysics*: Kiang (1966), Icke and van de Weygaert (1987), van de Weygaert and Icke (1989);

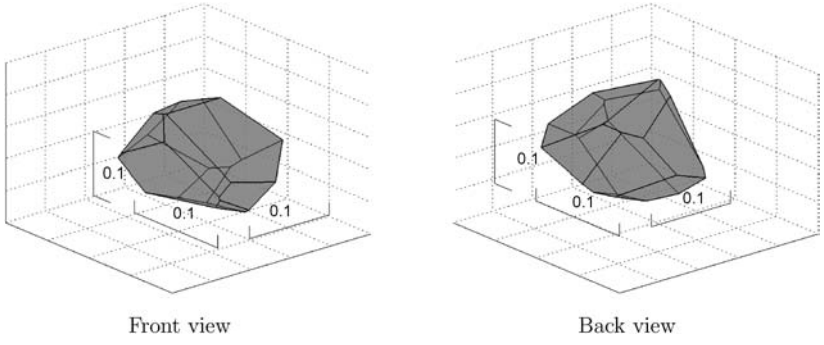

Figure 9.5 A cell of a spatial Poisson-Voronoi tessellation of intensity 1000 (so that the mean volume of the typical cell is 0.001). It has 15 facets, 39 edges and 26 vertices.

- *cell biology*: Lewis (1946), Honda (1983);
- *communication theory*: Shannon (1949);
- *crystal growth and aggregates*: Kolmogorov (1937), Meijering (1953), Gilbert (1962);
- *geography*: Mardia *et al.* (1977), Boots (1987);
- *geology*: Stiny (1929), Smalley (1966), Crain (1976), Gray *et al.* (1976), Stoyan and Stoyan (1980a);
- *metallography*: Boots (1982, 1984), Hermann (1991);
- *physics*: Rahman (1966), Ogawa and Tanemura (1974), Finney (1979), Boots (1982);
- *protein structure*: Richards (1974), Finney (1975), Poupon (2004);
- *zoology and ecology*: Hamilton (1971), Hasegawa and Tanemura (1976, 1980), Hasegawa *et al.* (1981).

See also the detailed description of applications of the Voronoi tessellation in the encyclopedic work of Okabe *et al.* (2000).

In addition to physical objects, the Voronoi tessellation can also model abstract concepts in the so-called conceptual spaces (Gärdenfors, 2000; Douven *et al.*, 2013) of cognitive science.

Voronoi tessellations are also used in data analysis of geometrical structures. Finney (1979) uses the term 'polyhedral statistics'. Sibson (1980, 1981) describes the use of Voronoi tessellations relative to given point patterns as a basis for 'natural neighbour interpolation': interpolating a smooth function for data located at irregularly distributed points. Thiessen and Alter had this in mind when they originally suggested this tessellation. A modern method in this spirit is discussed in Bernardeau and van de Weygaert (1996); see also Schaap and van de Weygaert (2000). Their 'Delaunay tessellation field estimator' yields estimates of random field values and point process intensities. To estimate a (unknown) bounded Borel set A in the unit cube $(0, 1)^d$ in $\mathbb{R}^d$, Khmaladze and Toronjadze (2001) and Penrose (2007) show that

the union of Voronoi cells

$$A_n = \bigcup_{x_i \in A \cap \Phi_n} C(x_i, \Phi_n) \cap (0,1)^d, \tag{9.2}$$

where Φ_n is a set of n i.i.d. random test points in the unit cube, is a consistent estimator of A in the sense that

$$\nu_d(A \triangle A_n) \to 0 \qquad \text{almost surely, as } n \to \infty. \tag{9.3}$$

The moments of this symmetric difference are studied in Heveling and Reitzner (2009) and Reitzner *et al.* (2012). For the problem of estimating, in the presence of noise, the region where the intensity of a bounded planar point process is nonzero, such as the detection of a minefield from an aerial image containing mine locations corrupted by clutter, Allard and Fraley (1997) propose a nonparametric maximum likelihood estimator in the form of a union of Voronoi polygons. In planar point pattern analysis, Voronoi tessellation is also used to estimate contour lines of the density of points (Picard and Bar-Hen, 2000), and to construct uniformity measures (Ong *et al.*, 2012) and test statistics for complete spatial randomness (Chiu, 2003).

Delaunay tessellations

If $\mathscr{V}(\Phi)$ has the property that almost surely each vertex is touched by exactly three cells (in the planar case) or by exactly four cells (in the spatial case) then a further tessellation can be constructed, the *Delaunay tessellation* $\mathscr{D}(\Phi)$. In the planar case it is constructed out of the triangles formed by points of whose cells share the same vertex. Section 9.7.4 contains some formulae for $\mathscr{D}(\Phi)$ in the Poisson process case. Early applications of the Delaunay network (the system of edges of the Delaunay tessellation) are given in Priolo *et al.* (1992) and Kumar and Singh (1995) in the context of conductance in disordered matter.

The construction of Voronoi and Delaunay tessellations, which is basic for simulations of random tessellations, is discussed in Preparata and Shamos (1985), Guibas and Stolfi (1988) Aurenhammer (1991), Fortune (1992), van de Weygaert (1994), Boissonnat and Yvinec (1998), Okabe *et al.* (2000), Hjelle and Dæhlen (2006) and de Berg *et al.* (2008). Efficient algorithms are implemented in, for example, the `R` packages such as `deldir` and `tripack` and the `MATLAB` toolbox `MPT`.

Generalised/higher-order Voronoi tessellations

The Voronoi tessellation has been generalised in many ways; see Okabe *et al.* (2000, Chapter 3) and Gavrilova (2008).

Miles (1970) defines the *generalised Voronoi tessellation*, also called *n^{th}-order Voronoi tessellation*, $\mathscr{V}_n$, which uses for cells the sets of positions sharing the same n nearest neighbours of the generating point pattern. Of course, $\mathscr{V}_1 = \mathscr{V}$, giving the original Voronoi tessellation. Details can be found in Miles (1970, 1972b) and Miles and Maillardet (1982). The latter paper gives pictures of $\mathscr{V}_n$ for $n = 4$, 16, 64 and 256 relative to a homogeneous Poisson process. See also Edelsbrunner (1987, Sections 13.3–13.5) and Okabe *et al.* (2000, Section 3.2) for elaborated surveys.

Non-Euclidean norm Voronoi tessellations

Ohser and Lorz (1994) show that interesting tessellations are obtained if the Euclidean norm is replaced by other norms for which the unit ball is not a ball in the usual sense but, for example, a square or cube. Scheike (1994) considers an elliptic distance to study tessellations resulting from elliptic, instead of spherical, growth.

Weighted Voronoi tessellations

Tessellations of particular interest are obtained if the Euclidean distance in (9.1) is replaced by other, weighted, distance definitions, where the weights may represent some measures of actual physical sizes of the nuclei. The following gives briefly two examples; the geometrical ideas behind are discussed in Okabe *et al.* (2000, Section 3.5.3).

The definitions are based on marked point processes. In the given context, the marks $w_i \in \mathbb{R}$ are called weights or powers, and instead of the set φ of unmarked nuclei, the set $\psi = \{[x_i; w_i]\}$ is considered.

(a) *Laguerre tessellations* The first definition of the distance between the marked point $[x_i; w_i]$ and a point y of $\mathbb{R}^d$ is as follows:

$$\mathrm{dist}_{\mathrm{L}}([x_i; w_i], y) = \|x_i - y\|^2 - w_i. \tag{9.4}$$

The corresponding cells are then defined as

$$C([x_i; w_i], \psi) = \{y \in \mathbb{R}^d : \mathrm{dist}_{\mathrm{L}}([x_i; w_i], y) \leq \mathrm{dist}_{\mathrm{L}}([x_j; w_j], y) \text{ for all } j \neq i\}. \tag{9.5}$$

The resultant tessellation is called *Laguerre tessellation*, also known as *radical* or *additively weighted power Voronoi tessellation*, or similarly. If the powers w_i are identical then the Laguerre tessellation coincides with the Voronoi tessellation.

The cells of a Laguerre tessellation are convex. It can happen that a nucleus produces an empty cell (therefore, it must be said more precisely that the tessellation consists of the non-empty cells only) or that a nucleus is not an element of its cell. There is a tendency for points x_i with large powers w_i to have large cells.

For a system of hard balls, the Laguerre tessellation relative to the ball centres marked by the respective radii has convex cells, where each ball is contained in a cell. A Voronoi tessellation is sufficient to obtain a tessellation the cells of which contain the balls only in the special case in which the balls have identical radii. Figure 9.6 and the book cover, respectively, show systems of hard discs in the plane and hard balls in space and the corresponding Laguerre tessellations; the weights are the ball radii.

(b) *Additively weighted Voronoi tessellations* The second definition considered here is

$$\mathrm{dist}_{\mathrm{aw}}([x_i; w_i], y) = \|x_i - y\| - w_i. \tag{9.6}$$

It leads to the so-called *additively weighted Voronoi tessellation*. Not all its cells are convex; in the planar case the boundaries are pieces of hyperbolic curves. An important related model is the Johnson–Mehl tessellation described below. In these cases the definition of tessellations has to be modified to include nonpolygonal cells, see Zähle (1988).

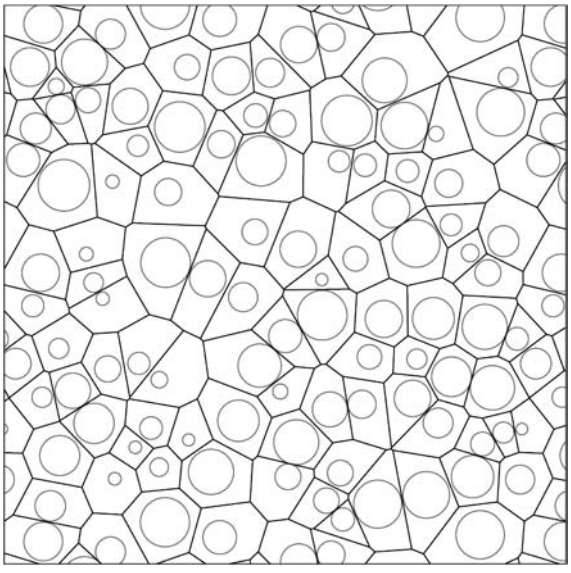

Figure 9.6 A Laguerre tessellation with respect to irregularly distributed discs. The Voronoi tessellation with respect to the disc centres does not have the property that all its cells contain the discs shown. Courtesy of C. Redenbach.

Voronoi S-tessellations

Of great practical interest is also the *Voronoi S-tessellation* in $\mathbb{R}^3$, where 'S' comes from 'distance to the surface'; see Medvedev (2000), Alinchenko *et al.* (2004) and Medvedev *et al.* (2006). As in the case of the Laguerre tessellation, this tessellation is defined relative to a system of hard balls, and each cell contains the respective ball. However, some of the cell boundaries are curved, hyperboloidal surfaces. Physicists consider this tessellation as more natural than the Laguerre tessellation, which they view only as a useful mathematical tool.

A definition which is more general than that for the case of hard balls is as follows. Let $K_1, K_2, \ldots$ be a sequence of non-overlapping convex bodies, for example balls with different radii. The cell containing K_i is defined by

$$C_i = \{x \in \mathbb{R}^d : d(x, K_i) \leq d(x, K_j) \text{ for all } j \neq i\}, \tag{9.7}$$

where $d(x, K)$ is the usual distance between the point x and the set K.

Clearly, if the K_i are identical balls, then the Voronoi S-tessellation coincides with the classical Voronoi tessellation.

The paper Medvedev *et al.* (2006) describes the construction of the Voronoi S-tessellation relative to given ensembles of balls of different radii.

Johnson–Mehl tessellations

The *Johnson–Mehl tessellation* is a structure resulting from a birth-and-growth process (see Section 6.6.4). This generalises the process leading to the Voronoi tessellation. Here 'nuclei' are generated by some birth process of intensity $\lambda(t)$, and each nucleus grows with speed $v(t)$ so that at time t after its birth it occupies all the previously vacant region within the ball of

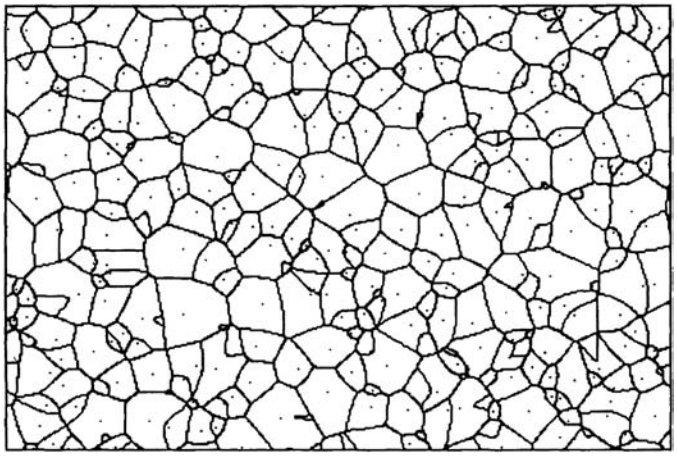

Figure 9.7 A realisation of a Johnson–Mehl model for a time-stationary birth process. This tessellation was drawn using a program which constructs the hyperbolic edges of the cells. Courtesy of L. Heinrich and E. Schüle.

radius $\int_0^t v(u)\mathrm{d}u$ centred on its original location. Consequently, if a nucleus is born at a point in the current cell of another nucleus then it vanishes immediately, while the survivors form cells that grow radially in each direction with age-dependent velocity $v(t)$, till cells of other nuclei are met in all directions. Figure 9.7 shows a part of a realisation of a planar Johnson–Mehl model. (A variant of the model considers that at time t after its birth, a nucleus born at time t_0 will occupy vacant region within a ball of radius $\int_{t_0}^{t+t_0} v(u)\mathrm{d}u$. However, it may lead to nuclei not lying in their own cells and/or disconnected cells and hence is less appealing.)

In the simplest Johnson–Mehl model, a homogeneous Poisson process of constant intensity in $\mathbb{R}^d \times [0, \infty)$ forms the nucleus birth locations and times. However, models with births that are stationary in space but not in time are also mathematically tractable (Møller, 1992). Section 9.9 contains some formulae for such time-inhomogeneous Johnson–Mehl models.

If all nuclei are born at the same instant then the resulting tessellation is Voronoi; but in the general model the cells of the tessellations are not even convex, and lens-shaped cells may appear. If the growth speed is a constant v then additively weighted Voronoi tessellations (generated by weighted nuclei $\{[x_i; -vt_i]\}$, where t_i is the birth time of the nucleus born at location x_i) are obtained.

The relevance of the model for crystal growth is discussed in Kolmogorov (1937), Johnson and Mehl (1939), Avrami (1939), Meijering (1953), Gilbert (1962) and Saltykov (1974). A nice application is given by Horálek (1988, 1990), who showed that the so-called ASTM model used in metallurgy can be well described by a particular Johnson–Mehl model.

The models of the DNA replication in Vanderbei and Shepp (1988) and Cowan *et al.* (1995) and the autoinhibited release of neurotransmitters at a synapse in Bennett and Robinson (1990) are also examples of applications of the Johnson–Mehl model.

Centroidal Voronoi tessellations

Voronoi tessellations in which the cell nuclei are also the centres of gravity of the cells are called *centroidal Voronoi tessellations*. Such tessellations exhibit a high degree of order.

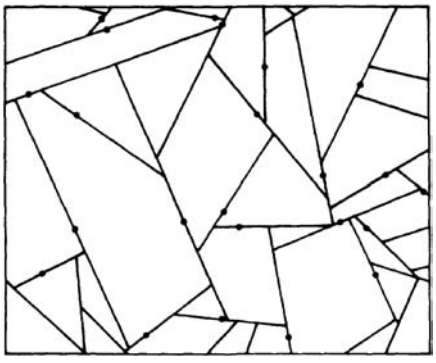

Figure 9.8 A sample of a Gilbert tessellation. The • are the starting points.

They can be computed iteratively using the Lloyd method (Du *et al.*, 1999) or a probabilistic algorithm (Ju *et al.*, 2002), and have applications in many different areas, such as image compression and meshless computing.

Crack and STIT tessellations

Many tessellations in nature are the result of fracture or crack processes. The tessellations are generated by edge growth instead of cell growth. Examples include crack systems on pottery surfaces, in rocks and in drying mud.

Typically, the edge systems of such tessellations show forms of hierarchical ordering, since the cracks appear step by step. In the planar case the vertices are typically T-shaped (π-vertices), and analogously in the spatial case. Such tessellations are not face-to-face.

A model of this class is the *Gilbert tessellation* (Gilbert, 1967), which is based on a marked point process in the plane, where each point is marked by the orientation of a line. The tessellation edges are generated by a growth process: each edge starts at one of the points and grows at constant rate in the two opposing directions specified by the mark of the point. Growth in a particular direction continues until further edges are hit. Figure 9.8 shows a part of a simulation of such a tessellation. Mathematically, the Gilbert tessellation is extremely difficult, even if the points form a homogeneous Poisson process and the orientations are uniform and independent; see Burridge *et al.* (2013).

A further crack tessellation extensively studied by mathematicians is the STIT tessellation, introduced by Nagel and Weiss (2003, 2004, 2005). The name comes from 'stability under iteration', a property which is related to a successive construction principle. Here the model construction (i.e. simulation) is described in the planar and isotropic case, following Mecke *et al.* (2008) and Calka (2010). For simplicity, the construction is described for the bounded planar case, that is, the tessellation is constructed in a bounded convex polygon W (e.g. $[0, 1]^2$). A definition for the unbounded and stationary case (for the whole $\mathbb{R}^d$) is possible; see Mecke *et al.* (2008) and Calka (2010). A further generalisation to the inhomogeneous case in $\mathbb{R}^2$ can be found in Mecke (2010).

The construction is dynamic and takes place over a time interval $[0, t]$. Intuitively, each polygon (including the window W itself) waits for an independent exponentially distributed time, with mean equal to its perimeter, and then divides into two polygons. The precise

definition is based on an i.i.d. sequence $\{\tau_i, g_i\}$ of exponentially distributed random times τ_i and random lines g_i from a motion-invariant Poisson line process as in Section 8.2.2 that hit W (and in later steps the smaller polygons resulting from divisions). The parameter of the exponential distribution of the τ_i is $L(W)$. (According to Formula (8.4) this is the motion-invariant measure of the set of all lines hitting W.) The construction steps are as follows.

1. If $\tau_1 > t$, the process does not start at all and the tessellation is simply the entire window W.

2. If $\tau_1 \leq t$, W is cut at time τ_1 by the line g_1 into two polygons W_+ and W_-. These are then treated in the following way, separately and independently.

   (a) If $\tau_1 + \tau_2 > t$, W_+ remains as it is and will be a part of the final tessellation.

   (b) If $\tau_1 + \tau_2 \leq t$, W_+ is divided into two new polygons by g_2 if it intersects W_+. If g_2 does not hit, the next potential division of W_+ happens at time $\tau_1 + \tau_2 + \tau_3$, provided that it is less than t, and so on, so that either W_+ is eventually divided into two new polygons by g_{i_0} at time $\tau_1 + \cdots + \tau_{i_0} \leq t$ for some $i_0 \geq 2$, or W_+ remains as it is because none of $g_2, \ldots, g_{i_0-1}$ hits W_+ and $\tau_1 + \cdots + \tau_{i_0} > t$.

   (c) Repeat steps (a) and (b) above for W_-, with the same τ_1 but $\{(\tau_2, g_2), \ldots, (\tau_{i_0}, g_{i_0})\}$ is replaced by an independent sequence $(\tau_{i_0+1}, g_{i_0+1})$, $(\tau_{i_0+2}, g_{i_0+2})$, $\ldots$, until either W_- is divided into two new polygons or the sum of the exponential times is greater than t.

3. If both W_+ and W_- are not divided further in step 2, the process stops and the resultant structure is the tessellation. Otherwise, the same procedure is applied to each new polygon with independent exponential times and random lines iteratively until no more division is possible (i.e. after time t).

Because the polygons do not overlap, dividing whether W_+ or W_- first does not matter. Figure 9.9 shows two steps of tessellating a square by the STIT algorithm.

Although τ_i have the same parameter, in step 2(b) it is clear that a smaller polygon has to wait for a longer time until it is hit and divided by a random line. An equivalent formulation is that each polygon waits for an exponential time with parameter given by the length of the polygon perimeter, and then, unless the time is up, it will be divided by a random line that hits the polygon.

In the given form the model depends only on time t, which controls the density of edges. In the anisotropic case the rose of directions of the lines is a further parameter. In the stationary case the line density L_A and specific surface S_V, respectively, serve as parameters.

Section 9.6 presents some results for stationary STIT tessellations, and in Section 10.6 a remark on the stereology of such tessellations is given, that is, on the tessellations resulting from planar sections of spatial STIT tessellations.

Generalisations of the STIT model are considered in Cowan (2010), in which STIT tessellations are the special case that employs the 'perimeter-weighted' cell selection rule.

Horgan and Young (2000) present a model for the geometry of two-dimensional crack growth in soil, producing patterns which resemble that in Figure 9.2 on p. 345. Also here the cracks appear stepwise; similar is the fragmentation model in Hernandez *et al.* (2012).

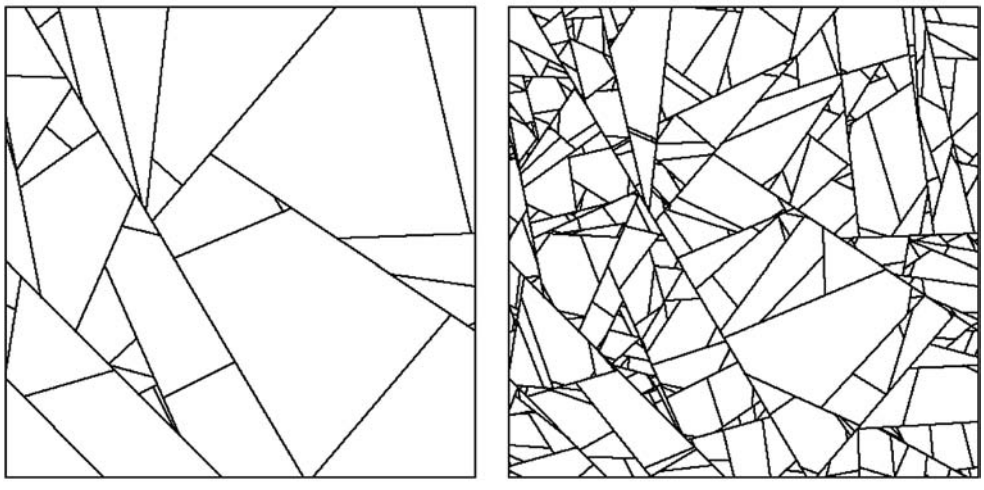

Figure 9.9 Two subsequent realisations of a planar isotropic STIT tessellation: (a) after 64 iterations; (b) after 1024 iterations. Courtesy of J. Ohser.

Further methods of tessellation construction

More complicated tessellations can be constructed by means of operations on simpler tessellations or by combining the methods given above. Here are two examples:

Superposition

If Θ_1 and Θ_2 are two tessellations with edge systems (regarded as fibre processes) E_{Θ_1} and E_{Θ_2}, respectively, then the union of E_{Θ_1} and E_{Θ_2} forms a further fibre process which defines the edge system of the *superposition tessellation*. For example, Mecke (1983, 1984b) and Santaló (1980, 1984) discuss this operation. Baccelli *et al.* (2000) study superpositions of planar Voronoi tessellations. Nagel and Weiss (2003) show that superpositions of suitably rescaled independent stationary planar tessellations converge to Poisson line tessellations.

Iterated division of cells

Each cell of a tessellation Θ may be divided into further cells by use of some rule of tessellation. There may be some form of hierarchy in the sense that the new cells tend to be an order smaller than the cells of Θ and are formed in close relation to them. For example, points may be scattered in the cells of Θ and then each cell itself is further tessellated by the Voronoi tessellation relative to the points in the cell, independent of the points in other cells, or as in Figure 9.10, each cell of a Voronoi tessellation is further subdivided by an independent line process. The process of division of cells may be continued in an iterative manner.

A general theory of iterated tessellations is presented in Maier and Schmidt (2003) and Maier *et al.* (2004). This includes formulae for intensities of face processes, for mean intrinsic volumes of typical faces and for the typical cell. Important components of these models are Poisson line and Poisson-Voronoi tessellations. A special construction principle leads to STIT tessellations.

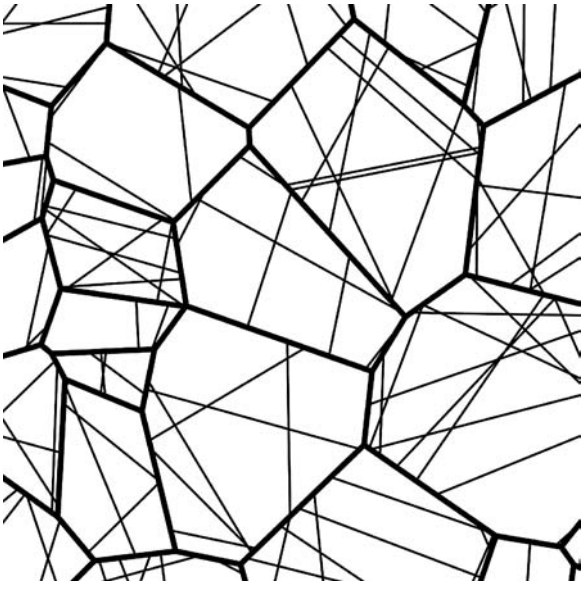

Figure 9.10 A part of a Voronoi tessellation in which the cells are further subdivided by independent line processes. Courtesy of V. Schmidt.

Iterated or nested tessellations are used as models of road networks in telecommunication; see Gloaguen *et al.* (2006). The edges of a first tessellation stand for the main roads. In its cells, independent copies of a secondary tessellation with higher edge density are drawn to obtain the secondary roads that end up in the main roads. The types of both tessellations are Voronoi, Delaunay, Poisson line or regular tessellations.

9.3 General ideas and results for stationary planar tessellations

This section is still restricted to tessellations consisting of polygonal cells only. However, some notions such as the typical cell and tessellation faces are also applicable under a more general definition of tessellations allowing cells with curved boundaries, see Section 9.9.

9.3.1 Point processes related to tessellations

Tessellations are often accompanied by other stochastic structures. For example, a planar tessellation Θ automatically produces a fibre process, more precisely a segment process, namely the edge system E_Θ. A corresponding important tessellation characteristic is the line density L_A as defined in Section 8.3.1, the mean total edge length per unit area.

Together with E_Θ, three planar point processes can be constructed from Θ: the *set of vertices* $\alpha_0(\Theta)$, the *set of edge midpoints* $\alpha_1(\Theta)$, and the *set of cell centroids* $\alpha_2(\Theta)$.

If Θ is stationary then so are E_Θ and the $\alpha_k(\Theta)$. The intensities of the $\alpha_k(\Theta)$ are denoted by λ_k for $k = 0$, 1 and 2; for example, λ_2 is the mean number of cells per unit area. It is assumed that L_A and all three λ_k are positive and finite.

Further characteristics of a tessellation can be obtained by marking the point processes $\alpha_k(\Theta)$ in various geometrically natural ways. The ensuing mark distributions or their means provide further valuable tessellation characteristics. The following are some useful real-valued marks. Note that, because some important models for tessellations (such as STIT) are not face-to-face, it is necessary to distinguish between k-faces of tessellation and k-faces of polygons. In the notation above and below, the number subscripts '$_0$', '$_1$' and '$_2$' indicate the dimension(s) of the k-faces of the tessellation, namely, vertices, edges and cells, whilst the letter subscripts '$_{\mathrm{C}}$' and '$_{\mathrm{S}}$' are used to indicate polygonal k-faces, that is, corners and sides.

Marks for a vertex x in $\alpha_0(\Theta)$:

$$n_{01}(x) = n_{02}(x) = \#\{C \in \Theta : x \in C\} = \text{the number of edges emanating from the vertex } x,$$
$$n_{0S}(x) = \text{the number of polygonal sides containing the vertex } x,$$
$$l_0(x) = \text{the total length of the edges emanating from the vertex } x.$$

Note that $n_{0\mathrm{S}}(x)$ and $n_{02}(x)$, even when there is no π-vertex, are not the same. See Formula (9.11) on p. 360 and the explanation below it.

Mark for an edge midpoint x in $\alpha_1(\Theta)$:

$$l_1(x) = \text{the length of the edge through } x.$$

Marks for a cell centroid x in $\alpha_2(\Theta)$:

$$n_{20}(x) = \text{the number of vertices on the boundary of } C,$$
$$c_2(x) = \text{the number of corners of } C,$$
$$l_2(x) = L_2(C), \text{ the perimeter of } C,$$
$$a_2(x) = \nu_2(C), \text{ the area of } C,$$

where $x \in C$ and $C \in \Theta$. Note that $c_2(x)$ must not be confused with the number of polygon corners *adjacent to* polygon C (which would be naturally denoted by $n_{2\mathrm{C}}(x)$). First, the π-vertices, if any, on the boundary of C are still corners of some other polygons and have to be counted in $n_{2\mathrm{C}}(x)$ but not in $c_2(x)$. Second, a corner of C is also a corner of other polygons adjacent to C and hence, similar to the case in $n_{0\mathrm{S}}(x)$, will be counted more than once in $n_{2\mathrm{C}}(x)$ but only once in $c_2(x)$.

Each of these marks leads to a mark distribution. Denote by n_{02}, n_{0S}, l_0, l_1, n_{20}, c_2, l_2, and a_2, respectively, random variables following these mark distributions. The corresponding means are denoted by $\overline{n}_{02}$, $\overline{n}_{0S}$, $\overline{l}_0$, $\overline{l}_1$, $\overline{n}_{20}$, $\overline{c}_2$, $\overline{l}_2$ and $\overline{a}_2$. Each random variable and its mean are referred to as the appropriate random value and the mean value of the corresponding geometrical quantity. For example, l_1 is the random length of the typical edge with mean $\overline{l}_1$, n_{20} the random number of vertices on the boundary of the typical cell with mean $\overline{n}_{20}$ (for the definition of typical vertex, edge and cell; see Section 9.3.2 below). Since the number of vertices on the boundary of a cell is equal to the number of edges, n_{20} is sometimes also denoted as n_{21}.

9.3.2 Typical vertex, edge and cell

Another marking of the points of $\alpha_2(\Theta)$ uses the corresponding cells, taken as compact convex sets. If x_n is a cell centre and C_n the corresponding cell, then $\{[x_n; C_n]\}$ is a germ–grain process

as in Section 6.5, which inherits stationarity from Θ. The typical grain of $\{[x_n; C_n]\}$ is called the *typical cell* of Θ in the tessellation context and is denoted by C^o. Analogously, for $\alpha_1(\Theta)$, the edge midpoints and the corresponding edges form another stationary germ–grain process, in which the typical grain is called the *typical edge*; for $\alpha_0(\Theta)$ with marks, the mark distribution is known as the distribution of the marks of the *typical vertex*.

Note that the typical cell should not be viewed as a particular cell selected from a given tessellation. Instead, it is a valuable distributional single-cell characteristic of a stationary tessellation. In a heuristic way the typical cell can be seen as the cell that is sampled from the population of all cells by random selection where all cells have the same chance to be chosen. The mean values for characteristics of the typical cell can be computed by using the Palm distribution of $\alpha_2(\Theta)$ or the ergodic theorem; see Section 9.3.4 below.

Simulating the typical cell of a stationary tessellation

The aim to simulate an i.i.d. sequence of the typical cell of a stationary tessellation can be achieved approximately by, first, simulating a sample of the tessellation in a large window, then enumerating the cells in some sequential order and finally selecting the cells uniformly at random. Clearly, this method is not very efficient; there may be stochastic dependence between the cells obtained from the same sample of the tessellation; and edge-effects may play a rôle.

In the case of Poisson-process-related tessellations considered in Sections 9.5 and 9.7 there are efficient special constructions. The Slivnyak–Mecke theorem is the basis for the simulations.

9.3.3 Zero cell

In addition to the typical cell, another single cell is associated with a tessellation, namely the *zero cell* C_0. This is that cell that contains the origin of the plane. Because of the different definitions, zero cell and typical cell are objects with different probability distributions. As will be shown in inequalities (9.30) and (9.31), the zero cell tends to be larger than the typical cell.

Another way to understand the distribution of the zero cell C_0 uses random sampling with points. Take a random test point and choose the cell containing the test point. By stationarity, the corresponding cell has the same distribution as the zero cell. It is plausible that the test point tends to fall with greater probability into larger cells. Thus the zero cell can be seen as a size-weighted version of a single cell.

Motivated by these heuristic considerations, Matheron (1975) used the terms *number law* for the distribution of the typical cell and *volume law* for that of the zero cell.

9.3.4 Mean-value relationships for stationary planar tessellations

There are a number of relationships holding between the above mean values in the case of stationary tessellations. They have been established by means of Palm distributions, as in Mecke (1980), and by means of limiting arguments involving ergodicity, as in Cowan (1978, 1980). As Mecke (1984b) showed, for almost surely face-to-face tessellations all these mean values can be expressed by the *three* parameters λ_0, λ_2 (or $\overline{n}_{02}$) and L_A. This follows from simple geometric reasoning, which has to be made precise in one of the two ways above. For general tessellations one more parameter is needed, namely, the proportion ϕ of π-vertices; see Weiss and Cowan (2011).

For example, the Euler formula for *finite* polyhedra in $\mathbb{R}^3$ states that

$$\#\text{corners} - \#\text{sides} + \#\text{facets} = 2.$$

Since a tessellation can be interpreted as the surface of an *infinite* polyhedron with the vertices and edges of the tessellation as the sides and corners of the polyhedron, limiting arguments show that

$$\lambda_0 - \lambda_1 + \lambda_2 = 0. \tag{9.8}$$

This basic formula exhibits a duality between vertices and cells; that is, between the indices '$_0$' and '$_2$'. Any tessellation is associated with a *dual* tessellation obtained by taking the original cell centroids as new vertex and connecting pairs of new vertices if they are centroids of neighbouring old cells. This association explains the duality between indices which underlies several of the formulae given below:

$$\overline{n}_{02} = 2 + \frac{2\lambda_2}{\lambda_0}, \tag{9.9}$$

$$\overline{n}_{20} = 2 + \frac{2\lambda_0}{\lambda_2}, \tag{9.10}$$

$$\overline{n}_{0S} = 2\overline{n}_{02} - \phi \tag{9.11}$$

$$\overline{c}_2 = \overline{n}_{20} - \frac{2\phi}{\overline{n}_{02} - 2}, \tag{9.12}$$

$$\overline{l}_0 = \frac{2L_A}{\lambda_0}, \tag{9.13}$$

$$\overline{l}_2 = \frac{2L_A}{\lambda_2}, \tag{9.14}$$

$$\overline{a}_2 = \frac{1}{\lambda_2}, \tag{9.15}$$

$$\overline{l}_1 = \frac{L_A}{\lambda_0 + \lambda_2} = \frac{L_A}{\lambda_1}. \tag{9.16}$$

Formula (9.11) can be understood easily from Figure 9.11, in which the three polygons meet at the π-vertex c; each of the polygons A and B has two sides adjacent to c, whilst polygon C has only one, which is the union of two edges. Now, first, note that for a non-π-vertex, each edge adjacent to it is a side (or part of a side) of two polygons adjacent to the vertex and will be counted twice, explaining the factor 2 in Formula (9.11). Then, consider the edges at a π-vertex, the two edges lying on the same polygonal side will together contribute not 4 but only 3 to the total count of sides adjacent to the π-vertex, leading to the term $-\phi$ in Formula (9.11).

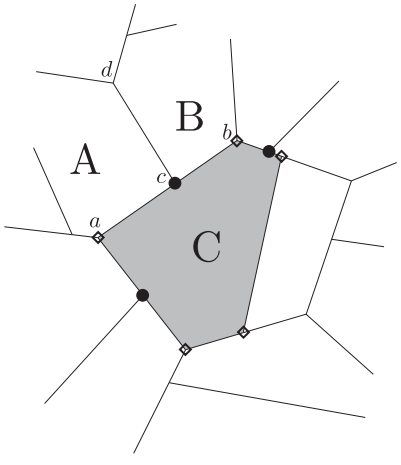

Figure 9.11 A tessellation with π-vertices. Polygon C has eight vertices, five of them are corners, represented by $\diamond$, and three are π-vertices, represented by $\bullet$. Vertex c is a corner of polygon A and polygon B. Polygon A has two sides adjacent to vertex c, namely (a, c) and (c, d); polygon B also has two sides adjacent to vertex c, namely (b, c) and (c, d); polygon C, however, has only one side adjacent to vertex c, namely (a, b). Thus, in total there are five polygonal sides adjacent to vertex c. On the other hand, three edges of tessellation are adjacent to vertex c, namely (a, c), (b, c) and (c, d).

From the above formulae others can be derived:

$$\frac{1}{\overline{n}_{02}} + \frac{1}{\overline{n}_{20}} = \frac{1}{2}, \tag{9.17}$$

$$\bar{l}_0 = \overline{n}_{02}\bar{l}_1, \tag{9.18}$$

$$\bar{l}_2 = \overline{n}_{20}\bar{l}_1, \tag{9.19}$$

$$\overline{n}_{20} = \frac{2\overline{n}_{02}}{\overline{n}_{02} - 2}, \tag{9.20}$$

$$\overline{c}_2 = \frac{2(\overline{n}_{02} - \phi)}{\overline{n}_{02} - 2}, \tag{9.21}$$

and finally

$$\overline{n}_{02}, \overline{n}_{20}, \overline{c}_2 \geq 3, \tag{9.22}$$

$$\overline{n}_{20} \leq 6 - \phi. \tag{9.23}$$

More mean-value formulae for tessellations with π-vertices can be found in Weiss and Cowan (2011). Most formulae related to polygonal k-faces involve two more parameters, namely, the second moment of n_{02} and the mean number of edges emanating from the typical π-vertex.

Kendall and Mecke (1987) study the range

$$0 < \lambda_0 \le 2\lambda_2, \qquad \lambda_2 \le 2\lambda_0, \qquad L_A \ge 0$$

and show that each triplet satisfying these relations can be derived from a suitable tessellation.

The ordinary equilibrium or *trivalent state* is an important special case. It holds when

$$n_{02}(x) \equiv 3 \qquad \text{for all } x \in \alpha_0(\Theta), \tag{9.24}$$

or

$$n_{02} \equiv 3 \qquad \text{almost surely}, \tag{9.25}$$

that is, when the number of edges emanating from each vertex is always three, as in Figures 9.1, 9.4 and 9.17. Tessellations with this property are also called *trivalent*. In this case simpler relationships hold:

$$\overline{n}_{20} = \overline{n}_{21} = 6, \qquad 2\lambda_2 = \lambda_0, \qquad \lambda_1 = 3\lambda_2. \tag{9.26}$$

Thus in the ordinary equilibrium state the mean number of edges (or vertices) of the typical cell is 6.

Formulae for the particular case of tessellations formed by line processes can be found in Section 9.5. In this case four edges emanate from each vertex and so the ordinary equilibrium state does not hold.

Proofs of some mean-value formulae

The following gives proofs for four examples of the general mean-value formulae. The proofs are taken from Mecke (1980) and are good examples of the use of the theorems concerning Palm distribution calculations. To begin with, some notation is introduced.

If Θ is a stationary planar tessellation with distribution P, then the $\alpha_k(\Theta)$ $(k = 0, 1, 2)$ and also $\beta(\Theta) = \alpha_0(\Theta) \cup \alpha_1(\Theta) \cup \alpha_2(\Theta)$ are stationary point processes. The intensity of $\beta(\Theta)$ is $\lambda' = \lambda_0 + \lambda_1 + \lambda_2$, and under the general assumptions on tessellations λ' is positive and finite.

The Palm distribution P_o of $\beta(\Theta)$ is

$$\lambda' P_o(Y) = \int_{\mathbb{T}} \sum_{x \in \beta(\theta) \cap [0,1]^2} \mathbf{1}_Y(\theta - x) P(\mathrm{d}\theta) \qquad \text{for } Y \in \mathscr{T}. \tag{9.27}$$

This may be regarded intuitively as the probability distribution of a tessellation configuration given that a point of $\beta(\Theta)$ is at o. It is useful to further distinguish which of the $\alpha_k(\Theta)$ contains this point, called the typical point. If $T_k = \{\theta \in \mathbb{T} : o \in \alpha_k(\theta)\}$, then let

$$E_k\big(g(\Theta)\big) = \frac{\lambda'}{\lambda_k} \int_{T_k} g(\theta) P_o(\mathrm{d}\theta) \qquad \text{for } k = 0, 1, 2,$$

where $g : T_k \to [0, \infty)$ is measurable. The quantity $E_k\big(g(\Theta)\big)$ can be interpreted as the mean of $g(\Theta)$ under the condition that o is a point of $\alpha_k(\Theta)$.

Proposition 9.1. *Let $C_0(\Theta) = C_0$ be the zero cell of the tessellation Θ. Then*

$$\mathbf{E}\big(f(\Theta)\big) = \lambda_2 E_2 \left(\int_{C_0(\Theta)} f(\Theta - x)\,\mathrm{d}x \right) \tag{9.28}$$

for every measurable nonnegative function f on $\mathbb{T}$.

Proof. The right-hand side, using the definition of $E_2(\cdot)$ and change of coordinates $x \to -x$, becomes

$$\lambda_2 E_2 \left(\int_{C_0} f(\Theta - x)\,\mathrm{d}x \right) = \lambda' \int_{\mathbb{T}} \int_{\mathbb{R}^2} \mathbf{1}_{T_2}(\theta)\mathbf{1}_{C_0(\theta)}(-x) f(\theta + x)\,\mathrm{d}x\, P_o(\mathrm{d}\theta).$$

A variation of Theorem 7.1(a) yields

$$\lambda' \int_{\mathbb{T}} \int_{\mathbb{R}^2} v(x,\theta)\,\mathrm{d}x\, P_o(\mathrm{d}\theta) = \int_{\mathbb{T}} \sum_{x \in \beta(\theta)} v(x, \theta - x) P(\mathrm{d}\theta),$$

which can be applied to show that the right-hand side above equals

$$\begin{aligned} &\int_{\mathbb{T}} \sum_{x \in \beta(\theta)} \mathbf{1}_{T_2}(\theta - x)\mathbf{1}_{C_0(\theta - x)}(-x) f(\theta) P(\mathrm{d}\theta) \\ &\quad = \int_{\mathbb{T}} f(\theta) \cdot \#\{x \in \alpha_2(\theta) : -x \text{ belongs to the cell of } \theta - x \text{ containing } o\} P(\mathrm{d}\theta). \end{aligned}$$

If o is in the edge system E_θ of θ then there can be no cells containing it. Otherwise there can be only one such cell, and so only one corresponding cell centroid x in $\alpha_2(\theta)$. Also by stationarity

$$\mathbf{P}(o \in E_\Theta) = 0.$$

Hence the right-hand side of the equation in the proposition must equal

$$\int f(\theta) P(\mathrm{d}\theta) = \mathbf{E}\big(f(\Theta)\big),$$

which proves the proposition. □

Formula (9.28) can be explained intuitively as follows. The left-hand side is a kind of average of $f(\theta - x)$ over all x using an ergodic-theoretic interpretation of expectation and the stationary nature of Θ. The right-hand side is a kind of sum over all cell centroids (hence the factor λ_2) of integrals of $f(\theta - x)$ for x in the zero cell. It is natural to suppose that the two sides should be equal. The proof above uses Palm distributions to make this argument rigorous.

Taking $f \equiv 1$, Formula (9.28) yields

$$1 = \lambda_2 E_2 \left(\int_{C_0(\Theta)} \mathrm{d}x \right) = \lambda_2 \bar{a}_2$$

which proves (9.15).

With $f(\Theta) = \nu_2\big(C_0(\Theta)\big)^{-1}$, Formula (9.28) yields

$$\mathbf{E}(\nu_2(C_0)^{-1}) = \lambda_2 = \frac{1}{\overline{a}_2}. \tag{9.29}$$

An application of Jensen's inequality for $\mathbf{E}\big(f(X)\big)$ with convex f for the special $f(x) = x^{-1}$ shows that

$$\overline{a}_2 \le \mathbf{E}\big(\nu_2(C_0)\big), \tag{9.30}$$

which says that the mean area of the typical cell is smaller than that of the zero cell. This is one aspect of the property that the zero cell is 'larger' than the typical cell. Mecke (1999) shows more: The distribution functions of the areas of C_0 and C^o, denoted by $F_0(a)$ and $F^o(a)$ respectively, satisfy the inequality

$$F_0(a) \le F^o(a) \qquad \text{for all } a \ge 0, \tag{9.31}$$

which implies that all moments of order greater than 1 of the area of the zero cell are greater than or equal to that of the typical cell, that is, the area of the zero cell is stochastically larger than that of the typical cell.

Mecke (1999) refines this inequality for Poisson line/plane tessellations and Poisson-Voronoi tessellations by showing that, by a suitable interpretation, the typical cell of each of two tessellations can be embedded as a subset in the zero cell of the tessellation, that is, there is $\subseteq_{\mathrm{st}}$-order as defined on p. 212.

Møller (1989, Theorem 5.1) (see also Schneider and Weil, 2008, p. 493, Theorem 10.4.1) proves the following equation, which was originally given in Miles (1974a) for Poisson line/plane tessellation:

$$\mathbf{E}\big(f(C_0)\big) = \mathbf{E}\big(f(C^o)\nu_2(C^o)\big)/\overline{a}_2 \tag{9.32}$$

for general stationary tessellations and any measurable, translation invariant function f that assigns to convex bodies nonnegative real numbers. An example is:

$$\mathbf{E}\big(N(C_0)\big) = \mathbf{E}\big(N(C^o)\nu_2(C^o)\big)/\overline{a}_2, \tag{9.33}$$

where $N(C)$ is the number of neighbours of C. It can be used to determine the mean number of neighbours of the zero cell. Equation (9.32) shows in a clear way that the zero cell can be interpreted, up to translations, an area-weighted version of the typical cell.

Stoyan (1986) shows by counter-examples that mean perimeter and mean number of neighbours of the zero cell are not necessarily larger than their counterparts for the typical cell.

Proposition 9.2. *For a tessellation θ let $S(\theta)$ be the set of all cell centroids of cells containing o. Then for a stationary tessellation Θ*

$$\lambda_2 E_2 \left(\sum_{x \in \alpha_0(\Theta) \cap C_0(\Theta)} h(x, \Theta) \right) = \lambda_0 E_0 \left(\sum_{x \in S(\Theta)} h(-x, \Theta - x) \right) \tag{9.34}$$

for any nonnegative measurable function h on $\mathbb{R}^2 \times \mathbb{T}$.

Proof. A variation on Theorem 7.1(c) yields

$$\lambda' \int_{\mathbb{T}} \sum_{x \in \beta(\theta)} v(x, \theta) P_o(\mathrm{d}\theta) = \lambda' \int_{\mathbb{T}} \sum_{x \in \beta(\theta)} v(-x, \theta - x) P_o(\mathrm{d}\theta) \tag{9.35}$$

for nonnegative measurable functions v on $\mathbb{R}^2 \times \mathbb{T}$. Taking

$$v(x, \theta) = \mathbf{1}_{T_2(\theta)} \cdot \mathbf{1}_{\alpha_0(\theta) \cap C_0(\theta)}(x) \cdot h(x, \theta)$$

and applying (9.35) to the left-hand side of (9.34) give

$$\begin{aligned}
\lambda_2 E_2 \left(\sum_{x \in \alpha_0(\Theta) \cap C_0(\Theta)} h(x, \Theta) \right) &= \lambda' \int_{T_2} \sum_{x \in \alpha_0(\theta) \cap C_0(\theta)} h(x, \theta) P_o(\mathrm{d}\theta) \\
&= \lambda' \int_{\mathbb{T}} \sum_{x \in \beta(\theta)} v(x, \theta) P_o(\mathrm{d}\theta) \\
&= \lambda' \int_{\mathbb{T}} \sum_{x \in \beta(\theta)} v(-x, \theta - x) P_o(\mathrm{d}\theta) \\
&= \lambda' \int_{\mathbb{T}} \sum_{x \in \alpha_2(\theta)} h(-x, \theta - x) \mathbf{1}_{\alpha_0(\theta - x)}(-x) \mathbf{1}_{C_0(\theta - x)}(-x) P_o(\mathrm{d}\theta).
\end{aligned}$$

If x is in $\alpha_2(\theta)$ then the summand above is zero unless o is a vertex, in which case $\theta \in T_0$, and x is the cell centroid of one of the cells in θ containing o as a vertex. Thus the last integrand above is equal to

$$\mathbf{1}_{T_0}(\theta) \sum_{x \in S(\theta)} h(-x, \theta - x)$$

from which (9.34) follows. □

As with (9.28), Formula (9.34) can be understood intuitively by an 'averaging' argument. The left-hand side averages over all cell centroids and counts the quantity $h(x, \theta)$ for x running through the vertices of the typical cell. The right-hand side averages over all vertices and counts the appropriate quantity for x running through the cell centroids of cells meeting at the typical vertex.

With $h \equiv 1$, Formula (9.34) yields

$$\lambda_2 \overline{n}_{20} = \lambda_0 \overline{n}_{02}.$$

If $h(x, \theta)$ is defined as the angle at the vertex x in $C_0(\theta)$ then the left-hand side of (9.34) involves the sum of all angles of the convex polygon $C_0(\theta)$. This sum equals

$$\big(\#\{\text{vertices of } C_0(\theta)\} - 2\big)\pi,$$

and so the left-hand side of (9.34) is $\lambda_2(\overline{n}_{20} - 2)\pi$. The right-hand side is $2\pi\lambda_0$ and so

$$\lambda_2 \overline{n}_{20} - 2\lambda_2 = 2\lambda_0.$$

Together these two derived equations yield the Formulae (9.9), (9.10) and (9.17).

9.3.5 The neighbourhood of the typical cell

The neighbourhood of the typical cell is of great importance, especially in materials science and physics. A well-known result is the so-called *Aboav's law* or *Aboav–Weaire's law* (Aboav, 1970, 1980; Weaire, 1974), which quantifies the relationship between n and $m(n)$, the mean number of edges of a randomly chosen neighbouring cell of the typical cell, under the condition that the typical cell has n edges. Lewis (1931) observed a remarkable tendency for few-edged cells to be in contact with many-edge cells and vice versa in epithelia. Aboav in a series of papers asserts empirically that

$$m(n) = 5 + \frac{8}{n}, \tag{9.36}$$

and Weaire suggested semi-empirically

$$m(n) = 5 + \frac{6 + \mathrm{var}(n_{20})}{n}, \tag{9.37}$$

where $\mathrm{var}(n_{20})$ denotes the variance of the number of edges (vertices) of the typical cell. Chiu (1994) showed mathematically that for trivalent tessellations,

$$m(n) = 5 + \frac{6}{n} + \frac{\mathrm{cov}\big(k(n, n_{20}), n_{20}\big)}{np_n} \qquad \text{for } n = 3, 4, \ldots, \tag{9.38}$$

where $p_n = \mathbf{P}(n_{20} = n)$ and $k(n, j)$ is the mean number of n-edged cells which belong to the complex of the typical cell, which has j edges. The complex of the typical cell is the union of the typical cell and its neighbouring cells, that is, its first shell. For an extensive review see Chiu (1995a) and Okabe *et al.* (2000, Section 5.5.3); for further work in this direction see Sahimi (2003, p.89) and Hilhorst (2006).

For unconditional mean-value formulae for the neighbourhood of the typical cell see Weiss (1995). She showed, in particular, that the mean area A_1 of the first shell of cells around the typical cell is given by

$$A_1 = \overline{a}_2 \mathbf{E}\big(N(C_0)\big), \tag{9.39}$$

where $\mathbf{E}\big(N(C_0)\big)$ is the mean number of edges of the zero cell. An analogous formula holds for tessellations in $\mathbb{R}^d$. Voloshin *et al.* (2010) study the volumes of the further shells by simulation. This is related to shell map analysis, see p. 393.

A heuristic proof of (9.39) goes as follows. Consider a large disc containing n cells. If edge-effects are ignored, the mean shell area A_1 is

$$A_1 = \frac{1}{n} \sum_i \sum_k a_{i,k},$$

where i goes over the cells and $a_{i,k}$ is the area of the k^{th} cell in the shell around cell i. Since the area a_i of cell i appears in the double-sum as many times as the total number n_i of its neighbours, the mean shell area can be rewritten as

$$A_1 = \frac{1}{n} \sum_i n_i a_i = \frac{\sum_i n_i a_i}{\sum_i a_i} \cdot \frac{\sum_i a_i}{n} = \mathbf{E}\big(N(C_0)\big)\overline{a_2}.$$

Note that the last equality comes from the interpretation that the zero cell is the size-weighted version of a single cell.

The neighbourhoods of typical k-faces of the Poisson-Voronoi tessellation are studied in Baumstark and Last (2007).

Other characteristics of an n-edged cell

Another empirically suggested relationship for an n-edge cell is the so-called *Lewis' law*, which asserts a linear relation for the mean area $A(n)$ of the typical cell under the condition that it has n edges, namely

$$A(n) = \bar{a}_2 \left(\frac{n - n_0}{6 - n_0} \right), \tag{9.40}$$

for some constant n_0. The value 6 in the denominator comes from the mean number of edges of the typical cell in a trivalent tessellation. Quine and Watson (1984) conjecture that for the Poisson-Voronoi tessellation $n_0 = 3/2$, leading to

$$A(n) = \bar{a}_2 \left(\frac{2n - 3}{9} \right), \tag{9.41}$$

see Chiu (1995a) and Okabe *et al.* (2000, p. 315) for more empirically suggested forms.

Rivier and Lissowski (1982) develop a maximum entropy argument to explain heuristically that the linear relationship might result from statistical equilibrium, which might also give rise to Aboav's law (Peshkin *et al.*, 1991). However, Chiu (1995b) proves that their argument could only lead to tautological statements.

Conditional mean perimeter and mean edge length of an n-edged cell in a Poisson-Voronoi tessellation have been studied by simulation; see Okabe *et al.* (2000, p. 319).

9.4 Mean-value formulae for stationary spatial tessellations

Mean-value relationships hold for stationary spatial tessellations and tessellations of $\mathbb{R}^d$ just as for stationary planar tessellations. The relationships are due to Mecke (1980) and Stoyan and Mecke (1983a, Chapter 8), to Radecke (1980) for $\mathbb{R}^3$, and to Møller (1989) for general $\mathbb{R}^d$, and are displayed below for $d = 3$.

Again the terminologies in Weiss and Cowan (2011) are adopted. For the spatial case, the 0-, 1-, and 2-faces of a tessellation are called *vertices*, *edges* and *plates*, whilst those of the constituent polyhedra are called *apices*, *ridges* and *facets*. The polyhedra (3-faces) themselves are still called *cells*. A cell k-face may be a union of some k-faces of the tessellation, and lower dimensional faces may be in the relative interior of higher dimensional faces. When an edge whose interior lies in the interior of a facet, it is called a *π-edge*, and a vertex lying in the interior of a facet is a *hemi-vertex*; the same as in the planar case, a vertex in the interior of a ridge is a π-vertex. A tessellation without π-vertices, π-edges and hemi-vertices is *face-to-face*. When there are always four cells meeting at a vertex and three cells meeting at an edge, the tessellation is said in *ordinary equilibrium state*.

The definition of a face-to-face and ordinary equilibrium planar tessellation can be extended to a tessellation in $\mathbb{R}^d$. A tessellation is face-to-face, or *regular*, if the set of all k-faces, $k = 0, 1, \ldots, d$, of the tessellation and that of the constituent cells coincide. If each k-face of a tessellation of $\mathbb{R}^d$ lies on the boundaries of $d - k + 1$ cells for $k = 0, 1, \ldots, d - 1$, the tessellation is said to be in ordinary equilibrium state or *normal*. The latter term is used because many real-life tessellations possess this property. Some authors additionally require that a normal tessellation has to be face-to-face (see e.g. Lautensack and Zuyev, 2008), but this book treats the two properties as separate.

Let λ_k be the intensity of the point process of centroids of k-faces ($k = 0, 1, 2, 3$), let L_V be the intensity of the fibre process of edges (1-faces of Θ), and let S_V be the intensity of the surface process of plates (2-faces of Θ). If the vertices and edges are marked or weighted with the numbers of adjacent cells then the vertex point process and edge fibre process can be used to produce stationary random measures (multiplying by the marks or weights) whose corresponding intensities are denoted by T_V and Z_V.

The following notation is used.

$$
\begin{aligned}
\overline{n}_{kl} &= \text{the mean number of } l\text{-faces of the tessellation adjacent to} \\
&\qquad \text{the typical } k\text{-face of the tessellation, for } k \text{ and } l \text{ in } \{0, 1, 2, 3\}, \\
\overline{l}_1 &= \text{mean length of the typical edge (1-face of the tessellation)}, \\
\overline{l}_2 &= \text{mean perimeter of the typical plate (2-face of the tessellation)}, \\
\overline{A}_2 &= \text{mean area of the typical plate (2-face of the tessellation)}, \\
\overline{l}_3 &= \text{mean total edge length of the typical cell}, \\
\overline{S}_3 &= \text{mean surface area of the typical cell}, \\
\overline{V}_3 &= \text{mean volume of the typical cell}, \\
\overline{\overline{b}_3} &= \text{mean average breadth of the typical cell}.
\end{aligned}
$$

The quantity $\overline{n}_{30}$ is sometimes called *mean coordination number*.

Mecke (1984c) showed that, when restricted to tessellation k-faces, all these means can be expressed by the *seven* parameters

$$\lambda_0,\ \lambda_3,\ \lambda = \lambda_1 + \lambda_2,\ L_V,\ S_V,\ T_V,\ \text{and } Z_V,$$

whilst Weiss and Cowan (2011) found that, when cell k-faces are also of interest, four more parameters are needed, namely,

$$
\begin{aligned}
\xi &= \text{proportion of } \pi\text{-edges}, \\
\kappa &= \text{proportion of hemi-vertices}, \\
\psi &= \text{mean number of ridge interiors adjacant to the typical vertex}, \\
\tau &= \text{mean number of plate-side interiors adjacent to the typical vertex}.
\end{aligned}
$$

The relevant formulae for k-faces of tessellation are as follows:

$$2\lambda_1 = \lambda + \lambda_0 - \lambda_3, \qquad 2\lambda_2 = \lambda + \lambda_3 - \lambda_0, \tag{9.42}$$

$$\lambda_0 \overline{n}_{03} = T_V, \qquad \lambda_3 \overline{n}_{30} = T_V, \tag{9.43}$$

$$\lambda_0 \overline{n}_{02} = T_V + \lambda - \lambda_0 - \lambda_3, \qquad \lambda_0 \overline{n}_{01} = \lambda + \lambda_0 - \lambda_3, \tag{9.44}$$

$$\lambda_3 \overline{n}_{32} = \lambda + \lambda_3 - \lambda_0, \qquad \lambda_3 \overline{n}_{31} = T_V + \lambda - \lambda_0 - \lambda_3, \tag{9.45}$$

$$\overline{n}_{12} = \overline{n}_{13} = 2\,\frac{T_V + \lambda - \lambda_0 - \lambda_3}{\lambda + \lambda_0 - \lambda_3}, \qquad \overline{n}_{21} = \overline{n}_{20} = 2\,\frac{T_V + \lambda - \lambda_0 - \lambda_3}{\lambda + \lambda_3 - \lambda_0}, \tag{9.46}$$

$$\overline{l}_1(\lambda + \lambda_0 - \lambda_3) = 2L_V, \qquad \overline{l}_2(\lambda + \lambda_3 - \lambda_0) = 2Z_V, \tag{9.47}$$

$$\overline{A}_2(\lambda + \lambda_3 - \lambda_0) = 2S_V, \tag{9.48}$$

$$\lambda_3 \overline{l}_3 = Z_V, \qquad \lambda_3 \overline{S}_3 = 2S_V, \tag{9.49}$$

$$\lambda_3 \overline{V}_3 = 1, \qquad 4\lambda_3 \overline{\overline{b}_3} = Z_V - 2L_V. \tag{9.50}$$

These formulae imply the relationships:

$$\lambda_0 - \lambda_1 + \lambda_2 - \lambda_3 = 0, \tag{9.51}$$

$$\overline{n}_{01} - \overline{n}_{02} + \overline{n}_{03} = 2, \qquad \overline{n}_{30} - \overline{n}_{31} + \overline{n}_{32} = 2, \tag{9.52}$$

$$\overline{n}_{12}\overline{n}_{01} = 2\overline{n}_{02}, \qquad \overline{n}_{21}\overline{n}_{32} = 2\overline{n}_{31}, \qquad \overline{n}_{02}\overline{n}_{30} = \overline{n}_{31}\overline{n}_{03}, \tag{9.53}$$

$$\overline{S}_3 = \overline{n}_{32}\overline{A}_2. \tag{9.54}$$

Some of these formulae display symmetries between the indices '$_0$' and '$_3$' and between the indices '$_1$' and '$_2$'.

For cell k-faces, denote by $\overline{m}_{kl}$ the mean number of cell l-faces of the typical k-face of the tessellation. The same as in the planar case, $\overline{m}_{kl}$ must not be confused the mean number of cell l-faces *adjacent to* the typical k-face, because when adjacency is of interest, each cell l-face will be counted as many times as the number of cells containing it. The relationships analogous to (9.21) are

$$\overline{m}_{30} = \overline{n}_{30} - \frac{2(\kappa + \psi)\overline{n}_{20}}{\overline{n}_{01}\overline{n}_{12} - \overline{n}_{20}(\overline{n}_{01} - 2)}, \tag{9.55}$$

$$\overline{m}_{31} = \overline{n}_{31} - \frac{(\xi\overline{n}_{01} + 2\psi)\overline{n}_{20}}{\overline{n}_{01}\overline{n}_{12} - \overline{n}_{20}(\overline{n}_{01} - 2)}, \tag{9.56}$$

$$\overline{m}_{32} = \overline{n}_{32} - \frac{(\xi\overline{n}_{01} - 2\kappa)\overline{n}_{20}}{\overline{n}_{01}\overline{n}_{12} - \overline{n}_{20}(\overline{n}_{01} - 2)}, \tag{9.57}$$

$$\overline{m}_{20} = \overline{m}_{21} = \overline{n}_{20}\left(1 - \frac{2\tau}{\overline{n}_{01}\overline{n}_{12}}\right). \tag{9.58}$$

More relationships involving cell k-faces can be found in Weiss and Cowan (2011).

Consider the special case of tessellations that are face-to-face. Then cell faces (apices, ridges and facets) and tessellation faces (vertices, edges and plates) coincide (and in this case the terms vertices, edges and cell facets are used in the literature). If in addition the tessellation is normal, that is, the random numbers n_{kl} of l-faces meeting at the typical k-face are

$$n_{03} \equiv 4 \text{ and } n_{13} \equiv n_{12} \equiv 3 \qquad \text{almost surely,} \tag{9.59}$$

then the intensities satisfy

$$\lambda_1 = 2\lambda_0, \qquad T_V = 4\lambda_0, \tag{9.60}$$

$$\lambda_2 = \lambda_0 + \lambda_3, \qquad Z_V = 3L_V, \tag{9.61}$$

and furthermore

$$\overline{n}_{01} = \overline{n}_{03} = 4, \qquad \overline{n}_{02} = 6. \tag{9.62}$$

Lautensack (2007) shows that each normal and face-to-face tessellation in $\mathbb{R}^d$ with $d > 2$ (the case $d = 2$ is excluded) can be considered as a Laguerre tessellation with some system of points and weights. Perhaps this striking fact will in due course find application in tessellation statistics and stereology.

For such tessellations the following important formula holds

$$\overline{n}_{32} = \frac{12}{6 - \overline{n}_{21}}. \tag{9.63}$$

It can be proved by means of Schneider and Weil (2008, Theorem 10.1.7).

Zero cells in the spatial case are also of interest. Their definition and properties are analogous to those in the planar case, and Inequality (9.31) remains true; see Mecke (1999).

It is not surprising that many of the mean-value relations are true under more general assumptions, since they have a topological character. That means that the convexity assumptions can be dropped. This is shown by Stoyan (1986), Zähle (1987b, 1988), Weiss and Zähle (1988) and Leistritz and Zähle (1992).

Formulae analogous to the Aboav law (9.37) and (9.38) for spatial tessellations have been suggested empirically in Aboav (1991, 1992) and established mathematically in Chiu (1994), respectively.

For planar sections of spatial tessellations see Section 10.6.

9.5 Poisson line and plane tessellations

Poisson line and plane processes produce tessellations in the plane and space respectively. Formulae for these tessellations are presented here; some results for higher-dimensional Poisson

hyperplane tessellations can be found in Miles (1971a, 1974a) and Matheron (1975, Chapter 6); see also Schneider and Weil (2000, Section 6.3; 2008, Section 10.3).

9.5.1 Poisson line tessellations

Let Ψ be a planar motion-invariant Poisson line process of intensity L_A; the line intersections form the vertices and a segment of a line in Ψ with vertices at both endpoints but no vertices in the segment interior is an edge. Thus, the Ψ induces an edge system of a motion-invariant tessellation, known as the *Poisson line tessellation*. Such tessellations have no π-vertices and hence are face-to-face. However, there are always four edges emanating from each vertex and so they are not trivalent.

The quantity

$$\varrho = P_L = \frac{2L_A}{\pi}, \tag{9.64}$$

the mean number of lines intersected by a test line segment of unit length, serves as a convenient model parameter. The intensity of the point process of intersection points on a fixed line g in Ψ is also given by ϱ. Hence

$$\bar{l}_1 = \frac{1}{\varrho}. \tag{9.65}$$

With probability one the Poisson line process is bundle-free (has no triplets of lines meeting in a single point) and so each vertex is formed by the intersection of precisely two lines. Thus

$$n_{02} \equiv 4 \qquad \text{almost surely}, \tag{9.66}$$

and the formulae in Section 9.3.4 yield

$$\lambda_0 = \frac{\pi\varrho^2}{4}, \quad \lambda_1 = \frac{\pi\varrho^2}{2}, \quad \lambda_2 = \frac{\pi\varrho^2}{4}, \quad \text{and} \quad \bar{l}_0 = \frac{4}{\varrho}. \tag{9.67}$$

Suppose the cell centroid process is marked by the corresponding polygons. The mark distribution is then the distribution of the typical polygon of the tessellation process, also known as the *Poisson polygon*.

Several formulae are known for the distributions of characteristics of this polygon. First and second moments are given in Table 9.1. They are mostly due to Miles (1964a,b, 1973). Information about density functions of area, perimeter, and breadth can be found in Miles (1964a,b, 1973), Matheron (1975), Crain and Miles (1976), Deng and Dodson (1994) and Stoyan and Stoyan (1994).

The distribution $\{p_n\}$ of the number of edges is partially known and tabulated in Table 9.2; see Tanner (1983a,b), who also gave the values of the corresponding third and fourth moments, and Calka (2003b), who obtained the distribution of the number of edges for the zero cell.

Crain and Miles (1976) noted that the distribution $\{p_n\}$ can be approximated by a shifted Poisson distribution of parameter 1:

$$p_n \approx \frac{\mathrm{e}^{-1}}{(n-3)!} \qquad \text{for } n = 3, 4, \ldots. \tag{9.68}$$

Table 9.1 Poisson polygon formulae.

		Area	Perimeter	Edge number
First moments:		A	L	N
		$4/(\pi\varrho^2) = \overline{a}_2$	$4/\varrho = \bar{l}_2$	$4 = \overline{n}_{20}$
Second moments:		A	L	N
	A	$8/\varrho^4$		
	l_2	$4\pi/\varrho^3$	$2(\pi^2+4)/\varrho^2$	
	N	$2\pi/\varrho^2$	$(\pi^2+8)/\varrho$	$(\pi^2+24)/2$

They found the values p_5, p_6, ... and much other distributional information by simulation of the Poisson polygon. The method was to generate the tessellation in a large circle and to determine the desired quantities by statistical estimation. George (1987) and Maier *et al.* (2004) described how to simulate single Poisson polygons. Michel and Paroux (2007) provided another fast simulation method.

The mean area and mean number of edges of the zero cell C_0, the cell containing the origin, are

$$\mathbf{E}\big(\nu_2(C_0)\big) = \frac{2\pi}{\varrho^2}, \quad \text{and} \quad \mathbf{E}\big(N(C_0)\big) = \frac{\pi^2}{2}, \tag{9.69}$$

which are larger than the corresponding parameters of the typical cell given in Table 9.1. Michel and Paroux (2007) estimate higher moments by simulation.

Note that mean-value formulae can also be derived in the *anisotropic* case. Let $\mathscr{R}$ be the rose of directions of the anisotropic Poisson line process generating the tessellation. Then

$$\overline{n}_{02} = 4 \tag{9.70}$$

Table 9.2 The distribution of the number of edges of the Poisson polygon and the zero cell.

Number of edges	Poisson polygon	zero cell
3	$2 - \pi^2/6 \approx 0.3551$	0.0767
4	$\pi^2 \ln 2 - 1/3 - 7\pi^2/36$ $-(7/2)(1 + 2^{-3} + 3^{-3} + \cdots) \approx 0.3815$	0.3013
5	0.1873	0.3415
6	0.0596	0.1905
7	0.0129	0.0682
8	0.0023	0.0155
⋮	⋮	⋮

and

$$\lambda_0 = \frac{L_A^2 \zeta}{2}, \tag{9.71}$$

where

$$\zeta = \int_{(0,\pi]} \int_{(0,\pi]} |\sin(\alpha - \beta)| \mathscr{R}(\mathrm{d}\alpha) \mathscr{R}(\mathrm{d}\beta)$$

(see Matheron, 1975; Thomas, 1984). From this one may conclude (using the mean-value formulae in Section 9.3.4) that

$$\overline{n}_{20} = 4, \tag{9.72}$$

$$\lambda_1 = L_A^2 \zeta, \qquad \lambda_0 = \lambda_2 = \frac{L_A^2 \zeta}{2}, \tag{9.73}$$

$$\bar{l}_1 = \frac{1}{L_A \zeta}, \qquad \bar{l}_0 = \bar{l}_2 = \frac{4}{L_A \zeta}, \tag{9.74}$$

$$\overline{a}_2 = \frac{2}{L_A^2 \zeta}. \tag{9.75}$$

The equation $\lambda_0 = \lambda_2$ is true for general stationary bundle-free line tessellations; for this the Poisson assumption is not necessary (Stoyan and Mecke, 1983a, Chapter 8).

9.5.2 Poisson plane tessellations

Let Ψ now be a motion-invariant Poisson plane process in $\mathbb{R}^3$ with intensity S_V. Similarly to the planar case, the quantity

$$\varrho = P_L = \frac{S_V}{2}, \tag{9.76}$$

the mean number of planes intersected by a test line segment of unit length, is used as the model parameter. With probability one there is no point of $\mathbb{R}^3$ in which more than three Poisson planes intersect and so

$$\overline{n}_{01} = 6, \qquad \overline{n}_{02} = 12, \tag{9.77}$$

$$\overline{n}_{03} = 8, \qquad \overline{n}_{12} = 4. \tag{9.78}$$

It is known that

$$\overline{n}_{30} = 8, \qquad \overline{n}_{31} = 12, \tag{9.79}$$

$$\overline{n}_{32} = 6, \qquad \overline{n}_{21} = 4, \tag{9.80}$$

(see Miles, 1974a; Mecke, 1984c). These are the same values as for a regular cubical tessellation. Basic parameters can be used, as in Section 9.4, to give

$$\lambda_0 = \lambda_3 = \frac{\pi\varrho^3}{6}, \tag{9.81}$$

$$\lambda = \pi\varrho^3, \qquad L_V = \frac{\pi\varrho^2}{2}, \tag{9.82}$$

$$S_V = 2\varrho, \qquad T_V = \frac{4\pi\varrho^3}{3}, \tag{9.83}$$

$$Z_V = 2\pi\varrho^2, \qquad \bar{l}_1 = \frac{1}{\varrho}. \tag{9.84}$$

Some formulae are known for the typical polyhedron (the *Poisson polyhedron* in the sense of the 'number law' of Matheron, 1975); these are mainly due to Miles (1972b). Table 9.3 gives first and second moments. Miles (1972b) and Matheron (1975) give further formulae. Also formulae are known for the zero polyhedron; see Matheron (1975) and Schneider and Weil (2008).

Several inequalities for Poisson line and plane tessellations are given in Mecke (1995) and Mecke *et al.* (1990, Section 3.12). Typically the isotropic case provides extremal values.

The Poisson polyhedron finds useful employment as a model component, for example as a typical grain for Boolean models, as in Serra (1982, p. 499) and Example 3.3 in this book.

In conclusion it should be noted that the planar section of a Poisson plane tessellation is a Poisson line tessellation defined on the plane of intersection with the same parameter ϱ as the plane tessellation. This allows application of the formulae of Section 9.5.1.

Mecke (1984c), see also Schneider (1987), showed that for general stationary bundle-free hyperplane tessellations in $\mathbb{R}^d$ the intensities satisfy

$$\lambda_k = \binom{d}{k} \cdot \lambda_0 \qquad \text{for } k = 1, 2, \ldots, d. \tag{9.85}$$

Table 9.3 Poisson polyhedron formulae; see Santaló (1976).

	Volume	Surface	Average breadth	Facet number
First moments:	V	S	$\overline{b}$	$N = n_{32}$
	$6/(\pi\varrho^3) = \overline{V}_3$	$24/(\pi\varrho^2) = \overline{S}_3$	$3/(2\varrho) = \overline{b}_3$	$6 = \overline{n}_{32}$

Second moments:	V	S	$\overline{b}$	N
V	$48/\varrho^6$			
S	$96/\varrho^5$	$240/\varrho^4$		
$\overline{b}$	$2(\pi^2+3)/(\pi\varrho^4)$	$(7\pi^2+12)/(\pi\varrho^3)$	$(13\pi^2+48)/(48\varrho^2)$	
N	$4(\pi^2+3)/(\pi\varrho^3)$	$2(7\pi^2+24)/(\pi\varrho^2)$	$(13\pi^2+120)/(12\varrho)$	$(13\pi^2+336)/12$

Note that a hyperplane process in $\mathbb{R}^d$ is said to be *bundle-free*, if each point of $\mathbb{R}^d$ is contained in at most d hyperplanes of the process.

Furthermore, the mean number $\overline{n}_{kj}$ of j-faces adjacent to the typical k-face satisfies

$$\overline{n}_{kj} = 2^{k-j}\binom{k}{j} \qquad \text{for } k = 1, 2, \ldots, d \text{ and } j = 0, 1, \ldots, k. \tag{9.86}$$

These formulae are the same as for a cubical tessellation generated by hyperplanes parallel to the coordinate hyperplanes.

The construction of the Poisson polyhedron, the typical cell of the Poisson hyperplane tessellation, is described in Calka (2010, pp. 149–50). He generalises a procedure of Miles (1973), which was originally made for the planar case.

For some applications the *twin flat model* may be of value, which can be found in Serra (1982). It is constructed from a Poisson plane process by duplicating each plane by a parallel plane separated from it by a random distance.

9.6 STIT tessellations

A stationary STIT tessellation is uniquely determined by its rose of directions and the densities L_A (planar case) and S_V (spatial case). Thus these appear in the formulae for tessellation characteristics.

Formulae for the typical cell are not given explicitly here since the interior of the typical cell of a stationary STIT tessellation has the same distribution as the interior of the typical cell of a Poisson line or plane tessellation with the same value of L_A and S_V, respectively; see Nagel and Weiss (2003). However, for the mean number of l-faces on the boundary of a cell, it is necessary to draw on the distinction between cell faces (corners and sides in the planar case; apices, ridges and facets in the spatial case) and tessellation faces (vertices and edges in the planar case; vertices, edges and plates in the spatial case). For example, in the planar case a π-vertex belongs to three polygons, but is not an apex of one of them, and in the spatial case all edges in STIT are π-edges. Weiss and Cowan (2011) worked out the following mean values for the spatial case:

$$\psi = 2, \qquad \kappa = \frac{2}{3}, \qquad \xi = 1, \qquad \tau = \frac{4}{3}, \tag{9.87}$$

$$\overline{n}_{30} = 24, \qquad \overline{n}_{31} = 36, \qquad \overline{n}_{32} = 14, \qquad \overline{n}_{21} = \overline{n}_{20} = \frac{36}{7}, \tag{9.88}$$

$$\overline{m}_{30} = 8, \qquad \overline{m}_{31} = 12, \qquad \overline{m}_{32} = 6, \qquad \overline{m}_{21} = \overline{n}_{20} = 4. \tag{9.89}$$

Recall that $\overline{n}_{kl}$ refer to tessellation k- and l-faces whilst $\overline{m}_{kl}$ to cell k- and l-faces. They also report mean values for the adjacency between a cell face and a tessellation face, for example the typical cell facet comprises, on average, 7/3 tessellation plates, 10 edges and 26/3 vertices, whilst the typical cell ridge contains two edges and three vertices.

Some further characteristics for isotropic STIT tessellations are, in the planar case,

$$\overline{l_1} = \frac{\pi}{3L_A}, \tag{9.90}$$

$$\lambda_0 = \frac{2L_A^2}{\pi}, \qquad \lambda_1 = \frac{3L_A^2}{\pi}, \qquad \lambda_2 = \frac{L_A^2}{\pi}, \tag{9.91}$$

and, in the spatial case,

$$L_V = \frac{\pi S_V^2}{4}, \tag{9.92}$$

$$\bar{l}_1 = \frac{1}{S_V}, \qquad \bar{l}_2 = \frac{36}{7S_V}, \tag{9.93}$$

$$\overline{A}_2 = \frac{48}{7\pi S_V^2}, \tag{9.94}$$

$$\lambda_0 = \frac{\pi S_V^3}{8}, \qquad \lambda_1 = \frac{\pi S_V^3}{4}, \qquad \lambda_2 = \frac{7\pi S_V^3}{48}, \qquad \lambda_3 = \frac{\pi S_V^3}{48}. \tag{9.95}$$

These formulae are proved in Nagel and Weiss (2006, 2008), in which the anisotropic case is also treated; see also Thäle and Weiss (2010).

9.7 Poisson-Voronoi and Delaunay tessellations

9.7.1 General

Voronoi and Delaunay tessellations generated by homogeneous Poisson processes have been studied extensively, both mathematically and by simulation. This section presents a brief overview of known results for the planar and spatial cases. Poisson-Voronoi tessellations in general $\mathbb{R}^d$ are studied in Møller (1989, 1994).

With probability one the Poisson-Voronoi and Poisson-Delaunay tessellations are normal and face-to-face. Thus, cell faces coincide with tessellation faces and in the literature one speaks only about vertices, edges and cell facets.

To the typical vertex of a Poisson-Voronoi tessellation there belong $d+1$ nearest generating points of the same distance Δ. The distribution of Δ is given in Muche (1996a); see also Muche (2005). Mecke and Muche (1995) show that the distributional behaviour of the pattern of the generating points outside the ball of radius Δ centred at the typical vertex is analogous to that of a Poisson process. Furthermore, they prove that the distribution of any characteristic of the typical edge coincides with that of a randomly chosen edge adjacent to the typical vertex; see also Baumstark and Last (2007).

Błaszczyszyn and Schott (2003, 2005) consider the case of inhomogeneous Poisson processes with piecewise constant intensity function, which allows a decomposition approximation of the distributions by a mixture of distributions in the homogeneous case.

9.7.2 Planar Poisson-Voronoi tessellations (Poisson-Dirichlet tessellations)

A Poisson-Dirichlet tessellation is the Voronoi tessellation relative to a planar homogeneous Poisson process.

Denote the intensity of the generating point process by ϱ, which will serve as the model parameter. Then, following Meijering (1953) and Miles (1970), the mean-value parameters of Section 9.3.4 are given by

$$\lambda_0 = 2\varrho, \tag{9.96}$$

$$\lambda_1 = 3\varrho, \tag{9.97}$$

$$\lambda_2 = \varrho, \tag{9.98}$$

$$L_A = 2\sqrt{\varrho}. \tag{9.99}$$

All other mean values follow from these by the mean-value formulae on pp. 360 and 361. For example, $\overline{n}_{20} = 6$, so that the typical cell of the Poisson-Dirichlet tessellation has an average of six vertices. At the typical vertex the interior angle has a mean of $2\pi/3$ and variance $\pi^2/9 - 5/6$.

Gilbert (1962) calculated the second moment of the area of the typical polygon of the Poisson-Dirichlet tessellation, obtaining the value of $1.280\varrho^{-2}$. In principle all other characteristics of this random polygon can also be obtained by numerical integration. For example, Miles and Maillardet (1982) derived integral formulae for the probabilities p_n of the polygon having exactly n vertices. Table 9.4 gives the first values of p_n, based on Calka (2003a); see also Hayen and Quine (2000a,b). Hilhorst (2008) discussed the behaviour of the p_n for large n and Hilhorst (2007) suggested a special simulation method for their determination.

For many characteristics of the Poisson-Dirichlet tessellation it is difficult to derive formulae or known formulae are complicated integrals, and hence it is helpful to determine such characteristics by simulation. Hinde and Miles (1980) obtained a series of estimates by simulating the typical cell directly rather than by dealing with the whole tessellation. Some of these results are given in Table 9.5 together with variances and correlations for geometrical characteristics of the typical cell (see also Okabe *et al.*, 2000, Table 5.5.1).

Hinde and Miles gave histograms for the distributions of perimeter, area, and inner angles (see also Quine and Watson, 1984, and Icke and van de Weygaert, 1987) and suggested approximations for perimeter and area distribution by the so-called generalised gamma densities, using the form

$$f(x) = \frac{\gamma\chi^{\nu/\gamma}}{\Gamma(\nu/\gamma)}\, x^{\nu-1} \exp(-\chi x^{\gamma}) \qquad \text{for } x \geq 0, \tag{9.100}$$

Table 9.4 The probability p_n that the typical Poisson-Dirichlet polygon has n vertices.

n	3	4	5	6	7	8	9	...
p_n	0.0112	0.1068	0.2595	0.2947	0.1988	0.0897	0.0295	...

Table 9.5 Second moments of characteristics of the typical cell of the Poisson-Dirichlet tessellation.

	Area	Perimeter	Edge number = n_{21}
Variance:			
	A	l_2	N
	$0.2802\varrho^{-2}$	$0.9455\varrho^{-1}$	1.7808

Correlation coefficient:	A	l_2	N
A	1		
l_2	0.953	1	
N	0.568	0.5021	1

where γ, ν and χ are positive parameters. (Good choices for the area case, supposing $\varrho = 1$, are the values $\gamma = 1.08$, $\nu = 3.31$ and $\chi = 3.03$; see Hinde and Miles, 1980.)

Note in passing that the classical gamma distribution with integer shape parameter (the Erlang distribution) appears frequently in stochastic geometry in the context of Poisson processes as model elements. This is systematically investigated in various papers; see Miles (1971a), Møller and Zuyev (1996), Zuyev (1999), Cowan *et al.* (2003), Cowan (2006) and Baumstark and Last (2009). There measures of so-called stopping sets are often sums of i.i.d. exponentially distributed random variables.

Mean-value characteristics of the zero cell of the Poisson-Dirichlet tessellation can be derived using equation (9.32) and the numerical values in Table 9.5. The mean edge number is greater than 6 (= the mean edge number of the typical cell), namely 6.40.

The point process of vertices of the Poisson-Dirichlet tessellation has also been considered. Its intensity is denoted by λ_0 as above. A study of its Palm distribution leads to the determination of the pair correlation function $g(r)$. It can be given in the form of a sum of four numerically tractable double-parameter integrals; see Heinrich and Muche (2008). The function $g(r)$ has a pole of order 1 at $r = 0$, that is, $g(r) = O(r^{-1})$ as $r \to 0$, and followed by a minimum and a subsequent maximum; see Figure 9.12(a). The pole results from very short edges.

The length of the typical edge was obtained by numerical integration in Brakke (1986a) (see also Muche, 1996b, and Schlather, 2000, for other integral expressions).

The semi-empirical Aboav law (9.37) for the mean number of edges of a random neighbour of a n-edged cell does not hold for the Poisson-Voronoi tessellation. This is discussed in detail in Hilhorst (2006).

Muche and Stoyan (1992) considered distributional characteristics of the random closed set of the edge system E_Θ and derived integral formulae for the linear and spherical contact distribution functions $H_l(r)$ and $H_s(r)$. Muche (2010) obtained formulae for the probability density functions of the linear contact distribution for the planar, as well as for the spatial case, see Figure 9.13, and also found formulae for the probability density functions of the spherical contact distributions and corresponding moments.

By Equation (6.66), an expression of $H_l(r)$ leads to one for the chord length distribution function $L(r)$, that is, the distribution function of the length of the chord generated by

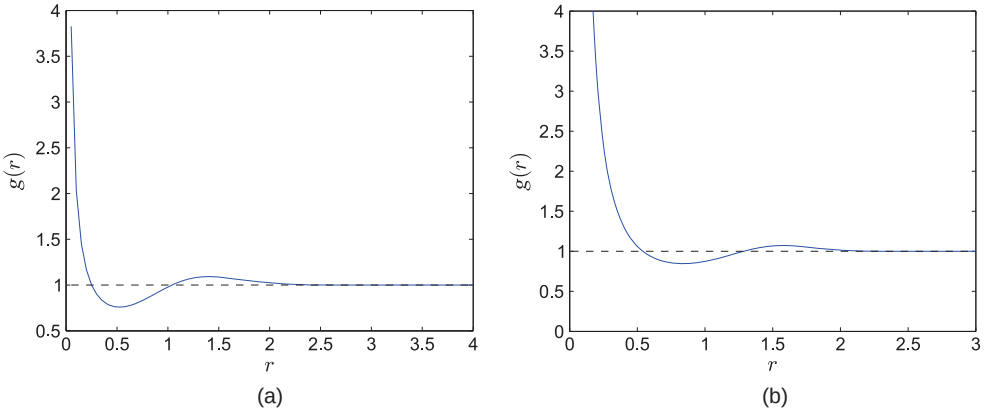

Figure 9.12 The pair correlation function of the point process of vertices of the Poisson-Voronoi tessellation in the (a) planar case and (b) spatial case, for $\varrho = 1$. Courtesy of L. Heinrich and L. Muche.

a random line with the typical Voronoi cell; see Figure 9.14. The corresponding first two moments are

$$\bar{l} = \frac{\pi}{4}\varrho^{-1/2} \qquad \text{and} \qquad \overline{l^2} \approx 0.806\varrho^{-1}. \tag{9.101}$$

Some more discussions will be found around (9.107) on p. 382.

Redenbach (2011) derived integral expressions for A_A and L_A of $E_\Theta \oplus B(0, r)$, the dilated set of the edge system of the Poisson-Voronoi tessellation.

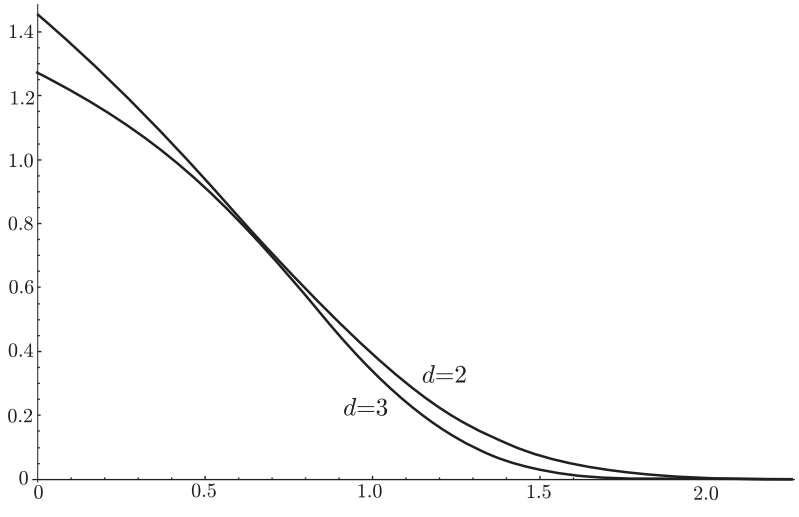

Figure 9.13 Probability densities of the linear contact distribution for the Poisson-Voronoi tessellation in $\mathbb{R}^2$ and $\mathbb{R}^3$ for $\varrho = 1$. Courtesy of L. Muche.

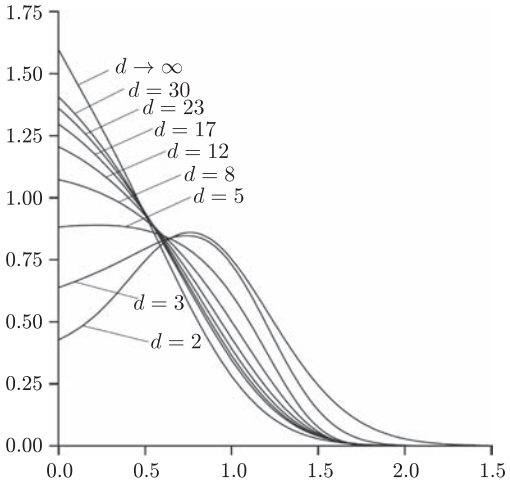

Figure 9.14 Probability density functions for the chord lengths through the typical Poisson-Voronoi polygon in $\mathbb{R}^d$ for $d = 2, 3, 5, 8, 12, 17, 23, 30, \infty$ with $\varrho = 1$. Reproduced from Muche (2010) with permission of the Applied Probability Trust.

9.7.3 Spatial Poisson-Voronoi tessellations

Here ϱ is the intensity of the underlying three-dimensional Poisson point process. The mean-value parameters of Section 9.4 take the values

$$\lambda_0 = \frac{24\pi^2\varrho}{35} \approx 6.768\varrho, \qquad \lambda_3 = \varrho, \qquad \lambda = 3\lambda_0 + \varrho, \tag{9.102}$$

$$L_V = \frac{16}{15}\Big(\frac{3}{4}\Big)^{1/3}\pi^{5/3}\Gamma\Big(\frac{4}{3}\Big)\varrho^{2/3} \approx 5.832\varrho^{2/3}, \tag{9.103}$$

$$S_V = 4\Big(\frac{\pi}{6}\Big)^{1/3}\Gamma\Big(\frac{5}{3}\Big)\varrho^{1/3} \approx 2.910\varrho^{1/3}, \tag{9.104}$$

$$T_V = 4\lambda_0, \qquad Z_V = 3L_V, \tag{9.105}$$

see Meijering (1953), Miles (1972b) and Møller (1989, 1994). All other mean values follow from these results. For example, the mean volume of the typical cell is given by

$$\overline{V}_3 = \varrho^{-1},$$

the mean surface area by

$$\overline{S}_3 = \left(\frac{256\pi}{3}\right)^{1/3}\Gamma\Big(\frac{5}{3}\Big)\,\varrho^{-2/3} \approx 5.821\varrho^{-2/3},$$

and the mean average breadth by

$$\overline{\overline{b_3}} = \frac{1}{5}\left(\frac{16\pi^5}{243}\right)^{1/3} \Gamma\left(\frac{1}{3}\right) \varrho^{-1/3} \approx 1.458\varrho^{-1/3}.$$

The mean total edge length and the means of numbers of apices (or vertices), ridges (or edges) and facets (or plates) of the typical cell are $17.496\ \varrho^{-1/3}$, 27.07, 40.61, and 15.54, respectively. For a table with numerical values and formulae of other means, see Okabe *et al.* (2000, pp. 316–7).

Gilbert (1962) uses numerical integration to estimate the second moment of the volume of the typical cell as $1.179\varrho^{-2}$. Muche and Ballani (2011) confirm this result (see also Brakke, 1986b) by analytical calculations, which leads to

$$\mathbf{E}(V^2) = \frac{1}{\varrho^2}\left(-\frac{4}{3^5}\left(16\mathscr{A} - 3\pi^2\right) + \frac{8}{27}\sqrt{3}\,\pi\right) \approx 1.179\varrho^{-2} \tag{9.106}$$

with the constant

$$\mathscr{A} \approx 3.4955.$$

Furthermore, they are able to give higher moments of the length of the typical edge of a k-dimensional section through a spatial Poisson-Voronoi tessellation.

The variances of surface area, total edge length and numbers of vertices, edges and facets of the typical cell have also been derived either by simulation or numerical integration, namely $2.19\varrho^{-4/3}$, $13.63\varrho^{-2/3}$, 43.98, 99.00, and 11.01, respectively.

Quine and Watson (1984) (see also Møller, 1994, Section 4.5) establish an important simulation method which produces i.i.d. realisations of the typical cell of the Poisson-Voronoi tessellation in $\mathbb{R}^d$ by means of the radial method, with an additional point at the origin o. For this point pattern the Voronoi tessellation is constructed, and the corresponding zero cell is a realisation of the typical cell in the Palm distribution sense. This allows a quick simulation of samples of typical cell realisations, which lead to precise estimates of distributional characteristics of the cells. Møller (1995) shows how to generalise this for Johnson–Mehl tessellations, and Lautensack (2007, Chapter 4) describes the procedure for Poisson-Laguerre tessellations.

The correlation coefficients in Table 9.6 are determined by simulation by Lorz and Hahn (1993). Their paper, as well as those by Quine and Watson (1984), van de Weygaert (1994) and Fleischer *et al.* (2009), also contains histograms of Voronoi cell (spatial as well as planar) characteristics. A gamma density function approximation such as (9.100) is possible; for the volume case good values are $\gamma = 1.41$, $\nu = 2.81$ and $\chi = 4.12$, for $\varrho = 1$, after Tanemura (1988); see also Andrade and Fortes (1988) and Kumar *et al.* (1992). A good survey can be found in Okabe *et al.* (2000, Section 5.5.4).

The mean-value characteristics of the zero cell of the Poisson-Voronoi tessellation can be derived using equation (9.32) and the numerical values in Table 9.6. The mean edge number is greater than 15.54 (the mean edge number of the typical cell), namely 16.58.

As mentioned in the previous section, Muche and Stoyan (1992) derive integral formulae for $H_l(r)$ and $H_s(r)$ of the system of cell facets; see Figure 9.13 for the former, which lead to

Table 9.6 Second moments of characteristics of the typical cell of the Poisson-Voronoi tessellation.

	Volume	Surface	Average breadth	Facet number
Variance:				
	V_3	S_3	$\overline{b}_3$	n_{32}
	$0.179\varrho^{-2}$	$2.19\varrho^{-4/3}$	$0.030\varrho^{-2/3}$	11.01

Correlation coefficients:	V_3	S_3	$\overline{b}_3$	n_{32}
V_3	1			
S_3	0.982	1		
$\overline{b}_3$	0.945	0.987	1	
n_{32}	0.736	0.711	0.671	1

the chord length distribution function $L(r)$, and the first two moments of it are

$$\bar{l} \approx 0.687\varrho^{-1/3} \qquad \text{and} \qquad \overline{l^2} \approx 0.631\varrho^{-2/3}. \tag{9.107}$$

Figure 9.14 shows the corresponding density functions for various dimensions d. The shape of these functions demonstrates what some statisticians have in mind when they say that the 'variability' of Poisson-Voronoi tessellations increases with increasing d. (The study of tessellations in $\mathbb{R}^d$ with $d > 3$ has practical applications: their intersections with three-dimensional subspaces are spatial tessellations; see Miles, 1986.)

Alishahi and Sharifitabar (2008) express the limits of $H_l(r)$ and $L(r)$ as $d \to \infty$ in terms of standard functions. The limiting value of the probability density function of $L(r)$ for $r = 0$ is given explicitly as 1.5968.

The probability density function of the spherical contact distribution is

$$h_s(r) = \frac{64\pi^2\varrho^2 r^5}{3} \int_0^1 \left(2t^2 \left(3 + t^2\right) \exp\left(-\frac{32\pi\varrho r^3}{3}(1 + t^2)\right) + \frac{(1+t)^3}{t^6} \exp\left(-\frac{4\pi\varrho r^3}{3}\frac{(1+t)^4}{t^3}\right)\right) \mathrm{d}t \qquad \text{for } r \geq 0, \tag{9.108}$$

see Muche (2010), in which formulae for the densities of the discoidal and bilinear contact distributions (i.e. the spherical contact distribution in a planar and a linear section, respectively) are also given; see Figure 9.15.

The fundamental ideas for obtaining these results are as follows. For a convex compact structuring element B containing the origin the contact distribution function $H_B(r)$ satisfies

$$H_B(r) = 1 - \mathbf{P}(rB \subset C_0) \qquad \text{for } r \geq 0,$$

where C_0 is the zero cell containing the origin. This cell is generated by the nearest neighbour of o in the Poisson point process of nuclei. The probability above is obtained by means of the void-probability of the nuclei process and use of geometrical properties of the cells.

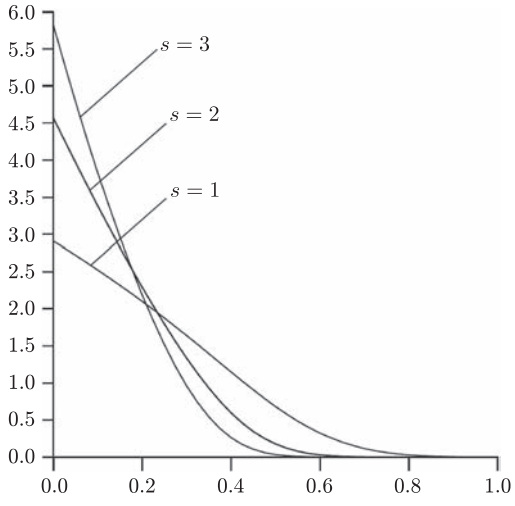

Figure 9.15 Probability density functions for the spherical contact distribution of an s-dimensional section of the spatial Poisson-Voronoi tessellation with $\varrho = 1$; $s = 3$, 2, and 1 correspond to the original spatial case, planar section and linear section, respectively. Reproduced from Muche (2010) with permission of the Applied Probability Trust.

Similar to the planar case, the distribution of the length of the typical edge is obtained by numerical integration in Brakke (1986b); see also Muche (1996b), Schlather (2000) and Muche (2005), who consider also higher dimensions. The corresponding density functions are similar but not equal to the density functions of Figure 9.14. Additionally, Muche (1996b, 1998, 2005) and Schlather (2000) give density functions for the angles between facets emanating in the same edge and between the typical edge and the line passing one of its endpoints and the nucleus of a neighbouring cell.

For the pair correlation function $g(r)$ of the vertices, the order of the pole of $g(r)$ is 2 (and in general is $d - 1$ in $\mathbb{R}^d$), followed by a minimum and then a maximum; see Figure 9.12(b). Thus the point process of vertices of the Poisson-Voronoi tessellation is not a suitable model for the galaxies in the universe since the pair correlation function for the latter follows the power law and has a pole of order around 1.8; see Snethlage *et al.* (2002) and Martínez *et al.* (2009).

Scheike (1994) gives mean-value formulae for scaled, that is, linearly transformed, Poisson-Voronoi tessellations. Such a tessellation can be interpreted as the result of a modified Voronoi construction with a metric which considers ellipses under the Euclidean metric as discs.

Intersections of spatial Poisson-Voronoi tessellations with planes are considered in Section 10.6.

9.7.4 Poisson-Delaunay tessellations

The dual graph of the Poisson-Voronoi tessellation, the *Poisson-Delaunay tessellation* has also been studied by many authors. With probability 1, it is face-to-face but not normal. Distributions and moments of characteristics of the typical Delaunay cell are more tractable than those of the typical Voronoi cell. For example, the k^{th} moments of the volume of the

typical cell of the Poisson-Delaunay tessellation in $\mathbb{R}^d$ are given by

$$\mathbf{E}(V_d^k) = \frac{\Gamma\left(\frac{d^2}{2}\right)\Gamma(d+k)\Gamma\left(\frac{d^2+dk+k+1}{2}\right)\Gamma^{d-k+1}\left(\frac{d+1}{2}\right)\prod_{i=2}^{d+1}\left(\Gamma\left(\frac{k+i}{2}\right)\Big/\,\Gamma\left(\frac{i}{2}\right)\right)}{\Gamma(d)\Gamma\left(\frac{d^2+1}{2}\right)\Gamma\left(\frac{d^2+dk}{2}\right)\Gamma^{d+1}\left(\frac{d+k+1}{2}\right)(2^d\pi^{(d-1)/2}\varrho)^k}, \tag{9.109}$$

for $k = 1, 2, \ldots$; see Miles (1974a) and Møller (1989, 1994). In particular, the mean area and volume are

$$\mathbf{E}(V_2) = \overline{a}_2 = \frac{1}{2\varrho}, \tag{9.110}$$

and

$$\mathbf{E}(V_3) = \overline{V}_3 = \frac{35}{24\pi^2\varrho}. \tag{9.111}$$

Size and shape of the Delaunay cells can also be characterised; see Miles (1974a) and D. G. Kendall (1983). In the case of a planar Poisson-Delaunay tessellation, each cell is a triangle, whose shape can be characterised by

$$\sqrt{\frac{4}{3}\left(\sin^2\vartheta_1 + \sin^2\vartheta_2 + \sin^2\vartheta_3\right)}, \tag{9.112}$$

where the ϑ_i are its interior angles (D.G. Kendall *et al.*, 1999, Section 8.3). The joint density of the interior angles was derived by Miles (1970):

$$f(\vartheta_1, \vartheta_2) = \frac{8}{3\pi}\big(\sin\vartheta_1 \sin\vartheta_2 \sin(\vartheta_1 + \vartheta_2)\big) \qquad \text{for } \vartheta_1, \vartheta_2 \geq 0 \text{ and } \vartheta_1 + \vartheta_2 < \pi, \tag{9.113}$$

which gives $\mathbf{E}(\vartheta_1) = \pi/3$. The maximum density at $\vartheta_1 = \vartheta_2 = \pi/3$ indicates that an equilateral triangle is the most likely shape. See Okabe *et al.* (2000, pp. 398–400) for the densities of the minimum, middle and maximum interior angles.

Rathie (1992) obtained a general expression for the distribution function of the volume of the typical Poisson-Delaunay cell in $\mathbb{R}^d$. The corresponding density function for the area $f_A(a)$ in the planar case is given by

$$f_A(a) = \frac{8}{9}\pi\varrho^2 a K_{1/6}^2\Big(\frac{2\pi\varrho a}{3\sqrt{3}}\Big) \qquad \text{for } a \geq 0, \tag{9.114}$$

where $K_{1/6}$ is the modified Bessel function of order $1/6$,

$$K_{1/6}(x) = \frac{\sqrt{\pi}\left(\frac{x}{2}\right)^{1/6}}{\Gamma(\frac{2}{3})}\int_1^\infty \exp(-xt)(t^2-1)^{-1/3}\,\mathrm{d}t.$$

However, for the spatial case Rathie's expression involves a sequence of complicated analytical functions. Muche (1996a) found numerically tractable forms for the area A in the planar case

and the volume V in the spatial case,

$$f_A(a) = \frac{\pi\varrho^2 a}{6}\int_0^{2\pi}\int_0^{2\pi-\theta_1} g(\theta_1,\theta_2)\exp\left(\frac{-\pi\varrho a g(\theta_1,\theta_2)}{2}\right)\mathrm{d}\theta_2\,\mathrm{d}\theta_1, \qquad \text{for } a \geq 0, \tag{9.115}$$

and

$$f_V(v) = \frac{35\varrho^2 v}{2}\int_0^{2\pi}\int_0^{2\pi-\theta_1}\int_0^{\pi} \sin\theta_3 \exp\left(\frac{-2\pi\varrho v g(\theta_1,\theta_2)}{(1+\cos\theta_3)\sin^2\theta_3}\right)\mathrm{d}\theta_3\,\mathrm{d}\theta_2\,\mathrm{d}\theta_1, \quad \text{for } v \geq 0, \tag{9.116}$$

in both of which

$$g(\theta_1,\theta_2) = \left(\sin\frac{\theta_1}{2}\sin\frac{\theta_2}{2}\sin\frac{\theta_1+\theta_2}{2}\right)^{-1}.$$

Collins (1968) (see also Sibson, 1980) gave the probability density function $f_{l_1}(r)$ of the length l_1 of the typical Poisson-Delaunay edge (equal to the distance between two 'neighbouring' points of the generating Poisson process) for the planar case, and Muche (1998) (see also Muche, 1996a) derived it for the spatial case:

$$f_{l_1}(r) = \begin{cases} \dfrac{\varrho\pi r}{3}\left(\sqrt{\varrho}r\exp\left(-\dfrac{\varrho\pi r^2}{4}\right) + \operatorname{erfc}\left(\dfrac{\sqrt{\varrho\pi}r}{2}\right)\right), & \text{for } d=2, \\ \dfrac{35\pi\varrho r^2}{2^8}\left(\left(\dfrac{\pi\varrho r^3}{4}+\dfrac{5}{2}\right)\exp\left(-\dfrac{\pi\varrho r^3}{6}\right)\right. & \\ \qquad \left. -\dfrac{\pi\varrho r^2}{3}\displaystyle\int_{r/2}^{\infty}\left(1+\frac{\pi\varrho r^2 t}{6}\right)\exp\left(-\frac{4\pi\varrho t^3}{3}\right)\mathrm{d}t\right), & \text{for } d=3, \end{cases} \tag{9.117}$$

for $r \geq 0$, where $\operatorname{erfc}(x)$ is the complementary error function, which is related to the standard Gaussian distribution function $\Phi(x)$ by

$$2\Phi(x\sqrt{2}) + \operatorname{erfc}(x) = 2. \tag{9.118}$$

Consequently

$$\mathbf{E}(l_1) = \bar{l}_1 = \begin{cases} \dfrac{32}{9\pi\sqrt{\varrho}} \approx 1.132\varrho^{-1/2}, & \text{for } d=2, \\ \dfrac{1715}{2304}\left(\dfrac{3}{4\pi\varrho}\right)^{1/3}\Gamma\left(\dfrac{1}{3}\right) \approx 1.237\varrho^{-1/3}, & \text{for } d=3. \end{cases} \tag{9.119}$$

The distributions of some other characteristics, such as the perimeter in the planar case and the surface area and the average breadth in the spatial case, are also known; see Okabe *et al.* (2000, Section 5.11) for more details. Some important mean values are:

for $d = 2$:

$$\bar{l}_2 = \frac{32}{3\pi\sqrt{\varrho}} \approx 3.395\varrho^{-1/2}, \tag{9.120}$$

$$\overline{n}_{02} = \overline{n}_{01} = 6; \tag{9.121}$$

for $d = 3$:

$$\overline{S}_3 = \frac{3500(3/4)^{2/3}\Gamma(2/3)}{243\pi^{5/3}\varrho^{2/3}} \approx 2.389\varrho^{-2/3}, \tag{9.122}$$

$$\overline{\overline{b}_3} \approx 1.118\varrho^{-1/3}, \tag{9.123}$$

$$\overline{n}_{03} = \frac{96\pi^2}{35} \approx 27.071, \quad \overline{n}_{02} = \frac{144\pi^2}{35} \approx 40.606, \quad \overline{n}_{01} = \frac{48\pi^2}{35} + 2 \approx 15.536, \tag{9.124}$$

$$\overline{n}_{12} = \overline{n}_{32} = \frac{144\pi^2}{24\pi^2 + 35} \approx 5.228. \tag{9.125}$$

9.8 Laguerre tessellations

Lautensack (2007) and Lautensack and Zuyev (2008) show that for the rather complicated Laguerre tessellations numerical results can be obtained not only by simulation. For the case of the Poisson-Laguerre tessellation they derived formulae for mean-value characteristics, which are, though containing complicated multiple integrals, useful in numerical calculations. A Poisson-Laguerre tessellation in $\mathbb{R}^d$ is the Laguerre tessellation with respect to an independently marked Poisson process in $\mathbb{R}^d$. The intensity of the Poisson process is denoted by ϱ and the distribution function of the (positive) marks is denoted by $F(w)$. The resulting tessellation is normal and face-to-face. The following presents some of the results from Lautensack and Zuyev (2008) for the planar case, just to show the reader some formulae and properties of planar Laguerre tessellations.

The characteristics of interest are λ_0, the mean number of vertices per area unit, and L_A, the mean edge length per area unit. It holds

$$\lambda_0 = \frac{\varrho^3}{12} \int_0^\infty \int_0^\infty \int_0^\infty \int_{-\min_i w_i^2}^\infty \exp\left(-\varrho\pi \int_0^\infty [t + w^2]^+ \mathrm{d}F(w)\right)$$

$$\times\, V_{2,0}\left(\sqrt{t + w_0^2}, \sqrt{t + w_1^2}, \sqrt{t + w_2^2}\right) \mathrm{d}t\, \mathrm{d}F(w_0)\, \mathrm{d}F(w_1)\, \mathrm{d}F(w_2), \tag{9.126}$$

where

$$[x]^+ = \max\{0, x\},$$

$$V_{2,0}(w_0, w_1, w_2) = 2\int_{S^1}\cdots\int_{S^1}\Delta(w_0u_0, \ldots, w_2v_2)\sigma(\mathrm{d}u_0)\cdots\sigma(\mathrm{d}u_2),$$

in which $\Delta(x_1, x_2, x_3)$ is the area of the triangle formed by the points x_1, x_2 and x_3, and σ is the length measure on the unit circle S^1. Furthermore,

$$L_A = \varrho^2\pi\int_0^\infty\int_0^\infty\int_{-\min_i w_i^2}^\infty \frac{2t + w_0^2 + w_1^2}{\sqrt{(t + w_0^2)(t + w_1^2)}}\int_0^\infty \exp\left(-\varrho\pi\int_0^\infty [t + s + w^2]^+\mathrm{d}F(w)\right)$$

$$\times\sqrt{s}\,\mathrm{d}s\,\mathrm{d}t\,\mathrm{d}F(w_0)\,\mathrm{d}F(w_1). \tag{9.127}$$

These two formulae provide the parameters that are required for determining all mean values of the tessellation characteristics using Mecke's mean-value formulae in Section 9.3.4. (Only two are sufficient because of the face-to-face and normality property of Poisson-Laguerre tessellations.) In particular, the probability p_0 that the typical point of the underlying Poisson process of intensity ϱ generates a non-empty cell is

$$p_0 = \frac{\lambda_0}{\varrho}. \tag{9.128}$$

The spherical and linear contact distributions and the chord length distribution have also been derived, which leads to the mean chord length:

$$\bar{l} = \frac{\pi}{2L_A} \tag{9.129}$$

(Lautensack, 2007, p. 63).

Example 9.1. *A planar Poisson-Laguerre tessellation with a two-atom mark distribution (Lautensack and Zuyev, 2008)*
The intensity of the Poisson process is $\varrho = 100$, the mark distribution is a two-atom distribution with the atoms $a = 0.01$ and $b = 0.01$, 0.05, 0.10, 0.15, 0.20, 0.25 and 0.30, with probability 0.5. The case $a = b$ corresponds to the Poisson-Voronoi tessellation.

The numerical values are summarised in Table 9.7. For comparison, the values for Poisson-Voronoi tessellations with intensity $\varrho = 100$ (this is the case of $a = b = 0.01$) and $\varrho = 50$ are included.

The values for $b = 0.30$ are close to those for the Poisson-Voronoi tessellation with $\varrho = 50$. This observation illustrates a limit theorem for Poisson-Laguerre tessellations: when there is an atom at the maximum weight b and when b is large, nearly all of the cells generated by the points with smaller weights will become empty and vanish, leaving the Voronoi tessellation generated by the points with the larger weight b.

Table 9.7 Mean values of cell characteristics for a two-dimensional Laguerre tessellation generated by a homogeneous Poisson process of intensity $\varrho = 100$ with weights independently drawn from a two-atom distribution taking the values 0.01 and b with probability 0.5 each. The columns PV_{100} and PV_{50} contain the values for Poisson-Voronoi tessellations of intensity $\lambda = 100$ and $\lambda = 50$, respectively. Courtesy of C. Redenbach and S. Zuyev.

		b						
	PV_{100}	0.05	0.10	0.15	0.20	0.25	0.30	PV_{50}
λ_0	200.000	192.406	148.398	110.968	101.050	100.043	100.001	100.000
λ_1	300.000	288.609	222.597	166.452	151.574	150.065	150.001	150.000
λ_2	100.000	96.203	74.199	55.484	50.525	50.022	50.001	50.000
L_A	20.000	19.203	16.283	14.529	14.173	14.143	14.142	14.142
$\overline{l_1}$	0.06667	0.06654	0.07315	0.08729	0.09351	0.09425	0.09428	0.09428
$\overline{a_2}$	0.0100	0.0104	0.0135	0.0180	0.0198	0.0200	0.0200	0.0200
$\overline{l_2}$	0.4000	0.3992	0.4389	0.5237	0.5610	0.5655	0.5657	0.5657

9.9 Johnson–Mehl tessellations

Johnson–Mehl tessellations, which are defined on p. 352, are more complicated than Voronoi tessellations, depend on more model parameters and do not necessarily have convex cells. General formulae are correspondingly more intricate. This section considers the birth locations and birth times of nuclei form a Poisson process in $\mathbb{R}^d \times [0, \infty)$ with time-dependent intensity function $\lambda(t)$, meaning that the births are homogeneous in space but not necessarily in time. The mean number of births per unit volume in the time interval $[s, t]$ is $\int_s^t \lambda(u)\mathrm{d}u$, but this is not the mean number of cells because some nuclei are born in occupied regions and then disappear without a trace.

An important characteristic is the volume fraction $V_V(t)$ of the space occupied by the Johnson–Mehl cells at time t. In the particular case of constant growth rate and time-homogeneous births

$$v(t) = v \qquad \text{and} \qquad \lambda(t) = \alpha,$$

it is given by

$$V_V(t) = 1 - \exp\left(-\frac{b_d \alpha v^d t^{d+1}}{d+1}\right) \qquad \text{for } t \geq 0. \tag{9.130}$$

A frequently considered time-inhomogeneous case is the one corresponding to the Weibull model of the underlying birth-and-growth process in Section 6.6.4:

$$\lambda(t) = ct^{m-1}, \qquad \text{where } c, m > 0,$$

considered in Horálek (1988, 1990) and Møller (1992).

In the general case with time-dependent birth rate $\lambda(t)$ and growth rate $v(t)$ the quantity $V_V(t)$ is given by

$$V_V(t) = 1 - \exp\left(-b_d \int_0^t \lambda(s) \left(\int_0^{t-s} v(u)\mathrm{d}u\right)^d \mathrm{d}s\right) \qquad \text{for } t \geq 0, \tag{9.131}$$

which differs from Formula (6.188) because, as explained on p. 353, in the Johnson–Mehl model the growth rate of a cell depends on the age of the cell (i.e. for a cell born at time s its growth rate at time t is equal to $v(t-s)$, see the age-dependent growth model (6.175)), whilst in Formula (6.188) the growth rate is a function of the time t (i.e. at time t the growth rate of all existing cells is $v(t)$).

These results can be obtained by means of the formulae for the Boolean model and were first given by Kolmogorov (1937), Johnson and Mehl (1939) and Avrami (1939). Formulae (9.130) and (9.131) are used for estimating $\lambda(t)$ statistically when $V_V(t)$ and $v(t)$ are observable.

The mean number of cells per volume unit of the resultant tessellation is

$$\lambda_d = \int_0^\infty V_V(t)\lambda(t)\,\mathrm{d}t. \tag{9.132}$$

Clearly, the mean volume of the typical cell is λ_d^{-1}.

Further mean-value formulae for fundamental characteristics of the cells and intersections are given in Møller (1992), who distinguishes between faces and *interfaces*. By faces he means the intersections of cells, the same as the tessellation faces introduced on p. 367, so that a k-face is the intersection of $d-k+1$ cells. However, in contrast to the convex case, a face may consist of several (connected) components, which Møller calls *k-interfaces*. A 0-interface is a vertex, while a $(d-1)$-interface is a facet between two cells.

Let λ_k be the intensity of k-interfaces, and let $\overline{n}_{kl}$ be the same mean as on p. 368 but with 'interface' instead of 'face'. Then in the particular case of $d = 2$ some of Møller's formulae are

$$\lambda_0 = 2\lambda_2, \qquad \lambda_1 = 3\lambda_2 \tag{9.133}$$

(generalising Formulae (9.96)–(9.98) from the Poisson-Dirichlet tessellation to this non-convex case), and

$$\overline{n}_{20} = \overline{n}_{21} = 6 \tag{9.134}$$

(again the same as in the Poisson-Dirichlet tessellation case).

The edges of planar Johnson–Mehl tessellations are hyperbolic arcs; this can be used for constructing such tessellations, as described in Møller (1995).

In the spatial case the situation may be more complicated. As Møller (1992) explained, it is possible there that 'a cell C_1 is surrounded by only two other cells C_2 and C_3. Then C_1 contains two facets and one edge but no vertices ... the edge is a closed curve and the 2-interface $C_2 \cap C_3$ is not simply connected, since C_1 causes a "gap" in the relative interior of $C_2 \cap C_3$. Such gaps represent the possibility of forming lenses'. This implies deviations from those mean-value relations which apply for tessellations with convex cells in the ordinary

equilibrium state. While in the convex ordinary equilibrium case

$$\lambda_2 = \lambda_0 + \lambda_3 \qquad \text{and} \qquad \lambda_1 = 2\lambda_0,$$

for a Johnson–Mehl tessellation the corresponding values satisfy

$$\lambda_2 \geq \lambda_0 + \lambda_3 \qquad \text{and} \qquad \lambda_1 \geq 2\lambda_0,$$

with equality holds only in the Voronoi tessellation case. The fragments are pieces of rotation hyperboloids. Møller (1995) showed how typical Johnson–Mehl cells can be simulated. Heinrich and Schüle (1995) discussed the simulation of the typical cell for some non-Poisson cases.

Mahin *et al.* (1980) studied planar sections of Johnson–Mehl tessellations by simulation.

Several asymptotic results for this tessellation observed in $[0, L]^d$, as $L \to \infty$, are mathematically tractable. Chiu (1995c) and Chiu and Yin (2000) derived the limiting distribution of the time of complete tessellation, which is the process duration for underlying birth-and-growth processes (see Section 6.6.4). Chiu (1997), Chiu and Quine (1997, 2001) and Chiu and Lee (2002) established central limit theorems for the total number of cells.

9.10 Statistics for stationary tessellations

9.10.1 Reconstruction

Frequently, the first step of statistics for tessellation data is their (re)construction. Empirical tessellations are often given only in a rough form, without clearly visible cells. It is a demanding task to transform such data to real tessellations. There are two established methods, which are applied to planar and spatial tessellations.

The first uses a chain of procedures of image processing and finally yields a unique tessellation. Typical references are Schwertel and Stamm (1997), Coster *et al.* (2005), Dillard *et al.* (2005), Lautensack and Sych (2006), Lautensack (2008) and Redenbach (2009b). The given image is binarised, then the Euclidean distance transform (see Ohser and Schladitz, 2009, p. 114ff.) is applied to yield for each pixel its distance to the system of cell boundaries, which leads to cell centres at the local maxima. Finally, the watershed algorithm divides the inverted distance image into cells; see Figure 9.16. As Lautensack (2008) reported, some manual polishing of the results obtained may be necessary.

The second method uses ideas of Bayesian statistics and MCMC algorithms. Typical references are Blackwell and Møller (2003), Skare *et al.* (2007) and Møller and Stoyan (2014), where one starts with patterns of points scattered around the invisible cell edges. The prior is a deformed Voronoi tessellation obtained by random perturbations of the vertices, the nuclei and the unknown parameters, whilst the posterior is product of the likelihood of the data modelled by the deformed tessellation and the prior. The MCMC output consists of many random reconstructions, which give the user some impression of the uncertainty of the result.

9.10.2 Summary characteristics

Given a sample of a tessellation, corresponding summary characteristics can be estimated. The methods for point processes, fibre and surface processes and random sets can be applied to the various related structures.

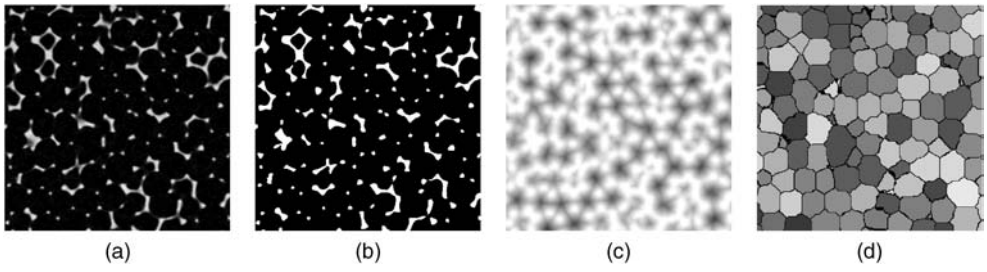

Figure 9.16 Reconstruction of a tessellation: (a) a section of the original three-dimensional image; (b) binarised image; (c) inverted distance image; (d) reconstructed cells. Reproduced from Redenbach (2009b) with permission of Società Editrice Esculapio Srl.

For example, the system of vertices is a point process and its intensity λ_0 can be easily estimated by the standard intensity estimator. The estimation of the intensity of the point process of cell centroids is more complicated; see the discussion below for the planar case. There the tessellations are viewed as particular germ–grain models.

The parameters L_A and L_V belong to the fibre process of edges and can be estimated by the methods described in Sections 8.3 and 8.4. Finally, the linear and spherical contact distributions $H_l(r)$ and $H_s(r)$ for the system of edges (planar case) and plates (spatial case) can be estimated by the methods of Section 6.4.5.

9.10.3 Statistics for planar tessellations

Some statistics for a tessellation are possible by considering the tessellation as a germ–grain model and applying the random-set methods of Chapter 6. This leads to information on the cells. Tessellation-specific methods start with the estimation of the four fundamental parameters: edge density L_A, intensity of vertices λ_0, the proportion ϕ of π-vertices, and mean number of cells per unit area (intensity of cell centroids) λ_2.

The characteristic L_A can be estimated by applying the methods for planar fibre processes as in Section 8.3, thus direct length measurement can be made or intersection methods may be applied. In the case of isotropy one can use Formula (8.38) in Section 8.3.2:

$$L_A = \frac{\pi}{2} P_L, \tag{9.135}$$

where P_L is the intensity of the intersection point processes on the test lines.

Estimation of the mean number λ_0 of vertices per unit area is a simple task: one has simply to estimate the intensity of a planar stationary point process, that of the vertices. Given a window of observation W of area $A(W)$ an unbiased estimator is clearly

$$\hat{\lambda}_0 = \frac{\#\{\text{vertices observed in } W\}}{A(W)}. \tag{9.136}$$

The proportion ϕ of π-vertices can be estimated analogously.

Unbiased estimation of λ_2, the mean number of cells per unit area, is a bit more complicated because of the influence of edge-effects. In the case of a trivalent tessellation, that is, $n_{02} \equiv 3$,

in each vertex precisely three edges emanate, Formula (9.26) can be used to obtain

$$\hat{\lambda}_2^{(\mathrm{tri})} = \frac{\hat{\lambda}_0}{2}, \tag{9.137}$$

and so the problem is reduced to estimation of λ_0.

If trivalence is not given, then in the case of isotropy,

$$\hat{\lambda}_2^{(\mathrm{iso})} = \frac{N(W) - 1 - \frac{1}{2}N_{\mathrm{e}}(W)}{A(W)} \tag{9.138}$$

for a convex window of observation W is an unbiased estimator. Here $N(W)$ is the number of cells intersecting the window W while $N_{\mathrm{e}}(W)$ is the number of edges intersecting the boundary of W.

Proof of unbiasedness of $\hat{\lambda}_2^{(\mathrm{iso})}$. Formula (6.123) yields

$$\lambda_2 A(W) = \mathbf{E}\big(N(W)\big) - 1 - \frac{L_A^*\, L_2(W)}{2\pi},$$

where L_A^* is the boundary length density of the germ–grain model formed by the cells. Clearly, $A_A = 1$ and $L_A^* = 2L_A$. Formula (9.135) can be used to replace L_A by $(\pi/2)P_L$, which is equal to $(\pi/2)\mathbf{E}\big(N_{\mathrm{e}}(W)\big)/L_2(W)$. This yields

$$\lambda_2 A(W) = \mathbf{E}\big(N(W)\big) - 1 - \frac{\mathbf{E}\big(N_{\mathrm{e}}(W)\big)}{2}. \qquad \square$$

If W is a rectangle and if the cells are so small that each intersecting cell produces at most two intersections of edges with the boundary of W then Formula (9.138) gives a third estimator, which is suggested by Saltykov (1974):

$$\hat{\lambda}_2^{(\mathrm{rect})} = \frac{z(W) + \frac{1}{2}w(W) + \frac{1}{4}u(W)}{A(W)}, \tag{9.139}$$

where

$$\begin{aligned} z(W) &= \text{number of cells completely in } W, \\ w(W) &= \text{number of cells that intersect the sides of } W \text{ but } \textit{not} \text{ its corners}, \\ u(W) &= \text{number of cells containing the corners of } W (= 4). \end{aligned}$$

In the trivalent case all three estimators $\hat{\lambda}_2^{(\mathrm{tri})}$, $\hat{\lambda}_2^{(\mathrm{iso})}$ and $\hat{\lambda}_2^{(\mathrm{rect})}$ yield the same result, and no isotropy assumption is necessary; see Hahn (1995).

Because of Formula (9.15) estimates of λ_2 also lead to estimates of mean cell area $\bar{a}_2$. Unbiased direct estimation of $\bar{a}_2$ is via the methods in Section 6.4. However, the following estimator $\hat{\bar{a}}_2$ for the mean cell area is biased, known to yield estimates too large:

$$\hat{\bar{a}}_2 = \frac{\sum\{\text{cell areas}\}}{\#\{\text{cells}\}}, \tag{9.140}$$

where the cells are a randomly chosen cell and the cells in some shells around; see Voloshin *et al.* (2010). Nevertheless, 'shell map analysis' as described by Aste *et al.* (1996) and Aste (1999) is a valuable tool in tessellation statistics.

Furthermore, it is a natural idea in the statistical analysis of an irregular system of hard objects, in particular balls, to construct the corresponding Voronoi, Voronoi S or Laguerre tessellation and to analyse these tessellations statistically. This idea is systematically treated in Medvedev (2000) and Alinchenko *et al.* (2004), under the name Voronoi–Delaunay method. Of particular interest are analyses of:

- Delaunay tetrahedra (shape and size);
- void geometry, using interstitial balls centred at the tessellation vertices;
- the whole network of edges of the tessellation.

If three-dimensional tessellation data are given, then at least in principle one can apply methods similar to those discussed above in the planar case. Until now such tessellations have been analysed mainly by means of particular models, as described in the next section.

9.10.4 Statistics for Voronoi, Laguerre and Johnson–Mehl tessellations

Planar Poisson-Voronoi tessellations

It would be convenient if one could conclude that each tessellation in the ordinary equilibrium state is interpretable as a Voronoi tessellation derivable from some point pattern. Unfortunately a simple degrees-of-freedom argument shows that this cannot be the case. So the Poisson-Voronoi tessellation is not more than a kind of 'benchmark'; the mean-value quantities extracted there provide values which (i) may serve as rough approximative estimates for the tessellation of interest or (ii) may give some impression about the deviations from a Poisson-Voronoi tessellation.

If it is known that a given tessellation is really a sample of a Poisson-Voronoi tessellation, then there are at least four possible procedures for estimation of the single model parameter ϱ, the mean number of cells per unit area in the planar case, based on determination of estimates of:

- λ_0 (intensity of vertices),
- λ_1 (intensity of edges),
- λ_2 (intensity of cells), and
- L_A (intensity of the fibre process of edges).

In the first and last case no edge-correction is necessary, while in the second and third case the ideas of Section 6.5.7 can be used, where the 'grains' are either the edges or the cells, respectively. The estimation variances for all four intensities above are smaller than the variance of estimation of the intensity of a Poisson process with parameter ϱ; see Hahn (1995). However, with increasing window size the difference vanishes. (The variances behave analogously if a Voronoi tessellation relative to another stationary point process is considered.)

If $\overline{n}_{02} = 3$ then the question may arise as to whether the analysed tessellation is Poisson-Dirichlet. One could first ask whether the tessellation is one that could be a Dirichlet

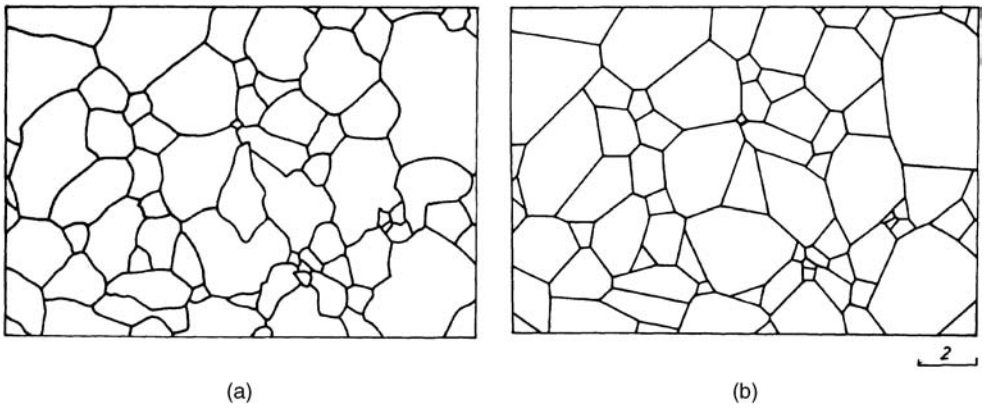

Figure 9.17 Planar section through a specimen of steel. (a) The lines correspond to smoothed boundaries of Austenite grains. (b) A tessellation with convex cells of the same topological form.

tessellation for any point process; if so then one might attempt to reconstruct the points of the generating pattern and test whether this pattern could be considered to be a Poisson process pattern. A quicker approach, and one that at any rate assesses the tessellation against the Poisson-Voronoi tessellation *qua* benchmark, uses estimates λ_0, λ_2 and L_A, and looks how these values satisfy the Poisson-Voronoi tessellation relationships

$$\lambda_0 = 2\lambda_2 \qquad \text{and} \qquad L_A = 2\sqrt{\lambda_2}.$$

Example 9.2. *Austenite grain boundaries in steel*
Figure 9.17 shows a planar tessellation which results from a naïve 'linearisation' of an image obtained by planar section through a sample of steel. (A better analysis of such tessellations should start with a reconstruction of the tessellation by the methods described in Section 9.10.1.) The edges correspond to the Austenite grain boundaries. From the tessellation the following values are obtained:

$$\text{total edge length} = 172.2, \quad \text{number of vertices} = 124, \quad N(W) = 78, \quad N_e(W) = 30.$$

These give

$$\hat{L}_A = 0.96 \qquad \text{and} \qquad \hat{\lambda}_0 = 0.69.$$

Since $\overline{n}_{02} = 3$, the mean number of cells per unit area λ_2 can be estimated in different ways. Formula (9.137) yields $\hat{\lambda}_2 = 0.34$, while (9.138) leads to $\hat{\lambda}_2 = 0.35$. Formula (9.139) is inappropriate here, as there are two cells hitting two sides of W each.

These estimates differ considerably from those expected for a Poisson-Dirichlet tessellation: from L_A the estimate $\hat{\lambda}_2 = 0.23$ is derived and this is considerably smaller than the direct estimates. But of course the sample is small, arguably too small for model tests to be realistic.

Note that it is methodologically doubtful to use a Poisson-Dirichlet tessellation in this application: probably there is no spatial tessellation of which the planar section is a

Poisson-Dirichlet tessellation. It is known that the planar section of a spatial Poisson-Voronoi tessellation is not a Dirichlet tessellation with respect to any point process (Chiu *et al.*, 1996); see also Section 10.6.

Further discussion of this example can be found in Example 10.4 on p. 442.

In the spatial case, stereological formulae given in Section 10.2 can be used.

Spatial Laguerre tessellations

As mentioned on p. 370, each normal and face-to-face spatial tessellation can be considered as a Laguerre tessellation with some system of points and weights. Many planar tessellations may also be well approximated by Laguerre tessellations. However, a marked Poisson process will produce Laguerre cells that are much more irregular in shape than those found in most empirical tessellations in the nature. It is more reasonable to consider Laguerre tessellations relative to centres of hard (non-overlapping) balls marked by the respective radii; see Figure 9.6 and the book cover. The balls are contained in their cells, which are more regular in shape. This approach is well demonstrated by Lautensack (2008), Lautensack *et al.* (2008) and Redenbach (2009a,b) for open and closed polymer, ceramic and metallic foams. Redenbach *et al.* (2012) consider modelling structures of strongly varying cell sizes.

The fitting involves two steps. The first step explores the type of the hard ball packing and that of the ball diameter distribution. Packings resulting from RSA or force-biased algorithms are successful examples. The ball volume distribution is heuristically assumed to be of the same type as that of the cell volumes, which is estimated from the given sample. Typical distributions for this purpose are gamma and lognormal. Their parameters can be estimated by classical methods of statistics.

In the second step the parameters of the ball volume distribution and the volume fraction of the corresponding hard ball packing are chosen to minimise the contrast d:

$$d = \sqrt{\sum_{i=1}^{8} \left(\frac{\hat{c}_i - c_i}{\hat{c}_i} \right)^2}, \tag{9.141}$$

with

$$\begin{aligned} c_1\ (c_2) &= \text{mean (variance) of cell volume,} \\ c_3\ (c_4) &= \text{mean (variance) of cell surface,} \\ c_5\ (c_6) &= \text{mean (variance) of number of facets per cell,} \\ c_7\ (c_8) &= \text{mean (variance) of average cell breadth.} \end{aligned}$$

The c_i are model characteristics under the respective parameters, obtained by simulation, while the $\hat{c}_i$ are the empirical counterparts. The Nelder–Mead algorithm may be used in this optimisation problem with simulated data. Instead of L^2-norm, it is also possible to use L^1- or L^∞-norm in (9.141).

Example 9.3. *Cell boundaries in thermal insulation materials (Lautensack, 2008)*
A three-dimensional grey value image of the cell boundaries in a material for the thermal insulation of buildings was taken. Then, the image processing procedures mentioned in

Section 9.10.1 were applied to yield the tessellation structure, from which the empirical model characteristics were

$$\hat{c}_1 = 0.0055532, \quad \hat{c}_3 = 0.17586, \quad \hat{c}_5 = 14.637, \quad \hat{c}_7 = 0.24698,$$
$$\hat{c}_2 = 0.0000065422, \quad \hat{c}_4 = 0.00280, \quad \hat{c}_6 = 10.378, \quad \hat{c}_8 = 0.00143.$$

Since $\hat{c}_5 = 14.637$, which is smaller than that (15.535) of the typical Poisson-Voronoi cell, but greater than that (13.0–14.2) in the typical Laguerre cell generated by dense packing of hard balls, systems of hard balls generated by RSA were considered.

By comparing the empirical cell volume distribution with the lognormal and gamma distributions, with parameters estimated by the maximum likelihood, the gamma distribution was chosen for the ball volumes. Finally, the volume fraction of the RSA and the parameters of the gamma distribution were chosen, by using large scale simulation to obtain the c_i, to minimise d in (9.141). The fitted model gave

$$c_1 = 0.0055532, \quad c_3 = 0.17507, \quad c_5 = 15.111, \quad c_7 = 0.24920,$$
$$c_2 = 0.0000072912, \quad c_4 = 0.00251, \quad c_6 = 11.079, \quad c_8 = 0.00115.$$

With the exception of the estimate for c_6, all these values are closer to the empirical $\hat{c}_i$ than the corresponding values of the Poisson-Voronoi tessellation of the same intensity.

Example 9.4. *Polyurethane foam (Lautensack et al., 2008)*
A three-dimensional grey value image of polyurethane foam was studied. In the reconstructed tessellations, the mean number of facets per cell was $\hat{c}_5 = 13.722$, suggesting that, unlike the sample in Example 9.3, a Laguerre tessellation relative to a system of random dense packing of hard balls could be tried. However, 2.9% of the cells were very small and irregular in shape, meaning that the cell sizes came from a mixture of two distributions. Thus, a system of two classes of balls, one having gamma distributed sizes and one having fixed small size, was used. First the force-biased algorithm was tried for both classes of balls, but it turned out that in the corresponding Laguerre tessellations the small cells were too regularly distributed. A better result was obtained by generating a dense packing of the gamma distributed balls by means of the force-biased algorithm, followed by distributing the small fixed size balls according to the RSA principle. A model-fitting procedure by minimising the contrast to estimate the gamma distribution parameters, the radius of the 2.9% small balls, and the volume fraction of balls was then applied.

Johnson–Mehl tessellations

For a Johnson–Mehl tessellation, even if a spatially homogeneous Poisson process for the birth locations is assumed, a sample of only the resultant tessellation has not enough information for the estimation of the time-dependent birth rate and the growth rate. In order to apply the estimation techniques given in last part of Section 6.6.4, temporal information, especially the birth times of nuclei, is essential.

Example 9.5. *Application to neurobiology (Quine and Robinson, 1992; Chiu et al., 2000; Molchanov and Chiu, 2000; Chiu et al., 2003)*
The terminal of a neuronal axon has branches consisting of strands. At a synapse an action potential triggers the release of neurotransmitter at randomly scattered sites on these strands.

Each quantum released would cause release of an inhibitory substance that diffuses along the terminal bi-directionally at a constant rate preventing further releases in the inhibited region. Thus, it can be modelled by the birth-and-growth process of a Johnson–Mehl tessellation on the line. The data are measurement of times of actual release of transmitters (birth times of nuclei) and amplitudes of releases that, after the inverse power transformation, could be served as surrogates for birth locations. Estimation methods described in Section 6.6.4 could then be applied to these 'tessellation' data.

9.11 Random geometrical networks

9.11.1 Introduction

This section considers random geometrical structures constructed by vertices, in this context often called 'nodes', and edges, which can be curves. The nodes are points randomly placed in $\mathbb{R}^d$.

Examples of geometrical networks

- *Edge systems of tessellations* are geometrical networks. A geometrical network is also obtained when some edges of the tessellation are removed, but so that isolated vertices are avoided. An example is the (planar) Arak polygonal Markov field, which is based on the Poisson line process; see Arak and Surgailis (1989), Arak *et al.* (1993) and van Lieshout (2012). It finds application in the context of image segmentation.
- *Crack or fracture networks* as in Figures 8.1 and 9.2 are geometrical networks, if the endpoints of free cracks (crack tips) are considered as vertices. A simple planar model is that the vertices are considered as primary elements and the straight line segment edges then are stepwise constructed.
- *Polymer networks* are an intensively studied class of geometrical networks. There the edges or 'chains' stand for chemical junctions. These networks exist in $\mathbb{R}^3$ and occupy there space. Edges connect preferably vertices close together. Physicists and chemists developed models of reaction kinetics that explain the generation of such networks, for which real images are not available. The classical theory of Flory (1953, 1976) uses an f-degree Cayley tree or Bethe lattice (a connected graph without cycles, see p. 403, where each vertex is connected with f other vertices, usually with $f = 3$ or 4; see Matoušek and Nešetřil, 2008), which is made incomplete and random, in the sense that only a fraction p of edges is realised. This theory has been generalised and extended; see Macosko and Miller (1991). This book has not the space to give even a sketch of these theories, which even use ideas from knot theory; see Michalke *et al.* (2001).
- *Communication networks* are of course an important class of networks, where also geometrical aspects play an important rôle; see Baccelli *et al.* (1997), Dousse *et al.* (2006), Gloaguen *et al.* (2006), Franceschetti and Meester (2008), Baccelli and Błaszczyszyn (2009a,b) and Zuyev (2010). A geometrical network of interest in this context is, for example, the 'connectivity network' or transmission range network, which connects all

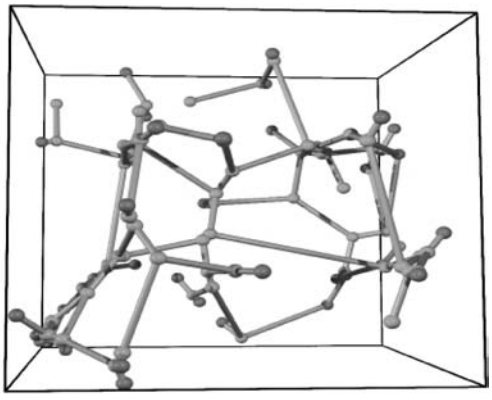

Figure 9.18 A small part of a three-dimensional geometrical network, showing the topology of a human forearm (radius) trabecular bone structure. It results from skeletonisation. Reproduced from Tscheschel and Stoyan (2003) with permission of Wiley.

antennas the SINR cells (see Section 6.5.4) of which each other overlap. However, in such networks many effects play rôles that are beyond the aims of the present section, which are only geometrical–topological.

- *Porous-media networks* result from geometrical constructions aiming at reducing porous media to networks that characterise their topological structure (perhaps to those of systems of balls and throats). Many geometrical networks could be considered as porous-media networks. The networks stand either for the system of pores or for its complement; in the latter case one speaks about 'skeletonisation'. Examples of such constructions can be found in Lowry and Miller (1995), Pothuaud *et al.* (2000), Sok *et al.* (2002), van Dalen *et al.* (2007), Thiedmann *et al.* (2009), Kadashevich and Stoyan (2010) and Stroeven *et al.* (2012). See Figure 9.18.
- *Traffic networks* (railway, highway) and *drainage networks*. See for example Gloaguen *et al.* (2006).

9.11.2 Formal definition of random geometrical networks

A *geometrical network* is a union δ of smooth curves, called edges, in $\mathbb{R}^d$ of finite length with two endpoints, which may coincide to allow self-loops. The system of curves is locally finite, that is, each compact set intersects δ in only a finite number of curves. Two arbitrary curves contact one another only in one common endpoint. Isolated points (segments of zero length) do not exist, but there may exist endpoints that are not connected with other curves ('dead ends'). The joint endpoints at the contacts of curves, the dead ends and the coincided endpoints in loops are all called vertices. Figure 9.18 shows a piece of a three-dimensional network.

Two edges are connected if they have a common vertex; two vertices are connected if there is a sequence of connected edges between them; a network is connected if every two vertices are connected.

A geometrical network is a closed subset of $\mathbb{R}^d$ and, if the curves are finite chains of straight line segments, polyconvex. Let $\mathbb{N}_{\mathrm{GN}}$ be the set of all geometrical networks of $\mathbb{R}^d$. The corresponding σ-algebra $\mathcal{N}_{\mathrm{GN}}$ is the trace of $\mathcal{F}$ on $\mathbb{N}_{\mathrm{GN}}$, where $\mathcal{F}$ is the hitting σ-algebra as introduced in Section 6.1.2. A *random geometrical network* Δ is a random element of $[\mathbb{N}_{\mathrm{GN}}, \mathcal{N}_{\mathrm{GN}}]$ or an $(\mathcal{A}, \mathcal{N}_{\mathrm{GN}})$-measurable mapping from a probability space $[\Omega, \mathcal{A}, \mathbf{P}]$ into $\mathbb{N}_{\mathrm{GN}}$. In short, Δ is a special random closed set the realisations of which are geometrical networks.

Stationarity and isotropy of random geometrical networks are defined analogously as for tessellations in Section 9.1. This section considers mainly stationary random geometrical networks. These have necessarily infinitely many edges. For random finite networks, it would be more natural to study them as random graphs; see Section 9.12.2.

Geometrical networks belong to the class of so-called *cell complexes* as studied by Zähle (1988). These consist not only of curves (edges), but also of lower-dimensional manifolds. For them generalisations of the formulae below hold.

9.11.3 Summary characteristics of stationary random geometrical networks

Characteristics for geometrical networks are often topological, except perhaps only those originated from fibre processes, such as L_A and L_V. The definitions and notations below are given for the spatial case, and those for the planar case are analogous.

As in the case of tessellations, the following characteristics of a random geometrical network Δ are natural:

- λ_0 = vertex intensity, the mean number of vertices per unit volume.
- λ_1 = edge intensity, the mean number of edge midpoints per unit volume.
- L_V = mean total length of edges per unit volume (and so the mean length of the typical edge is equal to L_V/λ_1).
- N_V = specific connectivity or specific Euler–Poincaré characteristic, which is defined as either

$$N_V = \lim_{K \uparrow \mathbb{R}^3} \frac{\mathbf{E}\big(\chi(\Delta \cap K)\big)}{\nu_3(K)}, \tag{9.142}$$

  where χ is the connectivity or Euler–Poincaré characteristic, or

  N_V = intensity of the 0-curvature measure corresponding to Δ.

- f = mean degree (also known as 'mean functionality' and 'mean coordination number') of the typical vertex, where the degree of a vertex is the number of edges emanating in it. This is the analogue to $\overline{n}_{01}$ for tessellations, and is defined as

$$f = \sum_{k=1}^{\infty} k p_k, \tag{9.143}$$

where p_k is the probability that the typical vertex has degree k. This probability is given by

$$p_k = v_k/\lambda_0, \tag{9.144}$$

in which v_k is the mean number per unit volume of k-degree vertices. Because each edge connects two vertices, it follows

$$\sum_{k=1}^{\infty} v_k = 2\lambda_1. \tag{9.145}$$

The characteristic f is considered dominant in the physical context (Chubynsky and Thorpe, 2001).

That some of the characteristics introduced above are closely related is shown by the formula

$$N_V = \lambda_0 - \lambda_1, \tag{9.146}$$

see Mecke and Stoyan (2001) for a proof. It is based on the fact that the 0-curvature measure is purely atomic and concentrated at the vertices, and assigns the mass $1 - k/2$ to every k-degree vertex.

Finally, it is the case that

$$N_V = (1 - f/2)\lambda_0. \tag{9.147}$$

Formula (9.147) shows that for connected networks the connectivity number is always nonpositive because $f \geq 2$, and becomes 'more negative' with increasing mean degree f. However, positive values are also possible. In the case of a rather strange 'network' that consists only of non-intersecting curves it is $N_V = \lambda_1$.

9.11.4 Statistics for networks

Tscheschel and Stoyan (2003) study a statistical problem for spatial networks, namely the estimation of their connectivity number N_V. By Equation (9.147) it can be reduced to a problem of point process statistics. While the network studied in this paper is a constructed porous-media network, the network in Lück *et al.* (2010, 2013) is a real filament network, the graph structure of which is reconstructed by means of image-analytical tools. Both papers characterise the network variability by pair correlation functions of associated point processes; see also Car and Parrinello (1988).

An example for network modelling and goodness-of fit testing is Thiedmann *et al.* (2009). The authors compare empirical and model minimum spanning trees, tortuosity characteristics and degree distributions.

Okabe and Sugihara (2012) study various statistical analyses of events (such as traffic accidents) on and alongside networks. Aldous and Shun (2010) consider statistics for planar networks which measure the trade-off between total graph length and efficiency of transportation.

9.11.5 Models of random geometrical networks

All tessellation models are of course models for geometrical networks. However, the converse is not true. In the following models, there is no guarantee that the construction procedures will lead to tessellations.

Nearest neighbour embracing graphs

Chiu and Molchanov (2003) introduced a model, now known as the *nearest neighbour embracing graph*, to incorporate not only distances but also directions. Each point of a given (finite or infinite) point process Φ in $\mathbb{R}^d$ is connected by a directed edge to its nearest neighbour, then to the second nearest, etc. From each point directed edges are continued to emanate until either the point is contained in the interior of the convex hull of the neighbours it connects to or the point is connected to all other points. If Φ is planar homogeneous Poisson, the mean out-degree (the number of directed edges going out) of the typical vertex is 5 (cf. the mean degree in the Poisson-Delaunay tessellation is 6). A realisation is shown in Figure 9.19(a). An optimal construction algorithm was proposed in Chan *et al.* (2004, 2006). The model has been used to study for example visibility and illumination in cameras and robot vision systems (Abellanas *et al.*, 2006, 2007a,b, 2009), limited range coverage of antennas (Matos, 2009), coverage boundary detection for wireless networks (Zhang *et al.*, 2006, 2009) and data depth in multivariate statistics (Cascos, 2007).

Poisson graphs with i.i.d. degrees

Deijfen *et al.* (2012) considered the following model, called the *polygamous Poisson process*. The starting point is a homogeneous Poisson point process in $\mathbb{R}^d$. Its points serve as vertices of the random geometrical network. The degrees of the vertices follow independently a prescribed distribution P. The first step is to attach an independent P-distributed random number of stubs

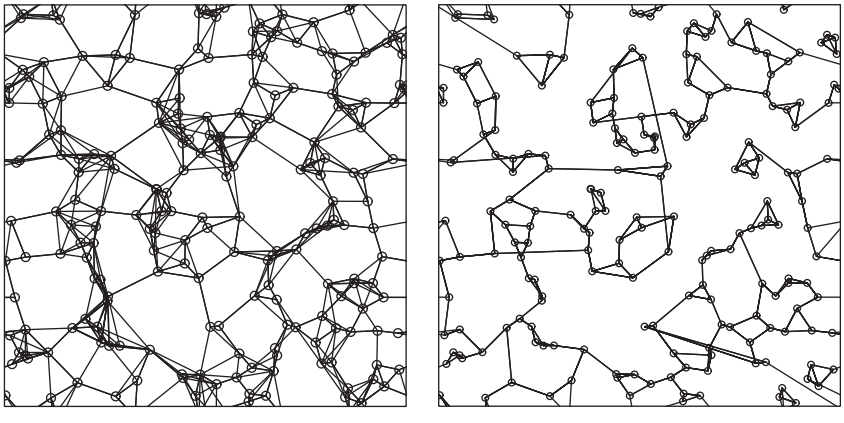

Figure 9.19 Realisations of network models for a sample of 200 i.i.d. uniform points in the unit square with periodic boundary conditions: (a) a nearest neighbour embracing graph; (b) a Poisson graph with degree 3 in all vertices. Note that some vertices in the graph seem as if they had degree 2; in fact these are pairs of vertices. Courtesy of T. Rajala.

to each vertex. Then a stable matching algorithm is applied to connect stubs. It has been shown that if the vertices form a homogeneous Poisson process, the matching scheme will be perfect and will lead to a stationary random network with the degree distribution P. Figure 9.19(b) shows a simulated sample of such a graph for the particular case that the degree of all vertices is 3, as it is in many tessellations. Clearly, the pattern is not a tessellation; it could be interpreted as a segment process. Deijfen (2009) studied the edge length distribution of such geometrical networks.

9.12 Random graphs

9.12.1 Introduction

This last section of the chapter presents some ideas of the theory of random graphs. Clearly, also tessellations and geometrical networks can be interpreted as graphs, but now only their topological nature matters, while locations and lengths are ignored. In many contexts such an approach is sufficient.

A *graph* consists of a nonvoid set of vertices connected by edges. Denote by $\mathbf{N}$ the set of all natural numbers. When vertices are labelled by natural numbers, a graph of n points can be represented by its $n \times n$ adjacency matrix, the (i, j)-entry of which indicates the number of edges connecting vertices i and j. Thus, $(\mathbf{N} \cup \{0\})^{\mathbf{N}\times\mathbf{N}}$ is the space of all graphs. A random element of this space endowed with the Borel σ-algebra with respect to the discrete topology (i.e. all subsets are open) is called a *random graph*.

Each tessellation and network can be interpreted as a graph by considering only the connectivity properties of vertices and edges.

A wide range of systems in nature and society can be described and hence modelled by random graphs. See Albert and Barabási (2002), Dorogovtsev and Mendes (2002) and Newman (2003) for details.

Classical random graph theory, initiated by Erdős and Rényi (1959), considers randomly chosen graphs among all possible N-edged graphs with n vertices. In later development, random graphs are considered as results of stochastic construction processes. A simple model of this type, $\mathbb{G}(n, p)$, proposed independently by Solomonoff and Rapoport (1951) and Gilbert (1959), is constructed as follows: there are given n vertices, and each pair of the vertices is connected independently with probability p. This model and its variants are (collectively) known as the *Erdős–Rényi model*, which is, however, too simple to model real-life situations.

In the following discussion, more complicated models and their properties are considered. For statistical analysis of empirical graphs, see for example Fienberg (2010a,b), in which the term 'network' is used instead of 'graph'.

The *typical vertex* of a random graph is just a randomly (uniformly) chosen vertex among the existing n vertices. Denote by p_k the probability that the typical vertex is of *degree* k, where the degree of a vertex is the number of edges connected to it. The distribution $\{p_k\}$ is called the *degree distribution*.

A *path* is a sequence of distinct vertices and corresponding edges, each of which connects two consecutive vertices in the sequence, and the length of a path is measured by the number of its edges. For a graph $\mathbb{G}_n$ with n vertices, let the *typical distance* H_n be the length of the shortest path between the members of a randomly and uniformly chosen pair of connected vertices. The *characteristic path length* L_n (also known as the *average path length* or simply

the *average distance*) is the mean of H_n, while the *diameter* $\mathrm{diam}(\mathbb{G}_n)$ is the maximum of the shortest path lengths between any pair of connected vertices.

A *cycle* is a 'closed path', that is, a path plus an edge connecting the two vertices with degree 1 (before adding the edge) in the path, and the cycle length is equal to its number of edges. A connected graph without any cycle is a *tree*.

9.12.2 Random graph models and their properties

Degree distributions

A *scale-free random graph* is one whose degree distribution satisfies the power law,

$$p_k \simeq Ck^{-\tau}, \tag{9.148}$$

for large k, where C and τ are some constants. Power-law degree distributions have been observed statistically in many natural graphs with τ around 2 to 3; see the examples tabulated in Albert and Barabási (2002, Table II) and Dorogovtsev and Mendes (2002, Table I).

The Erdős-Rényi model is not scale-free. For $\mathbb{G}(n, p)$ the degree distribution is binomial. As $n \to \infty$ such that $np \to \lambda$, its degree distribution converges to Poisson.

An important scale-free model is the Barabási–Albert model, also called *preferential attachment model*; see Barabási and Albert (1999). It starts with m_0 vertices without any edges and then grows by adding vertices, one by one sequentially, such that at each step the new vertex has m ($\leq m_0$) edges that link the new vertex to m different already existing vertices in such a way that an existing vertex is linked with probability proportional to its current degree. Thus, large degrees tend to become even larger and hence this model is also known as the *rich-get-richer model*.

The description given above is incomplete because it does not explain how the first edge is connected when all existing vertices are of zero degree, does not describe the dependencies between edges added at each step, and does not mention whether self-loops are allowed and whether the degrees should be intermediately updated. Since Bollobás *et al.* (2001) formulated the model rigorously many variants have been developed; see van der Hofstad (2010b, Chapter 8).

However, each of the different ways to make the model precise leads to '*the same asymptotic behaviour*' (Durrett, 2007, p. 90) and '*the results, in particular the occurrence of power laws and the power-law exponent, do not depend sensitively on the respective choices*' (van der Hofstad, 2010b, p. 166).

The limit of the degree distribution in the incomplete Barabási–Albert model above, as the number n of vertices goes to infinity, satisfies the power law

$$p_k \simeq \frac{2m^2}{k^3}, \tag{9.149}$$

for large k, giving $\tau = 3$ in Formula (9.148), which turns out to be true also in precise models; see Bollobás *et al.* (2001, Theorem 1) and Durrett (2007, Formula (4.1.1)).

To alter the value of the exponent τ, consider the *affine preferential attachment model* (van der Hofstad, 2010a,b) in which a third model parameter δ is introduced, which is larger than $-m$. The model starts with a single vertex with m self-loops. If $m = 1$, the new edge attached to the new vertex will be linked to the new vertex itself to form a self-loop with probability proportional to $1 + \delta$ and linked to an existing vertex with degree k with probability

proportional to $k + \delta$. If $m > 1$, the m new edges attached to a new vertex are linked to vertices sequentially with intermediate updating of their degrees. The i^{th} new edge is linked to an existing vertex with probability proportional to its current degree after linking the $(i-1)^{\text{st}}$ edge plus δ, and linked to the new vertex itself with probability proportional to its intermediately updated degree (with the convention that its degree is 1 when $i = 1$) plus $i\delta/m$. The special case $\delta = 0$ gives the preferential attachment model.

The power law given in (9.148) still holds with

$$\tau = 3 + \frac{\delta}{m} > 2, \tag{9.150}$$

see van der Hofstad (2010a, Theorem 6.16).

Whilst the preferential attachment model turned out to have a power-law degree distribution, Molloy and Reed (1995) constructed random graphs with other, specified degree distributions. Their model is described as follows.

Take a sequence $\{k_i\}_{i=1}^n$ of independent degrees from a given degree distribution $\{p_k\}$. Then k_i stubs, or half-edges, are attached to vertex i and pairs of stubs are chosen without replacement at random to be connected together. This model is called the *configuration model*, in which self-loops are possible. See Durrett (2007, Chapter 3) and van der Hofstad (2010b, Chapters 7 and 10) for details and references.

A generalisation of the Erdős–Rényi model $\mathbb{G}(n, p)$ is the *inhomogeneous random graph*, proposed by Bollobás *et al.* (2007) and further developed in van der Hofstad (2010b, Chapters 6 and 9). Instead of the same probability p for all pairs of vertices, the model has p_{ij} for the probability that the vertices i and j are connected. As an example, consider such a model with double randomness. Suppose $\{w_i\}$ are i.i.d. random weights assigned to the vertices. Conditionally on the weights $\{w_i\}$, letting

$$p_{ij} = 1 - \exp\left(-\frac{w_i w_j}{\sum_{k=1}^n w_k}\right)$$

gives the model in Norros and Reittu (2006). If the tail of the probability density of the w_i follows a power law with exponent $-\tau$, then the power law (9.148) holds (van der Hofstad, 2010a, Section 6.3.2.1).

Path lengths

The 'small-world phenomenon' is a famous, even notorious, term. It is popularly known as the *six degrees of separation* phenomenon, corresponding to the assertion that any two persons on this planet can be connected by a chain of at most six persons knowing each other on a first-name basis; see Newman *et al.* (2006). This happens when there is a relatively short path connecting any two vertices in a graph.

Many social networks of ten thousands to a million vertices have average distances of around 3 to 8 (Newman, 2003, Table 3.1), which loosely agrees with the six degrees of separation, while the average distance of the World Wide Web, with over 800 million vertices, was in the year 1999 around 19 (Albert *et al.*, 1999).

The *Watts–Strogatz model* is a one-parameter small-world model proposed by Watts and Strogatz (1998). Start from a one-dimensional lattice with periodic boundary conditions, that is, a ring lattice, with n vertices. Assume that there are k edges per vertex, and then each

edge may, with probability p independent of the others, be rewired by connecting one end of it to another vertex chosen at random.

In order to obtain a sparse but connected graph, one has to ensure that

$$n \gg k \gg \log n \gg 1,$$

in which the condition $k \gg \log n$ guarantees that during the rewiring process the random graph is connected at all times. Watts and Strogatz (1998) showed that in this model, the average distance L_n of a graph of n vertices has the following asymptotic properties:

$$\text{a large world:} \quad L_n \simeq \frac{n}{2k}, \qquad \text{as } p \to 0 \quad \text{(regular lattice)},$$

$$\text{a small world:} \quad L_n \simeq \frac{\log n}{\log k}, \quad \text{as } p \to 1 \quad \text{(entirely random)}.$$

The crossover (or phase transition) between the large world, in which L_n grows linearly with n, and the small world, in which L_n grows only logarithmically with n, is of some interest. For the Watts–Strogatz model and many of its variants, including models starting with a d-dimensional lattice, numerical simulations and analytical arguments led to the widely accepted general form

$$L_n \simeq \frac{n^{1/d}}{k} f(p \cdot k \cdot n), \tag{9.151}$$

where

$$f(u) = \begin{cases} \text{constant}, & \text{if } u \ll 1, \\ \dfrac{\log u}{u}, & \text{if } u \gg 1; \end{cases}$$

see Albert and Barabási (2002, Section VI) and Newman (2003, Section 6.3). Relation (9.151) shows that switching from a large world to a small world and vice versa can be achieved by changing just the single scaling variable, namely, the product $p \cdot k \cdot n$, which is the average number of rewiring of edges. However, the Watts–Strogatz model is not scale-free because the degree distribution decays exponentially (Albert and Barabási, 2002, Formula (77)).

For the affine preferential attachment model above, the random typical distance H_n, with probability tending to 1 as $n \to \infty$, satisfies

$$\begin{aligned} C_1 \log n \le H_n \le C_2 \log n, &\qquad \text{if } \tau > 3, \\ C_3 \log\log n \le H_n \le C_4 \log\log n, &\qquad \text{if } 2 < \tau < 3, \end{aligned} \tag{9.152}$$

and for the configuration model and inhomogeneous random graphs,

$$\begin{aligned} H_n \simeq C_5 \log n, &\qquad \text{if } \tau > 3, \\ H_n \simeq C_6 \log\log n, &\qquad \text{if } 2 < \tau < 3, \end{aligned} \tag{9.153}$$

for some positive constants C_i; the asymptotic equalities are in the sense of convergence in probability as $n \to \infty$ (van der Hofstad, 2010a, Theorems 6.18 and 6.19).

Note that the relations (9.152) and (9.153) do not include the case that $\tau = 3$. For the inhomogeneous random graph above where the p_{ij} are proportional to (deterministic) weight products $w_i w_j$, Chung and Lu (2003, Theorem 2.6) prove that

$$L_n \simeq C\frac{\log n}{\log\log n}, \qquad \text{for } \tau = 3, \tag{9.154}$$

for some positive constant C. However, it is not clear to what extent (9.154) is true also for other models.

Bollobás and Riordan (2004, Theorem 1) show that the preferential attachment model with a fixed $m \geq 2$ is also a small world in the sense that for any positive ϵ the probability that the following inequalities hold tends to 1 as $n \to \infty$:

$$(1-\epsilon)\frac{\log n}{\log\log n} \leq \operatorname{diam}(\mathbb{G}_n) \leq (1+\epsilon)\frac{\log n}{\log\log n}. \tag{9.155}$$

The lower bound in (9.155) also holds for $m = 1$, but the upper bound does not.

However, in general, the diameter is less robust and informative than the average and typical distance. Albert and Barabási (2002, p. 58) argue that the diameters of graphs of the same size n do not vary much and are concentrated around some value of order $O(\log n)$. This statement has been made more precise and proved for some special models (Chung and Lu, 2001 and Bollobás *et al.*, 2007, Theorem 3.16).

Clustering

In a small world, clustering is typical: it is likely that a friend of your friend is also your friend. A random graph is said to be *highly clustered* when two randomly chosen nearest neighbours of an arbitrary vertex are more likely to be connected to each other than an arbitrary pair of vertices. The *Barrat–Weigt clustering* (or *transitivity*) *coefficient* $C^{\mathbb{G}}_{BW}$ of a random graph $\mathbb{G}$ is defined as

$$C^{\mathbb{G}}_{\mathrm{BW}} = \frac{3 \times \text{number of triangles in } \mathbb{G}}{\text{number of connected triples}} = \frac{6 \times \text{number of triangles in } \mathbb{G}}{\text{number of paths of length two}}, \tag{9.156}$$

where the factors 3 and 6 arise in the numerators because each triangle has been counted three and six times, respectively, in the denominators (Barrat and Weigt, 2000). It tells the proportion of connected triples that form triangles and hence $0 \leq C^{\mathbb{G}}_{\mathrm{BW}} \leq 1$. A graph $\mathbb{G}$ containing n vertices and N edges is highly clustered if

$$C^{\mathbb{G}}_{\mathrm{BW}} \gg \frac{2N}{n(n-1)}. \tag{9.157}$$

It is also useful to define a local value to measure the strength of clustering of a vertex:

$$C^{\mathbb{G}}_i = \frac{\text{number of triangles in } \mathbb{G} \text{ connected to vertex } i}{\text{number of triples centred on vertex } i}, \tag{9.158}$$

with the convention that $0/0 = 0$ (Watts and Strogatz, 1998). This value reflects the extent to which neighbours of vertex i are also neighbours of each other. Define the *Watts–Strogatz*

Table 9.8 Barrat–Weigt clustering coefficients of random graph models.

Model of $\mathbb{G}$	Asymptotic $\mathbf{E}(C_{\mathrm{BW}}^{\mathbb{G}})$	Source
Erdős–Rényi	$\frac{\lambda}{n}$ (where $np \to \lambda$)	Watts and Strogatz (1998)
Watts–Strogatz	$\frac{3(k-1)}{2(2k-1)}(1-p)^3$	Barrat and Weigt (2000)
Barabási–Albert	$\frac{m-1}{8}\frac{(\log n)^2}{n}$ *	Bollobás and Riordan (2003)
Affine preferential attachment with $\delta > 0$	$C_{m,\delta}\frac{\log n}{n}$ †	Eggemann and Noble‡ (2011)

* This result is very different from the experimental value $O(n^{-0.75})$ in Albert and Barabási (2002).

† $C_{m,\delta}$ is an explicitly known value depending solely on m and δ.

‡ Their precise model is slightly different from the one given here.

clustering coefficient as

$$C_{\mathrm{WS}}^{\mathbb{G}} = \frac{\sum_{i\in\mathbb{V}} C_i^{\mathbb{G}}}{n}. \tag{9.159}$$

Both clustering coefficients above are of order n^{-1} for the Erdős–Rényi model. In contrast, for an empirical small-world network, it is suspected that either coefficient will tend to a non-zero limit as $n \to \infty$. Newman (2003, Table 3.1) reports that the mathematics coauthorship network of 253,339 vertices and 496,489 edges has $C_{\mathrm{BW}}^{\mathbb{G}} = 0.15$ and $C_{\mathrm{WS}}^{\mathbb{G}} = 0.34$. Table 9.8 summarises the known asymptotic formulae for the mean Barrat–Weigt clustering coefficients of various models.

Phase transition, percolation and resilience

One of the main questions in random graph theory is to determine critical model parameters ω_c for phase transition. The term *phase transition* (Janson *et al.*, 2000, Chapter 5) refers to the phenomenon that for some property, such as connectivity, graphs with model parameter $\omega < \omega_c$ are very unlikely (and almost surely not, as $n \to \infty$) to possess the property, whilst graphs with model parameter $\omega > \omega_c$ are very likely (and almost surely, as $n \to \infty$) to possess it.

Erdős and Rényi (1960) show that for the model $\mathbb{G}(n, p)$, the critical probability p for the existence of a giant component is $1/n$ and the critical probability p for connectivity is $\log n/n$. More precisely, with probability increasing to 1 as $n \to \infty$, when $p < 1/n$, none of the connected components has more than $O(\log n)$ vertices, whilst when $p > 1/n$, the number of vertices of the largest component is asymptotic to rn, for some $0 < r < 1$. In particular, at the critical probability $p = 1/n$, the largest component has a size of order $n^{2/3}$; for connectivity, if $p < \log n/n$, then $\mathbb{G}(n, p)$ contains isolated vertices and hence be disconnected, whilst if $p > \log n/n$, then $\mathbb{G}(n, p)$ is connected.

There is an analogous result for the existence of a giant component in scale-free models. The model parameter for such a phase transition is

$$\omega = \begin{cases} m & \text{for affine preferential attachment models,} \\ \dfrac{\mathbf{E}\big(K(K-1)\big)}{\mathbf{E}(K)} & \text{for the configuration model,} \\ \dfrac{\mathbf{E}(W^2)}{\mathbf{E}(W)} & \text{for inhomogeneous random graphs,} \end{cases} \tag{9.160}$$

where K denotes the random degree of the typical vertex, following the distribution $\{p_k\}$, and W denotes a random variable having the same distribution as the random weights w_i in inhomogeneous random graphs. For these three model classes, the critical values are the same and equal to $\omega_c = 1$ (see van der Hofstad 2010a,Theorem 6.17 for a precise statement).

One of the most striking features of natural networks is their high tolerance for errors. Albert *et al.* (2000) studied the effects of random errors (random removals of vertices, including edges attached to them) and targeted attack (targeted removal of high degree vertices) on the Internet and the World Wide Web, as well as the Erdős–Rényi model and the Barabási–Albert model. In the Erdős–Rényi model, the average path length increases monotonically with the percentage of removals, no matter random or targeted. In contrast, real networks and the Barabási–Albert model are robust to random errors but vulnerable to targeted attacks. The reader is referred to Albert and Barabási (2002, Section IX), Dorogovtsev and Mendes (2002, Section XI), Durrett (2007, Section 4.7) and Newman (2003, Sections 3.4 and 8.1) for detailed discussions and further references. Another way to study network vulnerability is to consider a stochastic model, such as a random walk, for the traffic of signals on a geometrical network; a vertex breaks down when the number of signals, such as random walkers, arrived exceeds a prescribed threshold; see for example Kishore *et al.* (2011, 2012).

The problem of connectivity of an infinite graph or network with random removals of vertices or edges is known as site percolation or bond percolation, respectively, which corresponds to the problem of the existence of a giant component in a finite graph. For a comprehensive review and further references; see van der Hofstad (2010a).

Some models for finite geometrical networks

While stationary random tessellations and geometrical networks are infinite graphs, finite geometrical networks are also important. In the following some popular models are briefly described.

Since physical measurements for these graphs are possible, the length of a path can be measured either topologically (i.e. still in terms of the number of edges) or physically (i.e. in terms of the sum of the Euclidean distances between consecutive vertices in the path). Of course properties of the topological length and of the physical length can be substantially different.

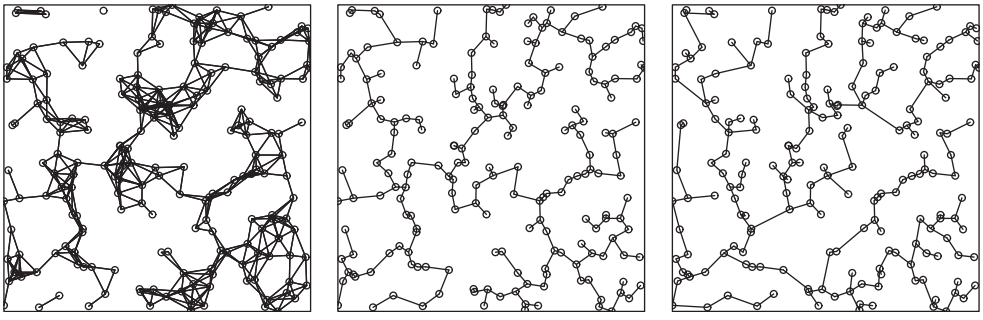

Figure 9.20 Realisations of finite random geometrical network models in the unit square for the same 200 points used in Figure 9.19: (a) a random geometric graph with $r = 0.1$; (b) a minimum spanning tree with the Euclidean distances as weights; (c) a radial spanning tree with respect to the origin, the lower left corner.

Random geometric graphs

The simplest model for a finite geometrical network is probably the *random geometric graph* $\mathbb{G}(\Phi_n, r)$, where $\Phi_n = \{x_1, \ldots, x_n\}$ denotes a finite point process in $\mathbb{R}^d$ and x_i and x_j are connected if the distance between them is not more than r; see Figure 9.20(a). Its infinite counterpart for the Poisson process case is discussed in Section 3.3.3. Probabilistic properties of $\mathbb{G}(\Phi_n, r)$ have been studied extensively in Penrose (2003). The existence of a giant component in $\mathbb{G}(\Phi_n, r)$ is related to the problem of continuum percolation, discussed in detail in Section 3.3.4 and Meester and Roy (1996), which plays an important rôle in studying actual percolation processes that occur in real heterogeneous materials (see e.g. Torquato, 2002, Section 9.2 and Chapter 10).

Special geometrical networks can be built based on $\mathbb{G}(\Phi_n, r)$. Examples include the restricted Delaunay triangulation (Chen, 2008), which contains the common edges in $\mathbb{G}(\Phi_n, r)$ and the Delaunay tessellation generated by Φ_n, and the geometric preferential attachment model (Flaxman *et al.*, 2006, 2007), where each new vertex in the preferential attachment model can only be connected to vertices within distance r. These models can be extended to the stationary case, using a homogeneous Poisson process instead of Φ_n.

Minimum spanning tree

A spanning tree of a finite set of vertices is a tree connecting all vertices. When each edge is given a weight, often but not necessarily its length, a *minimum spanning tree* is a spanning tree with the total weight not more than the total weight of any other spanning tree. Figure 9.20(b) shows a minimum spanning tree with Euclidean distance between vertices as weights.

Minimal directed spanning forests

Often edges in natural systems are directed. Such edges appear in the *minimal directed spanning forests* (Penrose and Wade, 2010), motivated by modelling communication and drainage networks. Suppose that the vertices have a partial order, such as the coordinatewise partial order. The distance between two vertices is measured by for example the L^p-norm. Each

non-minimal vertex under the partial order is connected to the nearest neighbouring vertex of lower order, or a randomly chosen one if there are more than one. Each connected component is a tree and the number of trees in this forest is equal to the number of minimal vertices. Different partial orders and distance functions lead to a rich variety of models. For example, the *radial spanning tree* is obtained if the partial order is the distance from the origin and the weights are the Euclidean distances; see Figure 9.20(c).

The minimal directed spanning forest must not be confused with the minimum spanning tree. The latter does not involve a partial order and its vertices are connected in such a way that the total weight is minimised; they do not necessarily connect to their nearest neighbours.