COUNTING FIGURES IN PLANAR RANDOM CONFIGURATIONS

A. M. KELLERER,* *University of Würzburg*

Abstract

Random configurations are considered that are generated by a Poisson process of figures in the plane, and a recent result is used to derive formulae for the estimation of the number of figures, and their mean area and perimeter. The formulae require merely the determination of the area, the perimeter, and the Euler–Poincaré characteristic of the random configurations in a fixed field of view. There are no similar formulae for the standard deviations of the estimates; their magnitudes in typical cases are therefore assessed by Monte Carlo simulations.

GEOMETRIC PROBABILITY; RANDOM CLUMPING; POISSON PROCESS; IMAGE ANALYSIS; PARTICLE COUNTING

1. Introduction

In a recent paper [4] formulae were given for the expectations of the area, α, the boundary length, β, and the Euler–Poincaré characteristic, χ, of random configurations formed within a specified field of view (window) by a Poisson process of figures, i.e. a random process that Matheron has termed a Boolean scheme [10], [11]. If window and figures are simply connected, the formulae contain only the area, A, and perimeter, S, of the window and, apart from the intensity, λ, of the Poisson process, the mean area, a, and mean perimeter, s, of the figures.

The present study examines the applicability of the results to the determination of λ, a, and s from observed random configurations. The need for such determinations arises in automatic image analysis, for example in track-etching dosimetry of heavy ions or in the automatic counting of bacterial colonies or mammalian cell colonies. This article is concerned with essentials of the problem, rather than technicalities of particular applications. In Section 2 the equations for the Minkowski functionals α, β, and χ are utilized to derive the formulae for estimating the number of figures and their mean area and

perimeter. No general relations exist for the variances of the estimates; computer-generated random configurations are, therefore, employed in Section 3 to assess the practicability of the procedure and the magnitude of the variances. A last section illustrates the method in terms of a practical example from track-etching dosimetry.

2. Formulae for the estimation of λ, a, and s

A Poisson point process in R^2 is considered, i.e., points are uniformly and independently distributed throughout the plane, with λ centers per unit area on the average. In line with the definition of a Boolean scheme, random figures are implanted with their centers of reference on the points. The figures are taken to be simply connected and sufficiently regular to be approximated by polygons. The varying types of figures are independently distributed, and isotropic and independent directional distribution is assumed.

In applications, such as the examples mentioned above, the figures are usually nearly circular, but it is also of interest to consider a somewhat more general case. The left panel of Figure 1 represents a configuration generated by repeated random selection from a set of 10 arbitrarily drawn figures. The right panel represents all boundaries, including those that are covered. Even with this full information there is no simple method of accounting rigorously for figures that are only partly contained in the window. Whatever counting procedure is chosen, it is desirable that it should lead, on the average, to the product, n, of the Poisson intensity, λ, and the area of the window. As a first approximation this could be achieved by counting partial figures half. However, this or a similar
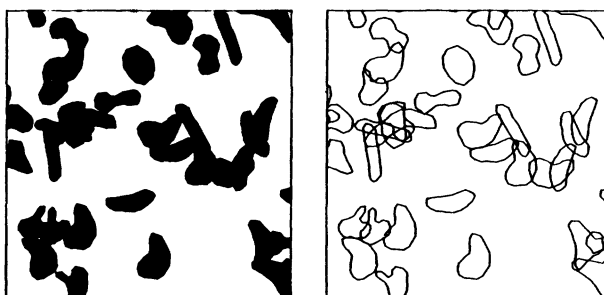

Figure 1. Random configuration generated at nominal coverage 0.3 from set A of 10 arbitrarily drawn figures (for details see Section 3). The left panel represents the configuration on a grid of 600 lines; the right panel gives the contours of all figures. Observed, estimated, and actual parameters:

$$\chi = 18 \qquad \beta = 11.19 \qquad \alpha = 0.284$$
$$\hat{n} = 34.0 \qquad \hat{s} = 0.456 \qquad \hat{a} = 0.00979$$
$$n = 30 \qquad s = 0.442 \qquad a = 0.01$$

procedure is merely an approximation. If, as in the left panel of the figure covered boundaries are invisible, the additional problem of overlap arises. In earlier work approximate formulae had been derived to correct for overlap (see [1], [6], [8], [9], [13], [14]). Such formulae can be satisfactory only if the nominal coverage, $\Psi = a\lambda$, and the resulting degree of overlap are sufficiently small. For an exact treatment, the more recent result can be utilized that accounts both for edge effects and for random overlap.

The random figures form a geometric object, I, in the window. I is the intersection of the union of all figures, i.e. of a realization of the Boolean scheme, with the window. Three fundamental parameters of the random configuration I can be observed. One parameter is its area, α, i.e. the covered area within the window. A second parameter is the length, β, of the boundary of I, i.e. the length of uncovered contours of figures within the window. For convenience those parts of the boundary of I that are formed by the frame of the window are not included in β. A third parameter is the Euler–Poincaré characteristic, χ, of I; it is equal to the number of clumps minus the number of enclosed voids. The characteristic is 1 for a simply connected figure without holes and it is 0 for a figure with one hole. More generally, it is $n - m$ for a configuration of n separate domains with a total of m enclosed voids. In the example of Figure 1 the characteristic is 18, as there are 21 clumps and 3 voids.

The average area of individual random figures is a, s is their mean perimeter and, as stated, the Euler–Poincaré characteristic of the figures is taken to be 1. The area and perimeter of the window are denoted by A and S. One has then the relations for the expectation values of α, β and χ [4]:

$$(1) \qquad \bar{\alpha} = A(1 - \exp(-\lambda a))$$

$$(2) \qquad \bar{\beta} = \lambda s A \exp(-\lambda a)$$

$$(3) \qquad \bar{\chi} = \exp(-\lambda a)(\lambda A + \lambda s S/2\pi - \lambda^2 A s^2/4\pi - 1) + 1.$$

Relations (1) and (2) are familiar, but Equation (3) is less evident. While it has been obtained without the results of Davy [2], [3], it can be derived also by adapting her formulae to a finite field of view.

By elementary rearrangement of Equations (1)–(3) one obtains three equations for the Poisson intensity, λ, and for the mean area, a, and mean perimeter, s, of the figures:

$$(4) \qquad \lambda = \frac{\bar{\chi} - 1}{A - \bar{\alpha}} + \frac{1}{A} - \frac{\bar{\beta} S}{2\pi A(A - \bar{\alpha})} + \frac{\bar{\beta}^2}{4\pi(A - \bar{\alpha})^2}$$

$$(5) \qquad a = -\ln(1 - \bar{\alpha}/A)/\lambda$$

$$(6) \qquad s = \bar{\beta}/(A - \bar{\alpha})\lambda.$$

The relations can be used to estimate λ, a, and s from observed images. For this purpose the unknown mean values $\bar{\alpha}$, $\bar{\beta}$, and $\bar{\chi}$ are substituted by the values α, β, and χ obtained from one image or by the averages of these quantities from several images. This procedure corresponds to the method of moments (see e.g. [7]) that leads to biased but consistent estimates.

On the basis of Davy's results or by different methods [5], [15] the equations for the Minkowski functionals of I can be obtained in their general form for R^n. The estimation procedure is therefore not restricted to R^2, although the equations for R^3 are already rather complex.

As stated at the outset, all considerations refer to the Boolean scheme, i.e. to random domains that are independently distributed. Effects of prior occupancy that cause interdependences between positions are not taken into account and can, in practice, restrict the applicability of the formulae. However, it will be noted that the result is otherwise general; it pertains to figures of varying shape in any field of observation, and it accounts equally for edge effects and random overlap of the figures. The extension to the anisotropic case is considered in Section 4.

There are no general relations for the variances of the estimates of λ, a, or s. Computer simulations are, therefore, required to assess the magnitude of the variances and the range of applicability of the procedure in typical cases.

3. Simulations

The simulations utilize as field of view a square window of unit area ($A = 1$, $S = 4$). Three sets of figures of different complexity are used. The mean area of the figures is scaled for all three sets to $a = 0.01$, so that the expected number, $n = A\lambda$, of centers in the field is always 100 Ψ. The simplest case is that of equal circles ($s = 0.3545$). The next case is that of a set, A, of 10 polygons (see Figure 1) with $s = 0.442$. The third case is that of a set, B, of somewhat more complex polygons (see Figure 2) with $s = 0.484$. The parameter $f = s^2/4\pi a$ has the value 1 for equal circles; for sets A and B the values are $f = 1.55$ and $f = 1.86$.

For the formation of images figures are selected randomly from the specified set of figures, random directions are chosen, and the centers are located randomly within the window and a sufficiently large surrounding frame. The technique of determining the parameters α, β, and χ in practical cases depends on particularities of the image-analysis equipment. For the present computer simulations the images were formed on a grid of 600 parallel lines (see Appendix).

Figure 3 gives, as further illustration, a random configuration obtained with set A at high nominal coverage, $\Psi = 1$. For all examples of random configurations the observed parameters α, β, and χ are listed together with the estimates $\hat{n}$, $\hat{a}$,

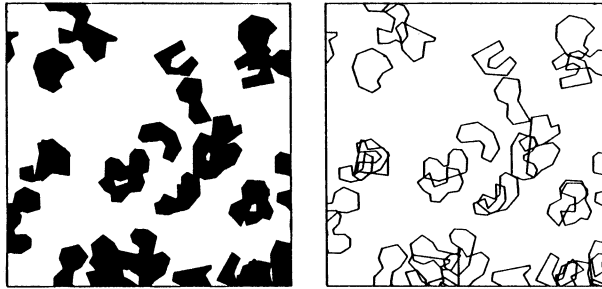Figure 2. Random configuration at nominal coverage 0.3 as in Figure 1, but obtained with set B of figures:

$\chi = 16$	$\beta = 11.03$	$\alpha = 0.283$
$\hat{n} = 31.0$	$\hat{s} = 0.497$	$\hat{a} = 0.0106$
$n = 30$	$s = 0.484$	$a = 0.01$
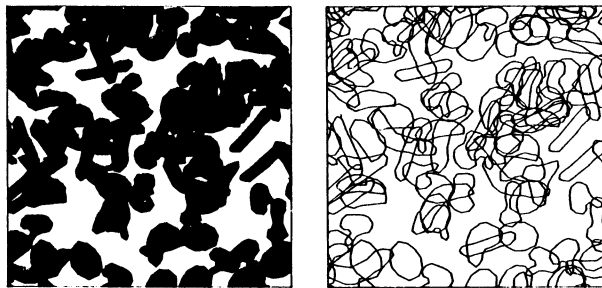
Figure 3. Random configuration as in Figures 1 and 2 but obtained with set A of figures at nominal coverage 1:

$\chi = -14$	$\beta = 15.92$	$\alpha = 0.641$
$\hat{n} = 87.4$	$\hat{s} = 0.506$	$\hat{a} = 0.0114$
$n = 100$	$s = 0.442$	$a = 0.01$

and $\hat{s}$ that result when the observed parameters are inserted into Equations (4)–(6) for $\bar{\alpha}$, $\bar{\beta}$, and $\bar{\chi}$.

Figure 4 represents for the different sets of figures the expected characteristic, $\bar{\chi}$, i.e. the expected number of clumps minus voids (see Equation (3)) and the averages obtained from series of 200 or more simulated images. No general relation is known for the variance of χ; the standard deviations obtained in the simulations are, therefore, given in Figure 5.

To estimate n, a, and s one can insert into Equations (4)–(6) averages of α, β, and χ from a sufficiently large number of configurations. The estimates from series of 200 images each were consistent with the actual values n, a, and s.
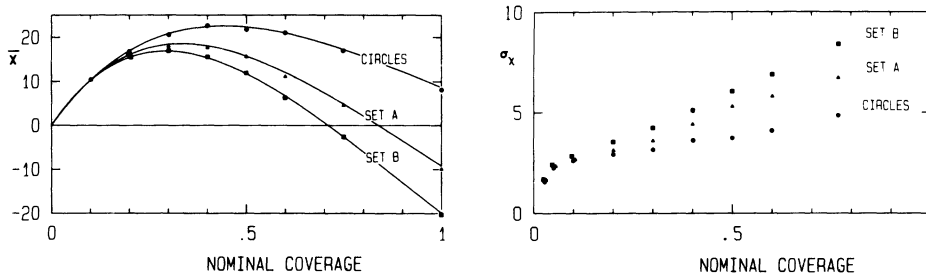
Figure 4. The mean characteristic, $\bar{\chi}$, (see Equation (3)) of the random configurations produced in a unit square by equal circles ($a = 0.01$, $s = 0.354$), and by the figures from set A ($a = 0.01$, $s = 0.442$) and set B ($a = 0.01$, $s = 0.484$). The points represent averages from series of at least 200 simulated images (equal circles (●), set A (▲), set B (■))

Figure 5. Standard deviation of the characteristic, χ, of random configurations produced in a unit square by equal circles (●), and by figures from set A (▲) and set B (■). Each point is based on a series of 200 or more random configurations

Deviations become evident at high nominal coverages ($\Psi > 1$), but they are a matter of the image representation on the line grid (see appendix) and they disappear in computations with finer grid spacing.

In applications one usually wants to obtain estimates from individual images, e.g. one may wish to have 'counts' for each observation. If, as in the present simulations ($A = 100a$), the field of view is of modest size the variances of α, β, and χ are substantial, and it is the purpose of the computer simulations to assess whether the estimation of n, a, and s remains practicable under such conditions.

Although α, β, and χ or their averages from several images are unbiased estimators of $\bar{\alpha}$, $\bar{\beta}$, and $\bar{\chi}$, the estimates of n, a, s are biased due to the non-linear form of Equations (4)–(6). The bias is most severe for estimates from individual images. In the present examples, the estimates from individual images were in the mean too large by about the same factor for n, a, and s and for the three sets of figures. These deviations increased with nominal coverages from roughly 1 per cent at $\Psi = 0.1$ to somewhat less than 3 per cent at $\Psi = 1$.

Of greater pragmatic importance are the standard deviations of the estimates from individual images. They were obtained from series of at least 200 images, and for the numbers, $n = A\lambda$, they are plotted in Figure 6 at different values of the nominal coverage. As one would expect, the standard deviations are largest for the more complex figures and smallest for equal circles. A notable result is the slow variation of the relative standard deviation with changing nominal coverage. When Ψ varies between 0.2 and 0.8 the fractional standard errors of n for individual images remain in the narrow range (0.21–0.23) for the circles, the case of greatest pragmatic importance. They lie in the range (0.25–0.27) and (0.28–0.30) for figures from sets A and B. One concludes, furthermore, that the
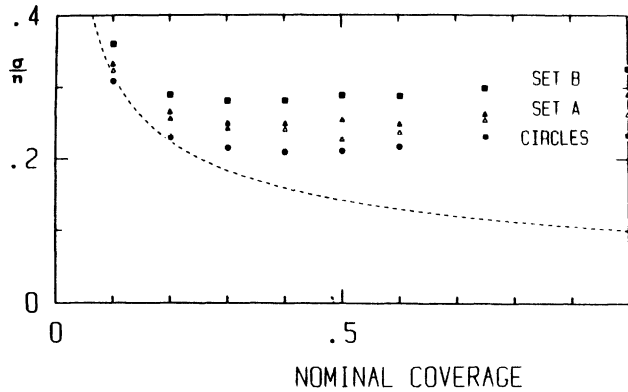
Figure 6. The fractional standard deviations, σ/n, of the estimates, $\hat{n}$, from individual random configurations with equal circles (●), and with figures from set A (▲) and set V (■). The open triangles (△) are results obtained with modified figures from set A (see Section 4 and Figure 7) in fixed direction. Each point is based on at least 200 estimates. The average number, $n = A\lambda$, of centers in the square window in 100 Ψ. The broken line represents the Poisson fluctuations, i.e. the fractional standard deviation, $1/\sqrt{n}$, of the number of centers in the window

estimation procedure remains practicable at high nominal coverage, where, as judged from the example of Figure 3, visual estimates begin to be difficult. A measure for the deterioration of the counts due to random overlap is the ratio, R, of the standard deviation to the Poisson deviations of the number of centers in the field of view (broken line in Figure 6). Without demonstration it may be stated that R remains nearly constant if the size of the window is increased. A condition is that the window is sufficiently large to contain far more figures entirely than partially. If this condition is not fulfilled, R can decrease substantially. It is evident that R tends to 0 if the window is a strip of vanishing width.

The standard deviations for the estimates of a and s from individual images are not given since they may be of less practical interest. But the general nature of the dependences on Ψ is similar to those for σ/n.

4. The case of anisotropy

The estimation procedure can be extended to figures of non-isotropic orientation. Equations (1) and (2) for the mean covered area, $\bar{a}$, and the mean boundary length, $\bar{\beta}$, of the configurations remain, as one can readily see, unaltered regardless of the directional distribution of the figures. However, as shown in the earlier article [4], the relation for the mean characteristic is modified:

(7) $\bar{\chi} = \exp(-\lambda a)(\lambda A + c_2 \lambda s S/2\pi - c_1 \lambda^2 A s^2/4\pi - 1) + 1.$

The coefficient c_1 depends on the distribution of the perimeter of the figures in direction; c_2 depends both on the distribution of the perimeter of the figures and the distribution of the boundary of the window in direction. Let $f(\phi)$ and $w(\phi)$ be the probability densities of the boundary in direction (traversal in mathematically positive sense) for the figures and the window, respectively. The distributions are assumed to be normalized to unity, and the subsequent relations are formulated for continuous distributions. In analogy to considerations given in the earlier article, one can derive the equations

$$(8) \qquad c_1 = \int_0^{2\pi} \int_0^{2\pi} f(\phi_1) f(\phi_2) \Delta\phi \, \sin(\Delta\phi) d\phi_1 d\phi_2$$

$$(9) \qquad c_2 = \int_0^{2\pi} \int_0^{2\pi} f(\phi_1) w(\phi_2) \Delta\phi \, \sin(\Delta\phi) d\phi_1 d\phi_2$$

with

$$\Delta\phi = \min\{|\phi_2 - \phi_1|, 2\pi - |\phi_2 - \phi_1|\}.$$

For a circular window one has $c_2 = 1$.

The change to the discrete case is straightforward. When the window is a square aligned with the coordinate system,

$$(10) \qquad c_2 = \int_0^{2\pi} f(\phi)(|\sin(\phi)| + |\cos(\phi)|)/2 \, d\phi.$$

The angular distribution of the boundary of the figures is likely to be unknown in actual cases. However, it is equal to the directional distribution of the uncovered boundary of the observed random configurations. This is so because the probability of a boundary element to be covered is independent of its orientation. From the directional distribution of the boundary, determined from one or more random configurations, one can derive the constants c_1 and c_2, and one can then utilize the modified equations for the estimation of λ, a, and s:

$$(11) \qquad \lambda = \frac{\bar{\chi} - 1}{A - \bar{\alpha}} + \frac{1}{A} - \frac{c_2 \bar{\beta} S}{2\pi A (A - \bar{\alpha})} + \frac{c_1 \bar{\beta}^2}{4\pi (A - \bar{\alpha})^2}$$

$$(12) \qquad a = -\ln(1 - \bar{\alpha}/A)/\lambda$$

$$(13) \qquad s = \bar{\beta}/(A - \bar{\alpha})\lambda.$$

To obtain a numerical example with a substantial degree of anisotropy, the figures of the random set A were expanded by a factor of 2 along an axis. As in the earlier examples, they were scaled to a mean area $a = 0.01$. The mean perimeter was then $s = 0.486$ ($f = 1.88$). Figure 7 gives a random configuration
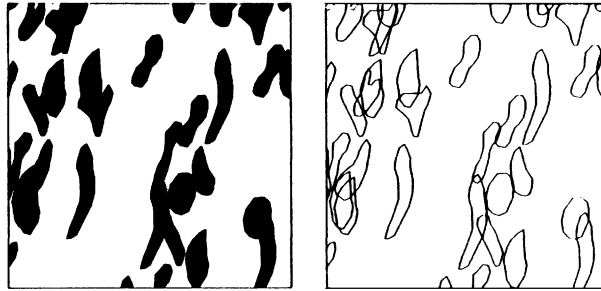
Figure 7. Random configuration as in Figure 1 obtained at nominal coverage 0.3 with modified figures from set A (expanded in one direction by a factor 2) with fixed direction ($a = 0.01$, $s = 0.486$):

$\chi = 23$	$\beta = 11.57$	$\alpha = 0.261$	$c_1 = 0.798$	$c_2 = 0.988$
$\hat{n} = 36.5$	$\hat{s} = 0.429$	$\hat{a} = 0.00829$		
$n = 30$	$s = 0.486$	$a = 0.01$		

with the figures in fixed orientation with nominal coverage $\Psi = 0.3$. The resulting standard deviations of the estimates $\hat{n}$ from individual images have been inserted into Figure 6.

5. Practical example and consideration of a simplified estimation procedure

The use of Monte Carlo simulations has made it possible, in Section 3, to establish basic characteristics of the estimation procedure without the need to assess the response characteristics of a particular image-analysis system. The technical implementation requires a separate study. It is, nevertheless, instructive to consider a practical example.

Figure 8 is a track-etching image of uranium ions (energy: 15 MeV/nucleon) normally incident on a glass sample. Such samples are used in heavy-ion dosimetry to determine the particle fluence; the fluence corresponds in this case to the Poisson intensity, λ. Routinely one limits the fluence sufficiently to avoid excessive overlap of the pits that are formed, during the etching process, around the points of incidence of the ions. However, over-exposures or over-etching can occur, and it is desirable that the automatic evaluation remain valid even in such cases. The sample in Figure 8 has been moderately over-etched; the etched pits have a diameter of about 3 μm instead of the usual value of about 2 μm. The estimate of the nominal coverage Ψ is 0.39.

The observed parameters χ, α, and β are listed with the image; so are the estimates $\hat{n}$, $\hat{a}$, and $\hat{s}$ that are obtained from Equations (4)–(6). The area of the window is $A = 50 \, \mu m \times 38 \, \mu m = 1900 \, \mu m^2$; its perimeter is $S = 176 \, \mu m$ ($F = 1.30$). The estimate $\hat{n} = 94.8$ is close to the visual count of 98 that is obtained by the attempt to enumerate the centers within the window.
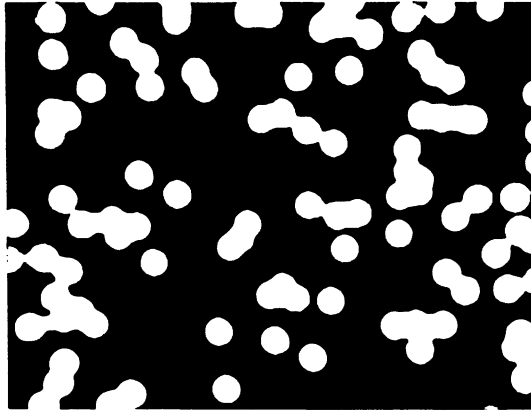
Figure 8. A 50 μm × 38 μm field on an etched glass sample that had been exposed to a parallel beam of high-energy uranium ions from the accelerator, UNILAC, of the GSI, Darmstadt. Observed and estimated parameters:

$$\chi = 45 \qquad \alpha = 617.5 \ \mu m^2 \qquad \beta = 680 \ \mu m$$

$$\hat{n} = 94.8 \qquad \hat{a} = 7.88 \ \mu m^2 \qquad \hat{b} = 10.6 \ \mu m$$

With $\hat{n} = 94.8$ the estimate of the Poisson standard error is 9.7. With the ratio $R = 1.3$ (see Figure 6) of the standard error for equal discs to the Poisson deviation, the standard error of $\hat{n}$ is estimated as 12.7. For the fluence, $\lambda = n/A$, one obtains therefore the estimate and standard error $\lambda = (4.99 \pm 0.66) \cdot 10^6/cm^2$. In actual dosimetry many fields are counted, and the standard error is obtained from the multiple counts. The number of pits in Figure 8 is larger than the number of figures per image in the simulations at the same nominal coverage; the bias of the estimate must therefore be less than the factor of roughly 1.015 obtained at $\Psi = 0.4$ in the simulations. A separate investigation would be required to assess the systematic errors caused by optical and electronic characteristics of the comparatively simple image-analysis equipment (Leitz Classimat) and by the relatively small number of 300 lines for the image in Figure 8.

Equipment presently available for the counting of bacterial colonies determines merely the characteristic, χ, and the total area, α, covered by the colonies. Estimates are then possible only, if either a or the ratio $f = s^2/4\pi a$ are separately obtained. Determination of a is often subject to considerable uncertainties. The parameter f tends to be less affected by technical factors, and its utilization will therefore be considered. Cultures of bacteria or mammalian cells contain nearly circular clones of varying size. One obtains then the parameter

(14)
$$f = \frac{1}{1 + (\sigma_r/\bar{r})^2}$$

where $\sigma_r/\bar{r}$ is the relative standard deviation of the radii. For instance, with bacterial colonies equidistributed in radius between 0.5 mm and 1 mm one has $f = 0.965$. In the example of Figure 8 one deals with discs of nearly equal size, hence $f = 1$.

To obtain the simplified estimation one uses the substitutions:

$$(15) \qquad \Psi = -\ln(1 - \bar{\alpha}/A), \quad f = s^2/4\pi a, \quad F = S^2/4\pi A$$

to derive from Equation (3):

$$(16) \qquad \frac{\bar{\chi} - \bar{\alpha}/A}{1 - \bar{\alpha}/A} - (1 - f\Psi)A\lambda = 2\sqrt{fF\Psi A\lambda}$$

and the quadratic equation for $\mu = \sqrt{\lambda}$

$$(17) \qquad (1 - f\Psi)A\mu^2 + 2\sqrt{fF\Psi A}\,\mu - \frac{\bar{\chi} - \bar{\alpha}/A}{1 - \bar{\alpha}/A} = 0.$$

The expression for λ is therefore

$$(18) \qquad \lambda = \left[\left(\pm \sqrt{fF\Psi A + \frac{\bar{\chi} - \bar{\alpha}/A}{1 - \bar{\alpha}/A}(1 - f\Psi)A} - \sqrt{fF\Psi A} \right) \Big/ (A(1 - f\Psi)) \right]^2.$$

Only positive values of μ are meaningful; the term within the square bracket must therefore not be negative, and this determines the choice of the sign with the first square root. Figure 9 illustrates the conditions by representing, for assumed values $f = 0.9$ and $F = 1$, the set of relations between $\bar{\chi}$ and Ψ for
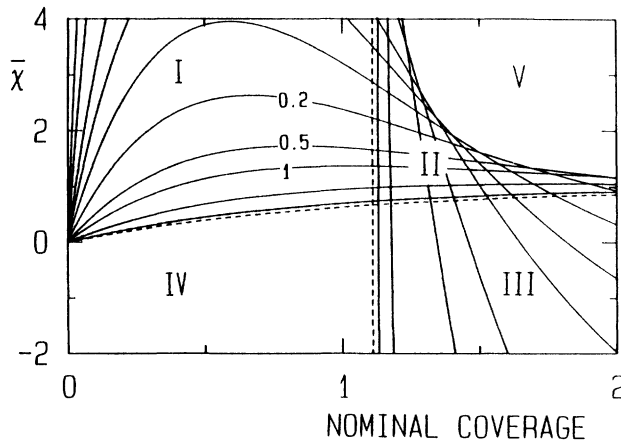

Figure 9. The mean characteristic, $\bar{\chi}$, as a function of the nominal coverage, Ψ, for $F = 1$ (circular window) and $f = 0.9$. Curves are given for different ratios a/A (100, 5, 1, 0.5, 0.2, 0.1, 0.06, 0.04, 0.02, 10^{-2}, 10^{-3}, 10^{-4}); the parameter is given for some of the curves. The broken lines correspond to the limits $a \rightarrow 0$ and $a \rightarrow \infty$. The different regions referred to in the text are indicated (I to V)

different mean areas, a. For $\Psi \leq 1/f$ and $\bar{\chi} > \bar{a}/A$ the positive sign must be chosen (region I); for $\Psi > 1/f$ and $\bar{\chi} \leq \bar{a}/A$ the negative sign must be taken (region III); for $\Psi > 1/f$ and $\bar{\chi} > \bar{a}/A$ both signs apply, i.e. the solution is not unique (region II).

Combinations with $\Psi \leq 1/f$ and $\bar{\chi} \leq \bar{a}/A$ do not occur (region IV), nor are there combinations with $\Psi > 1/f$ and $\bar{\chi} > \bar{a}/A + fF\Psi(1 - \bar{a}/A)/(f\Psi - 1)$ (region V). When observed values lie in region IV or V, the argument for the first square root is negative, and no solutions exist. This possibility and the absence of a unique solution in region II are, however, of little pragmatic concern. The simplified estimation is mainly of interest at low nominal coverages ($\Psi < 1/f$) and for small figures ($a \ll A$). Observed values in region IV are then unlikely.

Applying the simplified estimation to the example of Figure 8 one obtains the estimate $\hat{n} = 89.5$. This is in fair agreement with the estimate based on all three observed variables. The simplified method may be the optimal procedure with existing instruments. However, modified instruments are desirable that include the determination of the perimeter, β, and permit the general estimate.

Acknowledgement

Appendix. Numerical evaluation

For the computer simulations random images are formed and analysed on a square grid of 600 parallel lines. With this line number systematic errors of the estimates became apparent at high nominal coverage ($\Psi > 1$) where the sizes of some holes become comparable to the line spacing. With larger line numbers computations were feasible even at $\Psi = 2$.

The covered area, α, of the window is estimated as the fraction of total line length covered by the figures. The boundary length, β, of the random configuration I (excluding, as stated, the parts formed by the boundary of the window) is estimated from the number, ν, of border points of covered line segments in the interior of the window (endpoints on the boundary of the window are not included):

(A.1) $$\hat{\beta} = \nu \pi A / 2L$$

where L is the total line length and A the area of the grid (in the present examples $L = 600$, $A = 1$). Equation (A.1) reflects the fact that the probability of

a randomly oriented line element, $d\beta$, to intersect a grid with line length, l, per unit area is $2 \cdot l \cdot d\beta/\pi$([12], [14], see also [4]).

In the case of anisotropy the boundary length, β, is determined by summation of the distances between all neighbouring endpoints of covered segments on adjacent lines. Evaluation of these boundary elements provides also the directional distribution of the boundary. While this method has been used for the present computer simulations, other procedures may be more convenient in work with image-analysis equipment.

The determination of the characteristic, χ, i.e. the number of clumps minus voids, requires the comparison of segments on adjacent lines. As common with image-analysis equipment, the scoring is performed by adding up what one may call the number of *emergence* points and subtracting the number of *convergence* points. These notions are illustrated in the diagram of Figure 10. An emergence point is scored if a covered segment on a line is entirely contained within an open segment of the preceding line. A convergence point is registered if a covered interval on a line contains an open interval on the preceding line. This counting procedure yields the characteristic; e.g., it counts 1 for a simply connected object and 0 for a connected domain with one hole.

At high nominal coverages errors can arise in the scoring of emergence and convergence points; they can be particularly serious for closely adjacent, nearly parallel boundaries. Figure 11 indicates the nature of such errors that can lead to spurious negative counts. To reduce this type of error one can nominally increase the length of the open intervals in the scoring of convergence points. In the present computations it has been found adequate to widen the interval at each end by a distance that equals the spacing of the lines in the grid. Some
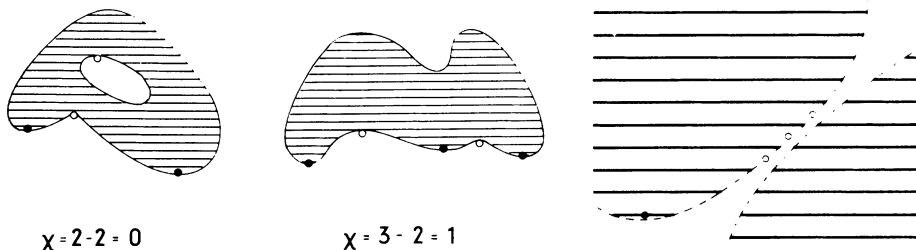

$$\chi = 2-2 = 0 \qquad\qquad \chi = 3 - 2 = 1$$

Figure 10. The notion of emergence points ($\bullet$) and convergence points ($\circ$). The scanning begins with the bottom lines. The characteristic is equal to the number of emergence minus convergence points

Figure 11. Example of a configuration where spurious convergence points ($\circ$) are obtained with the usual method of automatic scoring. The error can be reduced by a suitable enlargement of the open intervals

systematic errors will, nevertheless, persist with this simple method of counting clumps minus holes. The standard deviations obtained in the numerical examples may therefore be somewhat too large at high nominal coverage.

References

[1] ARMITAGE, P. (1949) An overlap problem arising in particle counting. *Biometrika* **36**, 257–266.

[2] DAVY, P. J. (1976) Projected thick sections through multi-dimensional particle aggregates. *J. Appl. Prob.* **13**, 714–722.

[3] DAVY, P. J. (1978) Stereology — A Statistical Viewpoint. Thesis, Australian National University.

[4] KELLERER, A. M. (1983) On the number of clumps resulting from the overlap of randomly placed figures in a plane. *J. Appl. Prob.* **20**, 126–135.

[5] KELLERER, H. G. (1984) Minkowski functions of Poisson processes. *Z. Wahrscheinlichkeitsth.* **67**, 63–84.

[6] KENDALL, M. G. AND MORAN, P. A. P. (1963) *Geometrical Probability*. Griffin, London.

[7] LINDGREN, B. W. (1976) *Statistical Theory*, 3rd edn. The Macmillan Company, New York.

[8] MACK, C. (1954) The expected number of clumps when convex laminae are placed at random and with random orientation on a plane area. *Proc. Camb. Phil. Soc.* **50**, 581–585.

[9] MACK, C. (1956) On clumps formed when convex laminae or bodies are placed at random in two or three dimensions. *Proc. Camb. Phil. Soc.* **52**, 246–256.

[10] MATHERON, G. (1967) *Eléments pour une théorie des milieux poreux*. Masson, Paris.

[11] MATHERON, G. (1974) *Random Sets and Integral Geometry*. Wiley, New York.

[12] MORTON, R. R. A. (1966) The expected number and angle of intersections between random curves in a plane. *J. Appl. Prob.* **3**, 559–562.

[13] ROACH, S. A. (1968) *The Theory of Random Clumping*. Methuen, London.

[14] SANTALÓ, L. A. (1976) *Integral Geometry and Geometric Probability*. Addison-Wesley, London.

[15] WEIL, W. AND WIEACKER, J. A. (1984) Densities for stationary random sets and point processes. *Adv. Appl. Prob.* **16**, 324–346.