Considerations about the Estimation of the Size Distribution in Wicksell's Corpuscle Problem

Joachim Ohser[1] and Konrad Sandau[2]

Abstract

Wicksell's corpuscle problem deals with the estimation of the size distribution of a population of particles, all having the same shape, using a lower dimensional sampling probe. This problem was originary formulated for particle systems occurring in life sciences but its solution is of actual and increasing interest in materials science. From a mathematical point of view, Wicksell's problem is an inverse problem where the interesting size distribution is the unknown part of a Volterra equation. The problem is often regarded ill-posed, because the structure of the integrand implies unstable numerical solutions. The accuracy of the numerical solutions is considered here using the condition number, which allows to compare different numerical methods with different (equidistant) class sizes and which indicates, as one result, that a finite section thickness of the probe reduces the numerical problems. Furthermore, the relative error of estimation is computed which can be split into two parts. One part consists of the relative discretization error that increases for increasing class size, and the second part is related to the relative statistical error which increases with decreasing class size. For both parts, upper bounds can be given and the sum of them indicates an optimal class width depending on some specific constants.

1 Introduction

Wicksell's corpuscle problem is one of the classical problems in stereology (Wicksell, 1925; Bach, 1963; Weibel, 1980). A set of spheres having an unknown distribution of diameters is hit by a section, see Figures 1a) to 1d). The unknown distribution has to be estimated using the observable distribution of the diameters of the section profiles. A large number of papers have been published since the first paper of Wicksell in 1925. Different types of

[1]Institute of Industrial Mathematics, Erwin-Schrödinger-Straße, D-67663 Kaiserslautern, Germany
[2]Faculty of Mathematics and Science, University of Applied Sciences, Schöfferstr. 3, D-64295 Darmstadt, Germany

solutions and different kinds of generalizations are studied in those papers. The reviews of Exner (1972) and Cruz-Orive (1983) can give a first impression, reflecting also the historical development and the amount of publications. The various methods are summarized in the book of Stoyan *et al.* (1995) where also more recent developments are considered. One kind of generalization deals with the thickness of the probes and with the dimension of sampling. We will distinguish here sections of a given thickness (thin sections) and ideal sections (planar sections). The cases of sampling are shown in Table 1.

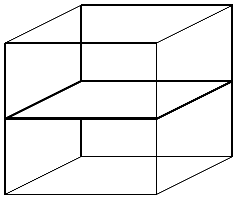
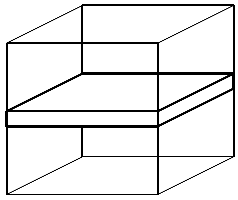
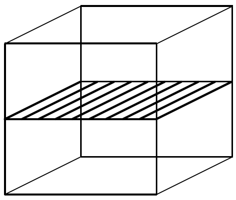
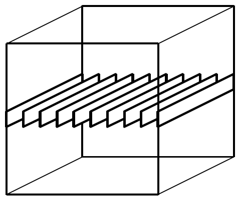
	Planar section	Thin section
Planar sampling		
Linear sampling		

Table 1: Survey of sectioning and sampling used for stereological estimation of particle size distribution. Using planar sampling design, a microstructure is observed in a planar window while linear sampling design uses test segments (commonly a system of parallel equidistant segments).

Another kind of generalization concerns the shape of the objects (or particles). In materials science it is convenient to use particle models, more often than in life sciences. The following models are in use:

- Spheres, spheroids, ellipsoids,
- cubes, regular polyhedra, prisms,
- lamellae, cylinders, and
- irregular convex polyhedra (with certain distribution assumptions).

A survey of stereological methods for systems of non-spherical particles is given in Ohser and Mücklich (2000).

Assuming that only one of the mentioned shapes appears in a particle population but with varying size, it is possible to completely transfer the principles of the solution, see Ohser and

Nippe (1997) and Mehnert *et al.* (1998), which are shown below for spheres. It should be mentioned that the section profiles of all other shapes are more informative than the sections of a sphere. In particular, if the particles are cubic in shape then the section profiles form convex polygons, and from the size and shape of any section polygon with more than three vertices the edge length of the corresponding cube can be determined, see Ohser and Nippe (1997) and Nippe and Ohser (1999).

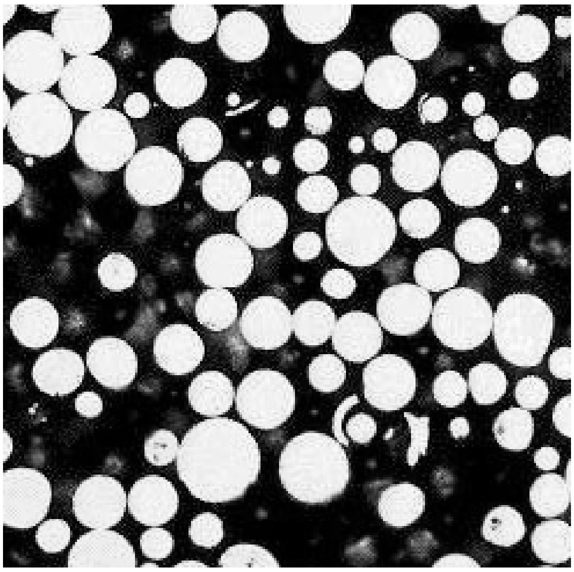
Stereological methods always depend on the assumptions of randomness. In materials analysis the spatial particle system is assumed to be homogeneous and isotropic. The principle behind this assumption of randomness is called the model-based approach. In this case information about the spatial structure can be obtained from an arbitrarily chosen planar or thin section. We use this model-based approach here instead of the design-based approach. The latter requires a deterministic aggregate of non-overlapping particles sampled by means of a planar section randomized in position and direction, see Jensen (1984) and Jensen (1998).

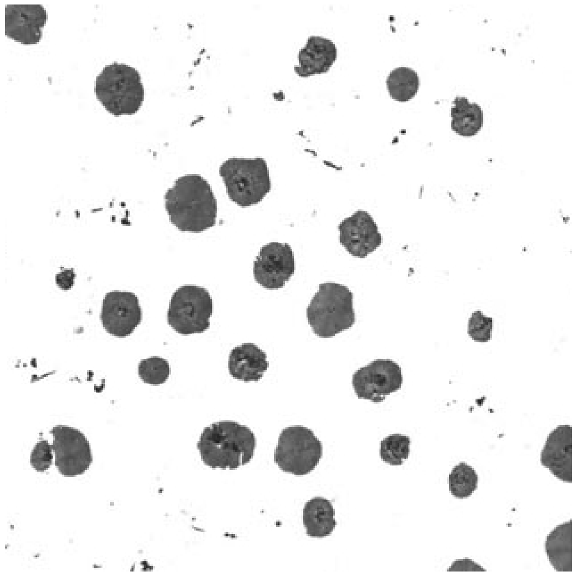
In the last 15 years other stereological methods, working without assuming a particular shape were often preferred in applications. A recently published book about stereology does not even mention Wicksell's classical problem (Howard and Reed, 1998). However, if stronger assumptions on the particle system are justified, they can help to get comparable results with less effort. Methods, which estimate the size distribution of particles without shape assumption need considerably more effort in measuring than methods assuming spherical shape of the particles. Even probing and sampling can be simplified if stronger assumptions are fulfilled.

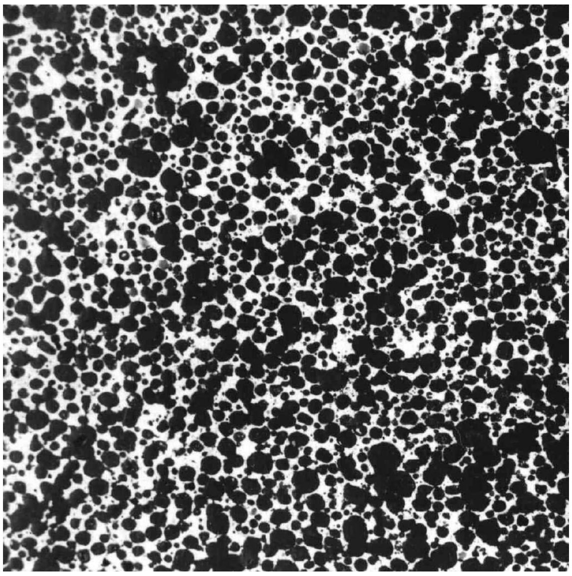
There is a considerable number of mathematically orientated papers dealing with Wicksell's corpuscle problem. This becomes understandable, looking to the interesting mathematical part of the problem in more detail. Considering the probability of appearing profiles in the section one yields an integral equation, determining the wanted distribution of the particle size, but only in an implicit form. We call this the direct equation. Wicksell's problem is often called an inverse problem because inversion of the equation leads to an explicit integral representation of the diameter distribution of the spheres. However, the kernel of the integration has now a singularity, and because of the resulting numerical problems it is called ill-posed or improperly posed. Jakeman and Anderssen (1975) characterize this property in the following way: *Small perturbations of the data can correspond to arbitrarily large perturbations of the solution.*

The solution given here, follows in a first step the proposal of Saltykov (1967), who solved the implicit equation numerically. To do this, the range of the size must be partitioned into classes. The size of the classes plays an important role. The smaller the class size, the better the approximation of the wanted distribution by a step function. The larger the class size, the more each class is occupied with data, thereby improving the statistical estimation of the class probability. Finally, the larger the class size, the better also the numerical stability. Obviously, the advantages go into different directions and thus we have to seek an optimum size.

In this paper we want to show some relations, illuminating in which way this optimum is affected. Discretization and regularization of the integral changes the direct equation to a


a)


b)


c)

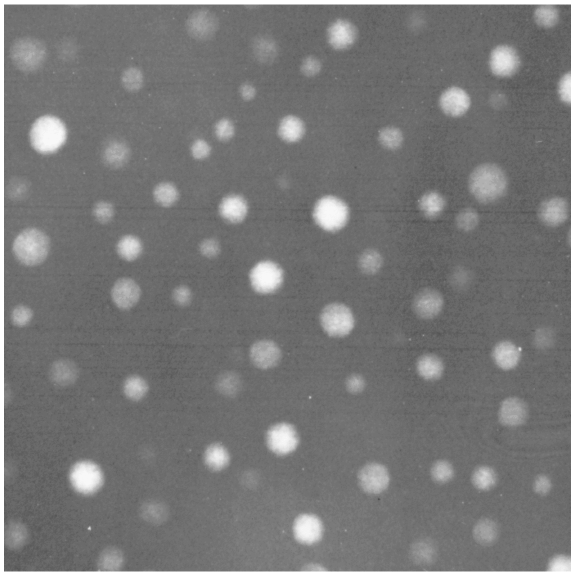

d)

Figure 1: Systems of spherically shaped objects observed in a planar or thin section. a) A metal powder embedded in a resin matrix. The section profiles of the spherical particles are observed in a planar section. b) Cast iron with spherolitic graphite, planar section. c) Pores in gas concrete, planar section. d) Particles in the nickel-base superalloy Nimonic PE16, orthogonal projection of a thin section.

matrix equation i.e. a system of linear equations. The relative deviation of the estimation of the outcome has an upper bound, depending on the product of relative deviation of the input and the so-called condition number. The condition number is a specific number for the matrix of a system of linear equations indicating the behavior of the system under small deviations of the input vector (i.e. the right-hand side of the linear equation). A smaller condition number indicates an outcome with smaller deviation. We use the condition number to evaluate different methods and different class sizes. The results we show in §4 are in contradiction to the numerical supposition that a regularization with a quadrature rule of higher order should be preferred (Mase, 1995). This is demonstrated here for rectangular and trapezoidal quadrature rule. The results using the condition number also quantify the reduction of instability with respect to growing section thickness.

An extension of this classical method is offered by the EM algorithm, see Little and Rubin (1987), which is first applied in Stereology by Silverman *et al.* (1990). This algorithm is presented here as an alternative method for the solution of the discretized stereological equation. In the EM algorithm, an approximation is calculated not only with respect to numerical but also with respect to the statistical aspects. Practical experience shows that the EM algorithm works even under weak assumptions (Stoyan *et al.*, 1995, p. 359). The results based on the condition number can be transferred in a similar way to the EM algorithm. In conclusion, these results suggest that the recommended method is in particular well suited for practical use.

2 The stereological equation

In the following we confine ourselves to the classical case where the particles form a system of spheres having random diameters. Other shape assumptions are touched on only briefly. Figures 1a) to 1d) show typical examples of spherically shaped particles (or pores) occurring in materials science. The Cu-35Sn powder shown in Figure 1a) has been embedded in an opaque resin in order to estimate the sphere diameter distribution by means of stereological methods. Cast iron with spherolitic graphite, see Figure 1b), is the playground for the application of stereology in materials science. The graphite form occurs by inoculation with metals such as manganese and cerium. The spherolites consist of a number of crystallites which grow outward from a common center. The microstructure of gas concrete, see Figure 1c), can be described by a system of overlapping spheres. The spherically shaped pores were infiltrated with black resin. Since the spherical pores touch each other, it is a delicate problem of image analysis to isolate the section circles. Anyway, if a representative sample of circle diameters is available, the sphere diameter distribution can be estimated by means of stereological methods. Figure 1d) shows a dark-field transmission electron micrographs of coherent γ'-particles in the nickel-base superalloy Nimonic PE16 (undeformed state). The image can be considered as an orthogonal projection of a thin section through a system of opaque spheres lying in a matrix which is transparent with respect to the applied radiation. For a sphere with its center lying in the slice, the diameter is equal to the diameter of its projected circle. In the other case, when the sphere center is outside the slice, a section circle is observed. In the projection image the two cases cannot be distinguished.

We are interested in the stereological determination of the following two spatial characteristics:

N_V – the mean number of spheres per unit volume (density of spheres),

F_V – the distribution function of the diameters of the spheres, $F_V(u) := \mathbb{P}(\text{diameter} \leq u)$.

It is possible to estimate these two quantities using data sampled in images of planar or thin sections, and in both cases planar or linear sampling designs can be applied. As a first step, one usually estimates the following quantities:

λ – the density of the section profiles,

F – the size distribution function of the section profiles, $F(s) := \mathbb{P}(\text{profile size} \leq s)$.

The second step consists of the solution of a stereological equation which can be given in the form

$$\lambda\,[1 - F(s)] = N_V \int_s^\infty p(u,s)\,\mathrm{d}F_V(u), \quad s \geq 0. \tag{1}$$

This integral equation is a Volterra equation of the first kind. Here $F_V(u)$ is the unknown distribution while the left-hand side $\lambda\,[1 - F(s)]$ is known from measurement in a planar or thin section. The function of two variables $p(u,v)$ is called the *kernel* of the integral equation.

The analytical expression for the kernel $p(u,s)$ of (1) depends on the applied sampling design, see Table 2. The kernel also carries information about the shape of the objects, which are in our case spheres. Kernels for non-spherical particles are given e.g. in Cruz-Orive (1976), Gille (1988), and Ohser and Nippe (1997). The function $F(s)$ in (1) stands for the distribution function $F_A(s)$ of the circle diameters and the distribution function $F_L(s)$ of the chord lengths, respectively. In the case of planar sampling design, the density λ is the mean number of circles per unit area, $\lambda = N_A$, and for linear sampling design the density λ is the mean number of chords per unit length.

	Planar section	Thin section
Planar sampling	$\sqrt{u^2 - s^2}$	$\sqrt{u^2 - s^2} + t$
Linear sampling	$\frac{\pi}{4}\,(u^2 - s^2)$	$\frac{\pi}{4}\,(u^2 - s^2) + t\,\sqrt{u^2 - s^2}$

Table 2: The kernel $p(u,s)$ of the stereological integral equations for $0 \leq s \leq u$ corresponding to the four sampling methods applied in stereology, see Table 1, in the case of spherically shaped particles. The parameter t is the thickness of the slice.

The problem of estimation of F_V is said to be an *inverse problem*, and when using this term one is tempted to ask for the *direct problem*. Consider for example planar sampling design in a planar section. Let $\lambda = N_A$ be the density of the section circles (the mean number of section circles per unit area). Using

$$N_A[1 - F_A(s)] = N_V \int_s^\infty \sqrt{u^2 - s^2}\, \mathrm{d}F_V(u), \quad s \geq 0, \tag{2}$$

the circle diameter distribution function F_A can be straightforwardly computed from F_V. This is the direct problem. In the given case an analytical solution of the inverse problem is known:

$$N_V[1 - F_V(u)] = N_A \frac{2}{\pi} \int_u^\infty \frac{\mathrm{d}F_A(s)}{\sqrt{s^2 - u^2}}, \quad u \geq 0, \tag{3}$$

i.e. we obtain analytically F_V from F_A. Since the formulation of one problem involves the other one, the computation of F_A from F_V and the computation of F_V from F_A are inverse to each other. Frequently, the solutions of such inverse problems do not depend continuously on the data. Furthermore, given any distribution function F_A then the formal solution F_V of the inverse problem is not necessarily a distribution function. Mathematical problems having these undesirable properties are referred to as *ill-posed problems*. Indeed, the stereological estimation of sphere diameter distribution is often criticized because of the ill-posedness of the corresponding stereological integral equation.

Applying a kernel to a function, as on the right-hand side of (1), is generally a smoothing operation (convolution). Therefore, numerical solution (deconvolution) which requires inverting the smoothing operator will be sensitive to small errors (noise) of the experimentally determined left-hand side, and thus integral equations such as Fredholm equations of the first kind are often ill-conditioned. However, the Volterra equation (1) tends not to be ill-conditioned since the lower limit of the integral yields a sharp step that destroys the smoothing properties of the kernel function. Hence, stereological estimation of sphere diameter distribution is a 'good-natured' ill-posed problem. As we will see, the condition numbers of the operators which correspond to the kernels in Table 2 are relatively small.

3 Numerical solutions of the stereological equation

There is a close correspondence between linear integral equations which specify linear relations among distribution functions and linear equations which specify analogous relations among vectors of relative frequencies. Thus, for the purpose of numerical solution, the stereological integral equation (1) is usually transformed into a matrix equation

$$y = P\,\vartheta \tag{4}$$

where ϑ is the vector of frequencies of the sphere diameters and y is the vector of frequencies of the profile sizes. Comparing with (1), we see that the kernel of the Volterra integral equation corresponds to a matrix $P = (p_{ki})$ that is upper triangular, with zero entries below the diagonal. This matrix equation is straightforwardly soluble by backward substitution.

To facilitate reconstruction, we introduce a discretization of the size s, i.e. the estimation of the distribution function $F(s)$ is carried out in histogram form. For simplification we consider a discretization where the bins (classes) are equally spaced. Given a constant class width Δ such that $y_k = \lambda[F(k\Delta) - F((k-1)\Delta)]$, $k = 1, \ldots, n$, from (1) one can easily derive the formula

$$y_k = N_V \int_s^\infty [p(u, (k-1)\Delta) - p(u, k\Delta)] \, \mathrm{d}F_V(u), \quad k = 1, \ldots, n. \tag{5}$$

The next step in the solution of (1) is a discretization of $F_V(u)$. A possible discretization is based on the Nystrom method which requires the choice of some approximate quadrature rule of the kind $\int f(x)\,\mathrm{d}x = \sum w_i f(x_i)$, where the w_i are the weights of the quadrature rule. This method involves $O(n^3)$ operations, and so from the numerical point of view the most efficient methods tend to use high-order quadrature rules like Simpson's rule or Gaussian quadrature to keep n as small as possible. Mase (1995) suggests using the Gauss-Chebychev quadrature rule. For smooth, nonsingular problems, Press *et al.* (1992, p. 792) concluded that nothing is better than Gaussian quadrature. However, as pointed out by Nippe and Ohser (1999), for many stereological problems concerned with non-spherical particles, analytical expressions for the corresponding kernel are not available, and the coefficients p_{ki} of P have to be computed numerically, see also Ohser and Mücklich (1995). For this reason, a geometric or probabilistic interpretation of the p_{ki} may be helpful. This is in fact possible only for the rectangular quadrature rule. Since in practical application experimental noise dominates, the accuracy of approximation is not as important as in numerical mathematics, and thus low-order quadrature rules are preferred to solve stereological integral equations. Furthermore, as shown in the following, low-order quadrature rules involve some kind of regularization. Two simple examples of low-order methods are given in the following, the repeated rectangular quadrature rule and the repeated trapezoidal quadrature rule.

3.1 Repeated rectangular quadrature rule

We will focus attention on one particular interval $[(i-1)\Delta, i\Delta)$. The sphere diameter is assumed to be a discrete random variable, i.e. the quadrature rule described here is based on the assumption that $F_V(u)$ is constant within each bin,

$$F_V^{app}(u) = a_i, \qquad (i-1)\Delta \le u < i\Delta, \; i = 1, \ldots, n, \tag{6}$$

with $a_0 = 0$, $a_{i-1} < a_i$ and $a_n = 1$. Let ϑ_i be the mean number of spheres per unit volume with diameters equal to $i\Delta$, $\vartheta_i = N_V\,[a_i - a_{i-1}]$. Then (1) can be rewritten as a linear equation system

$$y_k = N_V \sum_{i=1}^n p_{ki}\,[a_i - a_{i-1}] = \sum_{i=1}^n p_{ki}\,\vartheta_i, \quad k = 1, \ldots, n, \tag{7}$$

whose matrix analog is as (4). The coefficients p_{ki} are given by

$$p_{ki} = p(i\Delta, (k-1)\Delta) - p(i\Delta, k\Delta), \quad i \geq k,$$

and $p_{ki} = 0$ otherwise. Note that for $p(u,s) = \sqrt{u^2 - s^2}$ the well-known Scheil-Saltykov-Schwartz method is obtained, see e.g. Saltykov (1974), while $p(u,s) = (\pi/4)(u^2 - s^2)$ yields the Bockstiegel method (Bockstiegel, 1966).

3.2 Repeated trapezoidal quadrature rule

Now we turn to linear interpolation of $F_V(u)$ in the interval $[(i-1)\Delta, i\Delta)$,

$$F_V^{app}(u) = a_i + b_i\,[u - (i-1)\Delta],$$

i.e. the sphere diameter distribution is assumed to be uniform within the single histogram class. Furthermore, we suppose that the distribution function $F_V(u)$ is continuous, even at the points $i\Delta$, i.e. $a_i + b_i\Delta = a_{i+1}$. Hence $\vartheta_i = N_V\,[a_{i+1} - a_i]$ and

$$N_V\, b_i = N_V\, \frac{a_{i+1} - a_i}{\Delta} = \frac{\vartheta_i}{\Delta}.$$

In this particular case the integral equation will be solved with a repeated trapezoidal quadrature rule:

$$y_k = N_V \sum_{i=1}^{n} b_i \int\limits_{(i-1)\Delta}^{i\Delta} [p(u,(k-1)\Delta) - p(u,k\Delta)]\; \mathrm{d}u = \sum_{i=1}^{n} p_{ki}\,\vartheta_i, \quad k = 1,\ldots,n.$$

We obtain an equation system analogous to (7), but now the coefficients p_{ki} are given by

$$p_{ki} = \frac{1}{\Delta} \int\limits_{(i-1)\Delta}^{i\Delta} [p(u,(k-1)\Delta) - p(u,k\Delta)]\; \mathrm{d}u, \quad i \geq k,$$

and $p_{ki} = 0$ otherwise. With $q(u,s) = \int_0^u p(u',s)\mathrm{d}u'$ this formula can be rewritten as

$$p_{ki} = \frac{1}{\Delta}\Big[q\Big(i\Delta,(k-1)\Delta\Big) - q\Big(i\Delta,k\Delta\Big) - q\Big((i-1)\Delta,(k-1)\Delta\Big) + q\Big((i-1)\Delta,k\Delta\Big)\Big]$$

for $i \geq k$. Notice that the repeated trapezoidal quadrature rule was first suggested for the special case of planar sampling in a planar section by Blödner *et al.* (1984). For this particular case the function $q(u,s)$ is given by

$$q(u,s) = \frac{1}{2}\left(u\sqrt{u^2 - s^2} - s^2 \operatorname{arccosh}\frac{u}{s}\right), \quad 0 \leq s \leq u,$$

and $q(u,s) = 0$ otherwise.

3.3 Application of the EM algorithm

The stereological estimation of the vector ϑ of relative frequencies of sphere diameters entails both statistical and numerical difficulties. A natural statistical approach is by maximum likelihood, conveniently implemented using the EM algorithm. The EM algorithm was first applied in stereology by Silverman *et al.* (1990). In the EM approach, y is said to be the vector of *incomplete data* obtained by planar or linear sampling, while ϑ represents the parameter vector to be estimated. The EM algorithm is an iterative procedure which increases the log-likelihood of a current estimate ϑ^λ of ϑ.

Each iteration step consists of two substeps, an E-substep and an M-substep, where E stands for 'expectation' and M for 'maximization'. Let c_{ki} be the number of events occurring in bin i which contribute to the count in bin k; thus (c_{ki}) is called the matrix of complete data. The E-substep yields the expected value of the c_{ki}, given the incomplete data y, under the current estimate ϑ^ν of the parameter vector ϑ. The M-substep yields a maximum likelihood estimate of the parameter vector ϑ using the estimated complete data c from the previous E-substep.

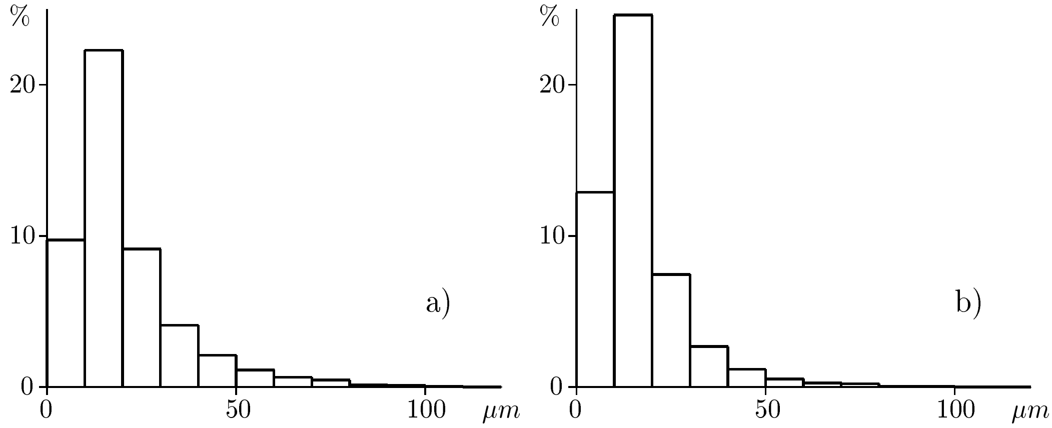


Figure 2: Stereological estimation of the sphere diameter distribution for the cast iron with spherolitic graphite shown in Figure 1b): a) Histogram of the sizes of section profiles observed in a planar section, b) histogram of the particle size. The sample size was 6995 and the total area of the sampling windows was $A(W) = 18.207\,\mathrm{mm}^2$. (The data have been obtained from 144 sampling windows.) This yields $N_A = 384\,\mathrm{mm}^{-2}$, and the particle density N_V is obtained from $N_V = \sum_i \vartheta_i = 17\,545\,\mathrm{mm}^{-3}$. The coefficients of the discretized stereological equation system have been obtained from the repeated rectangular quadrature rule, and the equation system has been solved by means of the EM-algorithm (32 EM-steps).

Define q_i and r_k to be $q_i = \sum_{k=1}^{i} p_{ki}$ and $r_k = \sum_{i=k}^{n} p_{ki}\,\vartheta_i^\nu$, respectively. Then the E-substep is $c_{ki} = \vartheta_i^\nu\, y_k\, p_{ki}/r_k$ for each i and k, and the M-substep is given by $\vartheta_i^{\nu+1} = \frac{1}{q_i}\sum_{k=1}^{i} c_{ki}$, $i = 1, \ldots, n$. Combining these two substeps we obtain an EM-step given by the updating formula

$$\vartheta_i^{\nu+1} = \frac{\vartheta_i^\nu}{q_i} \sum_{k=1}^{i} \frac{p_{ki}}{r_k} y_k, \quad i = 1, \ldots, n. \tag{8}$$

This formula yields a sequence $\{\vartheta^\nu, \nu = 0, 1, \ldots\}$ of solutions for the vector ϑ. The convergence of the algorithm is guaranteed in theory, but the solution depends on the chosen initial value ϑ_i^0. We propose using $\vartheta_i^0 = y_i$ for $i = 1, \ldots, n$, see Ohser and Mücklich (1995). The choice of non-negative initial values ensures that each of the ϑ^ν is automatically non-negative. Furthermore, when ν is not too big the EM algorithm includes some kind of regularization.

An example of application of the EM algorithm is shown in Figure 2.

We remark that the M-substep can be theoretically justified as a maximization step when the sphere diameters are independent and the sphere centers form a Poisson point field. In practice these assumptions are frequently unrealistic. However, the EM algorithm has been found by long experience to be reasonably adequate for practical purposes (cf. also the remarks in Stoyan 1995, p. 359).

4 Bounds of stereological estimation

The question most frequently asked of a statistican by the experimental researcher is: *How large a sample size do I need?* Vice versa, given the sample size, one may ask the question: *What is the optimal bin size?* In fact, the answer can be given if the estimation error is known. Unfortunately, there is no simple way to derive an expression for the error in the stereological estimation. A practical method for the error estimation is to divide the sample into two subsamples of equal size, say, and treat the difference between the two estimates as a conservative estimate of the error in the result obtained for the total sample size. However, this method cannot be used to compare properties of stereological estimation methods in general.

4.1 The condition number of the operator P

In numerical mathematics it is common to describe the behavior of a linear equation system like (4) by the condition number of the matrix P. We attempt to give a statistical interpretation: Let $\widehat{y}$ be an estimator of the vector y. The difference $\widehat{y} - y$ is the experimental measurement noise, and $\mathbb{E}|\widehat{y} - y|/|y|$ is the relative statistical error of $\widehat{y}$. The estimator $\widehat{\vartheta}$ of ϑ may be obtained via $\widehat{\vartheta} = P^{-1}\widehat{y}$ where P^{-1} is the inverse matrix of P. A well-known estimate of the relative statistical error $\mathbb{E}|\widehat{\vartheta} - \vartheta|/|\vartheta|$ of $\widehat{\vartheta}$ is given by

$$\frac{\mathbb{E}|\widehat{\vartheta} - \vartheta|}{|\vartheta|} \le \operatorname{cond}(P)\, \frac{\mathbb{E}|\widehat{y} - y|}{|y|}. \tag{9}$$

The condition number is defined as $\text{cond}(P) = |P| \cdot |P^{-1}|$ where the matrix norm corresponds to the Euclidean vector norm. It is an essential quantity giving an indication of the accuracy of numerical solutions of stereological problems; it describes the accuracy of numerical solutions independently of assumptions made for the sphere diameter distribution, see Kanatani and Ishikawa (1985), Gerlach and Ohser (1986), and Nagel and Ohser (1988).

n	$t_0 = 0$	$t_0 = \frac{1}{4}$	$t_0 = \frac{1}{2}$	$t_0 = 1$
4	3.10	2.04	1.69	1.41
8	5.68	2.53	1.91	1.50
12	8.28	2.77	2.00	1.53
16	10.88	2.92	2.05	1.55
20	13.48	3.01	2.09	1.56
24	16.08	3.08	2.11	1.57
28	18.68	3.14	2.13	1.58
32	21.29	3.18	2.14	1.58

n	$t_0 = 0$	$t_0 = \frac{1}{4}$	$t_0 = \frac{1}{2}$	$t_0 = 1$
4	10.82	6.40	5.19	4.30
8	38.30	14.53	10.77	8.42
12	82.69	23.05	16.42	12.55
16	143.98	31.72	22.10	16.69
20	222.19	40.47	27.79	20.83
24	317.30	49.25	33.49	24.97
28	429.31	58.06	39.19	29.12
32	558.25	66.88	44.90	33.26

Table 3: The condition number of the matrix P for the repeated rectangular quadrature rule, depending on the class number n and a normalized slice thickness $t_0 = t/d_V$ where d_V is the mean sphere diameter. The left table gives numerical values of the condition number for planar sampling design, and the right one presents those for linear sampling design.

The Tables 3 and 4 present numerical values of the condition number of the matrix P based on the spectral norm $|P| = \max_{x \neq 0} \frac{|Px|}{|x|}$. To compare the statistical accuracy of several numerical methods, the data space is divided into n bins of constant width Δ. There are significant differences in the condition numbers:

1. With decreasing class width Δ (increasing number n of classes) the statistical accuracy decreases.
2. The trapezoidal quadrature rule has no clear advantages over the rectangular one: the approximation is surely improved but the statistical error of the estimates increases rapidly. Hence, for a small sample size the repeated rectangular quadrature rule is preferable to higher-order quadrature rules. Obviously, one cannot improve approximation and regularization simultaneously.
3. With increasing slice thickness t the statistical accuracy increases. Formally, the slice thickness t can be considered as a parameter of regularization. However, in the orthogonal projection of a 'thick' slice, overlapping cannot be avoided and overlapping destroys the information available from the slice.
4. In general, the condition numbers corresponding to linear sampling design are greater than those corresponding to planar sampling design. Thus, stereological estimation based on linear sampling design is more unstable than that based on planar sampling design (cf. also the discussion in Ripley 1981, p. 212).

n	$t_0 = 0$	$t_0 = \frac{1}{4}$	$t_0 = \frac{1}{2}$	$t_0 = 1$
4	5.49	2.39	1.81	1.44
8	11.22	2.86	2.01	1.52
12	16.92	3.05	2.08	1.55
16	22.60	3.16	2.12	1.57
20	28.28	3.22	2.14	1.58
24	33.96	3.27	2.16	1.58
28	39.63	3.30	2.17	1.59
32	45.30	3.32	2.18	1.59

n	$t_0 = 0$	$t_0 = \frac{1}{4}$	$t_0 = \frac{1}{2}$	$t_0 = 1$
4	33.67	13.53	10.11	7.97
8	156.38	32.35	22.55	17.07
12	374.37	51.40	34.97	26.13
16	691.13	70.55	47.39	35.16
20	1 108.98	89.76	59.82	44.19
24	1 629.63	108.99	72.24	53.21
28	2 254.70	128.23	84.66	62.24
32	2 985.27	147.49	97.08	71.26

Table 4: The condition number of the matrix P for the repeated trapezoidal quadrature rule, depending on the class number n and a normalized slice thickness $t_0 = t/d_V$ where d_V is the mean sphere diameter. The left table gives numerical values of the condition number for planar sampling design, and the right one presents those for linear sampling design.

In comparison with other ill-conditioned equations which occur in practice, the condition numbers in Table 3 and Table 4 are rather small and thus the numerical methods for solving the stereological integral equations presented above can be considered as relatively stable. If experimental measurement noise does not dominate then one can expect that the variances of stereological estimators of sphere diameter distribution function are also small.

Finally, we remark that the consideration of weighted sphere diameter distributions can provide smaller condition numbers than the classical (number-weighted) sphere diameter distribution considered in this paper, see Nagel and Ohser (1988).

4.2 The optimal discretization parameter

Consider now the discretization error of stereological estimation. The discretization error can be expressed in terms of the $\mathcal{L}^2$-norm of the difference between the probability densities of the distribution functions $F_V^{app}(u)$ and $F_V(u)$, respectively, where $F_V^{app}(u)$ is a discretization of the sphere diameter distribution. The statistical error and the discretization error – considered as functions of the bin size Δ – are schematically shown in Figure 3. The errors behave in an opposite way. The relative total error of statistical estimation is the sum of the relative statistical error and the relative discretization error. It depends on the sample size and the unknown sphere diameter distribution.

Without loss of generality, let $[0, d]$ be the interval covered by the n bins, i.e. $d = n\Delta$, and let w be the size of the sampling window (the area of a planar window or the length of a test segment). Then an upper bound of the relative total error can be given by the expression

$$\frac{d}{\sqrt{3}}\, c\, \Delta + \left(1 + \frac{d}{\sqrt{3}}\, c\, \Delta\right) \operatorname{cond}(P) \sqrt{\frac{2\, d}{\pi\, \lambda\, w \Delta}},$$

see Appendix, where c is a constant, $c > 0$. The graphs of the relative errors shown in Figure 3 are obtained for $\lambda w = 1000$, $d = 1$, and $c = 8$.

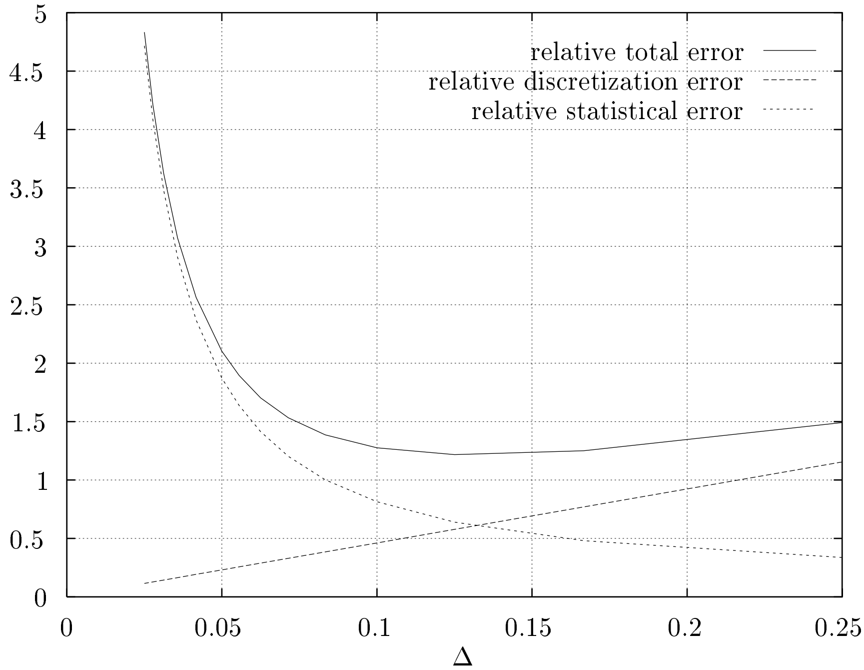


Figure 3: Errors of stereological estimation of particle size distribution based on the upper bounds computed in the Appendix. If the bin size Δ becomes too small, the relative statistical error increases. Vice versa, if Δ is too large, the relative discretization error increases. There is an 'optimal' discretization parameter $\Delta_{opt} \approx 0.125$ (8 bins) which can only be computed explicitly in special cases since it depends on unavailable information on the relative measurement error as well as the distribution function $F_V(u)$ to be estimated.

The choice of the optimal discretization parameter Δ_{opt} would minimize the total error, but even in optimal circumstances one always gets a loss of accuracy. Unfortunately, the optimal discretization parameter cannot be obtained from only the sample data. In conclusion, the results give some hints for designing the statistics and it turns out that the 'simple ways are good ways'. That means that the bin size should not be chosen to be too small and a low-order quadrature rule can be taken. The condition numbers confirm that planar sampling design gives more stable results than linear sampling design, and a positive slice thickness reduces the estimation error as well.

The computation of the upper bound of the error in the stereological estimation presented in the Appendix is based on the particular case of trapezoidal quadrature rule. These considerations can be transferred to other quadrature rules. Finally, we remark that the EM algorithm can be understood as a statistically motivated iterative method for solving the linear equation (4). The application of the EM algorithm does not change any properties of the operator P. Thus, in principle, the results of this section can also be extended to maximum likelihood estimators.

References

Bach, G. (1963) Über die Bestimmung von charakteristischen Größen einer Kugelverteilung aus der Verteilung der Schnittkreise. *Z. wiss. Mikrosk.* **65**, 285.

Blödner, R., P. Mühlig, and W. Nagel (1984) The comparison by simulation of solutions of Wicksell's corpuscle problem. *J. Microsc.* **135**, 61–64.

Bockstiegel, G. (1966) Eine einfache Formel zur Berechnung räumlicher Größenverteilungen aus durch Linearanalyse erhaltenen Daten. *Z. Metallkunde* **57**, 647–656.

Cruz-Orive, L.-M. (1976) Particle size-shape distributions: The general spheroid problem, I. Mathematical model. *J. Microsc.* **107**, 235–253.

Cruz-Orive, L. M. (1983) Distribution-free estimation of sphere size distributions from slabs showing overprojection and truncation, with a review of previous methods. *J. Microsc.* **131** 265–290.

Exner, H. E. (1972) Analysis of grain and particle size distributions in metallic materials. *Int. Metall. Rev.* **17**, 25.

Gerlach, W., and J. Ohser (1986) On the accuracy of numerical solutions such as the Wicksell corpuscle problem. *Biom. J.* **28**, 881–887.

Gille, W. (1988) The chord length distribution density of parallelepipeds with their limiting cases. *Experimentelle Technik in der Physik* **36**, 197–208.

Howard, C. V. and M. G. Reed (1998) *Unbiased Stereology, Three-dimensional Measurement in Microscopy.* Bios Scientific Publisher.

Jakeman, A. J. and R. S. Anderssen (1975) Abel type integral equations in stereology. I. General discussion. *J. Microsc.* **105**, 121–133.

Jensen, E. B. (1984) A design-based proof of Wicksell's integral equation. *J. Microsc.* **136**, 345–348.

Jensen, E. B. V. (1998) *Local Stereology.* World Scientific, Singapore, New Jersey, Hong Kong.

Kanatani, K. I., and O. Ishikawa (1985) Error analysis for the stereological estimation of sphere size distribution: Abel type integral equation. *J. Comput. Physics* **57**, 229–250.

Little, R. J. A., and D. B. Rubin (1987) *Statistical Analysis with Missing Data.* J. Wiley & Sons, New York.

Mase, S. (1995) Stereological estimation of particle size distributions. *Adv. Appl. Prob.* **27**, 350–366.

Mehnert, K., J. Ohser, and P. Klimanek (1998) Testing stereological methods for the estimation of grain size distributions by means of a computer-simulated polycrystalline sample. *Mat. Sci. Eng.* **A246**, 207–212.

Nippe, M. (1998) *Stereologie für Systeme homothetischer Partikel.* Ph.D. thesis, TU Bergakademie Freiberg.

Nagel, W., and J. Ohser (1988) On the stereological estimation of weighted sphere diameter distributions. *Acta Stereol.* **7**, 17–31.

Nippe, M., and J. Ohser (1999) The stereological unfolding problem for systems of homothetic particles. *Pattern Recogn.* **32**, 1649–1655.

Ohser, J., and F. Mücklich (1995) Stereology for some classes of polyhedrons. *Adv. Appl. Prob.* **27**, 384–396.

Ohser, J., and F. Mücklich (2000) *Statistical Analysis of Materials Structures.* J. Wiley & Sons, Chichester.

Ohser, J., and M. Nippe (1997) Stereology of cubic particles: various estimators for the size distribution. *J. Microsc.* **187**, 22–30.

Press, W. H., S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery (1994) *Numerical Recipes in C, 2nd Ed.* Cambridge University Press.

Ripley, B. D. (1981) *Spatial Statistics.* J. Wiley & Sons, Chichester.

Saltykov, S. A. (1967) The determination of the size distribution of particles in an opaque material from a measurement of the size distribution of their sections. In: Elias, H. (Ed.): *Proceedings of the Second International Congress for Stereology.*

Saltykov, S. A. (1974) *Stereometrische Metallographie.* Deutscher Verlag für Grundstoffindustrie, Leipzig.

Silverman, B. W., M. C. Jones, D. W. Nychka, and J. D. Wilson (1990) A smoothed EM approach to indirect estimation problems, with particular reference to stereology and emission tomography. *J. R. Statist. Soc.* **B52**, 271–324.

Stoyan, D., W. S. Kendall, and J. Mecke, (1995) *Stochastic Geometry and its Applications, 2nd Ed.* J. Wiley & Sons, Chichester.

Weibel, E. R. (1980) *Stereological Methods*, Vol. 1: *Practical Methods for Biological Morphology*, Vol. 2: *Theoretical Foundations.* Academic Press, London.

Wicksell, S. D. (1925) The corpuscle problem I. *Biometrica* **17**, 84–89.

Appendix

Let the probability density of F_V be f_V. We assume $f_V = 0$ outside the interval $[u_0, d]$, and f_V is Lipschitz with the Lipschitz constant $c > 0$. The interval $[0, d]$ is divided into n equally spaced bins of width Δ, so that $n\,\Delta = d$. Let b be a piecewise constant approximation of f_V,

$$b(u) = \sum_{i=1}^{n} b_i \, \mathbf{1}_{[(i-1)\Delta, i\Delta)}(u), \qquad u \geq 0,$$

where $\mathbf{1}_A(\cdot)$ is the indicator function of the set A, and b_i are suitable constants as described in Section 3 for F_V^{app}. Section 3 also shows the relation between the relative frequencies (class probabilities) ϑ_i with respect to the distribution function F_V, the relative frequencies y_i with respect to F and furthermore the relation $\vartheta_i = N_V \, \Delta \, b_i, \; i = 1, \ldots, n$.

A sample of sizes of section profiles gives an estimate $\widehat{y}$ of y which gives an estimate $\widehat{\vartheta}$ of ϑ where $\widehat{\vartheta} = P^{-1}\,\widehat{y}$, and $\widehat{\vartheta}$ gives an estimate $\widehat{b}(u)$ of the function $b(u)$.

We are interested to find upper bounds on the relative error. These bounds illustrate the interaction between the statistical error and the discretization error using the $\mathcal{L}^2$-norm $\|\cdot\|$. There is an correspondence between the piecewise constant function b, and the values b_i that form a vector $\underline{b}$ which can be considered with respect to the Euclidean vector norm $|\cdot|$. These cases shall be distinguished here. Clearly, for a function b as defined above, the relationship $\|b\| = \sqrt{\Delta}\,|\underline{b}|$ holds.

We get the following chain of inequalities

$$\begin{aligned}
\frac{\mathbb{E}\|f_V-\widehat{b}\|}{\|f_V\|} &\leq \frac{\|f_V-b\|}{\|f_V\|}+\frac{\mathbb{E}\|b-\widehat{b}\|}{\|f_V\|} = \frac{\|f_V-b\|}{\|f_V\|}+\frac{\|b\|}{\|f_V\|}\,\frac{\mathbb{E}\|b-\widehat{b}\|}{\|b\|} \\
&= \frac{\|f_V-b\|}{\|f_V\|}+\frac{\|b\|}{\|f_V\|}\,\frac{\mathbb{E}|\underline{b}-\widehat{\underline{b}}|}{|\underline{b}|} \\
&\leq \frac{\|f_V-b\|}{\|f_V\|}+\frac{\|b\|}{\|f_V\|}\,\frac{\mathbb{E}|\vartheta-\widehat{\vartheta}|}{|\vartheta|} \\
&\leq \frac{\|f_V-b\|}{\|f_V\|}+\left(1+\frac{\|f_V-b\|}{\|f_V\|}\right)\frac{\mathbb{E}|\vartheta-\widehat{\vartheta}|}{|\vartheta|}.
\end{aligned}$$

Now the two parts are considered separately.

For the term $\mathbb{E}|\vartheta-\widehat{\vartheta}|/|\vartheta|$ an upper bound is obtained from (9). To find a global bound of $\mathbb{E}|y-\widehat{y}|/|y|$, we give bounds of numerator and denominator separately.

A lower bound of the denominator can be obtained in the following way:

$$|y| = \sqrt{\sum_{i=1}^{n} y_i^2} \geq \sqrt{\sum_{i=1}^{n}\left(\frac{\lambda}{n}\right)^2} = \frac{\lambda}{\sqrt{n}}$$

where λ is the density of the section profiles, $\lambda=\sum_{i=1}^{n} y_i$.

An upper bound of the numerator can be get using the extreme case of only one class ($n=1$). Then it is

$$\mathbb{E}\,|y-\widehat{y}| \leq \mathbb{E}\,\sqrt{(\widehat{\lambda}-\lambda)^2} = \mathbb{E}\,|\widehat{\lambda}-\lambda|.$$

Using Stirling's formula one gets the approximation

$$\mathbb{E}\,|\widehat{\lambda}-\lambda| \approx \sqrt{\frac{2}{\pi}}\,\frac{\sqrt{w\,\lambda}}{w} = \sqrt{\frac{2\lambda}{\pi\,w}},$$

see Nippe (1998, p. 45). Here, w is the size of the sampling window W (for planar sampling design the area of W, and for linear sampling design the length of W).

From the above relationships we get

$$\frac{\mathbb{E}\,|y-\widehat{y}|}{|y|} \lesssim \frac{\sqrt{n}}{\lambda}\sqrt{\frac{2\lambda}{\pi\,w}} = \sqrt{\frac{2n}{\pi\,\lambda\,w}}.$$

The second term $\|f_V-b\|/\|f_V\|$ is considered using the Lipschitz condition and the elementary property that there is a $\xi_i\in((i-1)\Delta, i\Delta)$ with $f_V(\xi_i)=b_i$. Then

$$\|f_V-b\|^2 = \int_0^d [f_V(u)-b(u)]^2\,\mathrm{d}u = \sum_{i=1}^{n}\int_{(i-1)\Delta}^{i\Delta}[f_V(u)-f_V(\xi_i)]^2\,\mathrm{d}u$$

$$
\begin{aligned}
&\leq \sum_{i=1}^{n} \int_{(i-1)\Delta}^{i\Delta} c^2 [u - \xi_i]^2 \, \mathrm{d}u = \sum_{i=1}^{n} \frac{c^2}{3} \left[(i\Delta - \xi_i)^3 - ((i-1)\Delta - \xi_i)^3 \right] \\
&\leq \frac{c^2}{3} \Delta^3 n = \frac{c^2}{3} d \, \Delta^2
\end{aligned}
$$

and this results yields

$$
\frac{\|f_V - b\|}{\|f_V\|} \leq \sqrt{\frac{d}{3}} \, \frac{c}{\|f_V\|} \, \Delta
$$

which is a bound, depending on $\|f_V\|$. For an independent bound we are looking for a lower bound of $\|f_V\|$ which is obviously larger than the norm of the uniform distribution f_V^u on the closed interval $[0, d]$ because the Lipschitz condition must be satisfied. So we have for all probability densities f_V

$$
\|f_V\| \geq \|f_V^u\| = \frac{1}{\sqrt{d}}.
$$

Bringing all these results together we obtain

$$
\begin{aligned}
\frac{\mathbb{E} \, \|f_V - b\|}{\|f_V\|} &\leq \frac{\|f_V - b\|}{\|f_V\|} + \left(1 + \frac{\|f_V - b\|}{\|f_V\|} \right) \frac{\mathbb{E} \, |\vartheta - \widehat{\vartheta}|}{|\vartheta|} \\
&\leq \sqrt{\frac{d}{3}} \, \frac{c}{\|f_V\|} \Delta + \left(1 + \sqrt{\frac{d}{3}} \, \frac{c}{\|f_V\|} \Delta \right) \operatorname{cond}(P) \, \frac{\mathbb{E} \, |y - \widehat{y}|}{|y|} \\
&\lesssim \frac{d}{\sqrt{3}} \, c \, \Delta + \left(1 + \frac{d}{\sqrt{3}} \, c \, \Delta \right) \operatorname{cond}(P) \sqrt{\frac{2n}{\pi \, \lambda \, w}} \\
&= \frac{d}{\sqrt{3}} \, c \, \Delta + \left(1 + \frac{d}{\sqrt{3}} \, c \, \Delta \right) \operatorname{cond}(P) \sqrt{\frac{2 \, d}{\pi \, \lambda \, w \Delta}} \; .
\end{aligned}
$$