BULLETIN OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 84, Number 6, November 1978

QUASI-MONTE CARLO METHODS AND PSEUDO-RANDOM NUMBERS

BY HARALD NIEDERREITER[1]

> Nothing in Nature is random. . . . A thing appears random only through the incompleteness of our knowledge.
>
> Spinoza, *Ethics* I

CONTENTS

1. Introduction. The subject matter of this talk is at the crossroads of two areas which will turn out to have more than only an etymological kinship, namely numerical analysis and number theory. Like so many mixed breeds, it has its fascinations and attractions, but also its inherent dilemmas. A multitude of concepts and devices dear to numerical analysts and computer users are, in open or disguised form, of an arithmetic nature, and problems arising in the computational workshop, especially those requiring effective methods, are now treated quite frequently with the powerful tools of the number theorist. This provides for a vivid interplay and is a source of enrichment for both disciplines. Of course, the occasion only permits us to look at a certain segment in the broad spectrum of activities. The leitmotif in our discussion will be the simulation of procedures containing an element of randomness by judiciously chosen deterministic schemes, with number theory playing a

This is an expanded version of an invited address presented at the 83rd Annual Meeting of the Society in St. Louis, Missouri, on January 27, 1977, under the title, *Quasi-Monte Carlo methods and pseudo-random numbers*: *Some applications of number theory*; received by the editors September 23, 1977.

AMS (*MOS*) *subject classifications* (1970). Primary 65-02, 65C05, 65C10, 65D30, 10F40, 10K05; Secondary 10-02, 10A35, 10F10, 10F20, 10G05, 10K30, 12A15, 65D05, 65N05, 65R05, 68A55.

[1] The author gratefully acknowledges support received from NSF grant MCS 77-01699.

prominent part in the construction of such schemes.

Our exposition can be roughly divided into two parts, which will not preclude, however, some strong interrelations between these. §§2–5 are devoted to deterministic versions of Monte Carlo techniques that have come to be known under the collective term "quasi-Monte Carlo methods". The widest range of applications, and indeed the historical origin of these methods, is found in numerical integration, but related matters such as interpolation problems and the numerical solution of integral equations can also be dealt with successfully. It does not seem to be too well known among the craftsmen of the trade that quasi-Monte Carlo methods possess two big assets not shared by standard Monte Carlo techniques, namely effectiveness and fast convergence. It is therefore hoped that this talk will help in the dissemination and eventual adoption of these more efficient methods.

As we already indicated, the term "quasi-Monte Carlo" encompasses a variety of techniques. In their basic form, they all emerged sometime in the 1950s, and another common bond between them is their heavy reliance on number-theoretic concepts. These methods fall into three main categories which are characterized by the titles of §§3–5. For the sake of completeness, we will present the (at least from the standpoint of the specialist) classical foundations of these various techniques, but emphasize the refinements and extensions in scope achieved in recent years. This vigorous progress has been sustained by significant contributions from both the Russian school and the Western branch, with the Chinese doing more than just checking the balance.

In the second part, comprising §§6-11, we are concerned with the vital matter of pseudo-random number generation. The limelight will be on the most popular generators, namely Lehmer's linear congruential pseudorandom numbers. These appeared to fall into disfavor among theoreticians about ten years ago upon the disclosure of a certain undesirable lattice (or "crystalline") structure, but there is every indication that the practitioners did not care much about these squabbles and continued defiantly with their time-honored routines. Recent results of the author amount to a rehabilitation of these generators and justify *a posteriori* the woman (or man) at the computer in her (or his) preservative attitude. To sum up a complex matter succinctly, this research has shown that *a proper choice of parameters* in the generation procedure leads to a sequence of pseudo-random numbers enjoying excellent properties of statistical (almost-) independence among a given number of successive terms. The italicized part of the preceding sentence cannot be emphasized too much. Although not totally ignored, the matter of adequately selecting these parameters has all too frequently been left to chance or was based on insufficient evidence. If, as so often, generators with the said statistical independence properties are required, we can now provide for the first time reliable criteria for the choice of parameters which are buttressed by effective theoretical results.

Lest a contrary impression be created, it should be pointed out that the existence of linear congruential pseudo-random numbers with desirable statistical independence properties does not negate the results about their unfavorable lattice structure. The fact is simply that the lattice structure does not detract from the usefulness of these generators as long as we use the pseudo-random numbers for purposes in which only the statistical independence of successive terms or a good distribution behavior are relevant.[2] Whoever finds that there is still an irreconcilable dilemma here, may want to ponder the following morphological principle: if one attempts to distribute points in a well-planned and equitable manner over a given domain, one will, intentionally or inadvertently, provide the resulting collection of points with an intrinsic structure. Consider a very simple example in which we are challenged to distribute denumerably many points on the real axis in what we deem the fairest way. We will most likely end up with an equally spaced arrangement[3] and have thereby generated a point set with the structure of a one-dimensional lattice. Thus, a lattice structure may even be a virtue when good distribution properties are desired. More will be said about this in §§10 and 11.

The advance of our knowledge in this area of linear congruential pseudo-random numbers has gone hand in hand with progress in pure number theory, namely the establishment of nontrivial estimates for exponential sums with linear recurring arguments. We have found it convenient to relegate these number-theoretic matters to a separate section (see §8). Apart from the topics already mentioned, we shall discuss related work on Lehmer's pseudo-random numbers and similar generators as well as review their elementary properties.

As to the list of references, we have attempted to be fairly complete concerning the literature on quasi-Monte Carlo methods since no comprehensive survey of this area was available up to now. With respect to pseudo-random numbers, the prolific output in this discipline has been assessed periodically in review articles and bibliographies (cf. **[84]**, **[100]**, **[135]**, **[207]**, **[208]**, **[322a]**) and there are monographs covering the subject (cf. **[80]**, **[142]**, **[154]**, **[182]**, **[193]**, **[213]**), so that we have only listed those works having a direct bearing on our discussion.

Before we embark on our exposition proper, it is necessary to gain an understanding of the fundamental principle involved in the Monte Carlo method which is the progenitor of the great bulk of the theory to be expounded here. The *Monte Carlo method* (or "method of statistical trials") may be described in simple terms as a numerical method based on random sampling.[4] We give an illustration below. The history of the method has been charted often enough. Its birth is assumed to have taken place in 1949 with the publication of **[199]**, although it was definitely known earlier to a clandestine group[5] working on U. S. Defense projects. Statisticians have intuitively used its principle long before that (cf. **[334]**), but it was only the computer age that could turn it into a systematic and viable technique. A number of excellent monographs explore the whole range of the method (cf.

[2] Most applications in numerical analysis are of this type.

[3] A Rorschach test could be set up interpreting deviations from this arrangement in psychological terms.

[4] As Sobol' [**306**, p. 9] points out correctly, the method is not of much help in trying to win at roulette.

[5] Mainly J. von Neumann, N. Metropolis, S. M. Ulam, H. Kahn, and their collaborators at the Los Alamos Scientific Laboratory.

[25], [28], [66], [106], [309]). We also recommend the survey article [100], the recent bibliography [84], and the elementary introduction in [306]. The main reason for the popularity of the Monte Carlo method is its applicability to a never-ending variety of problems in numerical analysis, statistics, applied mathematics, particle physics, engineering, systems analysis, and so on. We refer to [28],[29], [70], [83], [106], [151], [200], [323], [340] for accounts of some of these applications. For the general area of simulation, see [17], [27], [69], [79], [190], [201], [209].

To present a simple example of a Monte Carlo calculation, we consider the problem of computing the area of a region E of complicated shape contained in the unit square $[0, 1] \times [0, 1]$. The idea is now to select at random N points from the unit square by performing N independent trials. In practice, N should be fairly large, say $N \approx 10^4$. If $\mathbf{x}_1, \ldots, \mathbf{x}_N$ are the points resulting from the sampling process, then the Monte Carlo approximation is

$$\text{area of } E \approx \frac{1}{N} \sum_{n=1}^{N} c_E(\mathbf{x}_n), \tag{1.1}$$

where c_E is the characteristic function of E. In other words, the fraction of the sample points falling into E is taken as an approximate value for the area of E. This procedure can be generalized immediately if one recognizes the left-hand side of (1.1) as the (Lebesgue) integral of c_E over the unit square (provided that E is Lebesgue-measurable, of course). Thus, for a given dimension $s \geqslant 1$ let $I^s = [0, 1]^s$ be the s-dimensional unit cube and let $f = f(\mathbf{t})$ be a bounded[6] Lebesgue-integrable function onI^s. Then the Monte Carlo approximation for the Lebesgue integral of f over I^s is

$$\int_{I^s} f(\mathbf{t})\, d\mathbf{t} \approx \frac{1}{N} \sum_{n=1}^{N} f(\mathbf{x}_n), \tag{1.2}$$

where $\mathbf{x}_1, \ldots, \mathbf{x}_N$ are random points from I^s obtained by N independent trials. Therefore, the basic principle of integration by the Monte Carlo method is to replace a continuous average by a discrete average over randomly selected points. In the same vein, if E is a Lebesgue-measurable subset of I^s, then we may say on the basis of (1.2) that

$$\int_E f(\mathbf{t})\, d\mathbf{t} = \int_{I^s} f(\mathbf{t}) c_E(\mathbf{t})\, d\mathbf{t} \approx \frac{1}{N} \sum_{n=1}^{N} f(\mathbf{x}_n) c_E(\mathbf{x}_n),$$

and so the Monte Carlo approximation is taken to be

$$\int_E f(\mathbf{t})\, d\mathbf{t} \approx \frac{1}{N} \sum_{\substack{n=1 \\ \mathbf{x}_n \in E}}^{N} f(\mathbf{x}_n), \tag{1.3}$$

where $\mathbf{x}_1, \ldots, \mathbf{x}_N$ are as in (1.2).

The strong law of large numbers guarantees that the numerical integration procedure (1.2) converges almost surely. Moreover, it follows from the central limit theorem that the expected integration error is $O(N^{-1/2})$. The remarkable feature here is that this order of magnitude does not depend on the

[6] This condition is adopted to avoid technicalities and can be relaxed.

dimension. This explains the interest in the Monte Carlo method for large dimensions, where classical techniques perform poorly. It should be mentioned, however, that the implied constant in this estimate depends on a certain variance factor through which the dimension may enter. The idea of reducing this variance by suitable transformations plays an important role in the Monte Carlo method (cf. **[66]**, **[100]**).

For the practical implementation of the Monte Carlo method, the fundamental question is, of course, how to produce a random sample. There is no ready-made answer since no satisfactory definition of randomness exists (see §6 for an elaboration on this point). Some people use tables of "random" numbers such as **[251]** or physical devices for generating random numbers such as white noise. But there is now an ever expanding school of thought which has come to realize that instead of trying to cope with the impalpable concept of randomness, one should select points according to a deterministic scheme that is well suited for the problem at hand. This is the underlying idea of a *quasi-Monte Carlo method*. For instance, in the area of numerical integration it turns out to be quite irrelevant whether the sample points or "nodes" are truly random; of primary importance is really the even distribution of the points over I^s (compare with §2). Thus, rather than worrying about random selection procedures, one should be concerned with finding sets of nodes having an optimally fair distribution (see §3).

If we want to emphasize the distinction between quasi-Monte Carlo methods and the standard Monte Carlo method, we will employ the term "statistical Monte Carlo method" for the latter. There is another differentiation in language that will occur in the course of our discussion and that is sometimes regarded as artificial, namely that between quasi-random points (or numbers) and pseudo-random numbers. Although no clear-cut line can be drawn, there is a subtle distinction here, in the sense that the use of quasi-random points is customarily restricted to numerical integration or very closely related applications and that, consequently, they only have to show an acceptable distribution behavior, whereas pseudo-random numbers are supposed to serve a multitude of purposes and should therefore perform well under a battery of statistical tests.

PART I. QUASI-MONTE CARLO METHODS

2. Quasi-Monte Carlo integration. We noted already that numerical integration by a quasi-Monte Carlo method depends on the judicious choice of nodes from the integration domain or a superset thereof. This poses the question as to the precise criteria according to which the nodes should be selected. To find out, let us first consider the case where the integration domain is $I^s = [0, 1]^s$. Here we use the Monte Carlo approximation

$$\int_{I^s} f(\mathbf{t})\, d\mathbf{t} \approx \frac{1}{N} \sum_{n=1}^{N} f(\mathbf{x}_n). \tag{2.1}$$

For the sake of this discussion, we adopt a simpler model by replacing the large set of nodes $\mathbf{x}_1, \ldots, \mathbf{x}_N$ by an infinite sequence $\mathbf{x}_1, \mathbf{x}_2, \ldots$ of points in I^s. Then as we increase N in (2.1), we obviously want the integration error to become negligible. Thus we require that

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} f(\mathbf{x}_n) = \int_{I^s} f(\mathbf{t})\, d\mathbf{t}. \tag{2.2}$$

This limit relation should hold for a reasonable class of integrands, say for all continuous functions f on I^s. The resulting condition on the sequence $\mathbf{x}_1, \mathbf{x}_2, \ldots$ is precisely one of the well-known criteria for this sequence to be uniformly distributed in I^s. This concept is usually defined as follows. The sequence $\mathbf{x}_1, \mathbf{x}_2, \ldots$ of points in I^s is called *uniformly distributed in* I^s if

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} c_J(\mathbf{x}_n) = |J|$$

holds for all subintervals J of I^s, where $|J|$ denotes the s-dimensional Lebesgue measure (= volume) of J. Intuitively, this means that the points $\mathbf{x}_1, \mathbf{x}_2, \ldots$ are spread out over the unit cube I^s according to the principle of proportional representation. A detailed treatment of uniformly distributed sequences can be found in the book of Kuipers and Niederreiter **[174]**.

In case the sequence $\mathbf{x}_1, \mathbf{x}_2, \ldots$ is uniformly distributed in I^s, the relation (2.2) actually holds for all Riemann-integrable functions f on I^s (cf. **[174**, pp. 3, 52**]**). On the other hand, (2.2) need not hold for arbitrary Lebesgue-integrable functions; e.g., it fails if f is the characteristic function of the set $\{\mathbf{x}_1, \mathbf{x}_2, \ldots\}$. In fact, (2.2) characterizes Riemann-integrability in the following sense: if f is a real-valued function on I^s such that the limit in (2.2) exists for all uniformly distributed sequences in I^s, then f must be Riemann-integrable on I^s (de Bruijn and Post **[54]**, Binder **[22]**). Therefore, numerical integration by a quasi-Monte Carlo technique should only be employed for a Riemann-integrable integrand since only in this case can we guarantee convergence. On the theoretical level, this is a significant difference as compared to a statistical Monte Carlo method, for which the strong law of large numbers affirms the almost sure convergence in (2.2) for any bounded Lebesgue-integrable f. But for practical purposes, the restriction to Riemann-integrable functions in quasi-Monte Carlo methods is not of a serious nature.

We turn now to the general case of an integration domain $E \subseteq I^s$. As we have seen in (1.3), the Monte Carlo approximation attains the form

$$\int_E f(\mathbf{t})\, d\mathbf{t} \approx \frac{1}{N} \sum_{\substack{n=1\\ \mathbf{x}_n\in E}}^{N} f(\mathbf{x}_n).$$

If we use again an infinite sequence as a model, then by what we have already learned, the sequence $\mathbf{x}_1, \mathbf{x}_2, \ldots$ should be uniformly distributed in I^s. Further inspection shows that the integration domain E cannot be quite arbitrary if we want convergence of the method. For even if we take a simple integrand such as the constant function $f \equiv 1$ (i.e., if we are asked to calculate the volume of E), we arrive at the convergence condition

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} c_E(\mathbf{x}_n) = \int_{I^s} c_E(\mathbf{t})\, d\mathbf{t},$$

which need only hold if c_E is Riemann-integrable, or, equivalently, if E itself is Jordan-measurable (= has an elementary volume). *In toto*, we get convergence in a quasi-Monte Carlo integration procedure if the integrand f is Riemann-integrable, the integration domain $E \subseteq I^s$ is Jordan-measurable, and the sequence $\mathbf{x}_1, \mathbf{x}_2, \ldots$ of nodes is uniformly distributed in I^s.

Returning from the question of convergence to the original setup, namely that of a finite collection of nodes, we realize that such a discrete set can never constitute a completely fair distribution over I^s since there will always be subintervals J of I^s (possibly of very small volume) which do not contain any one of the given points. Thus, the uniform distribution property is an idealization, and in actual practice we have to settle for an approximation. The model of an infinite sequence has provided the clue to the proper criterion for selecting nodes, viz. the even distribution of the nodes over I^s. Therefore, we shall introduce a quantity which measures the uniformity of distribution of a given set of nodes.

We consider first the one-dimensional case. Let $x_1, \ldots, x_N$ be N numbers in $I = [0, 1]$. If E is a subset of I, then

$$A(E; N) = \sum_{n=1}^{N} c_E(x_n)$$

counts the number of n, $1 \leqslant n \leqslant N$, with $x_n \in E$.

2.1. DEFINITION. The *discrepancy* D_N of the N numbers $x_1, \ldots, x_N$ in I is defined by

$$D_N = \sup_J \left| \frac{A(J; N)}{N} - |J| \right|, \tag{2.3}$$

where J runs through all subintervals of I and $|J|$ denotes the length of J.

It is immaterial whether one considers closed, open, half-open, or arbitrary intervals J since any one category leads to the same value of the supremum in (2.3), as can be seen from [**174**, p. 99]. A useful variant of the above definition is the following.

2.2. DEFINITION. The *discrepancy* D_N^* of the N numbers $x_1, \ldots, x_N$ in I is defined by

$$D_N^* = \sup_{0 \leqslant t \leqslant 1} \left| \frac{A([0, t); N)}{N} - t \right|.$$

Finite sequences of nodes with small discrepancy D_N resp. D_N^* provide a valid approximation to uniform distribution, in the sense that for an infinite sequence both $\lim_{N\to\infty} D_N = 0$ and $\lim_{N\to\infty} D_N^* = 0$ are equivalent to the sequence being uniformly distributed in I (cf. [**174**, Chapter 2, §1]). Here D_N resp. D_N^* stands for the discrepancy of the first N terms of the sequence.

In nonparametric statistics the method of measuring the deviation between the empirical distribution of $x_1, \ldots, x_N$ and the uniform distribution on I by the quantity D_N^* is also known as the two-sided Kolmogorov test. The discrepancy D_N^* may also be thought of as the supremum norm of the function

$$R_N(t) = N^{-1}A([0, t); N) - t, \qquad 0 \leqslant t \leqslant 1.$$

By taking various other norms of this function, one arrives at further concepts of discrepancy, the most commonly used among these being the L^2 *discrepancy* T_N given by

$$T_N = \left(\int_0^1 R_N(t)^2 \, dt \right)^{1/2}.$$

See [220] for further information about the L^2 discrepancy. One may also consider discrepancies with respect to distribution functions different from the uniform distribution ([125], [219]) and with respect to other summation methods ([121], [222], [223]).

The significance of the discrepancy stems from the fact that it occurs in the effective error estimates for quasi-Monte Carlo integration. The precise form of these estimates depends on the regularity of the integrand.

2.3. THEOREM (KOKSMA [156]). *If f is a function of bounded variation $V(f)$ on I and $x_1, \ldots, x_N$ are numbers in I with discrepancy D_N^*, then*

$$\left| \frac{1}{N} \sum_{n=1}^{N} f(x_n) - \int_0^1 f(t) \, dt \right| \leqslant V(f) D_N^*. \tag{2.4}$$

This error bound is particularly appealing since the influences on the integration error are clearly separated: the regularity of the integrand is reflected in the factor $V(f)$ and the uniformity of distribution of the nodes is controlled by the discrepancy D_N^*.

If f is of bounded variation and continuous on I, then Koksma's inequality can be proved very quickly using integration by parts. For

$$\begin{aligned} \int_0^1 R_N(t) \, df(t) &= \frac{1}{N} \sum_{n=1}^{N} \int_0^1 c_{[0,t)}(x_n) \, df(t) - \int_0^1 t \, df(t) \\ &= \frac{1}{N} \sum_{n=1}^{N} (f(1) - f(x_n)) - f(1) + \int_0^1 f(t) \, dt \\ &= -\frac{1}{N} \sum_{n=1}^{N} f(x_n) + \int_0^1 f(t) \, dt, \end{aligned}$$

and so

$$\begin{aligned} \left| \frac{1}{N} \sum_{n=1}^{N} f(x_n) - \int_0^1 f(t) \, dt \right| &= \left| \int_0^1 R_N(t) \, df(t) \right| \\ &\leqslant V(f) \sup_{0 \leqslant t \leqslant 1} |R_N(t)| = V(f) D_N^*. \end{aligned}$$

The general case is shown by a slight variation of this argument (see [174, p. 143]).

2.4. THEOREM (NIEDERREITER [217]). *If f is a continuous function on I with modulus of continuity M and $x_1, \ldots, x_N$ are numbers in I with discrepancy D_N^*, then*

$$\left|\frac{1}{N}\sum_{n=1}^{N} f(x_n) - \int_0^1 f(t)\,dt\right| \leqslant M(D_N^*).$$

A general error estimate valid for any Riemann-integrable integrand was established by Hlawka [**120**]. Analogues of the above inequalities can be shown for other summation methods ([**222**], [**223**, pp. 148–149], [**303**, Chapter 2]) and for other distribution functions ([**110**], [**125**], [**307**]). Estimates in terms of the L^2 discrepancy[7] are also available ([**292**], [**295**], [**366**]) and are obtained by applying the Cauchy-Schwarz inequality at a certain stage in the argument. For instance, if f has a continuous derivative on I and $x_1, \ldots, x_N$ are numbers in I with L^2 discrepancy T_N, then

$$\left|\frac{1}{N}\sum_{n=1}^{N} f(x_n) - \int_0^1 f(t)\,dt\right| \leqslant \left(\int_0^1 (f'(t))^2\,dt\right)^{1/2} T_N.$$

Error bounds of a different nature result from measuring the regularity of a periodic integrand by the size of its Fourier coefficients ([**140**], [**214**]).

We turn now to the multidimensional case in which quasi-Monte Carlo methods are usually applied. Let $I^s = [0, 1]^s$ be the s-dimensional unit cube. For N given points $\mathbf{x}_1, \ldots, \mathbf{x}_N$ in I^s and a subset E of I^s, we introduce the counting function

$$A(E; N) = \sum_{n=1}^{N} c_E(\mathbf{x}_n).$$

We use $|E|$ to denote the s-dimensional Lebesgue measure of E. To unify various definitions of discrepancy that we shall need, we start from a general concept.

2.5. Definition. Let $\mathfrak{M}$ be a nonempty family of Lebesgue-measurable subsets of I^s. Then the *discrepancy* $D_N(\mathfrak{M})$ of the N points $\mathbf{x}_1, \ldots, \mathbf{x}_N$ in I^s is defined by

$$D_N(\mathfrak{M}) = \sup_{E \in \mathfrak{M}} \left|\frac{A(E; N)}{N} - |E|\right|.$$

2.6. Definition. The *discrepancy* D_N of the N points $\mathbf{x}_1, \ldots, \mathbf{x}_N$ in I^s is defined by $D_N = D_N(\mathcal{J})$, where $\mathcal{J}$ is the family of all subintervals of I^s, and the *discrepancy* D_N^* is defined by $D_N^* = D_N(\mathcal{J}^*)$, where $\mathcal{J}^*$ is the family of subintervals of I^s of the form $[0, t_1) \times \cdots \times [0, t_s)$.

For an infinite sequence of points in I^s, we have as in the one-dimensional case that both $\lim_{N\to\infty} D_N = 0$ and $\lim_{N\to\infty} D_N^* = 0$ are equivalent to the sequence being uniformly distributed in I^s (cf. [**174**, Chapter 2, §1]).

In order to generalize Koksma's inequality, we have to set up an appropriate concept of total variation for functions of several variables. For a function f on I^s and an interval $J = [a_1^{(1)}, a_2^{(1)}] \times \cdots \times [a_1^{(s)}, a_2^{(s)}] \subseteq I^s$, we put

[7] More generally, the integration error can be estimated in terms of an appropriately defined L^p discrepancy [**303**, Chapter 2].

$$\Delta(f; J) = \sum_{\varepsilon_1=1}^{2} \cdots \sum_{\varepsilon_s=1}^{2} (-1)^{\varepsilon_1+\cdots+\varepsilon_s} f\left(a_{\varepsilon_1}^{(1)}, \dots, a_{\varepsilon_s}^{(s)}\right).$$

To define a *partition* $\mathcal{P}$ of I^s, we start from s finite sequences of the form $0 = \eta_0^{(j)} < \eta_1^{(j)} < \cdots < \eta_{m_j}^{(j)} = 1$ $(j = 1, 2, \dots, s)$. The partition then consists of all the intervals $[\eta_{i_1}^{(1)}, \eta_{i_1+1}^{(1)}] \times \cdots \times [\eta_{i_s}^{(s)}, \eta_{i_s+1}^{(s)}]$ with $0 \leqslant i_j < m_j$ for $j = 1, 2, \dots, s$.

2.7. Definition. For a function f on I^s, we set

$$V^{(s)}(f) = \sup_{\mathcal{P}} \sum_{J \in \mathcal{P}} |\Delta(f; J)|,$$

where the supremum is extended over all partitions $\mathcal{P}$ of I^s. If $V^{(s)}(f)$ is finite, then f is said to be of *bounded variation on I^s in the sense of Vitali.*

For functions $f = f(t_1, \dots, t_s)$ that are sufficiently regular, $V^{(s)}(f)$ can be represented by an integral; namely,

$$V^{(s)}(f) = \int_0^1 \cdots \int_0^1 \left| \frac{\partial^s f}{\partial t_1 \cdots \partial t_s} \right| dt_1 \cdots dt_s \tag{2.5}$$

whenever the indicated partial derivative is continuous on I^s. If f actually depends on less than s variables, then we always have $\Delta(f; J) = 0$, and so $V^{(s)}(f) = 0$. Since such a function f may still be highly irregular, we have to consider a more suitable notion of variation which is obtained by taking into account the behavior of f on the various faces of I^s.

2.8. Definition. Let f be a function on I^s. For $1 \leqslant k \leqslant s$ and $1 \leqslant i_1 < i_2 < \cdots < i_k \leqslant s$, we denote by $V^{(k)}(f; i_1, \dots, i_k)$ the k-dimensional variation in the sense of Vitali of the restriction of f to $I_{i_1 \cdots i_k}^s = \{(t_1, \dots, t_s) \in I^s$: $t_j = 1$ for $j \neq i_1, \dots, i_k\}$. If all variations $V^{(k)}(f; i_1, \dots, i_k)$ are finite, then f is said to be of *bounded variation on I^s in the sense of Hardy and Krause.*

Since $V^{(s)}(f) = V^{(s)}(f; 1, 2, \dots, s)$, a function of bounded variation in the sense of Hardy and Krause is automatically of bounded variation in the sense of Vitali.[8]

2.9. Theorem (Hlawka [111]). *If f is a function of bounded variation on I^s in the sense of Hardy and Krause and $\mathbf{x}_1, \dots, \mathbf{x}_N$ are points in I^s, then*

$$\left| \frac{1}{N} \sum_{n=1}^{N} f(\mathbf{x}_n) - \int_{I^s} f(\mathbf{t})\, d\mathbf{t} \right|$$

$$\leqslant \sum_{k=1}^{s} \sum_{1 \leqslant i_1 < i_2 < \cdots < i_k \leqslant s} V^{(k)}(f; i_1, \dots, i_k) D_N^*(i_1, \dots, i_k), \tag{2.6}$$

where $D_N^*(i_1, \dots, i_k)$ is the discrepancy in $I_{i_1 \cdots i_k}^s$ of the points obtained by orthogonal projection of $\mathbf{x}_1, \dots, \mathbf{x}_N$ onto $I_{i_1 \cdots i_k}^s$.

The above estimate is often called the Koksma-Hlawka inequality. A proof[9] of this result based on multidimensional integration by parts was given

[8] Information about these concepts of variation can be found in Hobson [127] and the literature in [174, p. 158].

[9] For sufficiently regular integrands, Hlawka [112] presents a simplified proof of the inequality, with the variations replaced by the corresponding integrals in (2.5).

by Zaremba [**366**] and is reproduced in [**174**, Chapter 2, §5].

One obtains a simplified form of (2.6) if one defines the variation $V(f)$ of f on I^s in the sense of Hardy and Krause to be

$$V(f) = \sum_{k=1}^{s} \sum_{1 \leqslant i_1 < i_2 < \cdots < i_k \leqslant s} V^{(k)}(f; i_1, \ldots, i_k).$$

Obviously, $V(f)$ is finite if and only if f is of bounded variation on I^s in the sense of Hardy and Krause. By using (2.5), $V(f)$ can be written in terms of integrals for sufficiently regular f. Since we always have

$$D_N^*(i_1, \ldots, i_k) \leqslant D_N^*(1, 2, \ldots, s) = D_N^*,$$

it follows from (2.6) that

$$\left| \frac{1}{N} \sum_{n=1}^{N} f(\mathbf{x}_n) - \int_{I^s} f(\mathbf{t})\, d\mathbf{t} \right| \leqslant V(f) D_N^*. \tag{2.7}$$

This inequality is completely analogous to (2.4).

The Koksma-Hlawka inequality can be generalized to more abstract settings ([**214**], [**278**]). Error estimates for other than equal-weight formulas are also known ([**222**], [**223**, pp. 148–149]). Inequalities for Riemann-integrable functions ([**120**]) and error bounds in terms of an appropriately defined L^2 discrepancy T_N ([**292**], [**366**], [**101**]) or an L^1 discrepancy ([**292**], [**303**, Chapter 8]), and more generally for an L^p discrepancy ([**303**, Chapter 8]), have been established as well. See [**126**] for the case of a nonuniform distribution function.

A measure for the evenness of distribution different from the discrepancy has also been considered in the literature. To introduce this concept, we need some preliminaries. By a *dyadic interval*, we mean a subinterval of I of the form $[j2^{-m}, (j+1)2^{-m})$ with integers $m \geqslant 1$ and $0 \leqslant j < m$, and with the stipulation that the interval be closed if its right endpoint is 1. In the s-dimensional case, a *dyadic box* is meant to be a cartesian product of any s dyadic intervals. Given a dyadic box B in I^s, we split it into two parts as follows. For the moment, we move the origin of the coordinate system to the center of B and denote by $\xi_1, \ldots, \xi_s$ the new coordinates. Then the union of those "quadrants" of B in which $\operatorname{sgn}(\xi_1 \cdots \xi_s) = (-1)^s$ is called the positive part B^+ of B, whereas the remaining portion of B is the negative part B^-. Let now $\mathbf{x}_1, \ldots, \mathbf{x}_N$ be N given points in I^s. Then we define the *s-dimensional nonuniformity* of these points to be

$$\sup_B |A(B^+; N) - A(B^-; N)|,$$

where the supremum is extended over all possible dyadic boxes B in I^s. Furthermore, we project the given points orthogonally onto the various k-dimensional faces of I^s ($1 \leqslant k < s$) and calculate the k-dimensional nonuniformity of the projected points in the respective face. Finally, the largest value among the s-dimensional nonuniformity and all these lower-dimensional nonuniformities is said to be the *nonuniformity* $\varphi_\infty(N)$ of the points $\mathbf{x}_1, \ldots, \mathbf{x}_N$ (Sobol' [**291**]). It is always a positive integer $\leqslant N$. The motivation for this definition comes from the theory of Haar functions (compare with

[303]). For $1 < q < \infty$, an L^q analogue $\varphi_q(N)$ has also been introduced (Sobol′ [289]). If $\mathbf{x}_1, \mathbf{x}_2, \ldots$ is a sequence of points in I^s, then $\varphi_q(N)$, $1 < q \leqslant \infty$, is the appropriate nonuniformity of the first N terms of the sequence. For any q, $1 < q \leqslant \infty$, we have the criterion that a sequence is uniformly distributed in I^s if and only if $\lim_{N\to\infty}\varphi_q(N)/N = 0$ ([289], [294]). A comparison with the definition of discrepancy leads easily to the inequality

$$\varphi_\infty(N) \leqslant 2^s N D_N \tag{2.8}$$

for any N points in I^s.

Estimates for integration errors can also be established in terms of nonuniformities. For instance, if $f(\mathbf{t}) = f(t_1, \ldots, t_s)$ is such that all its mixed partial derivatives

$$\frac{\partial^k f}{\partial t_{i_1} \cdots \partial t_{i_k}}, \qquad 1 \leqslant i_1 < i_2 < \cdots < i_k \leqslant s, 1 \leqslant k \leqslant s,$$

exist and are continuous on I^s, then

$$\left| \frac{1}{N} \sum_{n=1}^{N} f(\mathbf{x}_n) - \int_{I^s} f(\mathbf{t})\, d\mathbf{t} \right| \leqslant C(f) \frac{\varphi_\infty(N) \log^s N}{N}, \tag{2.9}$$

where $C(f)$ is a constant depending on f that reflects, as usual, the amount of oscillation of the integrand and $\varphi_\infty(N)$ is the nonuniformity of the nodes $\mathbf{x}_1, \ldots, \mathbf{x}_N$ (Sobol′ [291], [303, Chapter 4]). Analogous estimates using the nonuniformities $\varphi_q(N)$, $1 < q < \infty$, are also available ([289], [303, Chapter 4]).

So far, we have only considered the case where the integration domain is I^s, or what amounts to the same (modulo a simple change of variable), where the integration domain is an interval. We shall now study a much wider class of integration domains, namely that of bounded Jordan-measurable sets (= sets having a finite elementary volume). By applying, if necessary, a translation and a contraction, we can assume that such an integration domain is contained in I^s. We have already seen in (1.3) what the Monte Carlo approximation should look like in this case. We also noted earlier that Jordan-measurable subsets of I^s form the most general category of integration domains for which a quasi-Monte Carlo technique can be applied successfully.

The integration error can again be estimated effectively in terms of a suitable notion of discrepancy. We need some preliminaries before we can enunciate this result. Let $d(\cdot, \cdot)$ be the standard Euclidean distance in $\mathbf{R}^s$. For a subset E of I^s and $\varepsilon > 0$, we define

$$E_\varepsilon = \{\mathbf{x} \in I^s: d(\mathbf{x}, \mathbf{y}) < \varepsilon \text{ for some } \mathbf{y} \in E\},$$

$$E_{-\varepsilon} = \{\mathbf{x} \in I^s: d(\mathbf{x}, \mathbf{y}) \geqslant \varepsilon \text{ for all } \mathbf{y} \in I^s \setminus E\}.$$

The set E_ε is a superset of E open in the relative topology in I^s, whereas $E_{-\varepsilon}$ is a closed subset of E. We classify now the Jordan-measurable subsets E of I^s according to the measure of what may be called the "ε-collars" of E.

2.10. Definition. Let $b = b(\varepsilon)$ be a positive nondecreasing function defined for all $\varepsilon > 0$ and satisfying $\lim_{\varepsilon\to 0+} b(\varepsilon) = 0$. Then $\mathfrak{M}_b$ is defined as the family of all Lebesgue-measurable subsets E of I^s for which the inequalities

$$|E_\varepsilon \setminus E| \leqslant b(\varepsilon) \quad \text{and} \quad |E \setminus E_{-\varepsilon}| \leqslant b(\varepsilon)$$

hold for all $\varepsilon > 0$.

It is easily seen that every $E \in \mathfrak{M}_b$ is actually Jordan-measurable and that, conversely, every Jordan-measurable subset of I^s belongs to some $\mathfrak{M}_b$ (cf. **[220**, pp. 168–169]). Therefore, the families $\mathfrak{M}_b$ provide a complete classification of all Jordan-measurable subsets of I^s. The larger the function b, the more irregularly shaped can be the sets that are allowed in $\mathfrak{M}_b$.

For an integration domain $E \in \mathfrak{M}_b$, an error bound for quasi-Monte Carlo integration can now be given in terms of a discrepancy $D_N(\mathfrak{M}_c)$ defined according to Definition 2.5, where c is a function closely related to b.

2.11. Theorem (Niederreiter **[220]**). *For a function f of bounded variation $V(f)$ on I^s in the sense of Hardy and Krause, an integration domain $E \in \mathfrak{M}_b$, and points $\mathbf{x}_1, \ldots, \mathbf{x}_N$ in I^s we have*

$$\left| \frac{1}{N} \sum_{\substack{n=1 \\ \mathbf{x}_n \in E}}^{N} f(\mathbf{x}_n) - \int_E f(\mathbf{t})\, d\mathbf{t} \right| \leqslant (V(f) + |f(1, \ldots, 1)|) D_N(\mathfrak{M}_c), \qquad (2.10)$$

where c is the function $c(\varepsilon) = b(\varepsilon) + 2s\varepsilon$ for $\varepsilon > 0$.

One might wonder why the term $|f(1, \ldots, 1)|$ appears in this estimate. The reason is that, in contrast to earlier inequalities, the left-hand side of (2.10) is not invariant under the shift from f to $f + C$, where C is a constant. Therefore, the upper bound in (2.10) depends also on the magnitude of f.

There are two important facts buttressing the usefulness of the families $\mathfrak{M}_b$. First, the discrepancy $D_N(\mathfrak{M}_b)$ can be estimated effectively in terms of the discrepancy D_N, as we shall see in Theorem 3.10, so that we ultimately have control over the size of $D_N(\mathfrak{M}_b)$ for the common choices of nodes. Secondly, the discrepancy $D_N(\mathfrak{M}_b)$ still satisfies $\lim_{N\to\infty} D_N(\mathfrak{M}_b) = 0$ for any infinite sequence uniformly distributed in I^s, and uniform conditions on the measure of the "ε-collars" of sets E in a family $\mathfrak{M}$ (such as those imposed on the members of $\mathfrak{M}_b$) are necessary to guarantee this property for $D_N(\mathfrak{M})$ by a result of Billingsley and Topsøe **[21]**. Thus, the approach based on the families $\mathfrak{M}_b$ is essentially the most general one in quasi-Monte Carlo integration.

The case of a convex integration domain deserves special attention since it occurs quite frequently. Information about this class of integration domains is, of course, implicitly contained in the above considerations since every convex subset of I^s belongs to $\mathfrak{M}_b$ with $b(\varepsilon) = 2s\varepsilon$ for $\varepsilon > 0$. By using special properties of convex sets, it is, however, possible to obtain stronger results.

2.12. Definition. The *isotropic discrepancy* J_N of the N points $\mathbf{x}_1, \ldots, \mathbf{x}_N$ in I^s is defined by $J_N = D_N(\mathcal{C})$, where $\mathcal{C}$ is the family of all convex subsets of I^s.

The use of J_N was suggested by Hlawka **[115]** and the term "isotropic discrepancy" was coined by Zaremba **[370]**. This type of discrepancy figures in the error estimate for convex integration domains.

2.13. THEOREM (ZAREMBA [370]). *If f is a function of bounded variation $V(f)$ on I^s in the sense of Hardy and Krause, E is a convex subset of I^s, and $\mathbf{x}_1, \ldots, \mathbf{x}_N$ are points in I^s with isotropic discrepancy J_N, then*

$$\left| \frac{1}{N} \sum_{\substack{n=1 \\ \mathbf{x}_n \in E}}^{N} f(\mathbf{x}_n) - \int_E f(\mathbf{t})\, d\mathbf{t} \right| \leqslant (V(f) + |f(1, \ldots, 1)|) J_N.$$

Up to now, all the integrands considered were bounded functions. Sobol′ [311] succeeded in constructing quasi-Monte Carlo methods for improper integrals in which the behavior of the integrand at the singularity is under control. In a sense, the paper of Hardy and Littlewood [108] may be considered the first contribution to this subject.*

The study of quasi-Monte Carlo integration on the infinite-dimensional unit cube was initiated by Čencov [33] and Sobol′ [293]. See also [279], [301], [303, Chapter 7], [313], [314].

The results appearing in this section as well as in the three following ones should be compared with the features of classical techniques for computing simple and multiple integrals. We refer to [53], [95], [173], [329] for expository accounts of these techniques. Summarizing in a nutshell what transpires from this comparison, one can say that conventional methods are without doubt superior in the one-dimensional case, whereas quasi-Monte Carlo methods should be considered more and more favorable as the dimension increases. For simple integrals with integrands of a low degree of regularity, the results are basically equivalent.

Quasi-Monte Carlo methods have been employed for other purposes besides numerical integration. We mention here only those works that are based on the general principles of such methods. In later sections, we will indicate further applications using the special techniques to be described.

There are, of course, several important problems in numerical analysis that can be reduced to the approximate calculation of integrals, and under such circumstances quasi-Monte Carlo methods may be applied readily. A case in point is the numerical solution of integral equations; see [114], [116], [124], [313]. Applications to integro-differential equations occur in [210], [212], to linear partial differential equations in [211], and to the theory of turbulence in [15]. A method of determining surface areas by means of uniformly distributed sequences is discussed in [175]. Interesting connections with complex analysis appear in [118], [123], regarding problems of interpolation and analytic continuation, respectively. Quasi-Monte Carlo methods can also be used to approximate the extreme values of a function ([3], [122], [232], [317]) and the greatest and smallest real part of the eigenvalues of a matrix [318]. On the basis of ideas in [34], applications to the simulation of Markov chains are studied in [312]–[314].

3. Quasi-random points. A scrutiny of the various error estimates in §: provides convincing evidence that what we should employ in a quasi-Monte Carlo integration is a set of nodes with very small discrepancy. Finite or

*See also R. S. Lehman, Pacific J. Math. 5 (1955), 93–102.

infinite sequences with this property are called *low-discrepancy sequences* and their terms are loosely referred to as *quasi-random points* (or *quasi-random numbers* in the one-dimensional case). Strictly speaking, we have dealt not only with one, but with several concepts of discrepancy, so that the expression "low-discrepancy sequence" would seem to call for further explanation. It turns out, however, that the various notions of discrepancy are related to D_N in a known manner. Thus, D_N^* and D_N are linked by the inequalities

$$D_N^* \leqslant D_N \leqslant 2^s D_N^* \tag{3.1}$$

in the s-dimensional case (see [**174**, p. 93]), and J_N as well as $D_N(\mathfrak{M}_b)$ can also be estimated in terms of D_N (see (3.21) and Theorem 3.10). Moreover, the L^2 discrepancy T_N of any N points in I^s satisfies

$$C_s D_N^{*(s+2)/2} \leqslant T_N \leqslant D_N^*$$

with a constant $C_s > 0$ only depending on s (cf. [**220**, Theorem 4.2]). Therefore, "low discrepancy" will be interpreted to mean low discrepancy D_N.

The discrepancy of the average sequence of points in I^s is under control because of the law of the iterated logarithm established by Chung [**40**] in the one-dimensional case and Kiefer [**150**] in the multidimensional case, according to which we have

$$\limsup_{N\to\infty} \frac{\sqrt{2N}\, D_N^*}{\sqrt{\log\log N}} = 1$$

with probability 1, in the sense of an appropriate product measure on the sequence space.[10] Together with (3.1) we get

$$D_N = O\left(N^{-1/2}(\log\log N)^{1/2}\right)$$

almost surely. In combination with the error estimates in §2 (in particular, Theorem 2.3 and (2.7)), this ties in rather nicely with the probabilistic error bound for the Monte Carlo method. It is to be expected, however, that clever constructions should produce sequences that behave much better than the average sequence. As a matter of fact, we shall exhibit examples of finite sequences of N points in I^s for which $D_N = O(N^{-1}(\log N)^{s-1})$. Quasi-Monte Carlo integration with such quasi-random points involves then effective error bounds that are considerably smaller than the probabilistic Monte Carlo bound $O(N^{-1/2})$. Thus, as far as numerical integration is concerned, quasi-Monte Carlo methods based on determinate low-discrepancy sequences are, on the whole, preferable to statistical Monte Carlo methods.

We discuss now in detail the relevant properties of the various notions of discrepancy. We start with the one-dimensional case in which all the basic questions have been settled. Here the discrepancy D_N^* can be calculated in an easy manner by arranging the numbers $x_1, \ldots, x_N$ in nondecreasing order of magnitude, which obviously does not affect the value of D_N^*. We obtain then the formula[11]

[10] See also [**30**], [**242**], [**243**], [**372**].

[11] See [**215**]. For generalizations of this formula referring to other distribution functions or other summation methods, see [**219**], [**222**].

$$D_N^* = \max_{1 \leqslant n \leqslant N} \max\left(\left|x_n - \frac{n}{N}\right|, \left|x_n - \frac{n-1}{N}\right|\right)$$
$$= \frac{1}{2N} + \max_{1 \leqslant n \leqslant N} \left|x_n - \frac{2n-1}{2N}\right|. \qquad (3.2)$$

We deduce from (3.2) that $D_N^* \geqslant 1/(2N)$ for any N numbers in I, and that $D_N^* = 1/(2N)$ exactly for the sequence

$$\frac{1}{2N}, \frac{3}{2N}, \ldots, \frac{2N-1}{2N} \qquad (3.3)$$

or one of its rearrangements. A simple argument shows that $D_N \geqslant 1/N$ for any N numbers in I (cf. [174, p. 90]), and $D_N = 1/N$ holds for the sequence (3.3) and its rearrangements, but also for some other sequences. A formula similar to (3.2) can be obtained for the L^2 discrepancy T_N, namely

$$T_N^2 = \frac{1}{12N^2} + \frac{1}{N} \sum_{n=1}^{N} \left(x_n - \frac{2n-1}{2N}\right)^2$$

provided that $x_1 \leqslant x_2 \leqslant \cdots \leqslant x_N$ (cf. [220, p. 135]). This shows that the L^2 discrepancy of any N numbers in I is $\geqslant 1/(\sqrt{12}\ N)$ and that the lower bound is attained precisely for the sequence (3.3) and its rearrangements.[12] It is interesting to note that the nodes in (3.3) are also used in a classical integration method, namely the N-point midpoint rule (see [53, p. 40]). The integration error implied by the use of sequences equal or close to (3.3) was studied by Chui [37], [38], [39].

The terms of the optimal sequence (3.3) depend on the chosen value of N. In computational practice, it is often convenient to be able to change the value of N without losing the previously obtained data. For this purpose, it is advantageous to work with an infinite sequence and then to take its first N terms whenever the value of N has been decided upon. In this way, we may increase N if we desire greater accuracy and still use the results of the earlier computation.

It requires a deeper analysis to determine the behavior of the discrepancy for infinite sequences. In the first place, the discrepancy D_N of an infinite sequence cannot be uniformly of the order of magnitude N^{-1}. An important theorem of W. M. Schmidt [282] has established that for any infinite sequence of numbers in I there exist infinitely many N such that $D_N \geqslant D_N^* \geqslant (\log N)/(100\ N)$. On the other hand, infinite sequences with $D_N = O(N^{-1} \log N)$ have been known for several decades. For instance, if α is an irrational number for which the partial quotients in the continued fraction expansion are uniformly bounded, then the discrepancy D_N of the sequence $\{\alpha\}, \{2\alpha\}, \ldots, \{n\alpha\}, \ldots$ of fractional parts[13] satisfies $D_N \leqslant CN^{-1}(1 + \log N)$ for all $N \geqslant 1$, where C is an explicit constant (compare with [220, Theorem 3.2], [174, p. 125]).

[12] A similar statement holds for the L^1 discrepancy according to a result in [292] that also appeared in an essentially equivalent form in [236, Chapter 10].

[13] The *fractional part* $\{t\}$ of $t \in \mathbf{R}$ is defined by $\{t\} = t - [t]$, where $[t]$ is the greatest integer $\leqslant t$. We always have $0 \leqslant \{t\} < 1$.

Another example, the so-called van der Corput sequence, is more important for two reasons: first, the known discrepancy estimate is better than for the above example, and second, this sequence consists only of dyadic fractions and is therefore perfectly well adapted for the use in binary computers. The sequence can be conveniently defined in terms of a "radical-inverse function". If $g \geqslant 2$ is an integer, then every nonnegative integer n has an expansion in the base g of the form

$$n = \sum_{i=0}^{k} a_i g^i \quad \text{with } a_i \in \{0, 1, \dots, g-1\} \text{ for } 0 \leqslant i \leqslant k, \tag{3.4}$$

and this representation is unique apart from adding on higher powers of g with zero coefficients. Now the *radical-inverse function* ϕ_g is given by the well-defined expression

$$\phi_g(n) = \sum_{i=0}^{k} a_i g^{-i-1}. \tag{3.5}$$

The action of this function may be described as follows. Write the expansion of n in the base g as a string of digits $a_k \cdots a_1 a_0$; then $\phi_g(n)$ results from this by reflection about the "decimal point", i.e., $\phi_g(n)$ is the g-adic fraction given by $0.a_0 a_1 \cdots a_k$. The condition on the a_i in (3.4) implies that $0 \leqslant \phi_g(n) < 1$ for all $n \geqslant 0$.

The *van der Corput sequence*, which made its first appearance in **[341]**, is now defined as the sequence $\phi_2(0), \phi_2(1), \dots, \phi_2(n), \dots$ in $[0, 1)$. According to unpublished results of Tijdeman, its discrepancy D_N^* satisfies

$$ND_N^* \leqslant \tfrac{1}{3} \log_2 N + 1 \quad \text{for all } N \geqslant 1$$

and

$$\limsup_{N \to \infty} \left(ND_N^* - \tfrac{1}{3} \log_2 N\right) \geqslant \tfrac{4}{9} + \tfrac{1}{3} \log_2 3,$$

where $\log_2$ denotes the logarithm to the base 2. The fact that the coefficient $\frac{1}{3}$ of $\log_2 N$ is best possible was already shown earlier by Haber **[92]**. For a quick proof of $D_N = O(N^{-1} \log N)$, see **[174**, p. 127]. At present, we know of no infinite sequence in I whose discrepancy is uniformly smaller than that of the van der Corput sequence.*

There are various ways of estimating the discrepancy of a one-dimensional sequence (see **[217]** for a survey). The most common procedure is to reduce the estimation of the discrepancy to the problem of estimating associated exponential sums. The theoretical basis for this technique is a general principle of quantitative Fourier inversion. We need the following notion: if F is a function of bounded variation on I, then its Fourier-Stieltjes transform $\hat{F}$ is given by

$$\hat{F}(h) = \int_0^1 e^{2\pi i h t}\, dF(t) \quad \text{for all integers } h.$$

*Added in proof. Better sequences have recently been constructed by R. Bejian and H. Faure. For further information about the van der Corput sequence, see the joint paper of these authors in C. R. Acad. Sci. Paris Sér. A **285** (1977), 313–316.

3.1. Theorem (Niederreiter and Philipp [234]). *Let F be a nondecreasing function on I with $F(0) = 0$ and $F(1) = 1$, and suppose the function G on I satisfies the Lipschitz condition $|G(u) - G(v)| \leqslant L|u - v|$ for $u, v \in I$, as well as $G(0) = 0$ and $G(1) = 1$. Then for every positive integer m we have*

$$\sup_{u,v \in I} |(F(v) - F(u)) - (G(v) - G(u))|$$
$$\leqslant \frac{4L}{m+1} + \frac{4}{\pi} \sum_{h=1}^{m} \left(\frac{1}{h} - \frac{1}{m+1}\right)|\hat{F}(h) - \hat{G}(h)|.$$

3.2. Corollary. *Let $x_1, \dots, x_N$ be any real numbers. Then the discrepancy D_N of the finite sequence of fractional parts $\{x_1\}, \dots, \{x_N\}$ satisfies*

$$D_N \leqslant \frac{4}{m+1} + \frac{4}{\pi} \sum_{h=1}^{m} \left(\frac{1}{h} - \frac{1}{m+1}\right)\left|\frac{1}{N} \sum_{n=1}^{N} e^{2\pi i h x_n}\right|$$

for every positive integer m.

The corollary results from Theorem 3.1 by setting $F(t) = A([0, t); N)/N$ and $G(t) = t$ for $0 \leqslant t \leqslant 1$. The inequality in Corollary 3.2, without the specified constants, is due to Erdös and Turán [64]. For a special class of sequences, there exists a simplified version of this inequality which will be useful later on. It is convenient here to introduce for an integer $m \geqslant 2$ the summation symbol $\Sigma_{h(\text{mod } m)}$ which designates a sum over the complete residue system mod m consisting of all integers h with $-m/2 < h \leqslant m/2$. The summation symbol $\Sigma^*_{h(\text{mod } m)}$ refers to the same sum, but with $h = 0$ deleted from the range of summation.

3.3. Lemma. *Let $m \geqslant 2$ and $y_0, \dots, y_{N-1}$ be integers. Then the discrepancy D_N of the finite sequence of fractional parts $\{y_0/m\}, \dots, \{y_{N-1}/m\}$ satisfies*

$$D_N \leqslant \frac{1}{m} + \sum_{h(\text{mod } m)}^{*} \frac{1}{m \sin(\pi|h|/m)} \left|\frac{1}{N} \sum_{n=0}^{N-1} e^{2\pi i h y_n/m}\right|.$$

To make the relationship with the earlier inequality even more transparent, we use $\sin \pi t \geqslant 2t$ for $0 \leqslant t \leqslant \frac{1}{2}$ and get

$$D_N \leqslant \frac{1}{m} + \sum_{h=1}^{[m/2]} \frac{1}{h} \left|\frac{1}{N} \sum_{n=0}^{N-1} e^{2\pi i h y_n/m}\right|.$$

It should not come as a surprise that we obtain smaller constants in the special case.

The proof of Lemma 3.3 nicely illustrates the principles involved in these types of inequalities. For an integer k, let $A(k; N)$ be the number of n, $0 \leqslant n \leqslant N - 1$, with $y_n \equiv k \pmod m$. Then we can write

$$A(k; N) = \sum_{n=0}^{N-1} c_k(y_n),$$

where c_k is the characteristic function of the coset $k + m\mathbf{Z}$ of $\mathbf{Z}/m\mathbf{Z}$. Now with the abbreviation $e(t) = e^{2\pi i t}$ for $t \in \mathbf{R}$ we have

$$c_k(y) = \frac{1}{m} \sum_{h (\operatorname{mod} m)} e(h(y-k)/m) \quad \text{for } y \in \mathbf{Z},$$

so that

$$A(k; N) = \frac{1}{m} \sum_{h (\operatorname{mod} m)} e(-hk/m) \sum_{n=0}^{N-1} e(hy_n/m)$$

and

$$A(k; N) - \frac{N}{m} = \frac{1}{m} \sideset{}{^*}\sum_{h (\operatorname{mod} m)} e(-hk/m) \sum_{n=0}^{N-1} e(hy_n/m). \tag{3.6}$$

Let $J = [u, v)$ be an arbitrary half-open subinterval of I. We choose the largest closed subinterval of J of the form $[a/m, b/m]$ with integers $a \leqslant b$, which we denote again by $[a/m, b/m]$. The case where no such subinterval exists can be dealt with easily, since we have then $A(J; N) = 0$ and $v - u < 1/m$, hence

$$|A(J; N)/N - |J|\,| = |J| < 1/m. \tag{3.7}$$

In the remaining case, we obtain

$$A(J; N) - N|J| = \sum_{k=a}^{b} \left(A(k; N) - \frac{N}{m}\right) + \frac{N}{m}(b-a+1) - N|J|$$

$$= \frac{1}{m} \sideset{}{^*}\sum_{h (\operatorname{mod} m)} \left(\sum_{k=a}^{b} e(-hk/m)\right)\left(\sum_{n=0}^{N-1} e(hy_n/m)\right) + N\left(\frac{b-a+1}{m} - |J|\right)$$

by using (3.6). It follows that

$$\left|\frac{A(J; N)}{N} - |J|\right| \leqslant \frac{1}{m} \sideset{}{^*}\sum_{h (\operatorname{mod} m)} \left|\sum_{k=a}^{b} e(hk/m)\right| \left|\frac{1}{N} \sum_{n=0}^{N-1} e(hy_n/m)\right| + \left|\frac{b-a+1}{m} - |J|\right|. \tag{3.8}$$

Now for $0 < |h| \leqslant m/2$ we have

$$\left|\sum_{k=a}^{b} e(hk/m)\right| = \frac{|e(h(b-a+1)/m) - 1|}{|e(h/m) - 1|} \leqslant \frac{2}{|e(h/m) - 1|} = \frac{1}{\sin(\pi|h|/m)}. \tag{3.9}$$

From the definition of a and b it follows that

$$a/m = u + \theta_1 \quad \text{with } 0 \leqslant \theta_1 < 1/m$$

and

$$b/m = v - \theta_2 \quad \text{with } 0 < \theta_2 \leqslant 1/m,$$

so that

$$\left|\frac{b-a+1}{m} - |J|\right| = \left|\frac{1}{m} - \theta_1 - \theta_2\right| < \frac{1}{m}.$$

By combining this with (3.8) and (3.9), we arrive at

$$\left|\frac{A(J;N)}{N}-|J|\right|<\frac{1}{m}+\sum_{h(\mathrm{mod}\, m)}^{*}\frac{1}{m\sin(\pi|h|/m)}\left|\frac{1}{N}\sum_{n=0}^{N-1}e(hy_n/m)\right|.$$

In view of (3.7), this inequality holds for all J, and forming the supremum over J on the left-hand side completes the proof of Lemma 3.3.

There is another inequality for D_N in terms of exponential sums that is more of theoretical interest. The following general inequality improves and corrects a result of Elliott [63].

3.4. Theorem (Niederreiter [222]). *Let F and G be two functions on I satisfying the conditions in Theorem* 3.1. *Then,*

$$\sup_{u,v\in I}|(F(v)-F(u))-(G(v)-G(u))|$$
$$\leqslant\left(\frac{6L}{\pi^2}\sum_{h=1}^{\infty}\frac{1}{h^2}\left|\hat{F}(h)-\hat{G}(h)\right|^2\right)^{1/3}.$$

3.5. Corollary (LeVeque [177]). *Let $x_1,\dots,x_N$ be any real numbers. Then the discrepancy D_N of the finite sequence of fractional parts $\{x_1\},\dots,\{x_N\}$ satisfies*

$$D_N\leqslant\left[\frac{6}{\pi^2}\sum_{h=1}^{\infty}\frac{1}{h^2}\left|\frac{1}{N}\sum_{n=1}^{N}e^{2\pi ihx_n}\right|^2\right]^{1/3}.$$

The difficulties in the multidimensional case start already with the fact that there is no analogue of the simple formula (3.2). One can only provide certain estimates which allow the approximate calculation of D_N^* in relatively few steps and with a reasonable degree of accuracy (cf. [215]). As a consequence, it is impossible to give an easy description of the finite sequences of N points in I^s with minimal discrepancy D_N^*, as was done in (3.3) for $s = 1$. Finding these sequences is a laborious computational task, which is nontrivial even for the limited range considered in [361], namely $s = 2$ and $1 \leqslant N \leqslant 6$.

With increasing dimension, the requirement on a sequence to have a small discrepancy becomes more stringent. This is reflected in the known lower bounds for the discrepancy, which depend on the dimension. Roth [261] established a general lower bound on the L^2 discrepancy T_N, and consequently on D_N^*. In detail, for any N points in I^s we have

$$D_N^*\geqslant T_N>C_sN^{-1}(\log N)^{(s-1)/2}\tag{3.10}$$

with a positive constant C_s only depending on s. According to [174, p. 102] we may take $C_s = 2^{-4s}((s-1)\log 2)^{(1-s)/2}$ for $s \geqslant 2$. An improvement on the order of magnitude of this bound for D_N^* is only known for $s = 2$, where W. M. Schmidt [282] showed[14]

[14] The constant appearing here is better than in the original paper and was obtained in [174, p. 109].

$$D_N^* > \frac{\log N}{(132 \log 4)\, N} \tag{3.11}$$

for any N points in I^2. For infinite sequences of points in I^s, it follows from (3.10) and a standard argument (cf. [**174**, p. 105]) that

$$D_N^* > C_s' N^{-1} (\log N)^{s/2} \tag{3.12}$$

for infinitely many N, where C_s' is an effective positive constant only depending on s. A sharper bound is only available for $s = 1$ and was mentioned in the discussion of the one-dimensional case.

The lower bound (3.11) for $s = 2$ is best possible as far as the order of magnitude is concerned. A sequence of N points in I^2 with $D_N^* = O(N^{-1} \log N)$ was, for instance, constructed by Roth [**261**]. It consists of the points $(n/N, \phi_2(n))$, $n = 0, 1, \ldots, N-1$, where ϕ_2 is the radical-inverse function from (3.5). In case N is a power of 2, this sequence is known as the *Roth sequence*. For such N, the constant implied in Roth's estimate was improved by Gabai ([**75**], [**76**]) and an exact formula was obtained by Halton and Zaremba [**103**], in which the leading term for ND_N^* is $(1/3)\log_2 N$ (see [**92**] for an estimate with this leading term).

In the 2-dimensional case, the lower bound for the L^2 discrepancy T_N in (3.10) is also best possible apart from the value of the constant. Curiously enough, this lower bound is not reached by the Roth sequence, for Vilenkin [**346**] gave an exact formula[15] for NT_N with leading term $(1/8)\log_2 N$ (see [**92**] for an estimate with this leading term). Finite sequences with $T_N = O(N^{-1}(\log N)^{1/2})$ were first constructed by Davenport [**50**]. His sequences are basically of the form $(n/N, \{n\alpha\})$, $n = 0, 1, \ldots, N-1$, where α is an irrational with bounded partial quotients in its continued fraction expansion. Examples of such sequences involving only dyadic fractions (as is the case in the Roth sequence) were first obtained by Vilenkin [**346**] by introducing a perturbation in the Roth sequence. His exact formula for NT_N has leading term $((1/24)\log_2 N)^{1/2}$. Another example (called the *Zaremba sequence*) yielding the even smaller leading term $((5/192)\log_2 N)^{1/2}$ was found by Halton and Zaremba [**103**], using a different perturbation of the Roth sequence.[16] The Zaremba sequence may be described as follows. Let N be a power of 2, say $N = 2^k$, $k \geqslant 1$; then the sequence consists of the N points

$$\left(\frac{a_k}{2} + \frac{a_{k-1}}{2^2} + \cdots + \frac{a_1}{2^k}, \frac{a_1'}{2} + \frac{a_2'}{2^2} + \cdots + \frac{a_k'}{2^k} \right) \in I^2,$$

where $a_1, \ldots, a_k$ take independently of each other the values 0 and 1, while $a_i' = a_i$ for i even, $a_i' = 1 - a_i$ for i odd. Not surprisingly, the discrepancy D_N^* of the Zaremba sequence is smaller than that of the Roth sequence; in fact, the exact formula in [**103**] for ND_N^* involves the leading term $(1/5)\log_2 N$. See also [**100**] for a summary of these results. White [**360**] considered analogues of the Roth and Zaremba sequences using an arbitrary base in the construction (rather than the base 2 in the original sequences) and gave formulas for their

[15] This formula was shown independently by Halton and Zaremba [**103**].

[16] See also Roth [**262**] and Vilenkin [**347**].

L^2 discrepancies (compare also with [357]).

The construction techniques described above can be used for an arbitrary dimension $s \geqslant 2$. The so-called *Hammersley sequence* (of order N) was introduced in [105] (see also [99]) and is composed of the N points $(n/N, \phi_{g_1}(n), \ldots, \phi_{g_{s-1}}(n))$, $n = 0, 1, \ldots, N-1$, in I^s, where the bases $g_1, \ldots, g_{s-1}$ of the radical-inverse functions are supposed to be pairwise relatively prime. Usually, one takes for $g_1, \ldots, g_{s-1}$ the first $s-1$ primes. The discrepancy D_N of the Hammersley sequence was estimated in [99], with the result that $D_N \leqslant CN^{-1}(\log N)^{s-1}$ for all $N \geqslant 2$, where the constant C only depends on $g_1, \ldots, g_{s-1}$. An infinite sequence with uniformly small discrepancy was constructed in [99] and is nowadays called the *Halton sequence*. Its terms are the points $(\phi_{g_1}(n), \ldots, \phi_{g_s}(n))$, $n = 0, 1, \ldots,$ in I^s, where $g_1, \ldots, g_s$ are pairwise relatively prime bases. A convenient algorithm for fast computer generation of these points is available in [102]. The discrepancy D_N of the Halton sequence has the order of magnitude $N^{-1}(\log N)^s$, as shown in [99]. Better constants were obtained by Meijer [195], and his estimate yields

$$ND_N < 2\sum_{j=1}^{s} g_j + (2\log N)^s \prod_{j=1}^{s} \frac{g_j - 1}{\log g_j} \quad \text{for } N \geqslant 1.$$

A higher-dimensional analogue of the Zaremba sequence was proposed in [357]. We may summarize the above results as follows.

3.6. Theorem (Halton [99]). *For any dimension $s \geqslant 1$, there exists an infinite sequence of points in I^s such that*

$$D_N = O\big(N^{-1}(\log N)^s\big). \tag{3.13}$$

In addition, for every $N \geqslant 2$ there exists a finite sequence of N points in I^s such that

$$D_N = O\big(N^{-1}(\log N)^{s-1}\big). \tag{3.14}$$

This result is significant since it guarantees that for any dimension there exist quasi-Monte Carlo techniques that perform substantially better than the statistical Monte Carlo method.

It is a widely held belief that the orders of magnitude in (3.13) and (3.14) are best possible. In (3.13) this is known only for $s = 1$ (see the result of W. M. Schmidt [282] quoted in the discussion of the one-dimensional case), and in (3.14) only for $s = 1$ (cf. [174, p. 90]) and $s = 2$ (by (3.11)). Otherwise, there remains the gap between (3.13) and (3.12) on the one hand and between (3.14) and (3.10) on the other.

K. F. Roth has informed the author about an important development concerning the L^2 discrepancy T_N for higher dimensions. In his forthcoming paper [263], Roth shows for any $N \geqslant 2$ the existence of a sequence of N points in I^3 with $T_N = O(N^{-1}\log N)$, and in [264] he establishes for $s \geqslant 4$ and any $N \geqslant 2$ that there are sequences of N points in I^s with $T_N = O(N^{-1}(\log N)^{(s-1)/2})$. Therefore, the lower bound for T_N in (3.10) is best possible. Calculations of L^2 discrepancies for interesting sequences in dimensions $2 \leqslant s \leqslant 9$ were carried out in [357]. As to the L^1 discrepancy, Sobol′ [299] has given a lower bound for any dimension that is of the order N^{-1}. This is best possible for $s = 1$. For $s > 1$, W. M. Schmidt [284a] recently obtained an improved lower bound of the order

$$N^{-1} (\log \log N)/\log \log \log N,$$

and in the same paper it is shown that in (3.10) the L^2 discrepancy T_N can be replaced by the L^p discrepancy for any $p > 1$.

The distribution behavior of sequences such as the van der Corput and Roth sequences has led I. M. Sobol′ to the development of a general theory of sequences possessing certain uniformity properties with respect to dyadic boxes. Here it is convenient to count also I as a dyadic interval, so that I may now be used in the formation of dyadic boxes (compare with §2). We consider first the case of a finite sequence.[17]

3.7. Definition (Sobol′ [298], [300]). For a nonnegative integer τ, a finite sequence of points in I^s is called a *P_τ-net* if it contains 2^k points, where $k > \tau$ is an integer, and if every dyadic box of volume $2^{\tau-k}$ contains exactly 2^τ points of the sequence.

For instance, every initial segment of the van der Corput sequence containing 2^k points, where $k \geqslant 1$, is a P_0-net in I, and the Roth sequence is a P_0-net in I^2. The parameter τ cannot be selected quite arbitrarily since its choice is limited in a certain way by the dimension s. For example, there exist arbitrarily long P_0-nets in I, I^2, and I^3, but no P_0-net in I^s, $s \geqslant 4$, containing more than 2 points (Sobol′ [300]). Let the least value of τ for which there exist arbitrarily long P_τ-nets in I^s be denoted by $\tau(s)$. Thus, $\tau(1) = \tau(2) = \tau(3) = 0$; moreover, $\tau(4) = 1$ and, in general, $\tau(s) = O(s \log s)$ according to [303, Chapter 6]. The number $\tau(s)$ may be thought of as a geometric characteristic of the s-dimensional unit cube. Explicit constructions of P_τ-nets in I^s for $\tau \geqslant \tau(s)$ can be found in [300], [303, Chapter 6]. These constructions involve only very simple digital operations with dyadic fractions, so that P_τ-nets may be generated easily in binary computers.

If τ is small, one would expect a P_τ-net to display an excellent distribution behavior. This is reflected in a general inequality for the nonuniformity $\varphi_\infty(2^k)$ of a P_τ-net in I^s containing 2^k points, namely

$$\varphi_\infty(2^k) < 2^{s-1+\tau}. \tag{3.15}$$

It should be emphasized that this bound does not depend on k. In other words, the nonuniformity of a P_τ-net is $O_\tau(1)$ in terms of the length of the net. In the light of the error estimate (2.9), this guarantees the usefulness of P_τ-nets for the purposes of numerical integration. We can have equality in (3.15); e.g., if $s = 1, 2$, or 3 and $k \geqslant s - 1$, then every P_0-net in I^s containing 2^k points satisfies $\varphi_\infty(2^k) = 2^{s-1}$ (cf. [303, Chapter 6]). There is also a discrepancy estimate for P_τ-nets, to the effect that the discrepancy D_N^* of a P_τ-net in I^s containing $N = 2^k$ points satisfies

$$ND_N^* < 2^\tau \sum_{j=0}^{s-1} \binom{k-\tau}{j} \tag{3.16}$$

[17] The Russian literature uses the concise term "net".

provided that $k \geqslant s - 1 + \tau$ ([**298**], [**300**]). Since k is of the order of magnitude $\log N$, the inequality (3.16) leads to the estimate $D_N^* = O_\tau(N^{-1}(\log N)^{s-1})$. In other words, the performance of P_τ-nets with small τ can be compared to that of the Hammersley sequence. In fact, if we take for a given dimension s a P_τ-ne^t in I^s with minimal τ, i.e., with $\tau = \tau(s)$, then an analysis of the constants implied by (3.16) and by the known discrepancy estimate for the Hammersley sequence reveals that, at least for sufficiently large s, such P_τ-nets actually "beat" the Hammersley sequence ([**300**], [**303**, Chapter 6]). The expected integration error implied by the use of P_τ-nets and an integrand chosen at random from a class of functions with rapidly convergent Haar series was estimated in [**308**], with an improvement in [**310**].

We turn now to the analogous theory for infinite sequences. We use the following terminology: if $\mathbf{x}_0, \mathbf{x}_1, \ldots$ is an infinite sequence, then by a *binary segment* thereof we mean a block consisting of those $\mathbf{x}_n$ with $h2^i \leqslant n < (h+1)2^i$, where $h \geqslant 0$ and $i \geqslant 1$ are fixed integers.

3.8. DEFINITION (SOBOL′ [**298**], [**300**]). For a nonnegative integer τ, an infinite sequence of points in I^s is called an *LP_τ-sequence* if each of its binary segments containing at least $2^{\tau+1}$ points is a P_τ-net.

The simplest example is provided by the van der Corput sequence, which forms an LP_0-sequence in I. It is clear from the definitions that LP_τ-sequences can only exist for $\tau \geqslant \tau(s)$. But more care has to be taken since we have, for instance, $\tau(3) = 0$, but there are no LP_0-sequences in I^3; on the other hand, I^2 still contains LP_0-sequences (cf. [**303**, Chapter 6]). As to the explicit construction of LP_τ-sequences, the principles pertaining to P_τ-nets also apply here. For every dimension s, concrete values of τ are known for which LP_τ-sequences in I^s exist.

On the basis of their definition, LP_τ-sequences should be very evenly distributed. More explicitly, every LP_τ-sequence in I^s satisfies

$$\varphi_\infty(N) \leqslant 2^{s-1+\tau} \quad \text{for all } N \geqslant 1, \tag{3.17}$$

so that $\varphi_\infty(N) = O(1)$ as a function of N (cf. [**298**], [**300**]). From the criterion for uniform distribution in terms of φ_∞ it follows then that every LP_τ-sequence is uniformly distributed. We also infer from (3.17) that the van der Corput sequence satisfies $\varphi_\infty(N) = 1$ for all $N \geqslant 1$, a fact already established in [**294**], and since we have $\varphi_\infty(N) \geqslant 2$ for any N points in I^2 with $N \geqslant 2$ ([**303**, Chapter 4]), there is a sequence in I^2 with $\varphi_\infty(N) = 2$ for all $N \geqslant 2$. Analogous questions for I^3 have been discussed in [**305**]. For the discrepancy D_N^* of an LP_τ-sequence in I^s one has the estimate

$$ND_N^* \leqslant 2^\tau \sum_{j=0}^{s-1} \binom{m-\tau+1}{j+1} + 2^\tau - 1 \quad \text{for } N \geqslant 2^{s-1+\tau}, \tag{3.18}$$

where $m = [\log_2 N]$ (cf. [**298**], [**300**]). In terms of orders of magnitude we obtain $D_N^* = O(N^{-1}(\log N)^s)$, as for the Halton sequence. If one chooses for a given dimension s the minimal τ for which an LP_τ-sequence in I^s is known to exist, then it turns out that the bound in (3.18) is asymptotically better than the corresponding one for the Halton sequence, in the sense that the implied constant is smaller for sufficiently large s (cf. [**300**], [**303**, Chapter 6]).

LP_τ-sequences in I^s possessing additional uniformity properties, e.g., with respect to the 2^s cubes into which I^s is divided by the hyperplanes $x_j = 1/2$ $(1 \leqslant j \leqslant s)$, have been considered in [315], [316].

A program for calculating LP_τ-sequences was written up in [317]. Tables for generating such sequences are available in [303, Chapter 6], [309, Appendix], [316] and extend now to the dimension $s = 51$. The only detailed expository account of the theory of P_τ-nets and LP_τ-sequences is in Russian (cf. [303]).

Quasi-Monte Carlo methods based on LP_τ-sequences have been employed in various contexts. We mention applications to search algorithms ([3], [309, Chapter 8], [317]), to the problem of bounding the real parts of the eigenvalues of a matrix [318], and to the simulation of systems [353].

Following this discussion of special sequences, we add some remarks concerning the estimation of discrepancy in the multidimensional case. Because of later applications and the fact that the points in many of the interesting finite sequences have rational coordinates, we state a generalization of Lemma 3.3 that is proved in the same way as before. For an integer $m \geqslant 2$, the summation symbol $\Sigma^*_{\mathbf{h} \, (\mathrm{mod}\, m)}$ designates a sum over all lattice points $\mathbf{h} = (h_1, \ldots, h_s) \in \mathbf{Z}^s$ with $-m/2 < h_j \leqslant m/2$ for $1 \leqslant j \leqslant s$ and $\mathbf{h} \neq \mathbf{0}$. For a lattice point in this range, we define

$$r(\mathbf{h}, m) = r(h_1, m) \cdots r(h_s, m), \tag{3.19}$$

where $r(h, m) = m \sin(\pi |h|/m)$ for $h \neq 0$ and $r(0, m) = 1$.

3.9. LEMMA. *Let $\mathbf{y}_0, \ldots, \mathbf{y}_{N-1}$ be s-dimensional lattice points. Then for any integer $m \geqslant 2$, the discrepancy D_N of the finite sequence of fractional parts*[18] *$\{(1/m)\mathbf{y}_0\}, \ldots, \{(1/m)\mathbf{y}_{N-1}\}$ satisfies*

$$D_N \leqslant \frac{s}{m} + \sum^*_{\mathbf{h}(\mathrm{mod}\, m)} \frac{1}{r(\mathbf{h}, m)} \left| \frac{1}{N} \sum_{n=0}^{N-1} e^{2\pi i \mathbf{h} \cdot \mathbf{y}_n / m} \right|,$$

where the dot denotes the standard inner product in $\mathbf{R}^s$.

There is also a multidimensional version of the general inequality stated in Theorem 3.1 (see [234]). The special case contained in Corollary 3.2 had been generalized earlier ([157], [331]); see also [174, p. 116]. Extensions of the results in Theorem 3.4 and Corollary 3.5 to higher dimensions are not known.

The isotropic discrepancy J_N introduced in Definition 2.12 has received some attention recently. A global lower bound was established by W. M. Schmidt [284], namely

$$J_N \geqslant C_s N^{-2/(s+1)} \tag{3.20}$$

for any N points in I^s, where C_s is a positive constant only depending on s. This is a considerable improvement on a result of Zaremba [370]. The expected order of magnitude of J_N for a sequence of points selected at random from I^s is also known. For $s = 2$, Philipp [244] proved a law of the iterated logarithm for the isotropic discrepancy, and so, in particular, $J_N = O(N^{-1/2}(\log \log N)^{1/2})$ almost surely. Stute [330] investigated the behavior

[18] For $\mathbf{t} = (t_1, \ldots, t_s) \in \mathbf{R}^s$, we define $\{\mathbf{t}\} = (\{t_1\}, \ldots, \{t_s\}) \in I^s$ to be its *fractional part*.

of J_N for higher dimensions and obtained the following metric theorems. For $s = 3$, we have $J_N = O(N^{-1/2}(\log N)^{3/2})$ almost surely, and for $s \geqslant 4$ we have

$$J_N = O\left(N^{-2/(s+1)}(\log N)^{2/(s+1)}\right) \text{ almost surely.}$$

These results indicate that the lower bound (3.20) is remarkably accurate. In fact, for $s \geqslant 3$ the exponent of N in (3.20) cannot be improved, and there is a very thin margin between the performance of sequences with minimal isotropic discrepancy and that of random sequences. The information currently available does not allow us to decide whether (3.20) is also best possible for $s = 2$. There is also an open question concerning the law of the iterated logarithm for J_N. Clearly, the lower bound (3.20) rules out such a law for $s \geqslant 4$, but it could conceivably hold for $s = 3$.

At present, one knows of no feasible method of estimating the isotropic discrepancy of interesting sequences of points in I^s, $s \geqslant 2$, in a direct manner. There is, however, a way of getting upper bounds for J_N via the discrepancy D_N. The best result in this direction is an inequality of Niederreiter and Wills [235] to the effect that

$$J_N \leqslant 4s\, D_N^{1/s} \tag{3.21}$$

for any N points in I^s. This improves earlier estimates in [111], [119], [206], [217]. An example of Zaremba [367] shows that the exponent $1/s$ in (3.21) is best possible. On the basis of a theorem of Niederreiter [215], Zaremba [376] presents an algorithm for calculating J_N in case $s = 2$. Computational experience with this algorithm is still lacking.

If $\mathfrak{M}_b$ is the family of subsets of I^s introduced in Definition 2.10, then the discrepancy $D_N(\mathfrak{M}_b)$ defined in accordance with Definition 2.5 can still be estimated in terms of D_N. We recall that this type of discrepancy occurs in the error bound for quasi-Monte Carlo integration over a Jordan-measurable subset of I^s (see Theorem 2.11). The following general result holds.

3.10. Theorem (Niederreiter and Wills [235]). *Let $b = b(\varepsilon)$ be a nondecreasing function on the positive reals with $b(\varepsilon) \geqslant \varepsilon$ for all $\varepsilon > 0$ and $\lim_{\varepsilon \to 0+} b(\varepsilon) = 0$. Then for any N points in I^s the discrepancies $D_N(\mathfrak{M}_b)$ and D_N satisfy*

$$D_N(\mathfrak{M}_b) \leqslant 4b\left(2\sqrt{s}\, D_N^{1/s}\right).$$

A somewhat better, but more complicated estimate, which holds also without the condition $b(\varepsilon) \geqslant \varepsilon$, is contained in the original paper. The first inequality of this type was shown in [220]. In the interesting case where $b(\varepsilon) = C\varepsilon$ for $\varepsilon > 0$, with $C > 0$ being a constant, the inequality in [235] yields

$$D_N(\mathfrak{M}_b) \leqslant (4C\sqrt{s} + 2C + 1)D_N^{1/s}.$$

For this special case, W. M. Schmidt [284] shows that a similar estimate holds with D_N replaced by the "cube discrepancy", i.e., the discrepancy $D_N(\mathfrak{W})$ with $\mathfrak{W}$ being the family of closed cubes in I^s with sides parallel to the coordinate axes.[19]

[19] W. M. Schmidt also studied the discrepancy with respect to other families of special sets such as triangles, balls, spherical caps, etc. We refer to [283] for an account of this work.

For certain purposes it has become necessary to construct sequences of quasi-random points in the infinite-dimensional unit cube I^∞, the cartesian product of denumerably many copies of I. Sobol′ [**293**] defined an infinite-dimensional version of the Halton sequence, and the same author [**301**] introduced LP_τ-sequences in I^∞. Both types of sequences are not only useful for numerical integration in I^∞ ([**293**], [**301**], [**303**, Chapter 7], [**309**, Chapter 7], [**313**], [**314**]), but also for solving integral equations [**313**], simulating Markov chains ([**312**]–[**314**]), and for problems arising in particle physics [**302**].

4. Good lattice points. The method of low-discrepancy sequences suffers from a drawback that we have not yet brought into focus, but which is quite apparent upon inspection of the error bounds in §2. The point is that once the integrand is sufficiently regular, say of bounded variation in the sense of Hardy and Krause, then any additional regularity of the integrand is not reflected in the error estimate. This is in marked contrast to classical integration methods which, as a rule, become more efficient the more regular the integrand is. It should be mentioned, though, that the said deficiency is, in principle, also shared by the statistical Monte Carlo method.

There is a special kind of quasi-Monte Carlo technique, the so-called *method of good lattice points*, in which the degree of regularity of the integrand is duly honored. But a price has to be paid for this luxury since this method only applies to integrands that are periodic of period 1 in each variable. If f is such an integrand defined on $\mathbf{R}^s$, then the approximation used in this method is of the form

$$\int_{I^s} f(\mathbf{t})\,d\mathbf{t} \approx \frac{1}{m}\sum_{n=1}^{m} f\left(\frac{n}{m}\,\mathbf{g}\right), \tag{4.1}$$

where $m \geqslant 2$ is a fixed (large) integer and $\mathbf{g} \in \mathbf{Z}^s$ is a suitably chosen s-dimensional lattice point. Although the nodes $(n/m)\mathbf{g}$ in (4.1) need not belong to I^s, the approximation (4.1) is really of the same type as those considered earlier since each node may be replaced by its fractional part $\{(n/m)\mathbf{g}\} \in I^s$ because of the periodicity property of f.

At first glance, the restriction to periodic integrands seems to render the method unfit for most practical applications. However, there are ways of transforming a given integrand into a periodic one while preserving regularity properties and the value of the integral, though this has to be done at the expense of further calculations. A simple device[20] is the replacement of a given function f on I^s by the function

$$h(t_1, \ldots, t_s) = 2^{-s}\sum_{\varepsilon_1=0}^{1} \cdots \sum_{\varepsilon_s=0}^{1} f\left(\varepsilon_1 + (-1)^{\varepsilon_1}t_1, \ldots, \varepsilon_s + (-1)^{\varepsilon_s}t_s\right)$$

for $(t_1, \ldots, t_s) \in I^s$. Then h is identical on opposite faces of I^s and can therefore be extended to a periodic function on $\mathbf{R}^s$ of period 1 in each variable. Furthermore, we have

[20] This symmetrization process may be seen as an analogue of the antithetic variate technique in the statistical Monte Carlo method (cf. [**28**, Chapter 2, §2]).

$$\int_{I^s} f(\mathbf{t})\,d\mathbf{t} = \int_{I^s} h(\mathbf{t})\,d\mathbf{t}.$$

More sophisticated methods depend on appropriate changes of variables ([109], [169], [275]) or on adding an expression involving Bernoulli polynomials to the integrand ([167], [169], [369]). Surveys of these methods can be found in [369], [373].

The lattice point $\mathbf{g} \in \mathbf{Z}^s$ should be chosen so as to yield a small error in (4.1). Let f be a function on $\mathbf{R}^s$ which is periodic with period 1 in each variable and can be expanded into an absolutely convergent multiple Fourier series

$$f(\mathbf{t}) = \sum_{\mathbf{h}} c_{\mathbf{h}} e^{2\pi i \mathbf{h}\cdot\mathbf{t}},$$

where the sum is extended over all lattice points $\mathbf{h} \in \mathbf{Z}^s$ and $\mathbf{h}\cdot\mathbf{t}$ denotes the standard inner product in $\mathbf{R}^s$. If we write again $e(t) = e^{2\pi i t}$ for $t \in \mathbf{R}$ and observe that $c_{\mathbf{0}} = \int_{I^s} f(\mathbf{t})\,d\mathbf{t}$, then we get

$$\begin{aligned}\frac{1}{m}\sum_{n=1}^{m} f\left(\frac{n}{m}\mathbf{g}\right) - \int_{I^s} f(\mathbf{t})\,d\mathbf{t} &= \frac{1}{m}\sum_{n=1}^{m}\sum_{\mathbf{h}} c_{\mathbf{h}} e\left(\frac{n}{m}\mathbf{h}\cdot\mathbf{g}\right) - c_{\mathbf{0}} \\ &= \frac{1}{m}\sum_{\mathbf{h}} c_{\mathbf{h}} \sum_{n=1}^{m} e\left(\frac{n}{m}\mathbf{h}\cdot\mathbf{g}\right) - c_{\mathbf{0}} \\ &= \frac{1}{m}\sum_{\mathbf{h}\neq\mathbf{0}} c_{\mathbf{h}} \sum_{n=1}^{m} e\left(\frac{n}{m}\mathbf{h}\cdot\mathbf{g}\right).\end{aligned}$$

Now the inner sum is equal to 0 if $\mathbf{h}\cdot\mathbf{g} \not\equiv 0 \pmod{m}$ and equal to m if $\mathbf{h}\cdot\mathbf{g} \equiv 0 \pmod{m}$, and so

$$\frac{1}{m}\sum_{n=1}^{m} f\left(\frac{n}{m}\mathbf{g}\right) - \int_{I^s} f(\mathbf{t})\,d\mathbf{t} = \sum_{\substack{\mathbf{h}\neq\mathbf{0} \\ \mathbf{h}\cdot\mathbf{g}\equiv 0 \ (\mathrm{mod}\ m)}} c_{\mathbf{h}}. \tag{4.2}$$

Thus the integration error depends on the magnitude of certain Fourier coefficients. As in one-dimensional harmonic analysis, there is a principle in operation which says that the more regular f is, the more quickly its Fourier coefficients $c_{\mathbf{h}}$ will tend to 0 as $\mathbf{h}$ moves away from the origin. We measure the distance of $\mathbf{h} = (h_1, \ldots, h_s) \in \mathbf{Z}^s$ from the origin by the expression

$$r(\mathbf{h}) = \prod_{j=1}^{s} \max(1, |h_j|). \tag{4.3}$$

In other words, $r(\mathbf{h})$ is the absolute value of the product of the nonzero coordinates of $\mathbf{h}$.

4.1. Definition. For real numbers $k > 1$ and $C > 0$, we say that $f \in \mathcal{E}^k(C)$ if

$$|c_{\mathbf{h}}| \leqslant C\, r(\mathbf{h})^{-k} \quad \text{for all } \mathbf{h} \neq \mathbf{0}, \tag{4.4}$$

and that $f \in \mathcal{E}^k$ if $f \in \mathcal{E}^k(C)$ for some $C > 0$, i.e., if (4.4) holds for some $C > 0$.

There is a sufficient condition for the validity of (4.4) that is easier to check. Let $k > 1$ be an integer and suppose all the partial derivatives

$$\frac{\partial^{q_1+\cdots+q_s} f}{\partial t_1^{q_1} \cdots \partial t_s^{q_s}} \quad \text{with } 0 \leqslant q_j \leqslant k-1 \text{ for } 1 \leqslant j \leqslant s$$

exist and are of bounded variation on I^s in the sense of Hardy and Krause; then $f \in \mathcal{E}^k(C)$ with a value of C which can be given explicitly (Zaremba [366]). Thus (4.4) may be thought of as a regularity condition on f.

If $f \in \mathcal{E}^k$, then it is seen easily that its Fourier series is automatically absolutely convergent. By combining (4.2) and (4.4), we arrive at the following result which will be basic for the sequel. To facilitate the writing, we set

$$P^{(k)}(\mathbf{g}, m) = \sum_{\substack{\mathbf{h} \neq \mathbf{0} \\ \mathbf{h} \cdot \mathbf{g} \equiv 0 \pmod{m}}} r(\mathbf{h})^{-k}.$$

4.2. Theorem (Korobov [161], Hlawka [113]). *For every $f \in \mathcal{E}^k(C)$, $\mathbf{g} \in \mathbf{Z}^s$, and integer $m \geqslant 2$, we have*

$$\left| \frac{1}{m} \sum_{n=1}^{m} f\left(\frac{n}{m}\mathbf{g}\right) - \int_{I^s} f(\mathbf{t})\, d\mathbf{t} \right| \leqslant C P^{(k)}(\mathbf{g}, m).$$

It is clear that in the sum $P^{(k)}(\mathbf{g}, m)$ the main contribution comes from the lattice points $\mathbf{h}$ close to $\mathbf{0}$. We introduce a simplified, and finite, sum taking into account only such lattice points. We make use of the summation symbol

$$\sum_{\mathbf{h} \pmod{m}}^{*} \tag{4.5}$$

which designates (as in Lemma 3.9) a sum over all $\mathbf{h} = (h_1, \ldots, h_s) \in \mathbf{Z}^s$ with $-m/2 < h_j \leqslant m/2$ for $1 \leqslant j \leqslant s$ and $\mathbf{h} \neq \mathbf{0}$.

4.3. Definition. For $\mathbf{g} \in \mathbf{Z}^s$ and an integer $m \geqslant 2$, we set

$$R(\mathbf{g}, m) = \sum_{\substack{\mathbf{h} \pmod{m} \\ \mathbf{h} \cdot \mathbf{g} \equiv 0 \pmod{m}}}^{*} r(\mathbf{h})^{-1}.$$

The sum $P^{(k)}(\mathbf{g}, m)$ can now be estimated in terms of $R(\mathbf{g}, m)$. In fact, by using the method in [174, pp. 156–157] one shows that for any $k > 1$, $\mathbf{g} \in \mathbf{Z}^s$, and integer $m \geqslant 2$ we have

$$P^{(k)}(\mathbf{g}, m) \leqslant (1 + 2\zeta(k))^s \left(m^{-k} + R(\mathbf{g}, m)^k\right), \tag{4.6}$$

where $\zeta(k) = \sum_{n=1}^{\infty} n^{-k}$ is the Riemann zeta-function.

Because of Theorem 4.2 and the estimate (4.6), the usefulness of the present method depends on our ability to find lattice points $\mathbf{g}$ for which $R(\mathbf{g}, m)$ is small. The following new result, which was so far only known for primes m ([113], [373]), guarantees the existence of such lattice points.

4.4. Theorem (Niederreiter [231]). *For any integer $m \geqslant 2$ and every dimension $s \geqslant 2$, there exist lattice points $\mathbf{g} \in \mathbf{Z}^s$ with coordinates relatively prime to m and*

$$R(\mathbf{g}, m) < m^{-1}\left(\tfrac{7}{5} + 2 \log m\right)^s. \tag{4.7}$$

A lattice point $\mathbf{g} \in \mathbf{Z}^s$ satisfying (4.7) is called a *good lattice point modulo m.*

Korobov [161], [163], [169] speaks of "optimal coefficients modulo m" when referring to the coordinates of good lattice points modulo m. By combining the above results, we obtain that for a good lattice point $\mathbf{g}$ modulo m the integration error in (4.1) is of the order of magnitude $O(m^{-k}(\log m)^{ks})$ for integrands $f \in \mathcal{E}^k$. Therefore, the error becomes smaller for larger k, i.e., for more regular integrands. Bahvalov [5] has shown that at least for primes m there exist lattice points $\mathbf{g}$ for which the exponent ks of $\log m$ can be replaced by $k(s-1)$, but it is also known (see Šarygin [276]) that the exponent cannot be reduced to anything less than $s-1$.

The expression defining $R(\mathbf{g}, m)$ does not lend itself to a convenient numerical treatment since it is a sum of many small numbers. A quantity that is much easier to calculate is the integer

$$\rho(\mathbf{g}, m) = \min_{\mathbf{h}} r(\mathbf{h}) \quad \text{for } \mathbf{g} \in \mathbf{Z}^s, s \geqslant 2, \tag{4.8}$$

where $\mathbf{h}$ ranges over the same lattice points as in the summation for $R(\mathbf{g}, m)$, i.e., $\mathbf{h} = (h_1, \ldots, h_s) \in \mathbf{Z}^s$ with $\mathbf{h} \neq \mathbf{0}$, $\mathbf{h} \cdot \mathbf{g} \equiv 0 \pmod m$, and $-m/2 < h_j \leqslant m/2$ for $1 \leqslant j \leqslant s$. It is trivial that

$$1/\rho(\mathbf{g}, m) \leqslant R(\mathbf{g}, m). \tag{4.9}$$

On the other hand, $R(\mathbf{g}, m)$ may also be estimated from above in terms of $\rho(\mathbf{g}, m)$. The following inequality improves upon earlier results of Hlawka [113] and Zaremba [373].

4.5. Theorem (Niederreiter [229]). *For any integer $m \geqslant 2$, any dimension $s \geqslant 2$, and any lattice point $\mathbf{g} \in \mathbf{Z}^s$ we have*

$$R(\mathbf{g}, m) < \frac{(2 \log m)^s + 4(2 \log m)^{s-1}}{(\log 2)^{s-1}\rho(\mathbf{g}, m)} + \frac{2^{s+1}(2^{s-2}-1)}{\rho(\mathbf{g}, m)} \binom{p+s-2}{s-1},$$

where $p = [\log_2 m]$.

It is now clear that, as far as the quantity $\rho(\mathbf{g}, m)$ is concerned, the desirable lattice points are those for which $\rho(\mathbf{g}, m)$ is large. Such lattice points are also labeled "good lattice points modulo m". Thus, the characteristic property of a good lattice point $\mathbf{g}$ modulo m may be informally described by saying that all lattice points $\mathbf{h} \neq \mathbf{0}$ solving the congruence $\mathbf{h} \cdot \mathbf{g} \equiv 0 \pmod m$ should be rather far removed from the origin.[21] By combining (4.7) and (4.9), we see that $\rho(\mathbf{g}, m)$ can be at least of the order of magnitude $m/\log^s m$. A direct approach, first used by Hlawka [116] for primes m, yields an even better result.

4.6. Theorem (Zaremba [375]). *For every dimension $s \geqslant 2$ and every sufficiently large integer m there exists at least one lattice point $\mathbf{g} = (1, g_2, \ldots, g_s) \in \mathbf{Z}^s$ such that*

$$\rho(\mathbf{g}, m) > (s-1)!\, m/(2 \log m)^{s-1}.$$

At least for primes m, this result can be made effective by using the method

[21] This could be thought of as a modulo m analogue of the notion of "badly approximable point" in simultaneous diophantine approximations (compare with §5).

in [229, §4]. On the other end of the scale, we have $\rho(\mathbf{g}, m) \leq m/2$ for any $\mathbf{g} \in \mathbf{Z}^s$, $s \geq 2$, and any integer $m \geq 2$. For if $\mathbf{g} = (g_1, \dots, g_s)$ with g_1 relatively prime to m, then $hg_1 + g_2 \equiv 0 \pmod{m}$ for some $h \in \mathbf{Z}$ with $-m/2 < h \leq m/2$, and so $\rho(\mathbf{g}, m) \leq \max(1, |h|) \leq m/2$. Otherwise, there exists a proper divisor d of m with $dg_1 \equiv 0 \pmod{m}$, and then $\rho(\mathbf{g}, m) \leq |d| \leq m/2$.

The quantity $P^{(k)}(\mathbf{g}, m)$ occurring in the error bound in Theorem 4.2 can be estimated directly in terms of $\rho(\mathbf{g}, m)$. First of all, it is clear from the definitions that for $s \geq 2$ we always have

$$P^{(k)}(\mathbf{g}, m) \leq P^{(2)}(\mathbf{g}, m)/\rho(\mathbf{g}, m)^{k-2} \quad \text{for } k \geq 2. \tag{4.10}$$

A detailed study of $P^{(2)}(\mathbf{g}, m)$ leads then to the following inequality.

4.7. Theorem (Zaremba [373]). *For every dimension $s \geq 2$, any integer $m \geq 4$ and any lattice point of the form $\mathbf{g} = (1, g_2, \dots, g_s) \in \mathbf{Z}^s$ we have*

$$P^{(2)}(\mathbf{g}, m) \leq \frac{8^s \pi^2}{6\rho(\mathbf{g}, m)^2}\left(\frac{1}{(s-1)!}\left(\log_2 \frac{m}{4}\right)^{s-1} + 2^{s-1}\left(\log_2 \frac{m}{2}\right)^{s-2}\right)$$
$$+ \frac{2\pi^2}{3m^2}\left(1 + \frac{\pi^2}{3}\right)^{s-1}.$$

In the 2-dimensional case, the size of the constants appearing in the above estimate can be reduced substantially (cf. [377]). Together with (4.10), we find that for $s \geq 2$ the quantity $P^{(k)}(\mathbf{g}, m)$ is of the order of magnitude $O(\rho(\mathbf{g}, m)^{-k}(\log m)^{s-1})$.

It must be pointed out that the existence theorems for good lattice points (Theorems 4.4 and 4.6) are by no means constructive, as they depend, respectively, on an averaging argument and a technique of eliminating "bad" lattice points. In general, the only way known so far of producing good lattice points is by an extensive search based on tabulations, preferably of the quantity $\rho(\mathbf{g}, m)$. This is a finite search problem since $\mathbf{g}$ only matters mod m, so that *a priori* there are m^s candidates in dimension s. Usually one restricts the attention to lattice points of the form $(1, g_2, \dots, g_s)$, which further limits the number of possibilities. The task of finding good lattice points is complicated by the fact that their coordinates depend strongly on the dimension. Thus, simplistic schemes such as taking a good lattice point $(g_1, \dots, g_s)$ modulo m and searching for an integer g_{s+1} so as to get a good lattice point $(g_1, \dots, g_s, g_{s+1})$ modulo m in dimension $s + 1$ are, as a rule, doomed to failure (see however [378]). The reason is, of course, that the nature of the solutions $\mathbf{h}$ of a congruence $\mathbf{h} \cdot \mathbf{g} \equiv 0 \pmod{m}$ changes if an additional coordinate is introduced. However, "going down" in the dimension is feasible to some extent because of the following principle. Let $\mathbf{g}^{(s)} = (g_1, \dots, g_s) \in \mathbf{Z}^s$ with $s \geq 3$ and consider the "truncated" point $\mathbf{g}^{(t)} = (g_1, \dots, g_t)$, where $2 \leq t < s$; then we have

$$R(\mathbf{g}^{(t)}, m) \leq R(\mathbf{g}^{(s)}, m) \quad \text{for any integer } m \geq 2, \tag{4.11}$$

since with any lattice point $\mathbf{h}^{(t)} = (h_1, \dots, h_t) \in \mathbf{Z}^t$ contributing to the sum defining $R(\mathbf{g}^{(t)}, m)$ the lattice point $\mathbf{h}^{(s)} = (h_1, \dots, h_t, 0, \dots, 0) \in \mathbf{Z}^s$

contributes to the sum defining $R(\mathbf{g}^{(s)}, m)$ and $r(\mathbf{h}^{(t)}) = r(\mathbf{h}^{(s)})$. For a similar reason we have

$$\rho(\mathbf{g}^{(t)}, m) \geqslant \rho(\mathbf{g}^{(s)}, m) \quad \text{for any integer } m \geqslant 2. \tag{4.12}$$

If $\mathbf{g}^{(s)}$ is now a good lattice point modulo m, say with $\rho(\mathbf{g}^{(s)}, m)$ large (or even as large as possible), then $\rho(\mathbf{g}^{(t)}, m)$ is large, which makes $\mathbf{g}^{(t)}$ an acceptable lattice point–but there is, of course, no guarantee that $\mathbf{g}^{(t)}$ will have retained the optimality property.

From the constructive point of view, the only case that has been dealt with in a satisfactory manner is $s = 2$. For $s \geqslant 3$ one has to take recourse to the currently available tables. The most useful table in the literature is the one of Maisonneuve [**185**] for dimensions $3 \leqslant s \leqslant 10$ and with the order of magnitude of the moduli m being as large as 10^5. An earlier table of Saltykov [**274**] is reprinted in [**169**] and covers certain prime moduli for dimensions $3 \leqslant s \leqslant 6$ and moduli that are products of two distinct primes for dimensions $3 \leqslant s \leqslant 10$. A short table for some selected prime moduli and dimensions $s = 3, 4, 5, 6, 8$, and 10 was compiled by Haber [**96**]. Kedem and Zaremba [**148**] extended the calculation of Maisonneuve for $s = 3$. Further efforts are reported in [**147**]. Some of these tables are restricted to good lattice points of the form $(1, g, g^2, \ldots, g^{s-1})$ which are easier to search for since there are initially only m candidates. The use of these special lattice points was suggested by Korobov [**165**], [**166**], [**169**]. They also play an important role in a different context to be expounded later (see §11). Integration errors involved in the calculation of typical 4-dimensional integrals by the method of good lattice points were tabulated by Brušlinskaja [**26**] (see also [**169**, Chapter 3]), and for 3- and 4-dimensional integrals by Maisonneuve [**185**].

As was already mentioned, an explicit construction of good lattice points is possible in the 2-dimensional case. It rests on observing and exploiting an intriguing relation with continued fractions. For an integer $m \geqslant 2$, we consider a lattice point $\mathbf{g} = (1, g)$ with $\gcd(g, m) = 1$. Let

$$\frac{g}{m} = [a_0; a_1, a_2, \ldots, a_q] = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \ddots + \cfrac{1}{a_q}}}$$

be the expansion of the rational g/m into a finite simple continued fraction, where we assume $a_q = 1$ for the sake of uniqueness. The integer a_0 is just the integral part of g/m and is actually not relevant since we may as well suppose that $1 \leqslant g < m$. The positive integers $a_1, a_2, \ldots, a_q$ are the partial quotients of g/m. Let

$$K(g/m) = \max(a_1, a_2, \ldots, a_q) \tag{4.13}$$

be the largest partial quotient. Then Zaremba [**365**] has shown that

$$\frac{m}{K+2} \leqslant \rho(\mathbf{g}, m) \leqslant \frac{m}{K} \quad \text{with } K = K\left(\frac{g}{m}\right). \tag{4.14}$$

Consequently, we obtain good lattice points modulo m by choosing rationals

g/m with small partial quotients. For somebody familiar with diophantine approximations, this should not come as a surprise since good lattice points were described earlier as analogues of "badly approximable" points, and this diophantine approximation property is governed by the size of partial quotients.[22]

One would expect to get "best" lattice points by selecting rationals g/m with $K(g/m) = 1$. These rationals are produced via the sequence of Fibonacci numbers, defined by $F_1 = F_2 = 1$ and $F_n = F_{n-1} + F_{n-2}$ for $n \geqslant 3$. We put then $g = F_{n-1}$ and $m = F_n$, where $n \geqslant 3$, and we have $K(g/m) = 1$ as desired. This choice of parameters goes back to Bahvalov [5] and Hua and Wang [130]. For the lattice point $\mathbf{g} = (1, F_{n-1})$ we have then $\rho(\mathbf{g}, F_n) = F_{n-2}$ according to a result of Zaremba [365]. Therefore,

$$\rho(\mathbf{g}, F_n) = (F_{n-2}/F_n)\, F_n \geqslant \tfrac{3}{8}\, F_n \quad \text{for } n \geqslant 5, \tag{4.15}$$

and in the same paper of Zaremba it is verified that for any fixed $n \geqslant 5$ the lattice point $\mathbf{g}$ yields the largest possible value of $\rho(\cdot, F_n)$. The upshot of this is that for infinitely many moduli m there are lattice points $\mathbf{g}$ such that $\rho(\mathbf{g}, m)$ is of the order of magnitude m. This is better than what is predicted by Theorem 4.6 for $s = 2$. It follows from (4.10) and Theorem 4.7 that for these lattice points $\mathbf{g}$ we get $P^{(k)}(\mathbf{g}, m) = O(m^{-k}\log m)$ for $k \geqslant 2$, and so the lower bound of Šarygin for the integration error (which we have mentioned earlier) is attained for $s = 2$. More concretely, Zaremba [377] has shown for $\mathbf{g} = (1, F_{n-1})$ that

$$P^{(k)}(\mathbf{g}, F_n) < 120\left(\frac{8}{3}\right)^{k-2} \frac{\log F_n}{F_n^k} \quad \text{for } k \geqslant 2 \text{ and } n \geqslant 5.$$

A detailed study of these lattice points $\mathbf{g}$ was undertaken in [371]. Two-dimensional lattice points arising from continued fraction expansions are also discussed in [98].

The question poses itself whether in the 2-dimensional case $\rho(\mathbf{g}, m)$ can reach the order of magnitude m for all moduli m and not just for the Fibonacci numbers. Because of (4.14) this is equivalent to the following problem about rational continued fractions. For $m \geqslant 2$ we set

$$K_m = \min_{\substack{g \\ \gcd(g,m)=1}} K\left(\frac{g}{m}\right)$$

and we ask whether K_m is uniformly bounded. Zaremba [373, p. 76] has put forth a conjecture which amounts to suggesting that $K_m \leqslant 5$ for all $m \geqslant 2$. This would mean that for every integer $m \geqslant 2$ there exists a reduced fraction g/m for which all partial quotients are bounded by 5. It would also imply that for every modulus m there is a lattice point $\mathbf{g}$ with $\rho(\mathbf{g}, m) \geqslant m/7$. Borosh [23] calculated K_m for $2 \leqslant m \leqslant 10^4$ and verified Zaremba's conjecture for this range. Actually, his data show that in this range we have $K_m = 5$ only for $m = 54$ and $m = 150$ and $K_m = 4$ for 23 values of m, the largest one being $m = 6234$. More recent calculations of Borosh and Niederreiter [24]

[22] For instance, for irrational numbers it is known that the smaller the partial quotients, the "worse approximable" the number.

yield $K_m \leqslant 3$ for $m = 2^\alpha$, $1 \leqslant \alpha \leqslant 35$, with $K_m = 2$ occurring frequently. There seems to be ample evidence to conjecture that $K_m \leqslant 3$ for all sufficiently large m and that the set of $m \geqslant 2$ with $K_m = 2$ has positive lower density. The best order of magnitude presently known is $K_m = O(\log m)$, which follows from Theorem 4.6 and (4.14).

The information available for $s = 2$ suggests that in the general case there may exist good lattice points $\mathbf{g}$ modulo m (at least for an infinite class of moduli m) such that $\rho(\mathbf{g}, m)$ is of the order of magnitude $m/(\log m)^{s-2}$. There is some empirical evidence supporting this conjecture, but nothing has been proved in this direction for $s \geqslant 3$.

Attempts have been carried out, however, to explicitly construct lattice points $\mathbf{g} \in \mathbf{Z}^s$, $s \geqslant 3$, for which $\rho(\mathbf{g}, m)$ is at least of a tolerable size. Taking as their cue the analogy between good lattice points and badly approximable points (in the sense of simultaneous diophantine approximations), Hua and Wang [134] employ the following procedure. Let $\boldsymbol{\alpha} = (\alpha_2, \ldots, \alpha_s) \in \mathbf{R}^{s-1}$ have coordinates which are real algebraic numbers such that $1, \alpha_2, \ldots, \alpha_s$ are linearly independent over the rationals. For a fixed integer $m \geqslant 2$, let $g_2, \ldots, g_s$ be integers such that we have the simultaneous diophantine approximation

$$|\alpha_j - g_j/m| \leqslant dm^{-s/(s-1)} \quad \text{for } 2 \leqslant j \leqslant s,$$

where d is a positive constant. Then it is shown that the lattice point $\mathbf{g} = (1, g_2, \ldots, g_s)$ satisfies

$$\rho(\mathbf{g}, m) > c(\boldsymbol{\alpha}, \varepsilon, d, s) m^{s/2(s-1)-\varepsilon} \tag{4.16}$$

for every $\varepsilon > 0$, where $c(\boldsymbol{\alpha}, \varepsilon, d, s)$ is a positive constant depending on the indicated parameters. As to a convenient choice of $\boldsymbol{\alpha}$, the authors already proposed in [132], [133] to consider, for a prime $2s + 1$, the numbers

$$\alpha_j = 2 \cos \frac{2\pi(j-1)}{2s+1} \quad \text{for } 2 \leqslant j \leqslant s,$$

which together with 1 form a basis for the real cyclotomic field $\mathbf{Q}(2 \cos 2\pi/(2s+1))$ of degree s over $\mathbf{Q}$. Calculations of Haber [96] used this particular point $\boldsymbol{\alpha}$ and led to promising results.

Another construction of Hua and Wang [134] is based on the properties of a special class of algebraic numbers. Let ω be a PV number (for Pisot-Vijayaraghavan number) of degree $s \geqslant 2$, i.e., ω is a real algebraic integer of degree s such that $\omega > 1$ and its remaining algebraic conjugates $\omega^{(2)}, \ldots, \omega^{(s)}$ all lie in the open unit disk. We may order these conjugates so that

$$|\omega^{(2)}| \leqslant |\omega^{(3)}| \leqslant \cdots \leqslant |\omega^{(s)}| < 1. \tag{4.17}$$

Suppose the minimal polynomial of ω over $\mathbf{Q}$ is $x^s - a_{s-1}x^{s-1} - \cdots - a_1 x - a_0 \in \mathbf{Z}[x]$, and let $\mathbf{b} = (b_0, b_1, \ldots, b_{s-1}) \in \mathbf{Z}^s$ be a lattice point $\neq \mathbf{0}$. Define the linear recurring sequence $Q_0, Q_1, \ldots$ of integers by $Q_i = b_i$ for $0 \leqslant i \leqslant s-1$ and the recurrence relation

$$Q_{n+s} = a_{s-1}Q_{n+s-1} + \cdots + a_1 Q_{n+1} + a_0 Q_n \quad \text{for } n = 0, 1, \ldots.$$

Then for sufficiently large n we have $|Q_n| \geqslant 2$ and the lattice point $\mathbf{g} = (1,$

$Q_{n+1}, \ldots, Q_{n+s-1}$) satisfies

$$\rho(\mathbf{g}, |Q_n|) > c(\omega, \mathbf{b}, \varepsilon)|Q_n|^{(1+\beta)/2-\varepsilon} \tag{4.18}$$

for every $\varepsilon > 0$, where $c(\omega, \mathbf{b}, \varepsilon)$ is a positive constant depending on the indicated parameters and

$$\beta = -\log|\omega^{(s)}|/\log \omega.$$

The stipulation (4.17) implies $\beta \leqslant 1/(s-1)$, and so the lower bound in (4.18) is not better than the one in (4.16). However, the present method has the advantage that it is more explicit than the previous one. It is suggested to take for ω the largest real root of the irreducible equation $x^s - x^{s-1} - \cdots - x - 1 = 0$ and $\mathbf{b} = (0, \ldots, 0, 1)$. The resulting linear recurring sequence $Q_0, Q_1, \ldots$ may then be thought of as a generalized Fibonacci sequence. For $s = 2$ we obtain the example involving Fibonacci numbers.

Good lattice points can be used to produce finite sequences with small discrepancy. Let $\mathbf{g} \in \mathbf{Z}^s$ be a lattice point and $m \geqslant 2$ a modulus. It follows then from Lemma 3.9 and the inequality

$$r(\mathbf{h}, m) \geqslant 2r(\mathbf{h}) \tag{4.19}$$

for lattice points $\mathbf{h} = (h_1, \ldots, h_s) \neq \mathbf{0}$ with $-m/2 < h_j \leqslant m/2$ $(1 \leqslant j \leqslant s)$ that the discrepancy D_m of the points $\{(1/m)\mathbf{g}\}, \{(2/m)\mathbf{g}\}, \ldots, \{(m/m)\mathbf{g}\}$ satisfies

$$D_m \leqslant s/m + \tfrac{1}{2}R(\mathbf{g}, m).$$

Thus, if $\mathbf{g}$ is chosen to be a good lattice point modulo m, we get a small discrepancy D_m. Theorem 4.4 implies the following result.

4.8. THEOREM. *For every dimension $s \geqslant 2$ and every integer $m \geqslant 2$, there exists a lattice point $\mathbf{g} \in \mathbf{Z}^s$ with coordinates relatively prime to m such that the discrepancy D_m of the points $\{(1/m)\mathbf{g}\}, \{(2/m)\mathbf{g}\}, \ldots, \{(m/m)\mathbf{g}\}$ satisfies*

$$D_m < \frac{s}{m} + \frac{1}{2m}\left(\tfrac{7}{5} + 2\log m\right)^s.$$

One conjectures that for certain s-dimensional lattice points one even has $D_m = O(m^{-1}(\log m)^{s-1})$, which would be the same order of magnitude as that of the discrepancy of the Hammersley sequence (see §3). This conjecture has only been verified for $s = 2$, where, predictably, one takes a Fibonacci number $m = F_n$ $(n \geqslant 3)$ and the lattice point $\mathbf{g} = (1, F_{n-1})$ and one obtains $D_m < (7/6)m^{-1}\log(15m)$ according to a result of Zaremba **[365]**. Sobol′ **[291]**, **[294]**, **[296]** investigated the behavior of sequences arising from good lattice points with respect to his nonuniformities φ_q and φ_∞ (see also **[303**, Chapter 5**]**). It follows from (2.8) that the finite sequence of points in Theorem 4.8 satisfies $\varphi_\infty(m) = O(\log^s m)$, a result shown in **[294]** for primes m. Proĭnov **[249]** considers the L^2 discrepancy of the points obtained from those in Theorem 4.8 after symmetrization with respect to all vertices of I^s.

A number of other questions about good lattice points have been studied in the literature. The effect of randomizing the lattice point $\mathbf{g}$ was considered by Bahvalov **[8]**, **[9]** and Stojancev **[324]**. The device of shifting the nodes by a random vector is studied in **[49]**. Solodov **[321]** demonstrated the applicability of the method of good lattice points for certain integration domains with smooth boundary. Further theoretical results and/or remarks about the subject can be found in [7], [11], [42], [97], [128], [162], [165], [229], [269], [319], [320], [322], [367], [374], [383]. For surveys of various aspects of the theory of good lattice points, see [10], [95], [131], [168], [169], [174, Chapter 2, §5], [303, Chapter 5], [356], [368], [373].

The method of good lattice points can be used in several areas of numerical analysis. A rather immediate application is to the approximate solution of integral equations. This was first considered by Korobov [164], with later efforts in this direction being contained in [114], [116], [124], [134], [166], [169, Chapter 4], [265]–[268], [275], [356], [381]. Interpolation techniques based on good lattice points were first studied by Rjaben'kiĭ [257] and Smoljak [287], and this was further pursued in [117], [166], [169, Chapter 4], [266], [270], [276], [288], [355]. An application of these interpolation formulas is made in [139]. Bahvalov [6] uses good lattice points in a Dirichlet problem; see also [326], [380], [382]. The papers [258] and [325] discuss similar applications to Cauchy problems. A relation between good lattice points and approximation theory occurs in [4]. Korobov [170] employs good lattice points to approximate the values of multiple sums with many terms.

Prior to his investigation of good lattice points, Korobov [160] considered nodes of the form $(n/m, n^2/m, \ldots, n^s/m)$, $n = 1, 2, \ldots, m$, where $m = p$ or p^2 with a prime $p > s$ (see also [78], [113], [116], [169, Chapter 2]). These do not perform as well as good lattice points since for an integrand $f \in \mathcal{E}^k$, $k > 1$, we only get an error bound of $O(m^{-1/2})$. The discrepancy of this finite sequence was estimated by Hlawka [113], [116], with the result that $D_m = O(m^{-1/2}\log^s m)$.

A set of nodes that is sometimes used in classical integration techniques consists of the points $(n_1/m, n_2/m, \ldots, n_s/m)$, where $m \geqslant 2$ is an integer and the n_j run independently through all integers from 0 to $m - 1$, so that there are altogether $N = m^s$ nodes. Korobov [162] notes that in a quasi-Monte Carlo integration with these points and for an integrand $f \in \mathcal{E}^k$, $k > 1$, the integration error is $O(N^{-k/s})$. See also [113], [116], [129], [169, Chapter 1], [289], [383]. For other sets of nodes related to this one or to good lattice points, we refer to [13], [51], [74], [128], [271], [272].

5. Application of diophantine approximations. In the method of good lattice points we use as nodes the multiples of a "rational" point $(1/m)\mathbf{g} \in \mathbf{R}^s$. A closely related method is obtained if one employs instead the multiples of an "irrational" point $\boldsymbol{\alpha} \in \mathbf{R}^s$, i.e., of a point $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_s)$ for which 1, $\alpha_1, \ldots, \alpha_s$ are linearly independent over the rationals. The periodicity conditions on the integrand will be the same as in §4. Thus, the approximation formula can be written as

$$\int_{I^s} f(\mathbf{t})\,d\mathbf{t} \approx \frac{1}{N} \sum_{n=1}^{N} f(n\boldsymbol{\alpha}). \tag{5.1}$$

In order to analyze the integration error, we assume again that f belongs to the function class $\mathcal{E}^k(C)$ for some $k > 1$ and $C > 0$ (see Definition 4.1). By expanding f into its absolutely convergent multiple Fourier series, an argument similar to the one in the preceding section leads to

$$\left|\frac{1}{N}\sum_{n=1}^{N} f(n\boldsymbol{\alpha}) - \int_{I^s} f(\mathbf{t})\,d\mathbf{t}\right| = \left|\frac{1}{N}\sum_{\mathbf{h}\neq\mathbf{0}} c_{\mathbf{h}} \sum_{n=1}^{N} e^{2\pi i\mathbf{h}\cdot(n\boldsymbol{\alpha})}\right|$$

$$\leqslant \frac{C}{N}\sum_{\mathbf{h}\neq\mathbf{0}} r(\mathbf{h})^{-k}\left|\sum_{n=1}^{N} e^{2\pi i n\mathbf{h}\cdot\boldsymbol{\alpha}}\right|$$

The inner sum is a finite geometric series with $\mathbf{h}\cdot\boldsymbol{\alpha} \notin \mathbf{Z}$ because of the condition on $\boldsymbol{\alpha}$. A straightforward estimate yields then

$$\left|\sum_{n=1}^{N} e^{2\pi i n\mathbf{h}\cdot\boldsymbol{\alpha}}\right| \leqslant \frac{1}{2\|\mathbf{h}\cdot\boldsymbol{\alpha}\|},$$

where

$$\|t\| = \min_{m\in\mathbf{Z}} |t-m| \quad \text{for } t \in \mathbf{R}$$

designates the absolute distance from t to the nearest integer. Altogether we have

$$\left|\frac{1}{N}\sum_{n=1}^{N} f(n\boldsymbol{\alpha}) - \int_{I^s} f(\mathbf{t})\,d\mathbf{t}\right| \leqslant \frac{C}{2N}\sum_{\mathbf{h}\neq\mathbf{0}} r(\mathbf{h})^{-k}\,\|\mathbf{h}\cdot\boldsymbol{\alpha}\|^{-1}. \tag{5.2}$$

The convergence of this infinite series depends on how small $\|\mathbf{h}\cdot\boldsymbol{\alpha}\|$ can become for lattice points $\mathbf{h}$ close to the origin. The study of this behavior belongs to the realm of simultaneous diophantine approximations. We use the following notion to classify irrational points with regard to their diophantine approximation character.

5.1. DEFINITION. For a real number η, an irrational point $\boldsymbol{\alpha} \in \mathbf{R}^s$ is said to be of *finite type* η if η is the infimum of all numbers σ for which there exists a positive constant $c = c(\boldsymbol{\alpha}, \sigma)$ such that

$$r(\mathbf{h})^{\sigma}\|\mathbf{h}\cdot\boldsymbol{\alpha}\| \geqslant c \tag{5.3}$$

holds for all lattice points $\mathbf{h} \in \mathbf{Z}^s$ with $\mathbf{h} \neq \mathbf{0}$. If no such number σ exists, then $\boldsymbol{\alpha}$ is said to be of *infinite type*.

It follows from the Minkowski linear forms theorem that we always have $\eta \geqslant 1$. Irrational points of the smallest possible type $\eta = 1$ may be called *badly approximable points*. Such points can be constructed explicitly. A fundamental theorem of W. M. Schmidt [**281**] shows that every algebraic irrational point[23] is of finite type $\eta = 1$. A result of A. Baker [**12**] implies that $\boldsymbol{\alpha} = (e^{r_1}, \ldots, e^{r_s})$ with distinct nonzero rationals $r_1, \ldots, r_s$ is of finite type $\eta = 1$.

We return now to the estimate (5.2). Using techniques from [**216**], one proves the convergence of the infinite series on the right-hand side under the assumption that k is larger than the type of $\boldsymbol{\alpha}$. Thus, we obtain the following error estimate.

[23] This is an irrational point with real algebraic numbers as coordinates.

5.2. THEOREM (HASELGROVE [109], NIEDERREITER [220]). *Let $\alpha \in \mathbf{R}^s$, $s \geqslant 1$, be an irrational point of finite type η. Then we have*

$$\frac{1}{N}\sum_{n=1}^{N} f(n\alpha) - \int_{I^s} f(\mathbf{t})\,d\mathbf{t} = O\left(\frac{1}{N}\right)$$

for every $f \in \mathcal{E}^k$ with $k > \eta$.

For greater efficiency one will choose a point α of the smallest possible type $\eta = 1$. The use of algebraic irrational points was proposed in [216] on the basis of the pertaining special case of Theorem 5.2. More specifically, one may, in case $2s + 3$ is prime, take the point

$$\alpha = \left(2\cos\frac{2\pi}{2s+3}, 2\cos\frac{4\pi}{2s+3}, \ldots, 2\cos\frac{2\pi s}{2s+3}\right)$$

arising from a basis of a real cyclotomic field (see §4); for any $s \geqslant 1$ one may choose $\alpha = (\xi, \xi^2, \ldots, \xi^s)$ with $\xi = p^{1/(s+1)}$ and p prime, or $\alpha = (\sqrt{p_1}, \ldots, \sqrt{p_s})$ with distinct primes $p_1, \ldots, p_s$. The first two possibilities are to be preferred since the coordinates stem from an algebraic number field of degree $s + 1$ over $\mathbf{Q}$ (which is the smallest degree that can occur for an algebraic irrational point), whereas the coordinates of the third point generate an algebraic number field of degree 2^s over $\mathbf{Q}$. Zinterhof [384] works with the irrational points of A. Baker mentioned above.

Historically, this method dates back to papers of Richtmyer [254] and Peck [240], the latter stating the result of Theorem 5.2 as a conjecture. Similar ideas were used later on in [5], [10], [14], [16], [18], [44], [45], [162], [169, Chapter 2], [197], [358].[24] Computational work based on this method was carried out in [52] and [259], and for a slightly modified method in [93]. For an application of this method to interpolation problems, see [384].

If we have a very regular integrand, say $f \in \mathcal{E}^k$ with k much larger than the type η of α, then the method, as it stands, does not honor the high degree of regularity of f. A way has been found to alleviate this deficiency, namely by abandoning the equal-weight formula (5.1) and working with weight distributions adapted to the specific regularity condition on f. This is the only instance in the theory of quasi-Monte Carlo integration where the use of nonequal weights leads to a demonstrable gain in accuracy. The idea of constructing special weight distributions to achieve better error bounds in results such as Theorem 5.2 occurs already in Bahvalov [5], Haselgrove [109], and Wang [356].

A general theorem along these lines is obtained as follows. For positive integers q and N, we define weights of the form $a_{Nn}^{(q)}N^{-q}$ with $0 \leqslant n \leqslant q(N-1)$. The $a_{Nn}^{(q)}$ are positive integers which are introduced most conveniently by means of a generating function. In detail, the $a_{Nn}^{(q)}$ are determined from the polynomial identity

$$\left(\sum_{j=0}^{N-1} z^j\right)^q = \sum_{n=0}^{q(N-1)} a_{Nn}^{(q)} z^n.$$

[24] There is also a result of Pjateckiĭ-Šapiro included in [77] and appearing as well in [169, Chapter 2].

With these weights, the following result holds.

5.3. Theorem (Niederreiter [220]). *Let q be a positive integer and $\boldsymbol{\alpha} \in \mathbf{R}^s$, $s \geqslant 1$, an irrational point of finite type η. Then we have*

$$\frac{1}{N^q} \sum_{n=0}^{q(N-1)} a_{Nn}^{(q)} f(n\boldsymbol{\alpha}) - \int_{I^s} f(\mathbf{t})\,d\mathbf{t} = O\left(\frac{1}{N^q}\right)$$

for every $f \in \mathcal{E}^k$ with $k > q\eta$.

Thus, while carrying out only about q times as many function evaluations, the error bound becomes approximately the qth power of what we had earlier, at least for sufficiently regular integrands. If we take for $\boldsymbol{\alpha}$ an algebraic irrational point, then the integration error is $O(N^{-q})$ for every $f \in \mathcal{E}^k$ with $k > q$. A related result was established by Hua and Wang [134].

It is an advantageous feature of this method that the point $\boldsymbol{\alpha}$ is chosen independently of N. Thus, if the value of N is increased, the previously calculated function values can be used again. This is not the case in the method of good lattice points, since the property of being a good lattice point modulo m is obviously relative to the choice of m. Of course, both methods share the drawback that a nonperiodic integrand first has to be "periodized" by one of the procedures mentioned in §4 before the integration technique can be applied. It should also be mentioned that in an actual machine calculation based on the method in this section one will have to replace the irrational $\boldsymbol{\alpha}$ by a rational approximation. This produces an additional inaccuracy which may, however, be classified as a roundoff error. A further limitation of the method was discussed by Tsuda [339] who used numerical data to show that it does not work so well if the integrand has a very high and narrow peak. The intrinsic reason is that the constant involved in the error term becomes too large in this case.

An analysis of the proofs of Theorems 5.2 and 5.3 reveals that the only noneffective constant is the one stemming from the diophantine inequality (5.3). In the multidimensional case, it is usually very difficult to get an effective value for $c(\boldsymbol{\alpha}, \sigma)$ as long as σ is close to the type η of $\boldsymbol{\alpha}$. For algebraic irrational points $\boldsymbol{\alpha}$, there are ways of obtaining an effective constant $c(\boldsymbol{\alpha}, \sigma)$ by considering larger values of σ (cf. [216]). A fuller discussion of the issue of effectiveness can be found in [220, §10].

The discrepancy of the sequence $\{\boldsymbol{\alpha}\}, \{2\boldsymbol{\alpha}\}, \ldots, \{n\boldsymbol{\alpha}\}, \ldots$ of fractional parts can be estimated in terms of the type of $\boldsymbol{\alpha}$. If $\boldsymbol{\alpha} \in \mathbf{R}^s$ is an irrational point of finite type η, then we have

$$D_N = O\left(N^{-1/((\eta-1)s+1)+\varepsilon}\right)$$

for every $\varepsilon > 0$ (cf. [220, §6], [174, pp. 131–132]). This is known to be essentially best possible for $s = 1$ (cf. [174, p. 124]). If $\boldsymbol{\alpha} \in \mathbf{R}^s$ is an algebraic irrational point, then the above estimate implies $D_N = O(N^{-1+\varepsilon})$ for every $\varepsilon > 0$, a result conjectured and heuristically supported by Richtmyer [255] and first shown by Niederreiter [217]. Further remarks on this case can be found in [256]. For $s = 1$, we already mentioned in §3 that there are irrationals for which $D_N = O(N^{-1} \log N)$. In fact, most irrationals will produce a discrepancy nearly as small as that, for if g is a positive nondecreasing function such that $\sum_{n=1}^{\infty} g(n)^{-1}$ converges, then $D_N = O(N^{-1}(\log N)\, g\,(\log\log N))$ for almost all $\alpha \in \mathbf{R}$ in the sense of Lebesgue measure (cf. [219], [174, p. 128]). In the multidimensional case, W. M. Schmidt [280] proved a metric result to the effect that for every $\varepsilon > 0$ we have $D_N = O(N^{-1}(\log N)^{s+1+\varepsilon})$ for almost all $\boldsymbol{\alpha} \in \mathbf{R}^s$ in the sense of s-dimensional Lebesgue measure. No individual $\boldsymbol{\alpha} \in \mathbf{R}^s$, $s \geqslant 2$, is known for which $D_N = O(N^{-1}(\log N)^{s+1})$.

PART II. PSEUDO-RANDOM NUMBERS

6. Random numbers *vs.* pseudo-random numbers. We have seen in §1 that the essential step in the statistical Monte Carlo method is the random sampling of points (or numbers) from a given set by independent trials. If we assume, for simplicity, that the set is the unit interval $I = [0, 1]$, then an "unbiased" execution of this sampling procedure should produce a sequence of "random numbers" in I. However, as long as no concrete definitions of "unbiased" and no equal opportunity regulations for numbers are adopted, the above concept of a sequence of random numbers is fictional, or at best an idealization. We are faced here with an obvious contradiction in terms, namely to try to define or characterize a supposedly haphazard process by a fixed set of rules known in advance. This dilemma shifts the problem from the mathematical to the metaphysical level, without resolving it, of course (to the relief of every gambling casino in the world). The question has indeed attracted the attention of philosophers of science ([152], [194], [246, Chapter 6]). Some authors[25] have also perceived the challenge of getting a handle on the vexing idea of randomness as being basic for the foundations of probability theory.

Many attempts at "defining" a sequence of random numbers in I recognize the following principles: (i) the sequence should satisfy certain distribution properties; (ii) these distribution properties should be invariant under certain selection rules for subsequences. This viewpoint is due essentially to von Mises ([350], [351, §1]), although (i) was acknowledged much earlier (see, e.g., [344, pp. 64–67]). On this basis, Knuth [154, §3.5] discusses in detail a hierarchy of proposed definitions for a random number sequence (see also [198]). The minimum requirement is that of uniform distribution in I. But even a very nicely distributed sequence can exhibit a distinctly nonrandom behavior, e.g., the van der Corput sequence (see §3) has the property that its terms alternately hit the intervals $[0, \frac{1}{2})$ and $[\frac{1}{2}, 1)$. Therefore, the statistical independence of successive terms has to be taken into account, and this leads eventually to the requirement of complete uniform distribution.[26] A sequence $x_1, x_2, \ldots$ of elements of I is called *completely uniformly distributed* (abbreviated CUD) if for every integer $s \geqslant 1$ the sequence of points $(x_n, x_{n+1}, \ldots, x_{n+s-1})$, $n = 1, 2, \ldots$, is uniformly distributed in I^s. Equivalently, the sequence $x_1, x_2, \ldots$ is CUD if for every $s \geqslant 1$ and every s-dimensional lattice point $(h_1, \ldots, h_s) \neq \mathbf{0}$ the sequence of fractional parts $\{h_1 x_n +$

[25] See [41], [43], [137], [158], [253], [348], [351] and the summaries in [192], [237, §1.5].

[26] The term "∞-distribution" is used in [154]. See [174, p. 205] for literature on this subject.

$h_2x_{n+1} + \cdots + h_sx_{n+s-1}\}$, $n = 1, 2, \ldots$, is uniformly distributed in I. CUD sequences have, in fact, been used as sources of random numbers for Monte Carlo-type calculations and simulations (see, e.g., [15], [34], [73], [309, Chapter 7]). Explicit constructions of CUD sequences have already been given in the older literature on the subject (cf. [247]), but these are too inconvenient for practical use. A CUD sequence consisting only of dyadic fractions was constructed by Knuth [153], and a principle of obtaining simple sequences that are in a sense approximately CUD was pointed out by Haber [94]. More recently, Rauzy [252] proved that if f is an entire function that is not a polynomial, attains real values on the real axis, and satisfies

$$\limsup_{r\to\infty} \frac{\log\log M(f;r)}{\log\log r} < \tfrac{5}{4},$$

where $M(f;r) = \sup_{|z|<r}|f(z)|$, then the sequence of fractional parts $\{f(1)\}$, $\{f(2)\}, \ldots, \{f(n)\}, \ldots$ is CUD. M. B. Levin [180] showed how to construct, for every transcendental[27] number $\theta > 1$, a number α such that the sequence of fractional parts $\{\alpha\theta\}$, $\{\alpha\theta^2\}, \ldots, \{\alpha\theta^n\}, \ldots$ is CUD. This is related to a metric result of Franklin [72] to the effect that $\{\theta\}$, $\{\theta^2\}, \ldots, \{\theta^n\}, \ldots$ is CUD for almost all transcendental numbers $\theta > 1$ (in the sense of Lebesgue measure).

So far, we have ignored the invariance principle (ii) mentioned above. If one insists that the sequence and *all* of its subsequences should be CUD, one arrives at a void concept, since any CUD sequence has a subsequence of terms in $[0, \frac{1}{2}]$, which is therefore not even uniformly distributed in I. Hence, some restriction has to be imposed on the admissible selection rules. If one allows only recursively enumerable selection rules, then the same argument as before leads to the interesting situation that every concretely given sequence fails this test for randomness. Further possibilities for selection rules are discussed in [154, §3.5], and one also finds there A. Wald's construction of a sequence enjoying remarkable invariance properties. A systematic study of the relationship between complete uniform distribution (or "normality", the analogous concept for sequences of digits) and selection rules is carried out in [145], [146].

For finite sequences of digits, an information-theoretic approach to a definition of randomness based on preparatory work of Kolmogorov [159] has been chosen by Chaitin [35] and Martin-Löf [191]. The idea here is to designate those sequences as random which, among all sequences of fixed length, maximize the shortest length of the program required to generate the sequence on a Turing machine.[28] A pleasant introduction to this concept is presented in [36], where one also finds an elaboration on the paradoxical statement that randomness in this sense cannot be verified for a specified sequence of digits. A detailed survey of this and other approaches to randomness is given in [285]. See also the recent papers [178], [179].

[27] It is easily seen that a sequence of the form $\{\alpha\theta\}$, $\{\alpha\theta^2\}, \ldots, \{\alpha\theta^n\}, \ldots$ cannot be CUD for algebraic θ.

[28] As Knuth [154, p. 149] points out, this is of course the worst concept from the viewpoint of practical random-number generation.

There is a way of generating "random numbers" (in the intuitive sense) that is very much in the spirit of the statistical Monte Carlo method, namely by using physical devices. Among the possibilities that have been considered are white noise produced by electronic circuits ([**28**, Chapter 6], [**90**]), counts of the emission of radiated particles ([**28**, Chapter 6], [**138**], [**203**], [**349**]), flipping coins [**354**], mechanical gadgets operating on the principle of the roulette wheel ([**104**], [**142**]), and automated versions of the latter ([**149**], [**251**]). Extensive tables recording these experiments or reproducing random numbers obtained from other data have been published ([**144**], [**149**], [**251**], [**335**]). See [**334**] for the early history of these tables.

The use of physical or tabulated random numbers is problematic because of the difficulties of storage, the necessity of frequent testing for randomness, and the fact that they are not generated in the computer, but by an external source. We have also seen that those concepts of randomness that sound convincing from an axiomatic standpoint suffer from the practical deficiency that no concrete sequence actually used in a calculation can verifiably satisfy the definition. Therefore, one has taken recourse to sequences that make no pretense of being "random" in any meaningful sense of the term, but which can be readily generated in the computer by simple arithmetic algorithms, while still passing an assortment of statistical tests for randomness. The terms of such sequences are collectively (and loosely) called *pseudo-random numbers* (abbreviated PRN). It should be emphasized right away that no such sequence can perform well under *all* imaginable tests for randomness. Rather, the user of PRN has to be aware of the specific statistical properties that are desirable in his Monte Carlo calculation and choose PRN that are known to pass these tests. This "relativity principle" will become amply clear given the results of §11. On the whole, PRN have a record of meeting any reasonably limited set of statistical requirements if adequately chosen for the particular purpose.

We will only discuss PRN simulating the uniform distribution on I. A variety of methods has been devised for transforming such PRN into others obeying a nonuniform distribution law, e.g., a normal distribution. We refer to [**2**], [**28**, Chapter 7], [**66**, Chapter 2], [**80**], [**142**], [**154**, §3.4], [**193**, Chapter 3], [**213**]. The bibliographies on pseudo-random number generation mentioned in §1 also cover these techniques.

As we pointed out, it should be a cardinal virtue of a sequence of PRN that it gets a stamp of approval from a collection of statistical tests for randomness. We review now some of the tests that have been considered frequently in the literature.[29] We assume throughout that $x_0, x_1, \dots$ is a (finite or infinite) sequence of numbers in I to be tested for "randomness".

A. *Equidistribution test.* Here we check the evenness of distribution in I of an initial segment $x_0, x_1, \dots, x_{N-1}$ of the sequence. This is done by calculating (or estimating) the discrepancy D_N (see §2) of these numbers. The test is passed if D_N is small. In an alternative method, sometimes called the *frequency test*, one partitions I into subintervals, counts the number of terms falling into the various subintervals, and performs a chi-square test on these data.

[29] This list is by no means exhaustive. See [**154**, §3.3] for more information.

B. *Gap test.* Let J be a fixed proper subinterval of I. If for some $n \geqslant 0$ we have $x_{n+j} \notin J$ for $0 \leqslant j \leqslant k-1$, but $x_{n+k} \in J$, we speak of a gap of length k. We choose a positive integer h and count the number of gaps of lengths 0, $1, \ldots, h-1$ and $\geqslant h$ until a large total number of gaps is reached. Then we apply a chi-square test using the probabilities $p_i = p(1-p)^i$ for $0 \leqslant i \leqslant h-1$ and $p_h = (1-p)^h$ for these respective categories, where p is the length of J.

C. *Run test.* A segment of the sequence satisfying $x_n < x_{n+1} < \cdots < x_{n+k-1}$, but $x_{n-1} \geqslant x_n$ and $x_{n+k-1} \geqslant x_{n+k}$, is called a "run up" of length k. We count the number of "runs up" of lengths $1, 2, \ldots, h$ and $\geqslant h+1$ in a given initial segment of the sequence. Since adjacent runs are not independent, we cannot use a straightforward chi-square test for these data. A more complicated statistic has to be computed (see [154, pp. 60–63]). A similar test can be performed for "runs down".

D. *Permutation test.* We choose an integer $s \geqslant 2$ and consider the s-tuples $(x_n, x_{n+1}, \ldots, x_{n+s-1})$, $0 \leqslant n < N$. There are $s!$ possible relative orderings among the entries of such an s-tuple. We count the number of times each ordering appears and perform a chi-square test, using the probability $1/s!$ for each ordering, or determine the maximal deviation from the expected number.

E. *Serial correlation.* This is a rather weak test for the interdependence between x_n and x_{n+1}. We calculate the serial correlation coefficient

$$\sigma_1 = \frac{N\sum_{n=0}^{N-1} x_n x_{n+1} - \left(\sum_{n=0}^{N-1} x_n\right)^2}{N\sum_{n=0}^{N-1} x_n^2 - \left(\sum_{n=0}^{N-1} x_n\right)^2}.$$

If x_n and x_{n+1} are almost independent, then $|\sigma_1|$ is very small. It is a deficiency of this test that the converse does not necessarily hold. One may, more generally, consider the serial correlation coefficient σ_h reflecting the interdependence between x_n and x_{n+h}.

F. *Spectral test.* This test was introduced in [48] (see also [47]) and depends on a Fourier transform technique. We assume that all PRN in the given sequence are rationals with common denominator m, as will be the case for the important method to be introduced in §7. For any s-dimensional lattice point $(h_1, \ldots, h_s)$ we define the limit

$$L(h_1, \ldots, h_s) = \lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} e(h_1 x_n + h_2 x_{n+1} + \cdots + h_s x_{n+s-1}),$$

which will exist, for instance, in the special case mentioned above. In an ideal case of randomness we should have $L(h_1, \ldots, h_s) = 1$ if $h_1 \equiv \cdots \equiv h_s \equiv 0 \pmod m$ and $L(h_1, \ldots, h_s) = 0$ otherwise. The deviation from this behavior is a measure for the nonrandomness of the given sequence of PRN. The difficulty here is to find a realistic way of putting this idea into a quantitative form. In [48] an intuitive analogy with the theory of harmonic oscillation was used (see also [154, §3.3.4]), although there is no mathematical basis for this procedure.

G. *Serial test.* This is the most reliable test for statistical independence of successive terms. For a fixed dimension $s \geqslant 2$, we consider the points[30] $\mathbf{x}_n = (x_n, x_{n+1}, \ldots, x_{n+s-1})$, $n = 0, 1, \ldots, N-1$, in I^s. For a truly random sequence these points will, in the long run, be very evenly distributed over I^s. Therefore, the quality of the PRN can be measured by the discrepancy D_N (see §2) of the points $\mathbf{x}_0, \ldots, \mathbf{x}_{N-1}$. This test has an added significance since information about D_N yields effective error bounds for quasi-Monte Carlo integrations using the nodes $\mathbf{x}_0, \ldots, \mathbf{x}_{N-1}$ (compare with §2). An older and less powerful version of this test operates like the frequency test mentioned under A, by starting from a partition of I^s into subintervals and proceeding as in A.

For the special class of PRN that will be introduced in the next section, there is also the possibility of studying the structure of the lattice generated by the points occurring in test G (see §10 for details).

7. Linear congruential pseudo-random numbers. There is conceivably an overwhelming multitude of techniques for generating PRN, some discovered and many still dormant, but a simple algorithm proposed by Lehmer [**176**] has easily become the most popular one and is now a traditional tool of the numerical analyst. This method can be described as follows. Let $m \geqslant 2$ and r be integers, let y_0 be an integer in the least residue system modulo m, i.e., $0 \leqslant y_0 < m$, and let λ be a positive integer relatively prime to m which may as well be taken from the least residue system modulo m. Then a sequence $y_0, y_1, \ldots$ of integers in the least residue system modulo m is generated by the recursion

$$y_{n+1} \equiv \lambda y_n + r \pmod{m} \quad \text{for } n = 0, 1, \ldots. \tag{7.1}$$

From this sequence we derive a sequence $x_0, x_1, \ldots$ of numbers in I by setting $x_n = y_n/m$ for $n = 0, 1, \ldots$, and this is already the desired sequence of so-called *linear congruential PRN* (abbreviated LCPRN). In this context, m is referred to as the *modulus*, λ as the *multiplier*, and r as the *increment*. In practice, m is taken to be a large prime or a large power of 2, the latter choice being particularly suitable for a binary computer. To rule out trivial cases, we suppose that $\lambda \not\equiv 1 \pmod{m}$ and $(\lambda - 1)y_0 + r \not\equiv 0 \pmod{m}$.[31] An explicit formula for x_n can easily be established by induction, namely

$$x_n = \left\{ \lambda^n x_0 + \frac{\lambda^n - 1}{\lambda - 1} \cdot \frac{r}{m} \right\} \quad \text{for } n = 0, 1, \ldots. \tag{7.2}$$

It is customary to distinguish the *homogeneous case* $r \equiv 0 \pmod{m}$ (also called the *multiplicative congruential method*) and the *inhomogeneous case* $r \not\equiv 0 \pmod{m}$ (also called the *mixed congruential method*). In the homogeneous case we can assume w.l.o.g. that y_0 is relatively prime to m, for otherwise the same sequence of LCPRN could be produced with a smaller modulus.

The parameters m, λ, r, and y_0 have to be chosen in such a way that the resulting PRN will pass appropriate statistical tests for randomness. The

[30] In [**154**, pp. 55–56] only nonoverlapping s-tuples are considered, but this restriction is not necessary.

[31] If $(\lambda - 1)y_0 + r \equiv 0 \pmod{m}$, then $y_1 \equiv y_0 \pmod{m}$, and the sequences $y_0, y_1, \ldots$ and $x_0, x_1, \ldots$ are constant.

modulus m is selected in accordance with machine capabilities, typical choices being $m = 2^{32}$ or $m = 2^{35}$. The initial value y_0 has practically no, and the increment r only little influence on the behavior of the PRN, so that the properties of the sequence are mainly governed by the choice of the multiplier λ. In the light of the simplicity of the generation method, it is always a surprise to find out how remarkably well these LCPRN perform under rather stringent tests once a multiplier has been selected intelligently.

We review now briefly the elementary properties of LCPRN. See [**135**], [**142**], [**154**, §3.2] for a more detailed analysis. Most importantly, the sequence $x_0, x_1, \ldots$ is always purely periodic. We shall use τ to denote its least period. In the homogeneous case, it is easily seen that τ is equal to the least positive integer n for which $\lambda^n \equiv 1 \pmod{m}$, i.e., τ is the exponent to which λ belongs modulo m. In the inhomogeneous case and for a prime or prime power modulus, τ can be determined quickly on the basis of (7.2). We first introduce some notation that will also be convenient later on. For $m = p^\alpha$, p prime, $\alpha \geqslant 1$, let κ be the largest integer such that p^κ divides $\lambda - 1$ and let ω be the largest integer such that p^ω divides $(\lambda - 1)y_0 + r$. Then the following result is contained in [**226**, Lemma 8] if we note that under our conditions we always have $\omega < \alpha$.

7.1. LEMMA. *Let $m = p^\alpha$, p prime, $\alpha \geqslant 1$. Then the period τ of a sequence of LCPRN with modulus m and multiplier λ is equal to the exponent to which λ belongs modulo $p^{\alpha-\omega+\kappa}$.*

If m has at least two distinct prime factors, the period can still be obtained from this result because of [**154**, p. 16, Lemma Q]. For the homogeneous case and $m = p^\alpha$, with $p = 2$, $\alpha \leqslant 2$, or p an odd prime, the largest possible period is $\tau = p^{\alpha-1}(p - 1)$, and it is attained precisely for multipliers λ that are primitive roots modulo m, whereas for $m = 2^\alpha$, $\alpha \geqslant 3$, the largest possible period is $\tau = 2^{\alpha-2}$, attained precisely for $\lambda \equiv \pm 3 \pmod{8}$ if $\alpha \geqslant 4$ and $\lambda \equiv 3, 5, 7 \pmod{8}$ for $\alpha = 3$. For general m, the maximum period is obtained from the quoted result in [**154**]. In the inhomogeneous case, we can actually have $\tau = m$, i.e., all fractions in $[0, 1)$ with denominator m are generated, and this happens precisely if the following requirements are met: (i) r is relatively prime to m; (ii) $\lambda \equiv 1 \pmod{p}$ for every prime p dividing m; (iii) $\lambda \equiv 1 \pmod{4}$ if 4 divides m. This criterion can be deduced from Lemma 7.1, but it follows also from other elementary considerations (cf. [**135**], [**154**, p. 15]).

We have dwelt on the question of the period since a good sequence of LCPRN should at least have a long period. One reason is, of course, that periodicity is not a typical feature of a truly random sequence. Therefore, a sequence with a short period can hardly pretend to be pseudo-random. Moreover, in an actual quasi-Monte Carlo calculation based on LCPRN we want to use a large number N of terms, but we must have $N \leqslant \tau$ so that the grossly nonrandom periodicity property does not come into play; hence τ should be large. Further indications why a large value of τ is preferable will be given later on (see, e.g., §9).

In interesting cases we may describe all fractions that are generated by the method. This is, of course, trivial whenever the period is m. If λ is a primitive root modulo m and we are in the homogeneous case, then we produce exactly all reduced fractions in $[0, 1)$ with denominator m. Again in the homogeneous case and for $m = 2^\alpha$ with $\alpha \geqslant 3$ and $\lambda \equiv 5 \pmod 8$, we obtain exactly all fractions in $[0, 1)$ of the form a/m with $a \equiv y_0 \pmod 4$; if $\lambda \equiv 3 \pmod 8$, we generate exactly all fractions in $[0, 1)$ of the form a/m with $a \equiv y_0$ or $3y_0 \pmod 8$. A thorough study is carried out in **[188]**, where it is observed, for instance, that a sequence of LCPRN is always made up of a block of less than m terms, followed by translates of that block. The frequent phenomenon of one sequence of LCPRN being a cyclic permutation of another is investigated in **[31]**.

Historically, Lehmer's method was not the first proposal for pseudo-random number generation. A few years earlier, J. von Neumann already used his "middle-square method" to produce pseudo-random numbers for Monte Carlo calculations (cf. **[154**, §3.1**]**, **[352]**). However, this technique has proved unsatisfactory since it tends to lead to short cycles and the expected distribution is nonuniform (cf. **[336]**). In more promising schemes that may be called *hybrid methods*, one starts from a sequence of LCPRN and adds a new twist or combines it with other sequences of PRN. An intriguing idea is the "shuffling" scheme of MacLaren and Marsaglia **[184]** (see also **[189]**). Here one takes a sequence of LCPRN and shuffles it in a way which is directed by a different sequence of PRN. Thus we get a sequence of PRN with a "pseudo-random shuffle" thrown in for good measure. Theoretical evidence that shuffling may improve the performance of LCPRN is presented in **[260]**. Other hybrid methods can be found, for instance, in **[32]**, **[82]**, **[88]**, **[202]**, **[273]**, **[290]**, **[297]**, **[359]**.

Another idea that has been brought up in the literature is to replace (7.1) by a higher-order recurrence relation. Here it is convenient to use a prime number p as a modulus. Then a congruence such as (7.1) can be viewed as a recurrence relation in the finite field $F_p = \mathbf{Z}/p\mathbf{Z}$. Let k be a positive integer which will serve as the order of the new recurrence relation. We note that the finite field F_{p^k} of p^k elements is an extension of F_p and that its multiplicative group $F_{p^k}^*$ is cyclic. A polynomial $f(x) = x^k - a_{k-1}x^{k-1} - \cdots - a_0 \in \mathbf{Z}[x]$ is called a *primitive polynomial modulo p* if the polynomial $\bar{f}(x) \in F_p[x]$ canonically associated with $f(x)$ is the minimal polynomial over F_p of a generator of $F_{p^k}^*$. With such a primitive polynomial modulo p, we set up the kth order homogeneous linear recurrence

$$y_{n+k} \equiv a_{k-1}y_{n+k-1} + \cdots + a_0 y_n \pmod p \quad \text{for } n = 0, 1, \ldots. \qquad (7.3)$$

Any sequence $y_0, y_1, \ldots$ of integers in the least residue system modulo p satisfying (7.3) with $(y_0, \ldots, y_{k-1}) \neq (0, \ldots, 0)$ is called a *maximal period sequence modulo p*. The reason for this terminology lies in the fact that the length of the period of a maximal period sequence modulo p is equal to $p^k - 1$, the largest possible period of any kth order homogeneous linear recurring sequence in F_p. A maximal period sequence modulo p is easily seen to be purely periodic (see **[379]** for these facts about linear recurring sequences).

From a maximal period sequence $y_0, y_1, \ldots$ modulo p we derive a sequence $x_0, x_1, \ldots$ of PRN in I by setting $x_n = y_n/p$ for $n = 0, 1, \ldots$. In practice, the prime p will be chosen very large, but $p = 2$ is sometimes used to produce a pseudo-random sequence of binary digits. The advantages offered by such sequences of PRN can be seen as follows. Consider the k-tuples $(y_n, y_{n+1}, \ldots, y_{n+k-1})$, $n = 0, 1, \ldots, p^k - 2$. They are all distinct (for otherwise the period of $y_0, y_1, \ldots$ would be less than $p^k - 1$) and $(0, \ldots, 0)$ cannot occur among them, hence they run exactly through all k-tuples $\neq (0, \ldots, 0)$ of elements in the least residue system modulo p. Therefore, the points $(x_n, x_{n+1}, \ldots, x_{n+k-1})$, $n = 0, 1, \ldots, p^k - 2$, show an excellent distribution in $(p^{-1}\mathbf{Z}/\mathbf{Z})^k$. Also, in a full period of $y_0, y_1, \ldots$ each nonzero residue modulo p occurs exactly p^{k-1} times and 0 occurs exactly $p^{k-1} - 1$ times; each pair $\neq (0, 0)$ of residues modulo p occurs exactly p^{k-2} times among (y_n, y_{n+1}), $n = 0, 1, \ldots, p^k - 2$, whereas $(0, 0)$ occurs exactly $p^{k-2} - 1$ times; and so on for s-tuples with $s < k$. Altogether, we can say that the sequence $x_0, x_1, \ldots$ of PRN leads to practically perfect distribution properties in $(p^{-1}\mathbf{Z}/\mathbf{Z})^s$ for all dimensions $1 < s < k$. Of course, we have to pay a price for this benefit, in the sense that the generation method (7.3) is more complicated than (7.1). We will report on further results about these generators in §§9, 10.

The idea of using higher-order recurrences for pseudo-random number generation can be traced back to [62], [81], [343]. The method (7.3) was proposed by Tausworthe [333] and endorsed by Knuth [154, §3.2.2]. We shall refer to the associated sequences of PRN as *Tausworthe generators*.[32] So-called "additive generators", for which the coefficients in the recurrence are either 0 or 1, emerged already in the early times of pseudo-random number generation. A standard example is the Fibonacci generator $y_{n+2} \equiv y_{n+1} + y_n \pmod{m}$ (cf. [332]), which was, however, soon discarded because it displays a notoriously bad behavior. The recursion $y_{n+k} \equiv y_{n+k-1} + y_n \pmod{m}$ is a decent generator for $k = 16$ according to [85].

After Tausworthe's paper, higher-order generators received more attention. Apart from the investigations on Tausworthe generators to be mentioned later, there are studies of higher-order recurrences modulo powers of 2 (cf. [245]) and of certain second-order recurrences modulo primes (cf. [58]). A special method of generating sequences of n-bit strings of binary digits by a second-order recurrence involving bit-by-bit addition modulo 2 and cyclic shifts was proposed in [250], and algebraic results about the periods of these sequences were shown in [277].

8. Exponential sums. The main tool in our treatment of the equidistribution test and serial test for LCPRN will be certain exponential sums[33] that are intimately connected with these numbers. That such sums should play a rôle is quite clear in the light of Lemmas 3.3 and 3.9. We consider the general case of a kth order linear recurring sequence $z_0, z_1, \ldots$ of integers satisfying the recurrence relation

$$z_{n+k} = a_{k-1}z_{n+k-1} + \cdots + a_0 z_n + r \quad \text{for } n = 0, 1, \ldots, \qquad (8.1)$$

[32] For a convenient computer implementation of these generators, see [238].

[33] In different contexts, special cases appeared in [171], [172], [204], [205], [328]. See [228] for an application to coding theory.

where $a_{k-1}, \ldots, a_0, r$ are fixed integers. If such a sequence is viewed modulo an integer $m \geqslant 2$, it looks like a sequence $y_0, y_1, \ldots$ used in the generation of PRN (see §7). In this case, we will assume, for simplicity, that a_0 and m are relatively prime.[34] This guarantees that $z_0, z_1, \ldots$ is purely periodic modulo m. Let $\tau = \tau(m)$ be the length of its least period modulo m, and let $\mu = \mu(m)$ be the length of the least period modulo m of the sequence $v_0, v_1, \ldots$ of integers satisfying

$$v_{n+k} = a_{k-1}v_{n+k-1} + \cdots + a_0 v_n \quad \text{for } n = 0, 1, \ldots,$$

with the initial values $v_0 = \cdots = v_{k-2} = 0$, $v_{k-1} = 1$ ($v_0 = 1$ if $k = 1$). The following is a special case of a more general result. As before, we write $e(t) = e^{2\pi i t}$ for $t \in \mathbf{R}$.

8.1. Theorem (Niederreiter [224], [225]). *Let $m \geqslant 2$ and b be relatively prime integers, let $z_0, z_1, \ldots$ be the kth order linear recurring sequence of integers satisfying* (8.1) *with $\gcd(a_0, m) = 1$, and let τ and μ be defined as above. Then,*

$$\left|\sum_{n=0}^{\tau-1} e\left(\frac{b}{m} z_n\right)\right| \leqslant \left(\frac{m^k\tau - \tau^2}{\mu}\right)^{1/2} \tag{8.2}$$

and

$$\left|\sum_{n=0}^{N-1} e\left(\frac{b}{m} z_n\right)\right| < \left(\frac{m^k\tau}{\mu}\right)^{1/2}\left(\frac{2}{\pi}\log\tau + \frac{2}{5}\right) + \frac{N}{\tau}\left(\frac{m^k\tau - \tau^2}{\mu}\right)^{1/2} \quad \textit{for } 1 \leqslant N \leqslant \tau. \tag{8.3}$$

The upper bounds in (8.2) and (8.3) represent slight improvements on the original inequalities, which are achieved by subtracting the contribution from the zero vector in [224, Equation (16)] (compare with the proof of Theorem 8.2). These improvements are of interest when τ is large.

If a sequence of LCPRN is generated by the multiplicative congruential method, we are led to consider the recurrence relation $z_{n+1} = \lambda z_n$, $n = 0, 1, \ldots$, which produces powers of λ, up to a constant factor. We present a proof for this case to indicate the ideas involved. Note that under the conditions of the subsequent result we have $\tau = \mu$.

8.2. Theorem (Korobov [172], Niederreiter [226]). *Let $m \geqslant 2$, b, λ, and c be integers with $\gcd(b, m) = \gcd(\lambda, m) = 1$ and λ belonging to the exponent τ modulo m. Then,*

$$\left|\sum_{n=0}^{\tau-1} e\left(\frac{b}{m}\lambda^n\right) e\left(\frac{cn}{\tau}\right)\right| \leqslant \begin{cases} (m-\tau)^{1/2} & \textit{if } c \equiv 0 \pmod{\tau}, \\ m^{1/2} & \textit{if } c \not\equiv 0 \pmod{\tau}, \end{cases} \tag{8.4}$$

and

$$\left|\sum_{n=0}^{N-1} e\left(\frac{b}{m}\lambda^n\right)\right| < m^{1/2}\left(\frac{2}{\pi}\log\tau + \frac{2}{5}\right) + \frac{N}{\tau}(m-\tau)^{1/2} \quad \textit{for } 1 \leqslant N \leqslant \tau. \tag{8.5}$$

[34] This is always satisfied in the applications to PRN.

PROOF. For an integer a, write

$$\sigma(a, c) = \sum_{n=0}^{\tau-1} e\left(\frac{a}{m}\lambda^n\right)e\left(\frac{cn}{\tau}\right).$$

The general term of this sum, considered as a function of n, is periodic with period τ. Therefore, for any integer $j \geqslant 0$ we have

$$\sigma(a, c) = \sum_{n=0}^{\tau-1} e\left(\frac{a}{m}\lambda^{n+j}\right)e\left(\frac{c(n+j)}{\tau}\right),$$

and so we obtain the transformation formula

$$|\sigma(a, c)| = \left|\sum_{n=0}^{\tau-1} e\left(\frac{a\lambda^j}{m}\lambda^n\right)e\left(\frac{cn}{\tau}\right)\right| = |\sigma(a\lambda^j, c)|. \tag{8.6}$$

Since the integers $b, b\lambda, \ldots, b\lambda^{\tau-1}$ are pairwise incongruent modulo m and not divisible by m, (8.6) implies that

$$\begin{aligned}\tau|\sigma(b, c)|^2 &= \sum_{j=0}^{\tau-1} |\sigma(b\lambda^j, c)|^2 \leqslant \sum_{a=1}^{m-1} |\sigma(a, c)|^2 \\ &= \sum_{a=0}^{m-1} |\sigma(a, c)|^2 - |\sigma(0, c)|^2 \\ &= \sum_{h,j=0}^{\tau-1} e\left(\frac{c(h-j)}{\tau}\right)\sum_{a=0}^{m-1} e\left(\frac{a}{m}(\lambda^h - \lambda^j)\right) - |\sigma(0, c)|^2 \\ &= m\tau - |\sigma(0, c)|^2.\end{aligned}$$

The inequalities in (8.4) are immediate consequences. To prove (8.5), we use the identity

$$\sum_{n=0}^{N-1} e\left(\frac{b}{m}\lambda^n\right) = \sum_{n=0}^{\tau-1} e\left(\frac{b}{m}\lambda^n\right)\sum_{y=0}^{N-1}\frac{1}{\tau}\sum_{c=1}^{\tau} e\left(\frac{c(n-y)}{\tau}\right) \qquad \text{for } 1 \leqslant N \leqslant \tau,$$

which follows from the fact that the sum over y is equal to 1 for $0 \leqslant n \leqslant N - 1$ and equal to 0 for $N \leqslant n \leqslant \tau - 1$. We rewrite this identity in the form

$$\sum_{n=0}^{N-1} e\left(\frac{b}{m}\lambda^n\right) = \frac{1}{\tau}\sum_{c=1}^{\tau}\left(\sum_{y=0}^{N-1} e\left(-\frac{cy}{\tau}\right)\right)\left(\sum_{n=0}^{\tau-1} e\left(\frac{b}{m}\lambda^n\right)e\left(\frac{cn}{\tau}\right)\right).$$

Then by (8.4),

$$\begin{aligned}\left|\sum_{n=0}^{N-1} e\left(\frac{b}{m}\lambda^n\right)\right| &\leqslant \frac{1}{\tau}\sum_{c=1}^{\tau}\left|\sum_{y=0}^{N-1} e\left(\frac{cy}{\tau}\right)\right|\left|\sum_{n=0}^{\tau-1} e\left(\frac{b}{m}\lambda^n\right)e\left(\frac{cn}{\tau}\right)\right| \\ &\leqslant \frac{1}{\tau}m^{1/2}\sum_{c=1}^{\tau-1}\left|\sum_{y=0}^{N-1} e\left(\frac{cy}{\tau}\right)\right| + \frac{N}{\tau}(m-\tau)^{1/2},\end{aligned}$$

and (8.5) is obtained by a straightforward estimation of the remaining sum (compare with [**226**, Lemma 2]).

The sums in (8.4) contain Gaussian sums to a prime modulus or to an odd prime power modulus as special cases. The second part of (8.4) reduces then to a familiar inequality for Gaussian sums, and for certain values of c we will even have equality (cf. [225, p. 60]). On the other hand, Gaussian sums can be used to prove Theorem 8.2, but yield nothing for the inhomogeneous case or for higher-order recurrences. As an illustration, we prove the following result on the basis of ideas in [224]. We use ϕ to denote Euler's totient function.

8.3. Theorem. *Let m, b, and λ be as in Theorem 8.2. Then,*

$$\left|\sum_{n=0}^{\tau-1} e\left(\frac{b}{m}\lambda^n\right)\right| \leqslant \sqrt{m} - \frac{\tau}{\phi(m)}(\sqrt{m}-1) \tag{8.7}$$

and

$$\left|\sum_{n=0}^{N-1} e\left(\frac{b}{m}\lambda^n\right)\right| < \sqrt{m}\left(\frac{2}{\pi}\log\tau + \frac{2}{5}\right) + N\left(\frac{\sqrt{m}}{\tau} - \frac{\sqrt{m}-1}{\phi(m)}\right)$$

$$\textit{for } 1 \leqslant N \leqslant \tau. \tag{8.8}$$

Proof. We observe that $e(bh/m)$, considered as a function on the integers h with $\gcd(h, m) = 1$, can be expanded into a finite Fourier series with respect to the Dirichlet characters ψ modulo m. Thus,

$$e\left(\frac{bh}{m}\right) = \sum_{\psi} c(\psi)\psi(h) \quad \text{for } \gcd(h, m) = 1 \tag{8.9}$$

with Fourier coefficients

$$c(\psi) = \frac{1}{\phi(m)} \sum_{j=1}^{m}{}' e\left(\frac{bj}{m}\right)\bar{\psi}(j),$$

where the prime signalizes that we only consider those j coprime to m. We note that

$$c(\psi) = \frac{\psi(b)}{\phi(m)} \sum_{j=1}^{m}{}' e\left(\frac{bj}{m}\right)\bar{\psi}(bj) = \frac{\psi(b)G(\bar{\psi})}{\phi(m)}, \tag{8.10}$$

where

$$G(\bar{\psi}) = \sum_{j=1}^{m}{}' e\left(\frac{j}{m}\right)\bar{\psi}(j)$$

is a Gaussian sum. From (8.9) we obtain

$$\sum_{n=0}^{N-1} e\left(\frac{b}{m}\lambda^n\right) = \sum_{\psi} c(\psi) \sum_{n=0}^{N-1} \psi(\lambda^n)$$

$$= N \sum_{\substack{\psi \\ \psi(\lambda)=1}} c(\psi) + \sum_{\substack{\psi \\ \psi(\lambda)\neq 1}} c(\psi)\frac{1-\psi(\lambda^N)}{1-\psi(\lambda)}. \tag{8.11}$$

Setting $N = \tau$ in (8.11), we get

$$\sum_{n=0}^{\tau-1} e\left(\frac{b}{m}\lambda^n\right) = \tau \sum_{\substack{\psi \\ \psi(\lambda)=1}} c(\psi), \tag{8.12}$$

and then (8.10) implies

$$\left|\sum_{n=0}^{\tau-1} e\left(\frac{b}{m}\lambda^n\right)\right| \leq \frac{\tau}{\phi(m)} \sum_{\substack{\psi \\ \psi(\lambda)=1}} |G(\bar{\psi})|.$$

The last sum contains $\phi(m)/\tau$ terms, and since $|G(\bar{\psi})| \leq 1$ for ψ trivial and $|G(\bar{\psi})| \leq \sqrt{m}$ for ψ nontrivial, we obtain

$$\left|\sum_{n=0}^{\tau-1} e\left(\frac{b}{m}\lambda^n\right)\right|$$
$$\leq \frac{\tau}{\phi(m)} + \frac{\tau}{\phi(m)}\left(\frac{\phi(m)}{\tau} - 1\right)\sqrt{m} = \sqrt{m} - \frac{\tau}{\phi(m)}\left(\sqrt{m} - 1\right),$$

and (8.7) is shown. For $1 \leq N \leq \tau$, we use (8.7), (8.10), (8.11), and (8.12) to get

$$\left|\sum_{n=0}^{N-1} e\left(\frac{b}{m}\lambda^n\right)\right| \leq \frac{N}{\tau}\left|\sum_{n=0}^{\tau-1} e\left(\frac{b}{m}\lambda^n\right)\right| + 2 \sum_{\substack{\psi \\ \psi(\lambda)\neq 1}} \frac{|c(\psi)|}{|1 - \psi(\lambda)|}$$
$$\leq N\left(\frac{\sqrt{m}}{\tau} - \frac{\sqrt{m} - 1}{\phi(m)}\right) + \frac{2}{\phi(m)} \sum_{\substack{\psi \\ \psi(\lambda)\neq 1}} \frac{|G(\bar{\psi})|}{|1 - \psi(\lambda)|}.$$

Now $|G(\bar{\psi})| \leq \sqrt{m}$ and each τth root of unity $e(j/\tau)$ occurs exactly $\phi(m)/\tau$ times as a value of $\psi(\lambda)$, so that

$$\left|\sum_{n=0}^{N-1} e\left(\frac{b}{m}\lambda^n\right)\right|$$
$$\leq N\left(\frac{\sqrt{m}}{\tau} - \frac{\sqrt{m} - 1}{\phi(m)}\right) + \frac{2\sqrt{m}}{\phi(m)} \cdot \frac{\phi(m)}{\tau} \sum_{j=1}^{\tau-1} \left|1 - e\left(\frac{j}{\tau}\right)\right|^{-1}$$
$$= N\left(\frac{\sqrt{m}}{\tau} - \frac{\sqrt{m} - 1}{\phi(m)}\right) + \frac{\sqrt{m}}{\tau} \sum_{j=1}^{\tau-1} \csc\frac{\pi j}{\tau},$$

and (8.8) follows from the argument in [**226**, p. 574].

In important special cases, the exponential sums in question can actually vanish. A general result in this direction is the following.

8.4. THEOREM (NIEDERREITER [**226**]). *Let $m = p^\alpha$, p prime, $\alpha \geq 1$, let $\gcd(b, m) = \gcd(\lambda, m) = 1$, and let $z_0, z_1, \ldots$ be a sequence of integers with $z_{n+1} = \lambda z_n + r$ $(n = 0, 1, \ldots)$ such that the period $\tau = \tau(m)$ modulo m satisfies $\tau = p\tau(p^{\alpha-1})$. Then*

$$\sum_{n=0}^{\tau-1} e\left(\frac{b}{m} z_n\right) e\left(\frac{cn}{\tau}\right) = 0 \quad \textit{for all integers } c \equiv 0 \pmod{p}.$$

Using this fact, one can then establish the following improvement on (8.3) and (8.5) in the case under consideration.

8.5. THEOREM (NIEDERREITER [226]). *Suppose the conditions of Theorem* 8.4 *are satisfied, and let* λ *belong to the exponent* μ *modulo* m. *Then,*

$$\left|\sum_{n=0}^{N-1} e\left(\frac{b}{m} z_n\right)\right| < \left(\frac{m\tau}{\mu}\right)^{1/2}\left(\frac{2(p-1)}{\pi p} \log \tau + \frac{3}{4}\right) \quad \textit{for } 1 \leqslant N \leqslant \tau.$$

The estimates in this section are best possible, apart from the logarithmic factors that occur in those cases where a sum over a part of the period is considered. This applies even to the most general estimate (8.3), since it is shown in [224] that there are instances in which this particular sum is at least of the order of magnitude $m^{k/2}$, while $\tau = \mu$.

9. Equidistribution test. To check the performance of a sequence of LCPRN, we subject it first to test A from §6, the equidistribution test. We will also investigate the behavior of PRN generated by higher-order recurrences with respect to this test. We recall that this test amounts to calculating (or estimating) the discrepancy D_N of an initial segment $x_0, x_1, \ldots, x_{N-1}$ of the given sequence of PRN.

If $x_0, x_1, \ldots, x_{\tau-1}$ is the full period of a sequence of LCPRN, then the discrepancy D_τ (or D_τ^*) can actually be evaluated in the most interesting cases. For instance, if we have $\tau = m$, as happens for $m = 2^\alpha$, $\alpha \geqslant 2$, $\lambda \equiv 1 \pmod 4$, and r odd (see §7), it is clear that $D_\tau = D_\tau^* = 1/m$. For the homogeneous case with $m = p^\alpha$, p an odd prime, $\alpha \geqslant 1$, λ a primitive root modulo m, we get $D_\tau^* = 1/m$ (cf. [196]) and $D_\tau = 2/m$. Again in the homogeneous case and for $m = 2^\alpha$, $\alpha \geqslant 3$, we have $D_\tau^* = 3/m$ and $D_\tau = 4/m$ if $\lambda \equiv 5 \pmod 8$ and $D_\tau^* = 5/m$ and $D_\tau = 6/m$ if $\lambda \equiv 3 \pmod 8$. The first results of this type (cf. [140]) were very weak, with an order of magnitude for D_τ even worse than $m^{-1/2}$. The fact that the correct order of magnitude is m^{-1} was pointed out in [218]. See also [57], [58] for special cases.

As was already mentioned in §7, in an actual calculation based on LCPRN we will only use an initial segment of the period since any influence of the periodicity property could prove ruinous. Therefore, the meaningful part of the equidistribution test is the estimation of D_N for $N < \tau$. This was first achieved by Niederreiter [221]. Subsequent improvements, simplifications, and generalizations are contained in [226]. All the theorems in this section stem from these two papers. We occasionally incorporate some small improvements.

In the case of a prime modulus, we can use Lemma 3.3 and the estimates in §8 directly, and we note that $\tau = \mu$ in the inhomogeneous case because of Lemma 7.1.

9.1. THEOREM. *For a sequence of LCPRN with a prime modulus m we have*

$$D_N < \frac{\sqrt{m}}{N}\left(\frac{2}{\pi}\log\tau + \frac{2}{5}\right)\left(\frac{2}{\pi}\log m + \frac{2}{5}\right)$$
$$+\left(\frac{\sqrt{m}}{\tau} - \frac{1}{\sqrt{m}+1}\right)\left(\frac{2}{\pi}\log m + \frac{2}{5}\right) + \frac{1}{m}$$

for $1 \leqslant N \leqslant \tau$ *in the homogeneous case and*

$$D_N < \frac{\sqrt{m}}{N}\left(\frac{2}{\pi}\log\tau + \frac{2}{5}\right)\left(\frac{2}{\pi}\log m + \frac{2}{5}\right)$$
$$+\frac{\sqrt{m}-\tau}{\tau}\left(\frac{2}{\pi}\log m + \frac{2}{5}\right) + \frac{1}{m}$$

for $1 \leqslant N \leqslant \tau$ *in the inhomogeneous case.*

An estimate based on Corollary 3.2 can also be established. The resulting bounds (cf. [**226**, Theorems 1 and 4]) are more complicated, but often somewhat better than those in Theorem 9.1, although the order of magnitude remains the same.

For a prime power m, say $m = p^\alpha$ with p prime and $\alpha \geqslant 2$, there is a difficulty with applying Lemma 3.3 since there will now appear exponential sums for which, in the notation of §8, we have $\gcd(b, m) > 1$. In this case, it is advantageous to use Corollary 3.2 which allows us to "cut out" many of these bothersome sums. We again need the numbers κ and ω introduced prior to Lemma 7.1, as well as one more parameter. With $\mu(q)$ denoting the exponent to which the multiplier λ belongs modulo q, we define a positive integer β as follows: if p is odd, let β be the largest integer such that p^β divides $\lambda^{\mu(p)} - 1$; if $p = 2$, let β be the largest integer such that 2^β divides $\lambda^{\mu(4)} - 1$. The number β is small for the common choices of multipliers. If p is odd and λ is a primitive root modulo m, then β attains its minimal value $\beta = 1$. If $p = 2$, then β attains its minimal value $\beta = 2$ for $\lambda \equiv 5 \pmod 8$, whereas $\beta = 3$ for $\lambda \equiv 3 \pmod 8$. We set $\gamma = \beta + \omega - \kappa$ and note that we always have $\gamma \geqslant 0$. We write μ for $\mu(m)$.

9.2. THEOREM. *For a sequence of LCPRN with modulus* $m = p^\alpha$, p *prime,* $\alpha \geqslant 2$, *which satisfies* $\gamma < \alpha$ *and*

$$p^\gamma < (0.24)\left(\frac{m\tau}{\mu}\right)^{1/2}\left(\frac{2(p-1)}{\pi p}\log\tau + \frac{3}{4}\right), \tag{9.1}$$

we have

$$D_N < \frac{p^{3/2}-p^{1/2}}{p^{3/2}-1}X\log\left[1 + \frac{4(p^{3/2}-1)}{(p^{3/2}-p^{1/2})X}\right] + \left(\frac{p^{3/2}}{p^{3/2}-1} + \frac{\log p}{p}\right)X$$

for $1 \leqslant N \leqslant \tau$, *where*

$$X = \frac{4}{\pi N}\left(\frac{m\tau}{\mu}\right)^{1/2}\left(\frac{2(p-1)}{\pi p}\log\tau + \frac{3}{4}\right).$$

In practical cases, m and τ are large, so that the condition (9.1) can be satisfied by choosing the parameters in such a way that $\gamma < \alpha/2$. For the homogeneous case, the above result simplifies somewhat since $\tau = \mu$ and $\omega = \kappa$. Consequently, we have then $\gamma = \beta$, and as the above information about β shows, the condition (9.1) is easily satisfied in all reasonable circumstances. In the inhomogeneous case, the common choice $m = 2^{\alpha}$, $\alpha \geqslant 3$, $\lambda \equiv 5 \pmod 8$, and r odd leads to $\gamma = 0$, and (9.1) holds trivially.

An inspection of the bounds in Theorems 9.1 and 9.2 shows that in terms of orders of magnitude we have $ND_N = O(m^{1/2}\log^2 m)$. Since $ND_N \leqslant N$ is trivial, the estimates are only of interest for N appreciably larger than $m^{1/2}\log^2 m$, say $N \approx m^{1/2+\varepsilon}$ with $\varepsilon > 0$. But $\tau \geqslant N$, and so τ should be at least of this order of magnitude. If this is so, then the PRN have a small discrepancy and thus pass the equidistribution test.[35] The larger we can choose N (i.e., the larger τ), the better the distribution behavior will be. This provides again support for the familiar rule of thumb that large periods are preferable.

The equidistribution test is not very selective since the results only depend on the values of γ, τ, and μ. Thus, the test does not allow a distinction between the various primitive roots λ we might want to use for an odd prime or prime power modulus in the homogeneous case, or between the various multipliers $\lambda \equiv 5 \pmod 8$ in the case where m is a power of 2 and r is odd. It is guaranteed, however, that with such choices of parameters the test will be passed with flying colors. We note also that the initial value y_0 and the increment r only play an indirect rôle, insofar as they influence the values of γ and τ.

The above results can be extended to arbitrary moduli (see [**221**, §5]). Furthermore, these discrepancy estimates yield effective error bounds for quasi-Monte Carlo integrations based on the nodes $x_0, x_1, \ldots, x_{N-1}$ because of the inequalities in §2. The estimate $ND_N = O(m^{1/2}\log^2 m)$ established above is nearly best possible, as the following lower bounds indicate. We may state these results from [**226**, §5] in terms of D_N^* since [**226**, Lemma 11] holds with D_N replaced by D_N^*. Because of $D_N^* \leqslant D_N$, we automatically get lower bounds for D_N as well.

9.3. THEOREM. *For any prime modulus $m \geqslant 3$, any multiplier λ which is a primitive root modulo m, and any increment r, there exists an initial value y_0 with $(\lambda - 1)y_0 + r \not\equiv 0 \pmod m$ such that the derived sequence of LCPRN satisfies $ND_N^* > \frac{1}{8}\, m^{1/2}$ for some N with $1 \leqslant N \leqslant \tau$.*

The additional condition $(\lambda - 1)y_0 + r \not\equiv 0 \pmod m$, which does not appear in this form in the original result, can be obtained by excluding from the sum S in [**226**, p. 593] the unique residue b with $(\lambda - 1)b + r \equiv 0 \pmod m$ rather than the residue $b = 0$. A similar modification of [**226**, Theorem 10] yields the following result.

[35] Some qualitative evidence for the good performance of LCPRN under the equidistribution test was already presented by Franklin [**71**] who set up a continuous model and showed that for any integer $\lambda > 1$ and any real θ the sequence defined by $x_{n+1} = \{\lambda x_n + \theta\}$, $n = 0, 1, \ldots,$ is uniformly distributed in I for almost all initial values $x_0 \in I$ (in the sense of Lebesgue measure). See also [**248**, Chapter 3] for a detailed discussion of the case $\theta = 0$.

9.4. Theorem. *Let $m = p^{\alpha}$, p prime, $\alpha \geqslant 2$, be a modulus, let the multiplier λ belong to the largest possible exponent modulo m, and let r be an increment. If p is odd, there exists an initial value y_0 with $(\lambda - 1)y_0 + r \not\equiv 0 \pmod{p}$ such that the derived sequence of LCPRN satisfies*

$$ND_N^* \geqslant \frac{(p^2-1)^{1/2}}{8p} m^{1/2}$$

for some N with $1 \leqslant N \leqslant \tau$. If $p = 2$ and $r = 0$, there exists an odd initial value y_0 such that the derived sequence of LCPRN satisfies

$$ND_N^* \geqslant (1/8\sqrt{2})m^{1/2}$$

for some N with $1 \leqslant N \leqslant \tau$. If $p = 2$ and r odd, there exists an initial value y_0 such that the derived sequence of LCPRN satisfies $ND_N^* \geqslant \frac{1}{8}m^{1/2}$ for some N with $1 \leqslant N \leqslant \tau$.

It is remarkable how close the discrepancy for parts of the period of a sequence of LCPRN is to that of a randomly chosen sequence (compare with §3).

The techniques employed in this section can also be applied to sequences of PRN generated by higher-order recurrences. Using the notation and the result of Theorem 8.1 as well as Lemma 3.3, we obtain the following information concerning the equidistribution test.

9.5. Theorem. *For any prime p and any sequence of PRN derived from a kth order linear recurrence relation* (8.1) *with $a_0 \not\equiv 0 \pmod{p}$, we have*

$$D_N < \frac{1}{p} + \frac{1}{N}\left(\frac{p^k\tau}{\mu}\right)^{1/2}\left(\frac{2}{\pi}\log\tau + \frac{2}{5}\right)\left(\frac{2}{\pi}\log p + \frac{2}{5}\right) + \frac{1}{\tau}\left(\frac{p^k\tau - \tau^2}{\mu}\right)^{1/2}\left(\frac{2}{\pi}\log p + \frac{2}{5}\right) \quad \text{for } 1 \leqslant N \leqslant \tau. \tag{9.2}$$

9.6. Corollary. *For any kth order Tausworthe generator modulo the prime p we have*

$$D_N < \frac{1}{p} + \frac{p^{k/2}}{N}\left(\frac{2}{\pi}\log\tau + \frac{2}{5}\right)\left(\frac{2}{\pi}\log p + \frac{2}{5}\right) + \frac{1}{\tau}\left(\frac{2}{\pi}\log p + \frac{2}{5}\right) \quad \text{for } 1 \leqslant N \leqslant \tau = p^k - 1. \tag{9.3}$$

The bound (9.2) is only meaningful if N, and so the period τ, is considerably larger than $p^{k/2}$. Thus, as in the first-order case, recurrences producing long periods are preferable. If $N \geqslant p^{(k/2)+1+\varepsilon}$, $\varepsilon > 0$, then the main term in (9.3) is $1/p$. This term is needed since all PRN are rationals with denominator p, so that we must have $D_N \geqslant 1/p$. On the other hand, the upper bound (9.3) differs then by very little from this trivial lower bound implied by the discreteness of the sequence, which confirms the notion that Tausworthe generators, say of order $k \geqslant 3$, lead to an extremely even distribution. For small values of k the second term on the right-hand side of

(9.3) is required, at least up to the logarithmic factors, because of the following result: for every primitive polynomial modulo p of degree k there exists a corresponding Tausworthe generator such that $ND_N^* > \frac{1}{8}\, p^{k/2}$ for some N with $1 \leqslant N \leqslant \tau$ (see [226, p. 596]).

10. Interdependence of successive terms. The results of the preceding section have shown that a sequence of LCPRN will pass the equidistribution test as soon as its period is large. To achieve a finer distinction between the various multipliers, more powerful tests have to be applied. In particular, one has to take a closer look at order properties and relations between successive terms in the sequence.

Although numerical data about the performance of various LCPRN under stringent tests had already been collected earlier (cf. [143], [332], and the survey article [135]), the first effective theoretical results were only obtained in the 1960s when the problem of calculating serial correlation coefficients (see §6, test E) was solved satisfactorily. Coveyou [46] considered the question in the context of a continuous model, whereas Greenberger [86] provided a good estimate for serial correlation coefficients. To minimize the resulting expressions, the choice $\lambda \approx m^{1/2}$ suggested itself, leading to proposals such as $m = 2^{35}$, $\lambda = 2^{18} + 3$. A few years later, Jansson [141] gave exact formulas for serial correlation coefficients. All these results refer to the full period and to cases in which either all residues or a known set of residues modulo m are generated. See also [1], [56], [60], [142, Chapter 6], [154, §3.3.3], [327]. The values of serial correlation coefficients can be conveniently expressed in terms of generalized Dedekind sums. An allusion to these sums occurs already in [86]. Algorithms for the calculation of serial correlation coefficients based on the reciprocity law for generalized Dedekind sums are presented in [56], [60], [154, §3.3.3]. On the basis of such calculations and an analogy with diophantine approximations, it is suggested in [1] that for the homogeneous case and m a power of 2, one should select a multiplier λ with $\lambda \equiv 5 \pmod 8$ and $\lambda \approx m\xi/4$, where $\xi = (\sqrt{5} - 1)/2$ is the golden section number, in order to get small serial correlation.

Numerical data on the run test (see §6, test C) for commonly used generators were already collected in the 1950s (cf. [332]). This test has also been applied to a special class of LCPRN obtained by taking the Mersenne prime $m = 2^{31} - 1$ as the modulus in the multiplicative congruential method, with the multiplier λ being a primitive root modulo m. This proposal is due to D. H. Lehmer and has the merit of allowing a quite convenient computer implementation (cf. [91], [183]). Particularly good results were achieved with $\lambda = 7^5$ (cf. [181]), but there can be problems with multipliers that are too small (cf. [67]). For further work on these generators, see [60], [61], [136], [239], [286].

A feasible implementation of the spectral test (see §6, test F) offers some challenges, so that numerical data were only recently compiled in a systematic form (for some sample calculations, see [154, §3.3.4]). We refer to [55] for dimensions $2 \leqslant s \leqslant 4$ and [107] for dimensions $2 \leqslant s \leqslant 5$.

Because of the simple nature of the recursion (7.1), it is clear that sequences of LCPRN will be endowed with some intrinsic structure. This was made explicit by the results of Marsaglia [**186**], [**187**] on the lattice (or "crystalline") structure induced by s-tuples of successive LCPRN. For notational reasons, it is more convenient here to work with the sequence of integers $y_0, y_1, \ldots$ generated by (7.1). We choose a dimension $s \geqslant 2$ and consider the lattice points $\mathbf{y}_n = (y_n, y_{n+1}, \ldots, y_{n+s-1})$, $n = 0, 1, \ldots, \tau - 1$, where τ is, as usual, the period of the sequence. Then we subtract from each of those points the vector $\mathbf{c} = (c_0, c_1, \ldots, c_{s-1})$ whose coordinates result from (7.1) by starting with $c_0 = 0$. In this way we obtain τ lattice points $\mathbf{y}_0', \mathbf{y}_1', \ldots, \mathbf{y}_{\tau-1}'$. We now inspect the crystalline structure of these points; it may happen, for instance, that they all lie on a rather "coarse" lattice.[36] A measure for the "coarseness" of a lattice is its unit-cell volume, obtained by calculating the absolute value of a determinant whose row vectors form a basis for the lattice. Marsaglia [**187**] shows that for any choice of parameters in (7.1) the points $\mathbf{y}_0'$, $\mathbf{y}_1', \ldots, \mathbf{y}_{\tau-1}'$ all lie on a lattice with unit-cell volume m^{s-1}, which he finds unacceptably large. However, the result is stated this way for greater effect. We must not forget that the PRN themselves are produced by dividing each y_n by m. Taking into account the resulting scaling factor m^{-s}, the actual unit-cell volume becomes m^{-1}, which is about 2^{-35} for the standard generators and thus quite reasonable.

Marsaglia [**188**, p. 270] promises methods for perturbing congruential generators which would prevent the above points from lying on lattices with unit-cell volume greater than 1. This can of course be achieved in a trivial manner. For instance, in the homogeneous case we take the sequence y_0, $y_1, \ldots$ with initial value $y_0 = 1$ and throw in the numbers $y_{-s+1} = y_{-s+2} = \cdots = y_{-1} = 0$, ignoring the recursion for these indices. Then the vectors $\mathbf{y}_{-s+1}, \mathbf{y}_{-s+2}, \ldots, \mathbf{y}_0$ span a parallelepiped of volume 1 and the desired property holds. But such juggling acts are evidently not worth the effort since it is hard to see why the new sequence should be closer to randomness than the original one.

This raises some fundamental doubts about whether the "lattice test" (to use Marsaglia's term) has the proper credentials of a meaningful statistical test. As the above example indicates, it does not satisfy a mandatory requirement on such tests, namely that of stability. A small perturbation of the data (such as changing some terms slightly or adding a small number of new values) should not affect the outcome in any significant way, and this is clearly the case in all the tests mentioned in §6A-G. The "lattice test", on the other hand, reflects purely arithmetic coincidences that can be changed on a whim.

Beyer [**19**] stressed the fact that the configuration of points $\mathbf{y}_0'$, $\mathbf{y}_1', \ldots, \mathbf{y}_{\tau-1}'$, when extended with period m in each coordinate to all of $\mathbf{R}^s$, need not form a lattice by itself. This will be true, however, if $\tau = m$, which happens in the inhomogeneous case for a suitable choice of parameters (see §7). In the homogeneous case with a multiplier belonging to the largest possible exponent modulo m, the periodically extended configuration can be viewed as the union of a limited number of translated versions of a fixed

[36] A *lattice* in $\mathbf{R}^s$ is a set of lattice points obtained by forming all integral linear combinations of s linearly independent lattice points.

lattice. This has led to attempts at analyzing this lattice with the aim of finding criteria for "optimal" generators. The viewpoint commonly espoused in this connection is that a generator is the better the closer its associated lattice is to a "cubic" lattice (cf. [**20**], [**188**]). This rather ill-advised notion is based on simplistic myths about the cubic lattice. In the realm of facts, the cubic lattice is of dubious merit. Consider the $N = m^s$ points $(n_1/m, \ldots, n_s/m)$, with the n_j running independently through the integers 0, $1, \ldots, m-1$. They form a pleasing arrangement on a cubic grid. However, since no point falls into the interval $(0, 1/m) \times I^{s-1}$ of volume m^{-1}, the discrepancy D_N of these points satisfies $D_N \geqslant m^{-1} = N^{-1/s}$, a pitiable performance. Even more striking evidence for the fallacy of the "cubic-lattice criterion" will be presented in §11, where we will prove on the basis of effective theoretical results that a multiplier which Marsaglia [**188**, p. 275] calls "a candidate for the best of all multipliers" because of the closeness of its 2-dimensional lattice to a cubic lattice actually shows a miserable behavior under the 2-dimensional serial test.

There is a way in which the lattice structure can be profitably studied to gain insight into the behavior of LCPRN. It is based on the observation of Marsaglia [**186**] that the points $\mathbf{y}_0, \mathbf{y}_1, \ldots, \mathbf{y}_{\tau-1}$ all lie on a limited number of parallel hyperplanes. If this number is too small, then there are large portions of the unit cube I^s devoid of such points, and thus the generator is unacceptable. Equivalently, one may measure the distances between neighboring hyperplanes containing these points, and if the distances are too large, the generator should be disqualified. In the latter form, this test is very strongly linked with the spectral test (cf. [**154**, p. 100]). In fact, the "hyperplane test" may be viewed as a geometric version of the spectral test.

Relations between successive terms have also been investigated in the case of Tausworthe generators (see §7). Results about the serial correlation can already be found in the early papers on these generators ([**333**], [**362**]). The run test was applied in [**337**] and the weaker version of the serial test (see §6, test G, second part) in [**68**], [**333**]. In [**338**] excellent multidimensional uniformity properties are reported for a 3-term recurrence of order 607 modulo 2. But, as pointed out in [**363**], Tausworthe generators also have to be used with care and the desired degree of statistical independence should be ascertained before adopting them. The lattice structure of Tausworthe generators was studied in [**89**].

11. Serial test. The discussion in §10 has not included the strongest test for the statistical independence of successive terms, which will now be analyzed in detail. To set up the serial test (see §6, test G) for a sequence $x_0, x_1, \ldots$ of LCPRN, we choose a dimension $s \geqslant 2$ and consider the points $\mathbf{x}_n = (x_n, x_{n+1}, \ldots, x_{n+s-1})$, $n = 0, 1, \ldots$, in I^s and their distribution behavior, measured quantitatively by the discrepancy D_N. In the easier case of the multiplicative congruential method, the points $\mathbf{x}_n$ may be represented explicitly by the compact formula

$$\mathbf{x}_n = \{\lambda^n x_0 \boldsymbol{\lambda}\} \quad \text{for } n = 0, 1, \ldots,$$

where $\boldsymbol{\lambda}$ denotes the lattice point $(1, \lambda, \lambda^2, \ldots, \lambda^{s-1}) \in \mathbf{Z}^s$. It is therefore not surprising that this lattice point should play a special rôle in our investigation, and it will also transpire that the very same lattice point is of equal importance in the mixed congruential method.

The first attempt at getting a grip on the multidimensional distribution properties of LCPRN was undertaken in a paper of Franklin [72]. Here a continuous model was developed by accepting as an initial value any number x_0 in the unit interval I and generating a sequence $x_0, x_1, \ldots$ by the recursion $x_{n+1} = \{\lambda x_n + \theta\}$, $n = 0, 1, \ldots$, where the integer $\lambda > 1$ and the real number θ are fixed. For a given dimension $s \geqslant 2$, we associate with this sequence the points $\mathbf{x}_0, \mathbf{x}_1, \ldots$ in I^s in the same way as above. It is easily seen that the sequence $\mathbf{x}_0, \mathbf{x}_1, \ldots$ can never be uniformly distributed in I^s. However, the sequence is, in a certain sense, almost always "asymptotically" uniformly distributed. To make this explicit, we have to adopt a more careful notation. We fix θ and emphasize the dependence on λ by writing $\mathbf{x}_n(\lambda)$ instead of $\mathbf{x}_n$. Since the initial value x_0 is independent of λ, we may retain the simpler symbol for it. Franklin shows now that for almost all $x_0 \in I$ (in the sense of Lebesgue measure) we have

$$\lim_{\lambda\to\infty} \lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} c_J(\mathbf{x}_n(\lambda)) = |J|$$

for all subintervals J of I^s, where c_J is the characteristic function and $|J|$ the Lebesgue measure of J. Furthermore, if f is a continuous function on I^s, then for almost all $x_0 \in I$ the relation

$$\lim_{\lambda\to\infty} \lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} f(\mathbf{x}_n(\lambda)) = \int_{I^s} f(\mathbf{t})\, d\mathbf{t}$$

is satisfied (see also [304] for related results). The regularity condition on f may be relaxed. For the case $\theta = 0$ corresponding to the multiplicative congruential method, Ermakov [65] established a quantitative refinement in the form of an "asymptotic" central limit theorem. Let f again be a continuous function on I^s and denote by

$$R_N(f, x_0, \lambda) = \frac{1}{N} \sum_{n=0}^{N-1} f(\mathbf{x}_n(\lambda)) - \int_{I^s} f(\mathbf{t})\, d\mathbf{t}$$

the integration error. Then we have

$$\lim_{N\to\infty} \lim_{\lambda\to\infty} \left|\left\{x_0 \in I\colon R_N(f, x_0, \lambda) < \frac{\sigma u}{\sqrt{N}}\right\}\right| = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{u} e^{-t^2/2}\, dt, \tag{11.1}$$

where the constant σ designates a certain standard deviation depending on f. The condition on f may be relaxed in the same way as in Franklin's theorem. These results suggest that, at least on probabilistic grounds, things should work out all right for a sequence of LCPRN with large modulus and large multiplier, but they do not provide any information about individual sequences of LCPRN. We note that the shuffling scheme of MacLaren and Marsaglia [184] was analyzed by Rosenblatt [260] for dimensions $s = 2$ and 3 in terms of an analogous probabilistic model.

It is, of course, considerably more difficult to get effective results about specific sequences of LCPRN. Recent research of the author has produced a satisfactory theory for any dimension $s \geqslant 2$. We let $x_0, x_1, \ldots$ be an arbitrary sequence of LCPRN and define the points $\mathbf{x}_0, \mathbf{x}_1, \ldots$ in I^s as before. To recognize clearly the dependence on the dimension, we denote by $D_N^{(s)}$ the discrepancy of the points $\mathbf{x}_0, \mathbf{x}_1, \ldots, \mathbf{x}_{N-1}$ in I^s. It will be convenient to use an expression that is closely related to the one introduced in Definition 4.3.

11.1. Definition. For a modulus m, a multiplier λ, a dimension $s \geqslant 2$, and a positive integer d we set

$$R^{(s)}(\lambda, m, d) = \sum_{\substack{\mathbf{h} \,(\mathrm{mod}\, m) \\ \mathbf{h}\cdot\boldsymbol{\lambda} \equiv 0 \,(\mathrm{mod}\, d)}}^{*} r(\mathbf{h})^{-1},$$

where $\boldsymbol{\lambda} = (1, \lambda, \lambda^2, \ldots, \lambda^{s-1}) \in \mathbf{Z}^s$, the summation symbol is the same as in (4.5), and $r(\mathbf{h})$ is defined by (4.3).

If $d = m$, then $R^{(s)}(\lambda, m, m)$ is the same as $R(\boldsymbol{\lambda}, m)$ from Definition 4.3. The subsequent estimates for $D_N^{(s)}$ will be in terms of $R^{(s)}(\lambda, m, d)$, where d depends in a known way on the parameters determining the specific sequence. The case $N = \tau$, i.e., when we consider the full period, is somewhat easier and will be treated first. The method rests on Lemma 3.9 and the estimates in §8. We simplify the upper bounds by using the inequality (4.19). For a prime modulus m we obtain the following.

11.2. Theorem (Niederreiter [227], [229]). *For a sequence of LCPRN with prime modulus m we have*

$$D_\tau^{(s)} < \frac{s}{m} + \frac{1}{\tau}(m - \tau)^{1/2}\left(\frac{2}{\pi} \log m + \frac{7}{5}\right)^s + \frac{1}{2} R^{(s)}(\lambda, m, m).$$

The second term in this upper bound is nonincreasing as a function of τ and so becomes minimal for the maximal value of τ, namely $\tau = m - 1$. Therefore, if we choose λ to be a primitive root modulo m, then we obtain

$$D_{m-1}^{(s)} < \frac{s}{m} + \frac{1}{m-1}\left(\frac{2}{\pi} \log m + \frac{7}{5}\right)^s + \frac{1}{2} R^{(s)}(\lambda, m, m). \qquad (11.2)$$

To get LCPRN for which s successive terms have a low degree of statistical dependence, we now select a primitive root λ modulo m yielding a small value of $R^{(s)}(\lambda, m, m)$. Since $R^{(s)}(\lambda, m, m) = R(\boldsymbol{\lambda}, m)$, this is equivalent to the desideratum that $\boldsymbol{\lambda}$ be a good lattice point modulo m (compare with §4). We have thus established an intriguing connection between pseudo-random number generation and the theory of good lattice points. As it stands, this theory does not indicate how small we can make $R^{(s)}(\lambda, m, m)$ since the requirement that λ be a primitive root modulo m is of no consequence there. It is therefore necessary to prove an existence theorem going beyond those in §4 (cf. [229]). In combination with (11.2), one obtains then the following result.

11.3. Theorem (Niederreiter [229]). *For any prime modulus m and any dimension s* $\geqslant 2$, *there exists a primitive root* λ *modulo m such that*

$$D_{m-1}^{(s)} < \frac{1}{m-1}\left(1 + \frac{(m-2)(s-1)}{\phi(m-1)}\right)\left(\frac{2}{\pi}\log m + \frac{7}{5}\right)^s, \tag{11.3}$$

where ϕ is Euler's totient function.

Since it is well known that $\phi(b)^{-1} = O(b^{-1}\log\log b)$ with an effective implied constant, we conclude that for a prime modulus m the parameters in (7.1) can be chosen in such a way that $D_{m-1}^{(s)} = O(m^{-1}(\log m)^s \log\log m)$ with a constant only depending on s. This should be compared with the result from §3 according to which the best distribution of $m - 1$ points in I^s currently known involves a discrepancy of the order of magnitude $m^{-1}(\log m)^{s-1}$. Therefore, sequences of appropriately selected LCPRN perform remarkably well under the serial test. Actually, the bound in (11.3) represents the expected magnitude of $D_{m-1}^{(s)}$ for a random primitive root modulo m. This conforms with an experience encountered by people working with empirical data, namely that a blindly chosen multiplier will tend to lead to good statistical properties.

In the case of a prime power modulus, the bounds for $D_\tau^{(s)}$ become even simpler, although we have to employ now the general expression introduced in Definition 11.1. We also need the number γ defined prior to Theorem 9.2.

11.4. THEOREM (NIEDERREITER [230], [233]). *For a sequence of LCPRN with modulus $m = p^\alpha$, p prime, $\alpha \geqslant 2$, and satisfying $\gamma < \alpha$ we have*

$$D_\tau^{(s)} < s/m + \tfrac{1}{2}R^{(s)}(\lambda, m, p^{\alpha-\gamma}). \tag{11.4}$$

For the homogeneous case this was already shown in [227], [229]; then we have of course $\gamma = \beta$ (see §9). It is perhaps worthwhile to list the values of γ for $m = 2^\alpha$, $\alpha \geqslant 3$, and the most common choices of the multiplier λ and the increment r. In the homogeneous case, $\lambda \equiv 5 \pmod 8$ leads to $\gamma = 2$ and $\lambda \equiv 3 \pmod 8$ has $\gamma = 3$. In the inhomogeneous case with r odd, $\lambda \equiv 5 \pmod 8$ yields $\gamma = 0$, whereas $\lambda \equiv 3 \pmod 8$ leads to $\gamma = 2$. Since small values of γ are preferable, this provides another reason why $\lambda \equiv 5 \pmod 8$ and r odd seems to be a good choice. On the whole, however, the controversy about whether the multiplicative or the mixed congruential method is superior is rather academic since the difference between the two cases is negligible if Theorem 11.4 is taken as indicative. There is other evidence for the essential equivalence of the two methods, and it can be safely said that a consensus has been reached on this issue (compare with [57, p. 857], [154, p. 20], [188, p. 251]).

Because of the connection with good lattice points, and also for aesthetic reasons, a "symmetric" expression of the form $R^{(s)}(\lambda, m', m')$ is more agreeable than the one appearing in (11.4). Fortunately enough, the latter expression is related in a fairly simple manner to one of the symmetric type. In fact, for $0 \leqslant \gamma < \alpha$ we have

$$R^{(s)}(\lambda, m, p^{\alpha-\gamma}) \leqslant (1 + 2\gamma\log p)^s R^{(s)}(\lambda, p^{\alpha-\gamma}, p^{\alpha-\gamma}) \\ + (1 + (2\gamma p^\gamma \log p)/m)^s - 1 \tag{11.5}$$

for p odd, according to [229, Equation (5.12)], and an almost identical result holds for $p = 2$ (cf. [229, Equation (5.20)]). In the cases of interest, γ is very small, and then $R^{(s)}(\lambda, m, p^{\alpha-\gamma})$ is of the same order of magnitude as $R^{(s)}(\lambda, p^{\alpha-\gamma}, p^{\alpha-\gamma})$. This amounts to saying that the arithmetic properties of λ are not so much relevant modulo $m = p^{\alpha}$ but modulo $p^{\alpha-\gamma}$.

We turn now to estimates for the discrepancy $D_N^{(s)}$ referring to parts of the period. The following result can be shown for a prime modulus.

11.5. THEOREM (NIEDERREITER [230], [233]). *For a sequence of LCPRN with prime modulus m we have*

$$D_N^{(s)} < \frac{s}{m} + \frac{\sqrt{m}}{N}\left(\frac{2}{\pi}\log\tau + \frac{7}{5}\right)\left(\frac{2}{\pi}\log m + \frac{7}{5}\right)^s + \frac{1}{2}R^{(s)}(\lambda, m, m)$$

$$\textit{for } 1 \leqslant N \leqslant \tau. \quad (11.6)$$

If one chooses a primitive root λ modulo m according to Theorem 11.3, then the third term on the right-hand side of (11.6) is $O(m^{-1}(\log m)^s \log\log m)$, and so $ND_N^{(s)} = O(m^{1/2}(\log m)^{s+1})$. We observe that in general we should not expect a smaller exponent of m for large N because of Ermakov's central limit theorem in (11.1). If N is comparable with m in size, then the order of magnitude of $ND_N^{(s)}$ resembles that of a "random" sequence of points in I^s (see §3).

For the statement of the estimate in the case of a prime power modulus we use the same notation as in Theorem 9.2.

11.6. THEOREM (NIEDERREITER [230], [233]). *For a sequence of LCPRN with modulus $m = p^{\alpha}, p$ prime, $\alpha \geqslant 2$, and satisfying $\gamma < \alpha$ we have*

$$D_N^{(s)} < \frac{s}{m} + \frac{1}{N}\left(\frac{m\tau}{\mu}\right)^{1/2}\left(\frac{2(p-1)}{\pi p}\log\tau + \frac{3}{4}\right)\left(\frac{2}{\pi}\log m + \frac{7}{5}\right)^s$$

$$+ \frac{1}{2}R^{(s)}(\lambda, m, p^{\alpha-\gamma}) \quad \textit{for } 1 \leqslant N \leqslant \tau. \quad (11.7)$$

In the homogeneous case, (11.7) can be simplified somewhat on the basis of the identity $\tau = \mu$. For the mixed congruential method with $m = 2^{\alpha}$, $\alpha \geqslant 3$, and r odd, we have $\tau/\mu = 4$ for $\lambda \equiv 5 \pmod 8$ and $\tau/\mu = 2$ for $\lambda \equiv 3 \pmod 8$. By using (11.5), we may, up to a constant, replace the last term in (11.7) by $R^{(s)}(\lambda, p^{\alpha-\gamma}, p^{\alpha-\gamma})$ in the cases of interest.

As we already observed in the discussion of the theory of good lattice points, a quantity such as $R^{(s)}(\lambda, m, m)$ is awkward to deal with numerically. Consequently, we adopt the same remedy as there and introduce an integer $\rho^{(s)}(\lambda, m)$ by setting

$$\rho^{(s)}(\lambda, m) = \rho(\boldsymbol{\lambda}, m), \quad (11.8)$$

where $\boldsymbol{\lambda} = (1, \lambda, \lambda^2, \ldots, \lambda^{s-1}) \in \mathbf{Z}^s$ and $\rho(\boldsymbol{\lambda}, m)$ is defined by (4.8). We can then rephrase Theorem 4.5 as follows.

11.7. LEMMA. *For any dimension $s \geqslant 2$ and any integers $m \geqslant 2$ and λ we have*

$$R^{(s)}(\lambda, m, m) < C_s(\log m)^s/\rho^{(s)}(\lambda, m)$$

with a constant C_s only depending on s.

It follows from the above considerations that a sequence of LCPRN passes the s-dimensional serial test if, in the case of a prime modulus m, we choose a multiplier λ which is a primitive root modulo m and yields a large value of $\rho^{(s)}(\lambda, m)$, and in the case of a prime power modulus $m = p^\alpha$ we select the parameters in such a way that both the period and $\rho^{(s)}(\lambda, p^{\alpha-\gamma})$ are large. That the size of the number from (11.8) is really the correct indicator for the performance of a sequence of LCPRN under the s-dimensional serial test is shown by the following result.

11.8. Theorem (Niederreiter [230], [233]). *For any sequence of LCPRN with modulus m and multiplier* λ, *we have*

$$D_N^{(s)} \geqslant C_s'/\rho^{(s)}(\lambda, m) \quad \textit{for } 1 \leqslant N \leqslant \tau,$$

where C_s' is a positive constant only depending on the dimension s.

We may take $C_s' = s^{-s}$ for $2 \leqslant s \leqslant 3$ and $C_s' = (\pi/2)(\pi + \frac{1}{2})^{-s}$ for $s \geqslant 4$. The lower bound may be improved in some important special cases. For instance, if we consider the multiplicative congruential method with $m = 2^\alpha$, $\alpha \geqslant 3$, and $\lambda \equiv 5 \pmod 8$, then

$$D_N^{(s)} \geqslant C_s'/\rho^{(s)}(\lambda, 2^{\alpha-2}) \quad \text{for } 1 \leqslant N \leqslant \tau.$$

For the most popular generators and the serial test for the full period, we can summarize all this information in a more succinct form by employing Vinogradov's notation $\ll$ to assimilate constants only depending on s. If we use a prime modulus m and a primitive root λ modulo m in either a multiplicative or a mixed congruential method, or if we use a mixed congruential method with $m = 2^\alpha$, $\alpha \geqslant 3$, $\lambda \equiv 5 \pmod 8$, and r odd, then we obtain[37]

$$\frac{1}{\rho^{(s)}(\lambda, m)} \ll D_\tau^{(s)} \ll \frac{(\log m)^s}{\rho^{(s)}(\lambda, m)}. \tag{11.9}$$

If we use a multiplicative congruential method with $m = 2^\alpha$, $\alpha \geqslant 3$, and $\lambda \equiv 5 \pmod 8$, then we have

$$\frac{1}{\rho^{(s)}(\lambda, 2^{\alpha-2})} \ll D_\tau^{(s)} \ll \frac{(\log m)^s}{\rho^{(s)}(\lambda, 2^{\alpha-2})}. \tag{11.10}$$

Thus the order of magnitude of $D_\tau^{(s)}$ is, up to logarithmic factors, completely determined by the order of magnitude of $\rho^{(s)}(\lambda, m)$ in (11.9) resp. $\rho^{(s)}(\lambda, 2^{\alpha-2})$ in (11.10). We take this as the justification for referring, in general, to the integer $\rho^{(s)}(\lambda, m)$ defined by (11.8) as the (s-dimensional) *figure of merit* of the multiplier λ (with respect to m).

We have with deliberate care emphasized in the notation that the figure of merit $\rho^{(s)}(\lambda, m)$ depends strongly on the dimension s. It is implicit in this fact that a multiplier which is excellent for a certain dimension may be unaccept-

[37] We apply the inequality $\rho^{(s)}(\lambda, m) \leqslant m/2$ shown immediately after Theorem 4.6 in order to incorporate secondary terms into the main term.

able for another dimension. The reason behind this phenomenon has been discussed in §4, and we shall present concrete examples later on. There is thus a *relativity principle* in operation, according to which the choice of an excellent (or even optimal) multiplier has to be made relative to the dimensionality of, or the desired number of statistically independent successors in, the Monte Carlo problem at hand. *There is no such thing as a universally optimal multiplier.* Every individual pseudo-random number generator becomes unsuitable for sufficiently high dimensions, even though it may perform well for small dimensions. This can actually be put in a concrete form. For suppose we generate a sequence of LCPRN with modulus m and multiplier λ and apply the s-dimensional serial test with $s \geqslant \log_2(m + 1)$. Consider all lattice points $\mathbf{h} = (h_1, \ldots, h_s)$ with the h_j taking independently the values 0 and 1. There are 2^s such lattice points, and since $2^s \geqslant m + 1$, there exist two distinct lattice points $\mathbf{h}^{(1)}, \mathbf{h}^{(2)}$ among these such that $\mathbf{h}^{(1)} \cdot \boldsymbol{\lambda} \equiv \mathbf{h}^{(2)} \cdot \boldsymbol{\lambda} \pmod{m}$. Then $(\mathbf{h}^{(1)} - \mathbf{h}^{(2)}) \cdot \boldsymbol{\lambda} \equiv 0 \pmod{m}$ and $r(\mathbf{h}^{(1)} - \mathbf{h}^{(2)}) = 1$, and so $\rho^{(s)}(\lambda, m) = 1$ for any multiplier λ, which is unacceptable. For instance, the modulus $m = 2^{35}$ certainly has to be ruled out for dimensions $s \geqslant 36$. In practice, the threshold is of course considerably lower, and this particular modulus is probably dubious for $s > 10$. However, this should not be construed as a suggestion that Lehmer's method breaks down for moderately high dimensions. On the contrary, results such as Theorem 11.3 guarantee that if we are willing to use a sufficiently large modulus, then there is always a sequence of LCPRN which performs well under the s-dimensional serial test for an arbitrarily given dimension s. Deficiencies that may occur are therefore not the fault of the generator, but of the limited computer capacity. These limits are of course constantly pushed upwards.

As to the variability of excellent multipliers with changing dimension, something of a positive nature can be said. It is based on the inequality (4.12) which we can rewrite as

$$\rho^{(t)}(\lambda, m) \geqslant \rho^{(s)}(\lambda, m) \quad \text{for dimensions } 2 \leqslant t < s. \tag{11.11}$$

We infer from (11.11) that if a multiplier λ is favorable for the dimension s, i.e., if it leads to a large value of $\rho^{(s)}(\lambda, m)$, then we can expect this multiplier to show an acceptable (though not necessarily optimal) behavior for all lower dimensions. Thus, if a generator is to be used for several purposes, it suffices to choose it in such a way that it satisfies the most stringent statistical independence condition desired in these applications. Viewed from a different angle, the inequality (11.11) tells us that if a sequence of LCPRN fails the serial test for a certain dimension, it will fail the test for all higher dimensions, which is intuitively quite obvious.

We observed in §4 that for the case $s = 2$ more information is available in the theory of good lattice points because of intimate connections with the continued fraction algorithm. The same holds for the 2-dimensional serial test, i.e., for the distribution of pairs of successive LCPRN. In the first place, the 2-dimensional figure of merit $\rho^{(2)}(\lambda, m)$ is related in a very simple manner with a number arising from the continued fraction expansion of the rational λ/m. This is obtained from (4.14) which can be rewritten in the form

$$m/(K+2) \leqslant \rho^{(2)}(\lambda, m) \leqslant m/K, \tag{11.12}$$

where $K = K(\lambda/m)$ is defined by (4.13) and represents the maximal partial quotient in the continued fraction expansion of λ/m. Therefore, the figure of merit $\rho^{(2)}(\lambda, m)$ can only be large if λ/m yields small partial quotients. This raises the question of how small we can make $K(\lambda/m)$ given the modulus m. The number K_m from §4 pertains to this question, but it is not necessarily appropriate here since we also have to take into account the fact that λ should produce a sequence of LCPRN with large period. For instance, in the case of a prime modulus m we actually have to consider

$$P_m = \min_{\lambda} K\left(\frac{\lambda}{m}\right),$$

where λ runs through all primitive roots modulo m. A theoretical estimate for P_m is available since it was shown in [229, §4] that for any prime m there exists a primitive root λ modulo m with

$$\rho^{(2)}(\lambda, m) \gg \frac{\phi(m-1)}{\log \phi(m-1)},$$

which together with (11.12) implies that

$$P_m \ll \frac{m \log \phi(m-1)}{\phi(m-1)} \ll (\log m) \log \log m.$$

This bound is certainly not best possible.

The 2-dimensional case also permits us to establish finer estimates for the discrepancy $D_\tau^{(2)}$ of the full period, at least for the generators of practical interest. This is achieved by a direct approach and was first carried out by Dieter [57], [58]. The method of Dieter is based on an explicit representation of the local deviations $A(J; \tau) - \tau|J|$ for subintervals J of I^2 in terms of generalized Dedekind sums. These sums can then be calculated by means of a reciprocity law and other arithmetic properties, and the resulting algorithm resembles the continued fraction algorithm (or, equivalently, the Euclidean algorithm). In this way, one obtains exact formulas for the various local deviations. We get information about $D_\tau^{(2)}$ by establishing a global bound on the local deviations. This can be done using either standard continued fractions (cf. [155]) or continued fractions to nearest integers, the alternative chosen by Dieter. Generalized Dedekind sums can also be applied to get a handle on the permutation test (see §6, test D) for triples ([58], [59]).

Apart from the fact that it only works for the full period, there is a further restriction on the above technique insofar as we have to know exactly which residues modulo m are generated by the recursion (7.1). This explains the additional conditions in the subsequent result. The bound for $D_\tau^{(2)}$ will be in terms of partial quotients arising from a continued fraction algorithm to nearest integers.[38] The new feature here is that these partial quotients may also be negative.

11.9. Theorem (Dieter [57], [58]). *For a sequence of LCPRN with modulus*

[38] See [241, Chapter 5] for the theory of such continued fraction algorithms.

m, multiplier λ, and either (i) $\tau = m$; *or* (ii) $m = 2^{\alpha}$, $\alpha \geqslant 3$, $\lambda \equiv 5 \pmod 8$, $r = 0$, *we have*

$$\tau D_{\tau}^{(2)} < \frac{1}{3} \sum_{i=0}^{k} |b_i| + k + \frac{17}{3}, \tag{11.13}$$

where $b_0, b_1, \ldots, b_k$ are the partial quotients in the continued fraction expansion to nearest integers of λ/m in case (i) *and of $\lambda/2^{\alpha-2}$ in case* (ii).

A slightly simpler bound is obtained if the discrepancy is only extended over dyadic squares contained in I^2 rather than over all subintervals of I^2 (cf. [57, Theorem 5.1]). This latter bound has been tabulated in [55], whereas values of local deviations relative to dyadic squares can be found in [57], [58]. The estimate (11.13) shows again that small partial quotients are the trademark of favorable multipliers. Case (ii) above is also in perfect harmony with (11.10), with both results stressing that in this case the arithmetic properties of λ are more relevant modulo $2^{\alpha-2}$ than modulo 2^{α}.

We shall demonstrate now that a result of the same type as Theorem 11.9 can be established without the technical apparatus of generalized Dedekind sums. Our estimate is in terms of standard continued fractions and its proof uses only elementary facts about continued fractions. The following auxiliary result is basic.

11.10. Lemma. *Let $m \geqslant 2$ and λ be integers with $\gcd(\lambda, m) = 1$ and let θ be a real number. Then the discrepancy D_N of the finite sequence of fractional parts $\{(\lambda n/m) + \theta\}$, $n = 0, 1, \ldots, N-1$, satisfies*

$$ND_N < \sum_{i=1}^{q} a_i \quad \text{for } 1 \leqslant N < m \tag{11.14}$$

and $mD_m = 1$, where $a_1, \ldots, a_q$ are the partial quotients in the continued fraction expansion of λ/m. We also have

$$ND_N < C(K)\log(N+1) \quad \text{for } 1 \leqslant N < m, \tag{11.15}$$

where $K = K(\lambda/m) = \max(a_1, \ldots, a_q)$ and $C(K) = 2/\log 2$ for $1 \leqslant K \leqslant 3$, $C(K) = (K+1)/\log(K+1)$ for $K \geqslant 4$.

Proof. If $N = m$, then the numbers $\{(\lambda n/m) + \theta\}$, $0 \leqslant n \leqslant m-1$, form a sequence of m equidistant points with distance $1/m$, and therefore we get $D_m = 1/m$. In the general case, we adapt a method from [220]. Let $1 = r_0 \leqslant r_1 < r_2 < \cdots < r_q = m$ be the denominators of the convergents to λ/m. Then an integer N with $1 \leqslant N < m$ can be represented in the form $N = \sum_{i=0}^{h} c_i r_i$ with $0 \leqslant h < q$, $c_h \geqslant 1$, and $0 \leqslant c_i \leqslant a_{i+1}$ for $0 \leqslant i \leqslant h$. We decompose the given sequence $\{(\lambda n/m) + \theta\}$, $n = 0, 1, \ldots, N-1$, into blocks, namely c_h blocks of length r_h, c_{h-1} blocks of length r_{h-1}, and so on. Consider such a block of length r_i, $0 \leqslant i \leqslant h$; it consists of elements of the form $\{(\lambda n/m) + \theta\}$, $n = n_0 + 1, \ldots, n_0 + r_i$. Let $p_i/r_i = [a_0; a_1, \ldots, a_i]$ be the ith convergent to λ/m. Then from the theory of continued fractions (cf. [241, p. 37]) we know that

$$\frac{\lambda}{m} = \frac{p_i}{r_i} + \frac{\delta_i}{r_i r_{i+1}} \quad \text{with } |\delta_i| \leqslant 1.$$

Writing $n = n_0 + j$ with $1 \leqslant j \leqslant r_i$, we get

$$\left\{ \frac{\lambda n}{m} + \theta \right\} = \left\{ \frac{\lambda n_0}{m} + \frac{jp_i}{r_i} + \frac{j\delta_i}{r_i r_{i+1}} + \theta \right\}$$
$$= \left\{ \frac{jp_i}{r_i} + u + \frac{j\delta_i}{r_i r_{i+1}} \right\},$$

where $u = (\lambda n_0/m) + \theta$. Since $\gcd(p_i, r_i) = 1$, the numbers $(jp_i/r_i) + u$, $1 \leqslant j \leqslant r_i$, considered modulo 1, form a sequence of r_i equidistant points with distance $1/r_i$, which therefore has discrepancy $1/r_i$. Because of $|j\delta_i/r_i r_{i+1}| \leqslant 1/r_{i+1}$ for $1 \leqslant j \leqslant r_i$, the finite sequence $\{(\lambda n/m) + \theta\}$, $n = n_0 + 1, \ldots, n_0 + r_i$, is obtained by cyclically shifting modulo 1 the elements $\{(jp_i/r_i) + u\}$, $1 \leqslant j \leqslant r_i$, either all to the right or all to the left by at most a distance $1/r_{i+1}$, where the direction of the shift depends only on the sign of δ_i. It is then easily seen that the discrepancy D_{r_i} of the finite sequence $\{(\lambda n/m) + \theta\}, n = n_0 + 1, \ldots, n_0 + r_i$, satisfies

$$D_{r_i} \leqslant 1/r_i + 1/r_{i+1}.$$

From the triangle inequality for discrepancies (cf. [**174**, p. 115]) and the way in which we decomposed the original sequence, it follows that the discrepancy D_N of the sequence $\{(\lambda n/m) + \theta\}, n = 0, 1, \ldots, N - 1$, satisfies

$$ND_N \leqslant \sum_{i=0}^{h} c_i\left(1 + \frac{r_i}{r_{i+1}}\right)$$
$$\leqslant \sum_{\substack{i=0\\c_i \neq 0}}^{h} \left(c_i + \frac{a_{i+1}r_i}{r_{i+1}}\right) \leqslant \sum_{\substack{i=0\\c_i \neq 0}}^{h} (c_i + 1), \tag{11.16}$$

where we used $c_i \leqslant a_{i+1}$ and $a_{i+1}r_i \leqslant r_{i+1}$ (cf. [**241**, p. 24]). The coefficients c_i obtained from the algorithm in [**220**, p. 148] have the properties[39] that $c_0 < a_1$ and that $c_i = a_{i+1}$ implies $c_{i-1} = 0$, and so the estimate (11.14) follows.

To obtain (11.15), it suffices to combine (11.16) with the inequality

$$\sigma(N) = \sum_{\substack{i=0\\c_i \neq 0}}^{h} (c_i + 1) \leqslant C(K)\log(N + 1) \quad \text{for } 1 \leqslant N < m, \tag{11.17}$$

where $\sigma(N)$ is well defined if we use the coefficients c_i produced by the algorithm mentioned above. We establish (11.17) by induction on h. If $r_0 < r_1$, then the smallest possible h is $h = 0$, and a corresponding N satisfies $1 \leqslant N < r_1 \leqslant K$. If $r_0 = r_1 = 1$, then the smallest possible h is $h = 1$ and a corresponding N satisfies $1 \leqslant N \leqslant r_2 - 1 \leqslant K$. Since $\sigma(N) = N + 1$ for these N, it suffices to show for the first step in the induction that

$$N + 1 \leqslant C(K)\log(N + 1) \quad \text{for } 1 \leqslant N \leqslant K. \tag{11.18}$$

[39] The first property follows from $r_1 = a_1$, and for the second property we note that if $r_i \leqslant N < r_{i+1}$, then $N - c_i r_i = N - a_{i+1}r_i < r_{i+1} - a_{i+1}r_i = r_{i-1}$.

But this follows from the fact that

$$C(K) = \max_{1 \leqslant x \leqslant K} \frac{x+1}{\log(x+1)}.$$

Now take an arbitrary N with $1 < r_h \leqslant N < r_{h+1}$ and write $N = c_h r_h + N_{h-1}$ with $0 \leqslant N_{h-1} < r_h$. Then $\sigma(N) = c_h + 1 + \sigma(N_{h-1})$, and the induction hypothesis yields $\sigma(N) \leqslant c_h + 1 + C(K)\log(N_{h-1} + 1)$, which holds also for $N_{h-1} = 0$. Now $N + 1 \geqslant c_h(N_{h-1} + 1) + N_{h-1} + 1 = (c_h + 1)(N_{h-1} + 1)$, and so

$$\sigma(N) \leqslant c_h + 1 + C(K)\log \frac{N+1}{c_h+1}.$$

Since $1 \leqslant c_h \leqslant a_{h+1} \leqslant K$, we can complete the argument by applying (11.18) with c_h in place of N.

11.11. THEOREM. *For a sequence of LCPRN with modulus m, multiplier λ, and either* (i) $\tau = m$; *or* (ii) $m = 2^\alpha$, $\alpha \geqslant 3$, $\lambda \equiv 5 \pmod 8$, $r = 0$, *we have*

$$\tau D_\tau^{(2)} \leqslant 1 + \sum_{i=1}^{q} a_i \qquad (11.19)$$

and

$$\tau D_\tau^{(2)} \leqslant 1 + C(K) \log \tau, \qquad (11.20)$$

where $a_1, \dots, a_q$ are the partial quotients in the continued fraction expansion of λ/m in case (i) *and of $\lambda/2^{\alpha-2}$ in case* (ii), *$K = \max(a_1, \dots, a_q)$, and $C(K)$ is as in Lemma* 11.10.

PROOF. In case (i), all residues modulo m are generated by (7.1). Therefore the points $\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_{m-1}$ are a permutation of the points

$$(n/m, \{(\lambda n + r)/m\}), \qquad n = 0, 1, \dots, m-1.$$

For a subinterval $J = [u_1, u_2) \times [v_1, v_2)$ of I^2, the number $A(J)$ of these points falling into J is equal to the number of integers n with $mu_1 \leqslant n < mu_2$ and $\{(\lambda n + r)/m\} \in [v_1, v_2)$. If the first condition is not satisfied by any integer n, then

$$|A(J) - m|J|\,| = m|J| < 1. \qquad (11.21)$$

Otherwise, the first condition is satisfied by the integers $n = N_1, N_1 + 1, \dots, N_2 - 1$, say, where $0 \leqslant N_1 < N_2 \leqslant m$. Thus $A(J)$ is equal to the number of integers j, $0 \leqslant j < N_2 - N_1$, with

$$\left\{\frac{\lambda(N_1 + j) + r}{m}\right\} = \left\{\frac{\lambda j}{m} + \theta\right\} \in [v_1, v_2),$$

where $\theta = (\lambda N_1 + r)/m$. Hence we have $A(J) = A([v_1, v_2); N_2 - N_1)$, with the second counting function referring to the sequence $\{(\lambda j/m) + \theta\}$, $j = 0, 1, \dots, N_2 - N_1 - 1$. If $D_{N_2 - N_1}$ is the discrepancy of this sequence, then

$$\begin{aligned}|A(J) - m|J|\,| &= |A([v_1, v_2); N_2 - N_1) - m(u_2 - u_1)(v_2 - v_1)| \\ &\leqslant |A([v_1, v_2); N_2 - N_1) - (N_2 - N_1)(v_2 - v_1)| \\ &\quad + |N_2 - N_1 - m(u_2 - u_1)|(v_2 - v_1) \\ &\leqslant (N_2 - N_1)D_{N_2 - N_1} + |N_2 - N_1 - m(u_2 - u_1)|.\end{aligned}$$

Now $N_i = mu_i + \theta_i$, $0 \leqslant \theta_i < 1$, for $i = 1, 2$, and so

$$|A(J) - m|J|\,| \leqslant (N_2 - N_1)D_{N_2 - N_1} + 1 \leqslant \sum_{i=1}^{q} a_i + 1$$

by (11.14). Because of (11.21), this inequality holds for any J, so that we obtain (11.19) for the case (i). The inequality (11.20) for the same case is shown by applying (11.15) instead of (11.14).

In case (ii), the recursion (7.1) generates all residues modulo m that are $\equiv g \pmod 4$, where g is the least residue of y_0 modulo 4. Therefore the points $\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_{\tau-1}$ are a permutation of the points

$$((4n + g)/m, \{\lambda(4n + g)/m\}), \qquad n = 0, 1, \dots, \tau - 1,$$

with $\tau = m/4$. By using the same method as above, we obtain the desired results for case (ii).

The method in the proof of Theorem 11.11 can clearly be applied whenever the set of residues modulo m generated by (7.1) is known explicitly and consists of an arithmetic progression or possibly a union of arithmetic progressions. To provide one more example, consider the homogeneous case with a prime modulus m and a multiplier λ which is a primitive root modulo m. Then $\tau = m - 1$, and we get

$$(m - 1)D_{m-1}^{(2)} \leqslant 2 + \sum_{i=1}^{q} a_i$$

and

$$(m - 1)D_{m-1}^{(2)} \leqslant 2 + C(K) \log m$$

with the notation from case (i) in Theorem 11.11.

At first glance, the bound in (11.13) appears to be better than those in Theorem 11.11 by a factor $\frac{1}{3}$. This superficial comparison ignores the fact that the $|b_i|$ are usually larger than the a_i. If we use, for instance, $m = 2^{32}$, $\lambda = 1812433253$, and r odd, then $K = K(\lambda/m) = 2$, so that the bound in (11.20) becomes 65, whereas the bound in (11.19) is 50 and the bound in (11.13) is $49\frac{1}{3}$.

Borosh and Niederreiter [24] have carried out a systematic search for those multipliers λ yielding a small value of $K = K(\lambda/m)$ with respect to moduli m used in practice. The results show, for instance, that for each $m = 2^\alpha$, $6 \leqslant \alpha \leqslant 35$, there exists a multiplier λ with $K(\lambda/m) \leqslant 3$, and even one for which $\lambda \equiv 5 \pmod 8$. In many cases we can actually obtain $K(\lambda/m) = 2$. An example with $m = 2^{32}$ was already given above, and another interesting example is $m = 2^{30}$, $\lambda = 657759677$. Such multipliers yield, of course, a huge figure of merit $\rho^{(2)}(\lambda, m)$ and an absolutely fantastic distribution of pairs.

Further results for the case $s = 2$ can be found in [229]. For instance, for every modulus $m = 2^\alpha$ with $\alpha \geqslant 31$ there exists a $\lambda \equiv 5 \pmod 8$ with $\rho^{(2)}(\lambda, m) > \bar{m}/\log \bar{m}$, where $\bar{m} = 2m/5$. The calculations mentioned above suggest that one may, in fact, achieve a lower bound of the order of magnitude m. A similar existence theorem holds for odd prime power moduli m and multipliers λ that are primitive roots modulo m. For a positive divisor d of an arbitrary modulus m, the expression $R^{(2)}(\lambda, m, d)$ introduced in Definition 11.1 can be estimated in terms of $K(\lambda/d)$.

With all this information for $s = 2$, it can be said that the distribution of pairs of successive LCPRN is now clearly understood. For the moduli of practical interest, the discrepancy over the full period $\mathbf{x}_0, \mathbf{x}_1, \ldots, \mathbf{x}_{\tau-1}$ is $O(\tau^{-1} \log \tau)$ with suitable multipliers that can be determined effectively (cf. [24]). In the light of (3.11), this order of magnitude is the lowest possible for any τ points in I^2, so that well-chosen LCPRN lead to a nearly optimal distribution of pairs. In order to pass the 2-dimensional serial test, the decisive factor in the proper selection of a multiplier is the continued fraction expansion of λ/m (or of a closely related rational) which should only involve small partial quotients.

Proposals for the choice of multipliers have been based on weaker criteria, notably those of small serial correlation. The estimate of Greenberger [86] for the serial correlation suggested to take $\lambda \approx m^{1/2}$, but this yields $\rho^{(2)}(\lambda, m) = O(m^{1/2})$ and thus a distribution of pairs of modest quality. The same objection can be raised against any multiplier that is too small in comparison with the modulus. Even a generator such as $m = 2^{35}$, $\lambda = 2^{24} + 5$, $r = 1$, for which Jansson [141] reported excellent serial correlation properties, is not very impressive in terms of its distribution of pairs since we have $\rho^{(2)}(\lambda, m) \leqslant \lambda \approx m^{2/3}$. It must also be pointed out that multipliers which are too close to a fairly large power of 2 may permit a fast generation procedure, but perform very poorly under the 3-dimensional serial test. For instance, Jansson's generator satisfies $\lambda^2 - 10\lambda + 25 \equiv 0 \pmod m$, and so $\rho^{(3)}(\lambda, m) \leqslant 250$, which is much too small.[40] Another generator of this type that has been used in some computer installations is $m = 2^{35}$, $\lambda = 2^{18} + 3$, $r = 0$. Here we have $\gamma = 3$ and $\lambda^2 - 6\lambda + 9 \equiv 0 \pmod{2^{\alpha-\gamma}}$, so that $\rho^{(3)}(\lambda, 2^{\alpha-\gamma}) \leqslant 54$, which is totally unacceptable. The fact that such multipliers yield a bad distribution of triples was already noted earlier on the basis of numerical evidence (cf. [48], [87], [184], [342], [345]). The proposal of Ahrens, Dieter, and Grube [1] mentioned in §10 is thoroughly reasonable, although it was based on the weaker argument of small serial correlation and does not necessarily yield optimal multipliers with respect to the distribution of pairs.

Specific multipliers have also been proposed on the basis of the "cubic-lattice criterion" (see §10). Marsaglia [188, p. 275] comes out strongly in favor of $m = 2^{32}$, $\lambda = 69069$, which yields an almost cubic 2-dimensional lattice in the sense that it has a unit cell for which the longer side is about 1.06 times the shorter side. However, this does not imply a good distribution of pairs. Concretely, we have $69069 \cdot 1 - 1 \cdot \lambda \equiv 0 \pmod m$, and so $\rho^{(2)}(\lambda, m) \leqslant$

[40] This is an illustration of the fact that "going up" in the dimension may change the performance of a multiplier completely.

69069. Thus, if we use this multiplier in the mixed congruential method with r odd, it follows from Theorem 11.8 with $C_2' = 1/4$ that $mD_m^{(2)} > 15545$. On the other hand, $m = 2^{32}$, $\lambda = 1589013525$, and r odd yields $mD_m^{(2)} < 48$ by (11.13).[41] These data demonstrate more eloquently than anything else the inutility of the "cubic-lattice criterion". In fact, one would be better off choosing multipliers blindly rather than using this criterion, for Yao and Knuth [**364**] proved that the expected value of the bound in (11.19) is $(6/\pi^2)\log^2 m$ plus lower-order terms, which still leads to a satisfactory value for $D_m^{(2)}$.

We recapitulate that the performance of a sequence of LCPRN under the s-dimensional serial test is governed by a certain s-dimensional figure of merit which must be large for a good generator. We should observe the relativity principle and the general rule that a multiplier not be related to the modulus in too simple a manner. A saving of time in the generation procedure may be punished by a serious distortion in the numerical work for which the LCPRN are used. If the parameters are chosen according to the criteria spelled out in this section, then the s-tuples of successive LCPRN show an excellent distribution behavior in I^s, and so we have statistical almost-independence among s successive terms. In view of the simplicity of the recursion (7.1), it is quite astonishing how close the distribution of s-tuples can be to an optimal one. Marsaglia's devastating verdict that "congruential random number generators are not suitable for precision Monte Carlo use" [**188**, p. 250] can be overruled on the basis of new evidence.

Even among supporters of Lehmer's method one may occasionally sense a lingering suspicion about statistical independence properties (cf. [**154**, p. 56]). The standard argument leading to such misgivings has it that (7.1) establishes such an intimate link between successive terms that it is futile to hope for independence among these terms. In order to understand why LCPRN nevertheless perform so well, the following analogy may be helpful. In parentheses, we provide the translation from the model to LCPRN. Consider a square pool table and a billiard ball moving on it with enormously high speed (= the multiplier λ is very large). The tangent of the angle between a boundary of the table and the initial direction of the motion is supposed to be an irrational number (= for a good generator, the fraction λ/m has approximately the arithmetic properties of an irrational number). In one time unit, the billiard ball undergoes several thousand reflections at the boundaries (= y_{n+1} is obtained from $\lambda y_n + r$ by subtracting a multiple of m that is usually very large). The path of the billiard ball is locally a straight line and thus predictable, but the position at time $n + 1$ is practically independent of that at time n. The high velocity and the many reflections "smear out" any correlation that might exist locally (= y_{n+1} is almost independent of y_n). It is a known fact (cf. [**174**, p. 87, Exercise 9.29]) that under the given conditions the billiard ball visits, in the long run, every rectangular region and, more generally, every Jordan-measurable region on the table for an amount of time proportional to the area of that region (= a sequence of LCPRN with large

[41] The simpler bound (11.19) shows $mD_m^{(2)} < 51$, whereas (11.20) leads to $mD_m^{(2)} < 65$ since we have $K = K(\lambda/m) = 2$.

period and a properly chosen multiplier has an excellent distribution behavior). Thus, the pseudo-randomness properties of a good sequence of LCPRN are comparable to those implied by the determinate, but nevertheless wildly erratic and for all practical purposes "pseudo-Brownian" motion of a "supersonic" billiard ball.

References

1. J. H. Ahrens, U. Dieter and A. Grube, *Pseudo-random numbers: A new proposal for the choice of multiplicators*, Computing **6** (1970), 121–138.

2. J. L. Altaber, *Représentations arithmétiques de grandeurs aléatoires*, Ann. Fac. Sci. Univ. Clermont-Ferrand **37** (1967), 1–61.

3. I. I. Artobolevskiĭ, M. D. Genkin, V. K. Grinkevič, I. M. Sobol′ and R. B. Statnikov, *Optimization in the theory of machines by an LP-search*, Dokl. Akad. Nauk SSSR **200** (1971), 1287–1290. (Russian)

4. K. I. Babenko, *Approximation by trigonometric polynomials in a certain class of periodic functions of several variables*, Dokl. Akad. Nauk SSSR **132** (1960), 982–985 = Soviet Math. Dokl. **1** (1960), 672–675.

5. N. S. Bahvalov, *Approximate computation of multiple integrals*, Vestnik Moskov. Univ. Ser. Mat. Meh. Astr. Fiz. Him.**1959**, no. 4, 3–18. (Russian)

6. ______, *Numerical solution of the Dirichlet problem for Laplace's equation*, Vestnik Moskov. Univ. Ser. Mat. Meh. Astr. Fiz. Him.**1959**, no. 5, 171–195. (Russian)

7. ______, *An estimate of the main remainder term in quadrature formulae*, Ž. Vyčisl. Mat. i Mat. Fiz. **1** (1961), 64–77 = U.S.S.R. Computational Math. and Math. Phys. **1** (1961), 68–82.

8. ______, *On a rate of convergence of indeterministic integration processes within the functional classes* $W_p^{(l)}$, Teor. Verojatnost. i Primenen. **7** (1962), 238 = Theor. Probability Appl. **7** (1962), 227.

9. ______, *On the convergence of indeterministic integration processes on slightly smooth functions*, Teor. Verojatnost. i Primenen. **7** (1962), 473–474 = Theor. Probability Appl. **7** (1962), 463.

10. ______, *Optimal convergence bounds for quadrature processes and integration methods of Monte Carlo type for classes of functions*, Ž. Vyčisl. Mat. i Mat. Fiz. **4** (1964), no. 4, suppl., 5–63. (Russian)

11. N. S. Bahvalov, N. M. Korobov and N. N. Čencov, *The application of number-theoretic nets to numerical analysis problems*, Proc. Fourth All-Union Math. Congr. (Leningrad, 1961), vol. 2, Izdat. "Nauka", Leningrad, 1964, pp. 580–587. (Russian)

12. A. Baker, *On some diophantine inequalities involving the exponential function*, Canad. J. Math. **17** (1965), 616–626.

13. C. T. H. Baker, *On the nature of certain quadrature formulas and their errors*, SIAM J. Numer. Anal. **5** (1968), 783–804.

14. J. Bass, *Nombres aléatoires, suites arithmétiques, méthode de Monte-Carlo*, Publ. Inst. Statist. Univ. Paris **9** (1960), 289–325.

15. ______, *Stationary functions and their applications to turbulence*. I. *Stationary functions*, II. *Turbulent solutions of the Navier-Stokes equations*, J. Math. Anal. Appl. **47** (1974), 354–399, 458–503.

16. J. Bass and J. Guilloud, *Méthode de Monte-Carlo et suites uniformément denses*, Chiffres **1** (1958), 151–155.

17. K. Bauknecht, J. Kohlas and C. A. Zehnder, *Simulationstechnik*, Springer-Verlag, Berlin and New York, 1976.

18. J.-P. Bertrandias, *Calcul d'une intégrale au moyen de la suite* $X_n = An$. *Evaluation de l'erreur*, Publ. Inst. Statist. Univ. Paris **9** (1960), 335–357.

19. W. A. Beyer, *Lattice structure and reduced bases of random vectors generated by linear recurrences*, Applications of Number Theory to Numerical Analysis (S. K. Zaremba, ed.), Academic Press, New York, 1972, pp. 361–370.

20. W. A. Beyer, R. B. Roof and D. Williamson, *The lattice structure of multiplicative congruential pseudo-random vectors*, Math. Comp. **25** (1971), 345–363.

21. P. Billingsley and F. Topsøe, *Uniformity in weak convergence*, Z. Wahrscheinlichkeitstheorie verw. Gebiete **7** (1967), 1–16.

22. C. Binder, *Über einen Satz von de Bruijn und Post*, Österreich. Akad. Wiss. Math.-Natur. Kl. S.-B. II **179** (1970), 233–251.

23. I. Borosh, *Rational continued fractions with small partial quotients* (preprint).

24. I. Borosh and H. Niederreiter, *Optimal multipliers for pseudo-random number generation by the linear congruential method* (to appear).

25. G. W. Brown, *Monte Carlo methods*, E. F. Beckenbach (ed.), Modern Mathematics for the Engineer, McGraw-Hill, New York, 1956, Chapter 12.

26. O. V. Brušlinskaja, *Practical applications of the method of optimal coefficients to the computation of multiple integrals*, Questions of Computational Mathematics and Computing Technology (L. A. Ljusternik, ed.), Gos. Naučno-Tehn. Izdat. Mašinostr. Lit., Moscow, 1963, pp. 45–48. (Russian)

27. N. P. Buslenko, *Mathematical modeling of production processes on digital computers*, Izdat. "Nauka", Moscow, 1964; German transl., *Simulation von Produktionsprozessen*, Teubner, Leipzig, 1971.

28. N. P. Buslenko, D. I. Golenko, I. M. Sobol′, V. G. Sragovič and Ju. A. Šreĭder, *The method of statistical trials* (*Monte Carlo method*), Fizmatgiz, Moscow, 1962 = *The Monte Carlo method*, Yu. A. Shreider (ed.), Pergamon Press, Oxford, 1966.

29. V. V. Bykov, *Digital simulation and statistical radio engineering*, Izdat. "Sov. Radio", Moscow, 1971. (Russian)

30. J. W. S. Cassels, *An extension of the law of the iterated logarithm*, Proc. Cambridge Philos. Soc. **47** (1951), 55–64.

31. G. Cenacchi and A. de Matteis, *Pseudo-random numbers for comparative Monte Carlo calculations*, Numer. Math. **16** (1970), 11–15.

32. ______, *Quasi-random sequences by power residues*, Numer. Math. **20** (1972), 54–63.

33. N. N. Čencov, *On quadrature formulae for functions of an infinitely large number of variables*, Ž. Vyčisl. Mat. i Mat. Fiz. **1** (1961), 418–424 = U.S.S.R. Computational Math. and Math. Phys. **1** (1961), 455–464.

34. ______, *Pseudorandom numbers for modeling Markov chains*, Ž. Vyčisl. Mat. i Mat. Fiz. **7** (1967), 632–643 = U.S.S.R. Computational Math. and Math. Phys. **7** (1967), no. 3, 218–233.

35. G. J. Chaitin, *On the length of programs for computing finite binary sequences*, J. Assoc. Comput. Mach. **13** (1966), 547–569.

36. ______, *Randomness and mathematical proof*, Sci. Amer. **232** (1975), no. 5, 47–52.

37. C. K. Chui, *A convergence theorem for certain Riemann sums*, Canad. Math. Bull. **12** (1969), 523–525.

38. ______, *Concerning rates of convergence of Riemann sums*, J. Approximation Theory **4** (1971), 279–287.

39. ______, *Convergence of certain quadrature processes*, Aequationes Math. **9** (1973), 242–244.

40. K. L. Chung, *An estimate concerning the Kolmogoroff limit distribution*, Trans. Amer. Math. Soc. **67** (1949), 36–50.

41. A. Church, *On the concept of a random sequence*, Bull. Amer. Math. Soc. **46** (1940), 130–135.

42. H. Conroy, *Molecular Schrödinger equation*. VIII: *A new method for the evaluation of multidimensional integrals*, J. Chemical Phys. **47** (1967), 5307–5318.

43. A. H. Copeland, *Admissible numbers in the theory of probability*, Amer. J. Math. **50** (1928), 535–552.

44. J. Couot, *Applications des suites $m\theta$ à l'intégration numérique*, C. R. Acad. Sci. Paris Sér. A **264** (1967), 183–186.

45. ______, *Applications des suites $m\theta$ à l'intégration multiple sur le tore*, C. R. Acad. Sci. Paris Sér. A **266** (1968), 131–134.

46. R. R. Coveyou, *Serial correlation in the generation of pseudo-random numbers*, J. Assoc. Comput. Mach. **7** (1960), 72–74.

47. ______, *Random number generation is too important to be left to chance*, Studies in Appl. Math., vol. 3, Soc. Industr. Appl. Math., Philadelphia, Pa., 1969, pp. 70–111.

48. R. R. Coveyou and R. D. MacPherson, *Fourier analysis of uniform random number generators*, J. Assoc. Comput. Mach. **14** (1967), 100–119.

49. R. Cranley and T. N. L. Patterson, *Randomization of number theoretic methods for multiple integration*, SIAM J. Numer. Anal. **13** (1976), 904–914.

50. H. Davenport, *Note on irregularities of distribution*, Mathematika **3** (1956), 131–135.

51. P. J. Davis, *On the numerical integration of periodic analytic functions*, On Numerical Approximation (Proc. Sympos. Math. Research Center, Madison, Wis., 1958), R. E. Langer, ed., Univ. of Wisconsin Press, Madison, Wis., 1959, pp. 45–59.

52. P. J. Davis and P. Rabinowitz, *Some Monte Carlo experiments in computing multiple integrals*, Math. Tables Aids Comput. **10** (1956), 1–8.

53. ______, *Methods of numerical integration*, Academic Press, New York, 1975.

54. N. G. de Bruijn and K. A. Post, *A remark on uniformly distributed sequences and Riemann integrability*, Nederl. Akad. Wetensch. Proc. Ser. A **71** (1968) = Indag. Math. **30** (1968), 149–150.

55. R. Devillers, J. J. Dumont and G. Latouche, *Tests de générateurs pseudo-aléatoires*, Acad. Roy. Belg. Bull. Cl. Sci. (5) **59** (1973), 703–724.

56. U. Dieter, *Autokorrelation multiplikativ erzeugter Pseudo-Zufallszahlen*, Operations Research-Verfahren **6** (1969), 69–85.

57. ______, *Pseudo-random numbers: The exact distribution of pairs*, Math. Comp. **25** (1971), 855–883.

58. ______, *Statistical interdependence of pseudo-random numbers generated by the linear congruential method*, Applications of Number Theory to Numerical Analysis (S. K. Zaremba, ed.), Academic Press, New York, 1972, pp. 287–317.

59. ______, *Pseudo-random numbers: Permutations of triplets*, unpublished manuscript.

60. U. Dieter and J. Ahrens, *An exact determination of serial correlation of pseudo-random numbers*, Numer. Math. **17** (1971), 101–123.

61. D. Y. Downham and F. D. K. Roberts, *Multiplicative congruential pseudo-random number generators*, Comput. J. **10** (1967), 74–77.

62. H. J. A. Duparc, C. G. Lekkerkerker and W. Peremans, *Reduced sequences of integers and pseudo-random numbers*, Report ZW 1953-002, Math. Centrum, Amsterdam, 1953.

63. P. D. T. A. Elliott, *On distribution functions* (mod 1): *Quantitative Fourier inversion*, J. Number Theory **4** (1972), 509–522.

64. P. Erdös and P. Turán, *On a problem in the theory of uniform distribution.* I, Nederl. Akad. Wetensch. Proc. **51** (1948), 1146–1154 = Indag. Math. **10** (1948), 370–378.

65. S. M. Ermakov, *Note on pseudorandom sequences*, Ž. Vyčisl. Mat. i Mat. Fiz. **12** (1972), 1077–1082 = U.S.S.R. Computational Math. and Math. Phys. **12** (1972), no. 4, 307–314.

66. ______, *The Monte Carlo method and related questions*, 2nd ed., Izdat. "Nauka", Moscow, 1975; German transl. of 1st ed., *Die Monte-Carlo-Methode und verwandte Fragen*, Oldenbourg Verlag, Munich-Vienna, 1975.

67. M. Esmenjaud-Bonnardel, *Un procédé de génération de nombres "pseudo-aléatoires" pour* CAB 500, Rev. Française Traitement Information **7** (1964), 185–197.

68. B. M. Fellen, *An implementation of the Tausworthe generator*, Comm. ACM **12** (1969), 413.

69. G. S. Fishman, *Concepts and methods in discrete event digital simulation*, Wiley, New York, 1973.

70. L. D. Fosdick, *The Monte Carlo method in quantum statistics*, SIAM Rev. **10** (1968), 315–328.

71. J. N. Franklin, *On the equidistribution of pseudo-random numbers*, Quart. Appl. Math. **16** (1958), 183–188.

72. ______, *Deterministic simulation of random processes*, Math. Comp. **17** (1963), 28–59.

73. ______, *Numerical simulation of stationary and nonstationary Gaussian random processes*, SIAM Rev. **7** (1965), 68–80.

74. K. K. Frolov, *Upper error bounds for quadrature formulas on function classes*, Dokl. Akad. Nauk SSSR **231** (1976), 818–821 = Soviet Math. Dokl. **17** (1976), 1665–1669.

75. H. Gabai, *On the discrepancy of certain sequences* mod 1, Nederl. Akad. Wetensch. Proc. Ser. A **66** (1963) = Indag. Math. **25** (1963), 603–605.

76. ______, *On the discrepancy of certain sequences* mod 1, Illinois J. Math. **11** (1967), 1–12.

77. I. M. Gel'fand, S. M. Feĭnberg, A. S. Frolov and N. N. Čencov, *On application of the method of random trials (Monte Carlo method) for the solution of a kinetic equation*, Proc. 2nd Internat. Conf. on the Peaceful Uses of Atomic Energy (Geneva, 1958), vol. 2, Atomizdat, Moscow, 1959, pp. 628–633. (Russian)

78. I. M. Gel'fand, A. S. Frolov and N. N. Čencov, *The computation of continuous integrals by the Monte Carlo method*, Izv. Vysš. Učebn. Zaved. Matematika **1958**, no. 5, 32–45. (Russian)

79. V. S. Gladkiĭ, *Probabilistic computational models*, Izdat. "Nauka", Moscow, 1973. (Russian)

80. D. I. Golenko, *Simulation and statistical analysis of pseudo-random numbers on electronic computers*, Izdat. "Nauka", Moscow, 1965. (Russian)

81. S. W. Golomb, *Sequences with randomness properties*, Glenn L. Martin Co. Report, Baltimore, Md., 1955.

82. I. J. Good and R. A. Gaskins, *Some relationships satisfied by additive and multiplicative recurrent congruential sequences, with implications for pseudorandom number generation*, Computers in Number Theory (A. O. L. Atkin and B. J. Birch, eds.), Academic Press, London, 1971, pp. 125–136.

83. L. K. Gorskiĭ, *Statistical algorithms for investigating reliability*, Izdat. "Nauka", Moscow, 1970. (Russian)

84. B. L. Granovskiĭ and S. M. Ermakov, *The Monte Carlo method*, Itogi Nauki i Tekhniki **13** (1976), 59–108 = J. Soviet Math. **7** (1977), 161–192.

85. B. F. Green, Jr., J. E. K. Smith and L. Klem, *Empirical tests of an additive random number generator*, J. Assoc. Comput. Mach. **6** (1959), 527–537.

86. M. Greenberger, *An a priori determination of serial correlation in computer generated random numbers*, Math. Comp. **15** (1961), 383–389; Corrigenda, ibid. **16** (1962), 126, 406.

87. ______, *Method in randomness*, Comm. ACM **8** (1965), 177–179.

88. J. A. Greenwood, *A fast machine-independent long-period generator for* 31*-bit pseudorandom integers*, Compstat 1976: Proceedings in Computational Statistics (J. Gordesch and P. Naeve, eds.), Physica-Verlag, Vienna, 1976, pp. 30–37.

89. A. Grube, *Mehrfach rekursiv-erzeugte Pseudo-Zufallszahlen*, Z. Angew. Math. Mech. **53** (1973), T223–T225.

90. V. S. Gubenko, N. E. Kirillov, K. A. Meškovskiĭ and A. I. Čerkunov, *Formation of pseudo-random uniformly distributed numbers from noise-like signals*, Izv. Akad. Nauk SSSR Tehn. Kibernet. **1969**, no. 1, 57–63. (Russian)

91. F. G. Gustavson and W. Liniger, *A fast random number generator with good statistical properties*, Computing **6** (1970), 221–226.

92. S. Haber, *On a sequence of points of interest for numerical quadrature*, J. Res. Nat. Bur. Standards Sect. B **70** (1966), 127–136.

93. ______, *A modified Monte Carlo quadrature*, Math. Comp. **20** (1966), 361–368.

94. ______, *Sequences of numbers that are approximately completely equidistributed*, J. Assoc. Comput. Mach. **17** (1970), 269–272.

95. ______, *Numerical evaluation of multiple integrals*, SIAM Rev. **12** (1970), 481–526.

96. ______, *Experiments on optimal coefficients*, Applications of Number Theory to Numerical Analysis (S. K. Zaremba, ed.), Academic Press, New York, 1972, pp. 11–37.

97. S. Haber and C. F. Osgood, *On a theorem of Piatetsky-Shapiro and approximation of multiple integrals*, Math. Comp. **23** (1969), 165–168.

98. ______, *On the sum* $\Sigma\langle n\alpha\rangle^{-t}$ *and numerical integration*, Pacific J. Math. **31** (1969), 383–394.

99. J. H. Halton, *On the efficiency of certain quasi-random sequences of points in evaluating multi-dimensional integrals*, Numer. Math. **2** (1960), 84–90; Berichtigung, ibid., 196.

100. ______, *A retrospective and prospective survey of the Monte Carlo method*, SIAM Rev. **12** (1970), 1–63.

101. ______, *Estimating the accuracy of quasi-Monte Carlo integration*, Applications of Number Theory to Numerical Analysis (S. K. Zaremba, ed.), Academic Press, New York, 1972, pp. 345–360.

102. J. H. Halton and G. B. Smith, *Algorithm* 247: *Radical-inverse quasi-random point sequence* [G5], Comm. ACM **7** (1964), 701–702.

103. J. H. Halton and S. K. Zaremba, *The extreme and* L^2 *discrepancies of some plane sets*, Monatsh. Math. **73** (1969), 316–328.

104. H. C. Hamaker, *A simple technique for producing random sampling numbers*, Nederl. Akad. Wetensch. Proc. **52** (1949), 145–150.

105. J. M. Hammersley, *Monte Carlo methods for solving multivariable problems*, Ann. New York Acad. Sci. **86** (1960), 844–874.

106. J. M. Hammersley and D. C. Handscomb, *Monte Carlo methods*, Methuen, London, 1964.

107. N. Harada, *Optimal multipliers for the spectral test of uniform random number generators*, Information Processing in Japan **14** (1974), 120–126.

108. G. H. Hardy and J. E. Littlewood, *Notes on the theory of series*. XXIV: *A curious power series*, Proc. Cambridge Philos. Soc. **42** (1946), 85–90.

109. C. B. Haselgrove, *A method for numerical integration*, Math. Comp. **15** (1961), 323–337.

110. G. Helmberg, *Gleichverteilte Folgen in lokal kompakten Räumen*, Math. Z. **86** (1964), 157–189.

111. E. Hlawka, *Funktionen von beschränkter Variation in der Theorie der Gleichverteilung*, Ann. Mat. Pura Appl. **54** (1961), 325–333.

112. ______, *Über die Diskrepanz mehrdimensionaler Folgen* mod 1, Math. Z. **77** (1961), 273–284.

113. ______, *Zur angenäherten Berechnung mehrfacher Integrale*, Monatsh. Math. **66** (1962), 140–151.

114. ______, *Lösung von Integralgleichungen mittels zahlentheoretischer Methoden*. I, Österreich. Akad. Wiss. Math.-Nat. Kl. S.-B. II **171** (1962), 103–123.

115. ______, *Discrepancy and uniform distribution of sequences*, Compositio Math. **16** (1964), 83–91.

116. ______, *Uniform distribution modulo* 1 *and numerical analysis*, Compositio Math. **16** (1964), 92–105.

117. ______, *Trigonometrische Interpolation bei Funktionen von mehreren Variablen*, Acta Arith. **9** (1964), 305–320.

118. ______, *Interpolation analytischer Funktionen auf dem Einheitskreis*, Number Theory and Analysis (P. Turán, ed.), Plenum Press, New York, 1969, pp. 97–118.

119. ______, *Zur Definition der Diskrepanz*, Acta Arith. **18** (1971), 233–241.

120. ______, *Discrepancy and Riemann integration*, Studies in Pure Mathematics (L. Mirsky, ed.), Academic Press, New York, 1971, pp. 121–129.

121. ______, *Über eine Methode von E. Hecke in der Theorie der Gleichverteilung*, Acta Arith. **24** (1973), 11–31.

122. ______, *Anwendung zahlentheoretischer Methoden auf Probleme der numerischen Mathematik* I, Österreich. Akad. Wiss. Math.-Nat. Kl. S.-B. II **184** (1975), 217–225.

123. ______, *Numerische analytische Fortsetzung in Polyzylindern*, Österreich. Akad. Wiss. Math.-Nat. Kl. S.-B. II **184** (1975), 307–331.

124. E. Hlawka and K. Kreiter, *Lösung von Integralgleichungen mittels zahlentheoretischer Methoden* II, Österreich. Akad. Wiss. Math.-Nat. Kl. S.-B. II **172** (1963), 229–250.

125. E. Hlawka and R. Mück, *A transformation of equidistributed sequences*, Applications of Number Theory to Numerical Analysis (S. K. Zaremba, ed.), Academic Press, New York, 1972, pp. 371–388.

126. ______, *Über eine Transformation von gleichverteilten Folgen* II, Computing **9** (1972), 127–138.

127. E. W. Hobson, *The theory of functions of a real variable and the theory of Fourier's series*, vol. 1, 3rd ed., Cambridge Univ. Press, London, 1927.

128. L. C. Hsu, *Concerning the numerical integration of periodic functions of several variables*, Acta Sci. Math. (Szeged) **20** (1959), 230–233.

129. ______, *Note on the numerical integration of periodic functions and of partially periodic functions*, Numer. Math. **3** (1961), 169–173.

130. L.-K. Hua and Y. Wang, *Remarks concerning numerical integration*, Sci. Record (N.S.) **4** (1960), 8–11.

131. ______, *Numerical integration and its applications*, Science Press, Peking, 1963. (Chinese)

132. ______, *On diophantine approximations and numerical integrations*. I, II, Sci. Sinica **13** (1964), 1007–1010.

133. ______, *On numerical integration of periodic functions of several variables*, Sci. Sinica **14** (1965), 964–978.

134. ______, *On uniform distribution and numerical analysis* (*Number-theoretic method*). I, II, III, Sci. Sinica **16** (1973), 483–505; **17** (1974), 331–348; **18** (1975), 184–198.

135. T. E. Hull and A. R. Dobell, *Random number generators*, SIAM Rev. **4** (1962), 230–254.

136. D. W. Hutchinson, *A new uniform pseudorandom number generator*, Comm. ACM **9** (1966), 432–433.

137. R. Iglisch, *Zum Aufbau der Wahrscheinlichkeitsrechnung*, Math. Ann. **107** (1932), 471–484.

138. M. Isida and H. Ikeda, *Random number generator*, Ann. Inst. Statist. Math. Tokyo **8** (1956), 119–126.

139. M. I. Israilov and T. S. Maksudov, *Cubature formulae for singular integrals with Hilbert kernel on the class of functions E_n^α*, Dokl. Akad. Nauk UzSSR **1974**, no. 8, 10–12. (Russian)

140. D. L. Jagerman, *Some theorems concerning pseudo-random numbers*, Math. Comp. **19** (1965), 418–426.

141. B. Jansson, *Autocorrelations between pseudo-random numbers*, Nordisk Tidskr. Informations-Behandling **4** (1964), 6–27.

142. ______, *Random number generators*, Almqvist & Wiksell, Stockholm, 1966.

143. D. L. Johnson, *Generating and testing pseudo random numbers on the IBM type* 701, Math. Tables Aids Comp. **10** (1956), 8–13.

144. M. Kadyrov, *Tables of random numbers*, Izdat. Sredne-Aziatkogo Gos. Univ., Taškent, 1936. (Russian)

145. T. Kamae, *Subsequences of normal sequences*, Israel J. Math. **16** (1973), 121–149.

146. T. Kamae and B. Weiss, *Normal numbers and selection rules*, Israel J. Math. **21** (1975), 101–110.

147. G. Kedem, *The search for good lattice points in N dimensions*, Technical Report no. 1570, Math. Research Center, Madison, Wis., 1975.

148. G. Kedem and S. K. Zaremba, *A table of good lattice points in three dimensions*, Numer. Math. **23** (1974), 175–180.

149. M. G. Kendall and B. Babington Smith, *Random sampling numbers*, Tracts for Computers, no. 24, Cambridge Univ. Press, London, 1939.

150. J. Kiefer, *On large deviations of the empiric d. f. of vector chance variables and a law of the iterated logarithm*, Pacific J. Math. **11** (1961), 649–660.

151. G. W. King, *The Monte Carlo method as a natural mode of expression in operations research*, J. Operations Res. Soc. Amer. **1** (1953), 46–51.

152. P. Kirschenmann, *Concepts of randomness*, J. Philos. Logic **1** (1972), 395–414.

153. D. E. Knuth, *Construction of a random sequence*, Nordisk Tidskr. Informations-Behandling **5** (1965), 246–250.

154. ______, *The art of computer programming*, vol. 2: *Seminumerical algorithms*, Addison-Wesley, Reading, Mass., 1969.

155. ______, *Notes on generalized Dedekind sums*, Computer Sci. Dept., Stanford Univ., Stanford, Ca., 1975; Acta Arith. **33** (1977), 297–325.

156. J. F. Koksma, *Een algemeene stelling uit de theorie der gelijkmatige verdeeling modulo* 1, Mathematica B (Zutphen) **11** (1942/43), 7–11.

157. ______, *Some theorems on diophantine inequalities*, Scriptum no. 5, Math. Centrum, Amsterdam, 1950.

158. A. N. Kolmogorov, *On tables of random numbers*, Sankhyā Ser. A **25** (1963), 369–376.

159. ______, *Three approaches to the definition of the concept "quantity of information"*, Problemy Peredači Informacii **1** (1965), no. 1, 3–11. (Russian)

160. N. M. Korobov, *Approximate calculation of multiple integrals with the aid of methods in the theory of numbers*, Dokl. Akad. Nauk SSSR **115** (1957), 1062–1065. (Russian)

161. ______, *The approximate computation of multiple integrals*, Dokl. Akad. Nauk SSSR **124** (1959), 1207–1210. (Russian)

162. ______, *On some number-theoretic methods for the approximate computation of multiple integrals*, Uspehi Mat. Nauk **14** (1959), no. 2, 227–230. (Russian)

163. ______, *Computation of multiple integrals by the method of optimal coefficients*, Vestnik Moskov. Univ. Ser. Mat. Meh. Astr. Fiz. Him. **1959**, no. 4, 19–25. (Russian)

164. ______, *On the approximate solution of integral equations*, Dokl. Akad. Nauk SSSR **128** (1959), 235–238. (Russian)

165. ______, *Properties and calculation of optimal coefficients*, Dokl. Akad. Nauk SSSR **132** (1960), 1009–1012 = Soviet Math. Dokl. **1** (1960), 696–700.

166. ______, *Application of number-theoretic nets to integral equations and interpolation formulas*, Trudy Mat. Inst. Steklov. **60** (1961), 195–210. (Russian)

167. ______, *On applications of number-theoretic nets*, Computational Methods and Programming, Izdat. Moskov. Gos. Univ., Moscow, 1962, pp. 80–102. (Russian)

168. ______, *On number-theoretic methods in approximate analysis*, Questions of Computational Mathematics and Computing Technology (L. A. Ljusternik, ed.), Gos. Naučno-Tehn. Izdat. Mašinostr. Lit., Moscow, 1963, pp. 36–44. (Russian)

169. ______, *Number-theoretic methods in approximate analysis*, Fizmatgiz, Moscow, 1963. (Russian)

170. ______, *Some problems in the theory of diophantine approximation*, Uspehi Mat. Nauk **22** (1967), no. 3, 83–118 = Russian Math. Surveys **22** (1967), no. 3, 80–118.

171. ______, *Trigonometric sums with exponential functions and the distribution of signs in repeating decimals*, Mat. Zametki **8** (1970), 641–652 = Math. Notes **8** (1970), 831–837.

172. ______, *On the distribution of digits in periodic fractions*, Mat. Sb. (N.S.) **89** (1972), 654–670 = Math. USSR-Sb. **18** (1972), 659–676.

173. V. I. Krylov, *Approximate calculation of integrals*, Gos. Izdat. Fiz.-Mat. Lit., Moscow, 1959; Macmillan, New York, 1962.

174. L. Kuipers and H. Niederreiter, *Uniform distribution of sequences*, Wiley, New York, 1974.

175. R.-D. Kulle and A. Reich, *Flächenmessung mit gleichverteilten Folgen*, Nachr. Akad. Wiss. Göttingen, II. Math.-Phys. Kl., **1973**, no. 12, 217–225.

176. D. H. Lehmer, *Mathematical methods in large-scale computing units*, Proc. 2nd Sympos. on Large-Scale Digital Calculating Machinery (Cambridge, Mass., 1949), Harvard Univ. Press, Cambridge, Mass., 1951, pp. 141–146.

177. W. J. LeVeque, *An inequality connected with Weyl's criterion for uniform distribution*, Proc. Sympos. Pure Math., vol. 8, Amer. Math. Soc., Providence, R.I., 1965, pp. 22–30.

178. L. A. Levin, *On the notion of a random sequence*, Dokl. Akad. Nauk SSSR **212** (1973), 548–550 = Soviet Math. Dokl. **14** (1973), 1413–1416.

179. ______, *Uniform tests of randomness*, Dokl. Akad. Nauk SSSR **227** (1976), 33–35 = Soviet Math. Dokl. **17** (1976), 337–340; Erratum, ibid. **231** (1976), 264.

180. M. B. Levin, *On the uniform distribution of the sequence* $\{\alpha\lambda^x\}$, Mat. Sb. (N.S.) **98** (1975), 207–222 = Math. USSR-Sb. **27** (1975), 183–197.

181. P. A. W. Lewis, A. S. Goodman and J. M. Miller, *A pseudo-random number generator for the System*/360, IBM Systems J. **8** (1969), 136–146.

182. T. G. Lewis, *Distribution sampling for computer simulation*, Lexington Books, Farnborough, 1975.

183. W. Liniger, *On a method by D. H. Lehmer for the generation of pseudo-random numbers*, Numer. Math. **3** (1961), 265–270.

184. M. D. MacLaren and G. Marsaglia, *Uniform random number generators*, J. Assoc. Comput. Mach. **12** (1965), 83–89.

185. D. Maisonneuve, *Recherche et utilisation des "bons treillis". Programmation et résultats numériques*, Applications of Number Theory to Numerical Analysis (S. K. Zaremba, ed.), Academic Press, New York, 1972, pp. 121–201.

186. G. Marsaglia, *Random numbers fall mainly in the planes*, Proc. Nat. Acad. Sci. U.S.A. **61** (1968), 25–28.

187. ______, *Regularities in congruential random number generators*, Numer. Math. **16** (1970), 8–10.

188. ______, *The structure of linear congruential sequences*, Applications of Number Theory to Numerical Analysis (S. K. Zaremba, ed.), Academic Press, New York, 1972, pp. 249–285.

189. G. Marsaglia and T. A. Bray, *One-line random number generators and their use in combinations*, Comm. ACM **11** (1968), 757–759.

190. F. F. Martin, *Computer modeling and simulation*, Wiley, New York, 1968.

191. P. Martin-Löf, *The definition of random sequences*, Information and Control **9** (1966), 602–619.

192. ______, *The literature on von Mises' Kollektivs revisited*, Theoria **35** (1969), 12–37.

193. J. Maurin, *Simulation déterministe du hasard*, Masson, Paris, 1975.

194. P. McShane, *Randomness, statistics and emergence*, Univ. of Notre Dame Press, Notre Dame, Ind., 1970.

195. H. G. Meijer, *The discrepancy of a g-adic sequence,* Nederl. Akad. Wetensch. Proc. Ser. A **71** (1968) = Indag. Math. **30** (1968), 54–66.

196. H. G. Meijer and H. Niederreiter, *Equirépartition et théorie des nombres premiers,* Répartition Modulo 1 (Colloque de Marseille-Luminy, 1974), Lecture Notes in Math., vol. 475, Springer-Verlag, Berlin and New York, 1975, pp. 104–112.

197. M. Mendès France, *Calcul des moyennes des fonctions aléatoires ou pseudo-aléatoires par échantillonnage,* Publ. Inst. Statist. Univ. Paris **11** (1962), 225–256.

198. ______, *Suites de nombres au hasard (d'après Knuth),* Sém. Théorie des Nombres 1974–1975, Univ. Bordeaux, Exp. 6.

199. N. Metropolis and S. M. Ulam, *The Monte Carlo method,* J. Amer. Statist. Assoc. **44** (1949), 335–341.

200. H. A. Meyer (ed.), *Symposium on Monte Carlo methods,* Wiley, New York, 1956.

201. G. A. Mihram, *Simulation: Statistical foundations and methodology,* Academic Press, New York, 1972.

202. O. Miyatake, *Generation of uniform random numbers of good quality,* Math. Japon. **17** (1972), 79–84.

203. O. Miyatake, H. Inoue and Y. Yoshizawa, *Generation of physical random numbers,* Math. Japon. **20** (1975), 207–217.

204. L. J. Mordell, *On the exponential sum* $\Sigma_{x=1}^{X}$ $\exp(2\pi i(ax + bg^x)/h)$, Mathematika **19** (1972), 84–87.

205. ______, *A new type of exponential series,* Quart. J. Math. **23** (1972), 373–374.

206. R. Mück and W. Philipp, *Distances of probability measures and uniform distribution* mod 1, Math. Z. **142** (1975), 195–202.

207. R. E. Nance and C. Overstreet, Jr., *Bibliography* 29: *A bibliography on random number generation,* Comput. Rev. **13** (1972), 495–508.

208. T. H. Naylor, *Bibliography* 19: *Simulation and gaming,* Comput. Rev. **10** (1969), 61–69.

209. T. H. Naylor, J. L. Balintey and D. S. Burdick, *Computer simulation techniques,* Wiley, New York, 1966.

210. H. Neunzert and J. Wick, *Die Theorie der asymptotischen Verteilung und die numerische Lösung von Integrodifferentialgleichungen,* Numer. Math. **21** (1973), 234–243.

211. ______, *Die Darstellung von Funktionen mehrerer Variabler durch Punktmengen,* Report no. 996-MA, Kernforschungsanlage Jülich (West Germany), 1973.

212. ______, *Die Approximation der Lösung von Integro-Differentialgleichungen durch endliche Punktmengen,* Numerische Behandlung nichtlinearer Integrodifferential- und Differentialgleichungen (R. Ansorge and W. Törnig, eds.), Lecture Notes in Math., vol. 395, Springer-Verlag, Berlin and New York, 1974, pp. 275–290.

213. T. G. Newman and P. L. Odell, *The generation of random variates,* Hafner, New York, 1971.

214. H. Niederreiter, *Diskrepanz in kompakten abelschen Gruppen* II, Manuscripta Math. **1** (1969), 293–306.

215. ______, *Discrepancy and convex programming,* Ann. Mat. Pura Appl. **93** (1972), 89–97.

216. ______, *On a number-theoretical integration method,* Aequationes Math. **8** (1972), 304–311.

217. ______, *Methods for estimating discrepancy,* Applications of Number Theory to Numerical Analysis (S. K. Zaremba, ed.), Academic Press, New York, 1972, pp. 203–236.

218. ______, *On the distribution of pseudo-random numbers generated by the linear congruential method,* Math. Comp. **26** (1972), 793–795.

219. ______, *Metric theorems on the distribution of sequences,* Proc. Sympos. Pure Math., vol. 24, Amer. Math. Soc., Providence, R.I., 1973, pp. 195–212.

220. ______, *Application of diophantine approximations to numerical integration,* Diophantine Approximation and Its Applications (C. F. Osgood, ed.), Academic Press, New York, 1973, pp. 129–199.

221. ______, *On the distribution of pseudo-random numbers generated by the linear congruential method.* II, Math. Comp. **28** (1974), 1117–1132.

222. ______, *Quantitative versions of a result of Hecke in the theory of uniform distribution* mod 1, Acta Arith. **28** (1975), 321–339.

223. ______, *Résultats nouveaux dans la théorie quantitative de l'équirépartition,* Répartition Modulo 1 (Colloque de Marseille-Luminy, 1974), Lecture Notes in Math., vol. 475, Springer-Verlag, Berlin and New York, 1975, pp. 132–154.

224. ______, *Some new exponential sums with applications to pseudo-random numbers*, Topics in Number Theory (Debrecen, 1974), Colloq. Math. Soc. János Bolyai, vol. 13, North-Holland, Amsterdam, 1976, pp. 209–232.

225. ______, *On the cycle structure of linear recurring sequences*, Math. Scand. **38** (1976), 53–77.

226. ______, *On the distribution of pseudo-random numbers generated by the linear congruential method*. III, Math. Comp. **30** (1976), 571–597.

227. ______, *Statistical independence of linear congruential pseudo-random numbers*, Bull. Amer. Math. Soc. **82** (1976), 927–929.

228. ______, *Weights of cyclic codes*, Information and Control **34** (1977), 130–140.

229. ______, *Pseudo-random numbers and optimal coefficients*, Advances in Math. **26** (1977), 99–181.

230. ______, *The serial test for linear congruential pseudo-random numbers*, Bull. Amer. Math. Soc. **84** (1978), 273–274.

231. ______, *Existence of good lattice points in the sense of Hlawka*, Monatsh. Math. (to appear).

232. ______, *A quasi-Monte Carlo method for the approximate computation of the extreme values of a function*, Paul Turán Memorial Volume (to appear).

233. ______, *The serial test for pseudo-random numbers generated by the linear congruential method* (in preparation).

234. H. Niederreiter and W. Philipp, *Berry-Esseen bounds and a theorem of Erdös and Turán on uniform distribution* mod 1, Duke Math. J. **40** (1973), 633–649.

235. H. Niederreiter and J. M. Wills, *Diskrepanz und Distanz von Massen bezüglich konvexer und Jordanscher Mengen*, Math. Z. **144** (1975), 125–134; Berichtigung, ibid. **148** (1976), 99.

236. S. M. Nikol'skiĭ, *Quadrature formulae*, Fizmatgiz, Moscow, 1958 = Hindustan Publ. Corp., Delhi, 1964.

237. O. Onicescu, *Nombres et systèmes aléatoires*, Editions Eyrolles, Paris, 1964.

238. W. H. Payne, *Fortran Tausworthe pseudorandom number generator*, Comm. ACM **13** (1970), 57.

239. W. H. Payne, J. R. Rabung and T. P. Bogyo, *Coding the Lehmer pseudo-random number generator*, Comm. ACM **12** (1969), 85–86.

240. L. G. Peck, *On uniform distribution of algebraic numbers*, Proc. Amer. Math. Soc. **4** (1953), 440–443.

241. O. Perron, *Die Lehre von den Kettenbrüchen*, vol. 1, 3rd ed., Teubner, Stuttgart, 1954.

242. W. Philipp, *Das Gesetz vom iterierten Logarithmus mit Anwendungen auf die Zahlentheorie*, Math. Ann. **180** (1969), 75–94; Corrigendum, ibid. **190** (1971), 338.

243. ______, *Mixing sequences of random variables and probabilistic number theory*, Mem. Amer. Math. Soc. no. 114, Amer. Math. Soc., Providence, R.I., 1971.

244. ______, *Empirical distribution functions and uniform distribution* mod 1, Diophantine Approximation and Its Applications (C. F. Osgood, ed.), Academic Press, New York, 1973, pp. 211–234.

245. Ju. G. Polljak, *On the analysis of pseudorandom numbers*, Avt. i Vyčisl. Tehn. **1968**, no. 5, 31–35. (Russian)

246. K. Popper, *Logik der Forschung: Zur Erkenntnistheorie der modernen Naturwissenschaft*, Springer, Vienna, 1935.

247. A. G. Postnikov, *Arithmetic modeling of random processes*, Trudy Mat. Inst. Steklov. **57** (1960) = Selected Transl. Math. Statist. Probability, vol. 13, Amer. Math. Soc., Providence, R.I., 1973, pp. 41–122.

248. ______, *Ergodic problems in the theory of congruences and of diophantine approximations*, Trudy Mat. Inst. Steklov. **82** (1966) = Proc. Steklov Inst. Math., vol. 82, Amer. Math. Soc., Providence, R.I., 1967.

249. P. D. Proĭnov, *The square discrepancy of symmetric lattices*, Vestnik Moskov. Univ. Ser. I Mat. Meh. **30** (1975), no. 2, 41–47 = Moscow Univ. Math. Bull. **30** (1975), no. 1/2, 105–109.

250. C. M. Rader, L. R. Rabiner and R. W. Schafer, *A fast method of generating digital random numbers*, Bell System Tech. J. **49** (1970), 2303–2310.

251. RAND Corporation, *One million random digits and* 100,000 *normal deviates*, Free Press, Glencoe, Ill., 1955.

252. G. Rauzy, *Fonctions entières et répartition modulo un*. II, Bull. Soc. Math. France **101** (1973), 185–192.

253. H. Reichenbach, *Axiomatik der Wahrscheinlichkeitsrechnung*, Math. Z. **34** (1932), 568–619.

254. R. D. Richtmyer, *On the evaluation of definite integrals and a quasi-Monte Carlo method based on properties of algebraic numbers*, Report LA-1342, Los Alamos Sci. Lab., Los Alamos, N.M., 1951.

255. ______, *A non-random sampling method based on congruences for Monte-Carlo problems*, AEC Research and Development Rep. NYO-8674, AEC Comp. Appl. Math. Center, New York Univ., New York, 1958.

256. R. D. Richtmyer, M. Devaney and N. Metropolis, *Continued fraction expansions of algebraic numbers*, Numer. Math. **4** (1962), 68–84.

257. V. S. Rjaben'kiĭ, *Tables and interpolation of a certain class of functions*, Dokl. Akad. Nauk SSSR **131** (1960), 1025–1027 = Soviet Math. Dokl. **1** (1960), 382–384.

258. ______, *A way of obtaining difference schemes and the use of number-theoretic nets for the solution of the Cauchy problem by the method of finite differences*, Trudy Mat. Inst. Steklov. **60** (1961), 232–237. (Russian)

259. P. Roos and L. Arnold, *Numerische Experimente zur mehrdimensionalen Quadratur*, Österreich. Akad. Wiss. Math.-Nat. Kl. S.-B. II **172** (1963), 271–286.

260. M. Rosenblatt, *Multiply schemes and shuffling*, Math. Comp. **29** (1975), 929–934.

261. K. F. Roth, *On irregularities of distribution*, Mathematika **1** (1954), 73–79.

262. ______, *On irregularities of distribution*. II, Comm. Pure Appl. Math. **29** (1976), 749–754.

263. ______, *On irregularities of distribution*. III, Acta Arith. (to appear).

264. ______, *On irregularities of distribution*. IV (to appear).

265. Ju. N. Šahov, *Approximate solution of second kind Volterra equations by means of iterations*, Dokl. Akad. Nauk SSSR **128** (1959), 1136–1139. (Russian)

266. ______, *The approximate solution of Volterra equations of the second kind by the method of iterations*, Dokl. Akad. Nauk SSSR **136** (1961), 1302–1305 = Soviet Math. Dokl. **2** (1961), 206–209.

267. ______, *On calculating the eigenvalues of a multidimensional symmetric kernel using number-theoretic nets*, Ž. Vyčisl. Mat. i Mat. Fiz. **3** (1963), 988–997 = U.S.S.R. Computational Math. and Math. Phys. **3** (1963), 1350–1362.

268. ______, *On the approximate solution of higher-dimensional linear Volterra equations of second kind by an iteration method*, Ž. Vyčisl. Mat. i Mat. Fiz. **4** (1964), no. 4, suppl., 75–100. (Russian)

269. ______, *The calculation of integrals of increasing multiplicity*, Ž. Vyčisl. Mat. i Mat. Fiz. **5** (1965), 911–916 = U.S.S.R. Computational Math. and Math. Phys. **5** (1965), no. 5, 184–192.

270. ______, *On the error made in recovering functions of a certain class on parallelepiped-type grids*, Mat. Zametki **15** (1974), 749–756 = Math. Notes **15** (1974), 448–452.

271. M. Saint-André, *Calcul de la moyenne d'une fonction presque-périodique: application au calcul d'intégrales*, Rev. Française Informat. Recherche Opérationnelle **4** (1970), Sér. R-3, 141–146.

272. ______, *Détermination d'un vecteur optimal pour le calcul d'intégrales (simples ou multiples)*, Rev. Française Informat. Recherche Opérationnelle **5** (1971), Sér. R-2, 141–149.

273. R. Salfi, *A long-period random number generator with application to permutations*, Compstat 1974: Proceedings in Computational Statistics (G. Bruckmann, F. Ferschl and L. Schmetterer, eds.), Physica-Verlag, Vienna, 1974, pp. 28–35.

274. A. I. Saltykov, *Tables for computing multiple integrals by the method of optimal coefficients*, Ž. Vyčisl. Mat. i Mat. Fiz. **3** (1963), 181–186 = U.S.S.R. Computational Math. and Math. Phys. **3** (1963), 235–242.

275. I. F. Šarygin, *The use of number-theoretic methods of integration in the case of nonperiodic functions*, Dokl. Akad. Nauk SSSR **132** (1960), 71–74 = Soviet Math. Dokl. **1** (1960), 506–509.

276. ______, *A lower estimate for the error of quadrature formulas for certain classes of functions*, Ž. Vyčisl. Mat. i Mat. Fiz. **3** (1963), 370–376 = U.S.S.R. Computational Math. and Math. Phys. **3** (1963), 489–497.

277. M. Sato, *On the periods of certain pseudorandom sequences*, Publ. Res. Inst. Math. Sci. **10** (1974/75), 77–89.

278. K. Schmidt, *Über die C-Gleichverteilung von Massen*, Z. Wahrscheinlichkeitstheorie und verw. Gebiete **17** (1971), 327–332.

279. K. Schmidt and P. Zinterhof, *Über Quadraturformeln auf T^{ω}*, Computing **6** (1970), 94–96.

280. W. M. Schmidt, *Metrical theorems on fractional parts of sequences*, Trans. Amer. Math. Soc. **110** (1964), 493–518.

281. ———, *Simultaneous approximation to algebraic numbers by rationals*, Acta Math. **125** (1970), 189–201.

282. ———, *Irregularities of distribution*. VII, Acta Arith. **21** (1972), 45–50.

283. ———, *Lectures on irregularities of distribution*, Lecture notes, Boulder, Co., 1973.

284. ———, *Irregularities of distribution*. IX, Acta Arith. **27** (1975), 385–396.

284a. ———, *Irregularities of distribution*. X, Number Theory and Algebra (H. Zassenhaus, ed.), Academic Press, New York, 1977, pp. 311–329.

285. C. P. Schnorr, *Zufälligkeit und Wahrscheinlichkeit*, Lecture Notes in Math., vol. 218, Springer-Verlag, Berlin and New York, 1971.

286. C. S. Smith, *Multiplicative pseudo-random number generators with prime modulus*, J. Assoc. Comput. Mach. **18** (1971), 586–593.

287. S. A. Smoljak, *Interpolation and quadrature formulas for the classes W_s^α and E_s^α*, Dokl. Akad. Nauk SSSR **131** (1960), 1028–1031 = Soviet Math. Dokl. **1** (1960), 384–387.

288. ———, *Quadrature and interpolation formulae on tensor products of certain function classes*, Dokl. Akad. Nauk SSSR **148** (1963), 1042–1045 = Soviet Math. Dokl. **4** (1963), 240–243.

289. I. M. Sobol′, *Multidimensional integrals and the Monte Carlo method*, Dokl. Akad. Nauk SSSR **114** (1957), 706–709. (Russian)

290. ———, *Pseudo-random numbers for the machine "Strela"*, Teor. Verojatnost. i Primenen. **3** (1958), 205–211 = Theor. Probability Appl. **3** (1958), 192–197.

291. ———, *An accurate error estimate for multidimensional quadrature formulae for the functions of the class S_p*, Dokl. Akad. Nauk SSSR **132** (1960), 1041–1044 = Soviet Math. Dokl. **1** (1960), 726–729.

292. ———, *An exact estimate of the error in multidimensional quadrature formulae for functions of the classes $\tilde{W}_1$ and $\tilde{H}_1$*, Ž. Vyčisl. Mat. i Mat. Fiz. **1** (1961), 208–216 = U.S.S.R. Computational Math. and Math. Phys. **1** (1961), 228–240.

293. ———, *On the evaluation of infinite-dimensional integrals*, Ž. Vyčisl. Mat. i Mat. Fiz. **1** (1961), 917–922 = U.S.S.R. Computational Math. and Math. Phys. **1** (1961), 1086–1091.

294. ———, *On the evaluation of multidimensional integrals*, Dokl. Akad. Nauk SSSR **139** (1961), 821–823 = Soviet Math. Dokl. **2** (1961), 1022–1025.

295. ———, *The use of the ω^2-distribution for error estimation in the calculation of integrals by the Monte Carlo method*, Ž. Vyčisl. Mat. i Mat. Fiz. **2** (1962), 717–723 = U.S.S.R. Computational Math. and Math. Phys. **2** (1962), 808–816.

296. ———, *The application of Haar series in the theory of quadrature formulae*, Questions of Computational Mathematics and Computing Technology (L. A. Ljusternik, ed.), Gos Naučno-Tehn. Izdat. Mašinostr. Lit., Moscow, 1963, pp. 31–35. (Russian)

297. ———, *On the periods of pseudo-random sequences*, Teor. Verojatnost. i Primenen. **9** (1964), 367–373 = Theor. Probability Appl. **9** (1964), 333–338.

298. ———, *Distribution of points in a cube and integration nets*, Uspehi Mat. Nauk **21** (1966), no. 5, 271–272. (Russian)

299. ———, *An integral encountered in quadrature formulae theory*, Ž. Vyčisl. Mat. i Mat. Fiz. **6** (1966), 1084–1089 = U.S.S.R. Computational Math. and Math. Phys. **6** (1966), no. 6, 189–196.

300. ———, *The distribution of points in a cube and the approximate evaluation of integrals*, Ž. Vyčisl. Mat. i Mat. Fiz. **7** (1967), 784–802 = U.S.S.R. Computational Math. and Math. Phys. **7** (1967), no. 4, 86–112.

301. ———, *The use of Haar series in estimating the error in the computation of infinite-dimensional integrals*, Dokl. Akad. Nauk SSSR **175** (1967), 34–37 = Soviet Math. Dokl. **8** (1967), 810–813.

302. ———, *A Monte Carlo method for critical calculation in multigroup approximation*, The Monte Carlo Method in Problems of Radiation Transfer, Atomizdat, Moscow, 1967, pp. 232–254. (Russian)

303. ———, *Multidimensional quadrature formulas and Haar functions*, Izdat. "Nauka", Moscow, 1969. (Russian)

304. ———, *On an approach to the computation of multiple integrals*, Voprosy Vyčisl. i Prikl. Mat. (Taškent) **1970**, no. 38, 100–111. (Russian)

305. ———, *The problem of the minimum of φ_∞ in the three-dimensional cube*, Voprosy Vyčisl. i Prikl. Mat. (Taškent) **1970**, no. 38, 112–115. (Russian)

306. ______, *The Monte Carlo method*, Izdat. "Nauka", Moscow, 1972 = Mir Publishers, Moscow, 1975.

307. ______, *A deterministic interpretation of goodness-of-fit tests and a test of pseudo-random numbers*, Operations Research and Statistical Modeling, vol. 1, Izdat. Leningrad. Univ., Leningrad, 1972, pp. 162–169. (Russian)

308. ______, *A probabilistic estimate of the error for nonrandom integration nets*, Voprosy Vyčisl. i Prikl. Mat. (Taškent) **1972**, no. 14, 5–11. (Russian)

309. ______; *Numerical Monte Carlo methods*, Izdat. "Nauka", Moscow, 1973. (Russian)

310. ______, *A probabilistic estimate of the integration error for P_τ-nets*, Ž. Vyčisl. Mat. i Mat. Fiz. **13** (1973), 1035–1037 = U.S.S.R. Computational Math. and Math. Phys. **13** (1973), no. 4, 259–262.

311. ______, *Calculation of improper integrals using equidistributed sequences*, Dokl. Akad. Nauk SSSR **210** (1973), 278–281 = Soviet Math. Dokl. **14** (1973), 734–738.

312. ______, *Pseudo-random numbers for constructing discrete Markov chains by the Monte Carlo method*, Ž. Vyčisl. Mat. i Mat. Fiz. **14** (1974), 36–44 = U.S.S.R. Computational Math. and Math. Phys. **14** (1974), no. 1, 36–45.

313. ______, *Infinite-dimensional uniformly distributed sequences in numerical mathematics*, Preprint no. 22, Inst. Prikl. Mat. Akad. Nauk SSSR, Moscow, 1974. (Russian)

314. ______, *On convergence of infinite-dimensional cubature and simulation of Markov chains*, Voprosy Vyčisl. i Prikl. Mat. (Taškent) **1975**, no. 32, 162–167. (Russian)

315. ______, *Uniformly distributed sequences with an additional uniformity property*, Ž. Vyčisl. Mat. i Mat. Fiz. **16** (1976), 1332–1337 = U.S.S.R. Computational Math. and Math. Phys. **16** (1976), no. 5, 236–242.

316. I. M. Sobol′ and Ju. L. Levitan, *Generation of points uniformly distributed in a multidimensional cube*, Preprint no. 40, Inst. Prikl. Mat. Akad. Nauk SSSR, Moscow, 1976. (Russian)

317. I. M. Sobol′ and R. B. Statnikov, *LP-search and problems of optimal design*, Problems of Random Search, vol. 1, Izdat. "Zinatne", Riga, 1972, pp. 117–135. (Russian)

318. I. M. Sobol′, R. B. Statnikov and N. F. Ovčinnikova, *Localization of the characteristic roots of a matrix*, Ž. Vyčisl. Mat. i Mat. Fiz. **13** (1973), 1581–1583 = U.S.S.R. Computational Math. and Math. Phys. **13** (1973), no. 6, 255–258.

319. V. M. Solodov, *On the calculation of multiple integrals*, Dokl. Akad. Nauk SSSR **127** (1959), 753–756. (Russian)

320. ______, *On the error involved in a numerical integration*, Dokl. Akad. Nauk SSSR **148** (1963), 284–287 = Soviet Math. Dokl. **4** (1963), 85–88.

321. ______, *Integration over regions different from the unit cube*, Ž. Vyčisl. Mat. i Mat. Fiz. **8** (1968), 1334–1341 = U.S.S.R. Computational Math. and Math. Phys. **8** (1968), no. 6, 198–207.

322. ______, *An application of the method of optimal coefficients to numerical integration*, Ž. Vyčisl. Mat. i Mat. Fiz. **9** (1969), 14–29 = U.S.S.R. Computational Math. and Math. Phys. **9** (1969), no. 1, 14–34.

322a. E. R. Sowey, *A chronological and classified bibliography on random number generation and testing*, Internat. Statist. Rev. **40** (1972), 355–371.

323. J. Spanier and E. M. Gelbard, *Monte Carlo principles and neutron transport problems*, Addison-Wesley, Reading, Mass., 1969.

324. V. T. Stojancev, *Indeterminate methods of integration with a finite number of feasible methods*, Ž. Vyčisl. Mat. i Mat. Fiz. **9** (1969), 1235–1246 = U.S.S.R. Computational Math. and Math. Phys. **9** (1969), no. 6, 1–16.

325. ______, *Solution of the Cauchy problem for a parabolic equation by a quasi-Monte Carlo method*, Z. Vyčisl. Mat. i Mat. Fiz. **13** (1973), 1153–1160 = U.S.S.R. Computational Math. and Math. Phys. **13** (1973), no. 5, 67–75.

326. ______, *Solution of the Dirichlet problem by a quasi-Monte Carlo method*, Uspehi Mat. Nauk **30** (1975), no. 1, 263–264. (Russian)

327. R. G. Stoneham, *On a new class of multiplicative pseudo-random number generators*, Nordisk Tidskr. Informations-Behandling **10** (1970), 481–500.

328. ______, *On the uniform ε-distribution of residues within the periods of rational fractions with applications to normal numbers*, Acta Arith. **22** (1973), 371–389.

329. A. H. Stroud, *Approximate calculation of multiple integrals*, Prentice-Hall, Englewood Cliffs, N.J., 1971.

330. W. Stute, *Convergence rates for the isotrope discrepancy*, Ann. Probability (to appear).

331. P. Szüsz, *On a problem in the theory of uniform distribution*, Compt. Rend. Premier Congrès Hongrois, Budapest, 1952, pp. 461–472. (Hungarian)

332. O. Taussky and J. Todd, *Generation and testing of pseudo-random numbers*, Symposium on Monte Carlo Methods (H. A. Meyer, ed.), Wiley, New York, 1956, pp. 15–28.

333. R. C. Tausworthe, *Random numbers generated by linear recurrence modulo two*, Math. Comp. **19** (1965), 201–209.

334. D. Teichroew, *A history of distribution sampling prior to the era of the computer and its relevance to simulation*, J. Amer. Statist. Assoc. **60** (1965), 27–49.

335. L. H. C. Tippett, *Random sampling numbers*, Tracts for Computers, no. 15, Cambridge Univ. Press, London, 1927.

336. K. D. Tocher, *The application of automatic computers to sampling experiments*, J. Roy. Statist. Soc. Ser. B **16** (1954), 39–61.

337. J. P. R. Tootill, W. D. Robinson and A. G. Adams, *The runs up-and-down performance of Tausworthe pseudo-random number generators*, J. Assoc. Comput. Mach. **18** (1971), 381–399.

338. J. P. R. Tootill, W. D. Robinson and D. J. Eagle, *An asymptotically random Tausworthe sequence*, J. Assoc. Comput. Mach. **20** (1973), 469–481.

339. T. Tsuda, *Numerical integration of functions of very many variables*, Numer. Math. **20** (1973), 377–391.

340. S. M. Ulam, *Monte Carlo calculations in problems of mathematical physics*, E. F. Beckenbach (ed.), Modern mathematics for the engineer, 2nd series, McGraw-Hill, New York, 1961, Chapter 11.

341. J. G. van der Corput, *Verteilungsfunktionen*. I, II, Nederl. Akad. Wetensch. Proc. **38** (1935), 813–821, 1058–1066.

342. A. van Gelder, *Some new results in pseudo-random number generation*, J. Assoc. Comput. Mach. **14** (1967), 785–792.

343. A. van Wijngaarden, *Mathematics and computing*, Proc. Sympos. Automatic Digital Computation (London, 1954), H. M. Stationery Office, London, 1954, pp. 125–129.

344. J. Venn, *The logic of chance*, Macmillan, London, 1876.

345. P. H. Verdier, *Relations within sequences of congruential pseudo-random numbers*, J. Res. Nat. Bur. Standards Sect. B **73** (1969), 41–44.

346. I. V. Vilenkin, *Plane nets of integration*, Ž. Vyčisl. Mat. i Mat. Fiz. **7** (1967), 189–196 = U.S.S.R. Computational Math. and Math. Phys. **7** (1967), no. 1, 258–267.

347. ———, *More on plane nets of integration*, Ž. Vyčisl. Mat. i Mat. Fiz. **13** (1973), 854–864 = U.S.S.R. Computational Math. and Math. Phys. **13** (1973), no. 4, 43–56.

348. J. Ville, *Étude critique de la notion de collectif*, Gauthier-Villars, Paris, 1939.

349. S. von Hoerner, *Herstellung von Zufallszahlen auf Rechenautomaten*, Z. Angew. Math. Physik **8** (1957), 26–52.

350. R. von Mises, *Grundlagen der Wahrscheinlichkeitsrechnung*, Math. Z. **5** (1919), 52–99.

351. ———, *Wahrscheinlichkeit, Statistik und Wahrheit*, Springer, Vienna, 1928.

352. J. von Neumann, *Various techniques used in connection with random digits*, NBS Appl. Math. Series, no. 12, U. S. Government Printing Office, Washington, D. C., 1951, pp. 36–38 = Collected Works, vol. 5, Pergamon Press, Oxford, 1963, pp. 768–770.

353. Ju. V. Voroncov and Ju. G. Polljak, *On the use of quasirandom sequences in the direct probabilistic simulation of systems*, Avt. i Vyčisl. Tehn. **1971**, no. 6, 23–27. (Russian)

354. J. E. Walsh, *An experimental method for obtaining random digits and permutations*, Sankhyā **17** (1957), 355–360.

355. Y. Wang, *A note on interpolation of a certain class of functions*, Sci. Sinica **10** (1961), 632–636.

356. ———, *On numerical integration and its applications* (*Number-theoretic method*), Shuxue Jinzhan **5** (1962), no. 1, 1–44. (Chinese)

357. T. T. Warnock, *Computational investigations of low-discrepancy point sets*, Applications of Number Theory to Numerical Analysis (S. K. Zaremba, ed.), Academic Press, New York, 1972, pp. 319–343.

358. Y. Watanabe, *An improvement for the Richtmyer-Haselgrove method*, Sci. Rep. Osaka **22** (1973), no. 1, 33–44.

359. W. J. Westlake, *A uniform random number generator based on the combination of two congruential generators*, J. Assoc. Comput. Mach. **14** (1967), 337–340.

360. B. E. White, *Mean-square discrepancies of the Hammersley and Zaremba sequences for arbitrary radix*, Monatsh. Math. **80** (1975), 219–229.

361. ______, *On optimal extreme-discrepancy point sets in the square*, Numer. Math. **27** (1977), 157–164.

362. J. R. B. Whittlesey, *A comparison of the correlational behavior of random number generators for the* IBM 360, Comm. ACM **11** (1968), 641–644.

363. ______, *On the multidimensional uniformity of pseudorandom generators*, Comm. ACM **12** (1969), 247.

364. A. C. Yao and D. E. Knuth, *Analysis of the subtractive algorithm for greatest common divisors*, Proc. Nat. Acad. Sci. U.S.A. **72** (1975), 4720–4722.

365. S. K. Zaremba, *Good lattice points, discrepancy, and numerical integration*, Ann. Mat. Pura Appl. **73** (1966), 293–317.

366. ______, *Some applications of multidimensional integration by parts*, Ann. Polon. Math. **21** (1968), 85–96.

367. ______, *Good lattice points in the sense of Hlawka and Monte Carlo integration*, Monatsh. Math. **72** (1968), 264–269.

368. ______, *The mathematical basis of Monte Carlo and quasi-Monte Carlo methods*, SIAM Rev. **10** (1968), 303–314.

369. ______, *A quasi-Monte Carlo method for computing double and other multiple integrals*, Aequationes Math. **4** (1970), 11–22.

370. ______, *La discrépance isotrope et l'intégration numérique*, Ann. Mat. Pura Appl. **87** (1970), 125–136.

371. ______, *A remarkable lattice generated by Fibonacci numbers*, Fibonacci Quart. **8** (1970), 185–198.

372. ______, *Sur la discrépance des suites aléatoires*, Z. Wahrscheinlichkeitstheorie und verw. Gebiete **20** (1971), 236–248.

373. ______, *La méthode des "bons treillis" pour le calcul des intégrales multiples*, Applications of Number Theory to Numerical Analysis (S. K. Zaremba, ed.), Academic Press, New York, 1972, pp. 39–119.

374. ______, *Good lattice points modulo primes and composite numbers*, Diophantine Approximation and Its Applications (C. F. Osgood, ed.), Academic Press, New York, 1973, pp. 327–356.

375. ______, *Good lattice points modulo composite numbers*, Monatsh. Math. **78** (1974), 446–460.

376. ______, *Computing the isotropic discrepancy of point sets in two dimensions*, Discrete Math. **11** (1975), 79–92.

377. ______, *L'erreur dans le calcul des intégrales doubles par la méthode des bons treillis*, Demonstratio Math. **8** (1975), 347–364.

378. ______, *On Cartesian products of good lattices*, Math. Comp. **30** (1976), 546–552.

379. N. Zierler, *Linear recurring sequences*, J. Soc. Industr. Appl. Math. **7** (1959), 31–48.

380. Ja. M. Žileĭkin, *Approximate solution of the Dirichlet problem for the Laplace equation*, Dokl. Akad. Nauk SSSR **155** (1964), 999–1002 = Soviet Math. Dokl. **5** (1964), 528–531.

381. ______, *On the approximate solution of integral equations*, Ž. Vyčisl. Mat. i Mat. Fiz. **4** (1964), 749–753 = U.S.S.R. Computational Math. and Math. Phys. **4** (1964), no. 4, 176–181.

382. ______, *An approximate method of solving the Dirichlet problem for the Laplace equation in a rectangular parallelepiped*, Ž. Vyčisl. Mat. i Mat. Fiz. **5** (1965), 345–347 = U.S.S.R. Computational Math. and Math. Phys. **5** (1965), no. 2, 246–249.

383. ______, *Quadrature formulae on classes of functions*, Ž. Vyčisl. Mat. i Mat. Fiz. **8** (1968), 507–516 = U.S.S.R. Computational Math. and Math. Phys. **8** (1968), no. 3, 1–14.

384. P. Zinterhof, *Einige zahlentheoretische Methoden zur numerischen Quadratur und Interpolation*, Österreich. Akad. Wiss. Math.-Nat. Kl. S.-B. II **177** (1969), 51–77.

Department of Mathematics, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801

Current address: Chair in Pure Mathematics, University of the West Indies, Kingston 7, Jamaica