EDMUND HLAWKA

Discrepancy and uniform distribution of sequences

Compositio Mathematica, tome 16 (1964), p. 83-91
<http://www.numdam.org/item?id=CM_1964__16__83_0>

Discrepancy and uniform distribution of sequences*

by

Edmund Hlawka

First I want to repeat some definitions: If we have a point x in R^s, we call his coordinates $p_j(x)$ $(j = 1, \ldots s)$. Two points x, x' are congruent modulo 1, if $x - x'$ has integral coordinates. Now if we are only interested in properties modulo 1, we can always suppose that x is in the unit cube $E = E^s : 0 \leqq p_j < 1$ $(j = 1, \ldots s)$ in other words we consider only the fractional part $\{x\}$ of x. Now we consider first infinite sequence $\omega : x_1, x_2, \ldots$ of points or more general $\omega : x_{11}; x_{21}, x_{22}; \ldots; x_N, \ldots x_{NN}; \ldots$ all lying in E^s. Now we consider an intervall $Q : \alpha_j \leqq p_j < \beta_j$ $(j = 1, \ldots s)$ in E^s. Let $\chi_Q(x)$ be the characteristic function of Q. Then $N(Q, \omega) = \sum_{k=1}^N \chi_Q(x_k)$ is the number of points x_n of the given sequence with $n \leqq N$ lying in Q. If we have a double sequence then we define $N(Q, \omega) = \sum_{k=1}^N \chi_Q(x_{Nk})$. The sequence ω is called uniformly distributed modulo 1 if for each interval Q of E^s, $\lim_{N\to\infty} N(Q, \omega)/N$ exists and is equal to the volume of $Q : \mu(Q) = \int_E \chi_Q(x)dx$. This means that [1]

$$(1) \qquad \lim_{N\to\infty} \lambda_N(Q, \omega) = \lim_{N\to\infty} \frac{\sum_{k=1}^{N} \chi_Q(x_k)}{N} = \int_E \chi_Q(x)dx.$$

Sometimes it is better to write it in another form. Let Q be an arbitrary interval $\alpha_j \leqq p_j < \beta_j$ in R^s, not necessarily lying in the unit cube but with $\beta_j - \alpha_j < 1$ $(j = 1, \ldots s)$ and counting now the number $F(Q, \omega, N) = \sum_{k=1}^N \chi'_Q(x_k)$, where χ'_Q is the function χ_Q periodically extended by translating O to all lattice points t, we see immediately that 1) $F(Q, \omega) = \sum_{k=1}^N \chi_Q(x_k)$ if Q is in the unit cube 2) F is periodic with period 1 and the sequence ω is uniformly distributed if and only if

$$(1') \qquad \lim_{N\to\infty} \frac{F(Q, \omega, N)}{N} = \mu(Q)$$

for each interval Q.

*) Nijenrode lecture.

In this definition it is not necessary to assume that the x_n are in the unit cube. Now it follows immediately from the definition of the Riemann integral for any function $f(x)$ (real or complex) defined in the unit cube or what is the same for any periodic function with period 1 and Riemann integrable.

$$(2) \qquad \lim_{N\to\infty} \lambda_N(f, \omega) = \lim_{N\to\infty} \frac{1}{N} \sum_{k=1}^{N} f(x_k) = \int_E f(x)dx = \mu(f)$$

and conversely. If we have any set M which has a content in the sense of Jordan we have setting $f(x) = \chi_M(x)$

$$(3) \qquad \lim_{N\to\infty} \frac{N(M, \omega)}{N} = \mu(M)$$

where $N(M, \omega)$ is the number of points $x_1, \ldots x_N$ of the sequence in M. All the formulas (1), (2) und (3) are equivalent. If we take [1] $f(x) = e^{2\pi i\langle hx\rangle}$, where h is an integral vector and $S_N(h, \omega) = 1/N \sum_{k=1}^{N} e^{2\pi i\langle hx_k\rangle}$ then we have

$$(4) \qquad \lim_{N\to\infty} S_N(h, \omega) = 0$$

(Criterion of Weyl), if only if ω is uniformly distributed 1.

It was I. Schoenberg, who generalized formula (4) in the one-dimensional case. He considered an arbitrary continuous distribution function $z(x)$ and defined: a sequence ω is asymptotically distributed modulo 1, with the distribution function $z(x)$, if $\lim \lambda_N(Q, \omega) = \int_E \chi_Q(x)dz(x)$ for any Q in E^s. If $z(x)$ has a density $\rho(x)$ we have $dz(x) = \rho(x)dx$. This definition was generalized to more than one dimension. It is now not necessary to speak more about this, because other members of this symposium will speak about these things. Now we will go back to (1), that means $\lim_N(\lambda_N(Q, \omega)-\mu(Q)) = 0$. Now we define: $N(\lambda_N(Q, \omega)-\mu(Q))$ is called the remainder term and we can ask, if we can say more about the sequence ω than $\lambda_N(Q)-\mu(Q) = o(1)$. Now we define the discrepancy: In the classical sense it is $D_N(\omega) = \sup_Q |\lambda_N(Q)-\mu(Q)|$ about all intervals Q in E^s. It has been shown by H. Weyl that for each uniformly distributed sequence ω, $\lim_N D_N(\omega) = 0$, that means that (1) is uniform in Q. It is obvious that $D_N \leqq 1$. Trivial is also that $ND_N \geqq 1$ and this was very improved by the important work of van Aardenne-Ehrenfest

[1] $\langle hx\rangle = \sum_{j=1}^{s} p_j(h)p_j(x)$.

and Roth; Cassels will speak about this. We have many results [2] in the one-dimensional case for special sequences so for the sequence $\{\alpha n\}$, α an irrational number. The results are depending on the finer arithmetical properties of α. If for instance there is a number γ such that $n^\gamma\{n\alpha\} \geqq C$ for all n, then $ND_N = O((\log N))$ if $\gamma = 1$, $O(N^{1-(1/\gamma)+\varepsilon})$ if $\gamma > 1$ ($\varepsilon > 0$); on the other hand $ND_N = \Omega(\log N)$ for all α. Well known is the work of Vinogradov [3] about polynomials $f(n)$. In the linear case there are elementary methods using continued fractions, but the easiest way is to work with Weyl sums [4]. The form of this connection is the formula of Erdös-Turan-Koksma: for any positive integer[2] M

$$D_N \leqq 30^s \left(\frac{1}{M} + \sum_{0<||h||\leqq M} R^{-1}(h)|S_N(\omega, h)|\right)$$

$$||h|| = \text{Max}\,(|h_1|, \dots |h_s|),$$

$$R(h) = \prod_{j=1}^{s} \text{Max}(|h_j|, 1).$$

Now I want to say some remarks on the definition of D. Some times is used another definition [5]: $D' = \sup_Q |(F(Q, N, \omega)/N) - \mu(Q)|$. It is easy to see that $D \leqq D' \leqq 2^s D$. One can also define the discrepancy D^* with the intervals $Q : 0 \leqq p_i < \beta_i < 1$. It is clear that $D^* \leqq D \leqq 2^s D^*$. If we have an asymptotic distribution ρ then we define

$$D_N(\rho; \omega) = \sup_Q \left|\lambda_N(Q) - \int_E \rho(x)\chi_Q(x)dx\right|. \tag{5}$$

If we know something about D, can we say something about the difference

$$\lambda_N(f) - \mu(f)$$

where f is a Riemann integrable function? For the one-dimensional case Koksma 1942 [6] has shown the following result: if f is of bounded variation, $V(f)$ the total variation then

$$|\lambda_N(f) - \mu(f)| \leqq D_N(\omega)V(f) \tag{6}$$

generalizing results of Behnke and his own Satz 1 chap. 9 § 98. I have generalized [7] it to more dimensions. The variation $V(f)$ is to take in the sense of Hardy and Krause. For $f = P_r(x)$ the Bernouilli polynomial of degree r you get estimates for the sums

[2] $p_j(h) = h_j$.

$$(8) \qquad N(\lambda_N(P_r, \omega)-\mu(P_r)) = \sum_{k=1}^{N} P_r(x_k) = R_r^*$$

the most studied case [8] is $R_1 = R_1^* = \sum_{k=1}^{N} (\{x_k\}-\frac{1}{2})$.

I want to make another application in the one-dimensional case: consider the function $f(x) = \chi_\gamma(1/x-[1/x])$ where $\chi_\gamma(x)$ is the characteristic function of the interval $0 \leqq x < \gamma$. This function is not of bounded variation, the integral $\int_E f(x)dx = \int_0^\gamma \tau(x)dx$ with $\tau(x) = \sum_{n=1}^{\infty} 1/(n+x)^2$.

We consider therefore for a natural number l $f_l(x) = f(x)$ for $1/l \leqq x < 1$, $f_l(x) = 0$ for $0 \leqq x < 1/l$. The variation of $f_l(x)$ is $2l+1$, therefore we have

$$\begin{aligned}
|\lambda_N(f)-\mu(f)| &= \left|\frac{1}{N}\sum_{k=1}^{N}\chi_\gamma\left(\frac{1}{x_k}-\left[\frac{1}{x_k}\right]\right) - \int_0^1 \chi_\gamma\tau(x)dx\right| \\
&= \left|\frac{1}{N}\sum_{x_k<1/l}\chi_\gamma\left(\frac{1}{x_k}-\left[\frac{1}{x_k}\right]\right) - \int_0^{1/l} + \frac{1}{N}\sum_{x_k>1/l}\chi_\gamma - \int_{1/l}^1\right| \\
&\leqq |\lambda_N(f_l)-\mu(f_l)| + \left|\frac{1}{N}\sum_{x_k<1/l} 1\right| + \frac{1}{l} \\
&\leqq |\lambda_N(f_l)-\mu(f_l)| + \left|\frac{1}{N}\sum_{x_k\leqq 1/l} - \int_0^{1/l} dx\right| + \frac{2}{l} \\
&\leqq (2l+2)D_N(\omega) + \frac{2}{l}\cdot
\end{aligned}$$

We take $l \sim D_N^{-\frac{1}{2}}$ and we get for each interval Q

$$\left|\frac{1}{N}\sum_{k=1}^{N}\chi_Q\left(\frac{1}{x_k}-\left[\frac{1}{x_k}\right]\right) - \int_0^1 \chi_Q(x)\tau(x)dx\right| \leqq 6\sqrt{D}$$

and therefore we have for the discrepancy D^+ of the sequence ω^+: $(1/x_k-[1/x_k])$

$$(7) \qquad D_N(\tau;\, \omega^+) \leqq 6\sqrt{D_N}(\omega).$$

If ω is uniformly distributed we get the result of I. Schoenberg [9], that ω^+ is asymptotically distributed with the distribution density $\tau(x)$. I have generalized this to more than one-dimension. If $x = (p_1, \dots p_s)$ is a point in R_s then I define $x^+ : p_j(x^+) = 1/p_j(x)-[1/p_j(x)]$ $(j = 1, \dots s)$. I consider with the sequence ω: $(x_1, x_2, \dots)$ the sequence $\omega^+ : (x_1^+, x_2^+, \dots)$ then I have [10]

$$(7') \qquad D_N(\rho;\, \omega^+) \leqq 6^s D_N^{1/(s+1)}(\omega)$$

$(\rho(x) = \tau(p_1) \dots \tau(p_s))$. Another application of (6) in the more dimensional case is the following: if x is a point with coordinates $p_1 \dots p_s$, then let be $z^1 \leqq z^2 \dots \leqq z^s$ these numbers in ascending order and we set $f^{(j)}(x) = z^{(i)}$, for instance $f^{(1)}(x) = \text{Min}\,(p_1, \dots p_s)$. If we have a sequence $\omega = (x_k)$ then we consider the one dimensional sequences $\omega^{(j)} = (f^{(j)}(x_k), (j = 1, \dots s)$ and the s dimensional sequences $\tilde{\omega} = (f^{(1)}(x_k), \dots f^{(s)}(x_k))$; it is easy to show and well known [11] in the theory of order statistics that if the sequence ω is uniformly distributed, then the $\omega^{(j)}$ are asymptotically distributed with a density $\rho_i(x) = s\binom{s-1}{i-1} x^{i-1}(1-x)^{s-i}$ and $\tilde{\omega}$ with the density $\tilde{\rho} = s!$ in $p_1 \leqq p_2 < \dots \leqq p_s < 1$, 0 otherwise in E^s. We can easily show with the help of (6) that

$$(8) \qquad D_N(\rho_j;\, \omega^{(j)}) \leqq C_s D_N(\omega),\ D_N(\tilde{\rho};\, \tilde{\omega}) \leqq 6^s D_N(\omega)$$

(C_s absolute constant).

To show it for $j = 1$ take in (6), where $\chi_\gamma\,(x)$ is defined as above,

$$f(x) = \sum_{k=1}^{s} \sum_{(i_1, \dots, i_k)} \prod_{j=1}^{k} \chi_\gamma(p_{i_j}) \prod_{i \neq i_j} (1 - \chi_\gamma(p_i)) = \chi_\gamma(f^{(1)}(x)).$$

A weaker result than (6) we get [12] for all Riemann integrable functions f. If Z is a partition of the unit cube E in intervals, with Norm $l(Z)$, $\sigma(f, Z)$ the mean oscillation of f for the partition Z, we have

$$(9) \qquad |\lambda_N(f) - \mu(f)| \leqq 6^s \textstyle\sum (f, D^{1/s})$$

where $\sum (f, k) = \sup_{l(Z) \leqq k} \sigma(f, Z)$.

If we have a set M in E with the property that any line parallel to the coordinate axes intersects M in a set of points which, if not empty, consists of at most h intervals and the same is true for any of the m dimensional regions ($m \leqq s$) obtained by projecting M on one of the coordinate spaces defined by equating a selection of $s-m$ of the coordinates to zero then we get [13].

$$(10) \qquad |\lambda_N(M, \omega) - \mu(M)| \leqq 20^s D_N^{1/s}(\omega)$$

This is true if M is convex; therefore if we consider a set $\{K\}$ of convex bodies in E we get

$$(11) \qquad D_N(\{K\}, \omega) = \sup_{\{K\}} |\lambda_N(K, \omega) - \mu(K)| \leqq 20^s D_N^{1/s}(\omega)$$

We can say $D_N(\{K\}, \omega)$ is the discrepancy of ω for this set $\{K\}$. Bergström [14] 1935 has considered such a discrepancy but I do not know, for what sets of bodies. Now I want to estimate the discrepancy in the classical sense from above with the help of D_N

$(\{K\}, \omega))$. This seems to be a difficult problem. Take for example the set $\{S\}$ of all spheres. Then I can only show [15]

$$D_N(\omega) \leqq C(S)/\log D_N^{-1}(\{S\}, \omega)) \tag{12}$$

where the constant $C(S)$ is only depending on s. The difficulty stems from the fact that the covering and packing with spheres has a density different of one.

With the sequence $\omega = (x_k)$ consider the sequences $\omega^{(h)} = (x_{k+h} - x_k)$ $(h = 1, 2, \ldots)$. The fundamental theorem of van der Corput says that ω is uniformly distributed, if for each h, $\omega^{(h)}$ is uniformly distributed. Is there a connection between the discrepancies $D_N(\omega)$ and $D_N(\omega^{(h)})$? In the one-dimensional case this question was first considered by Vinogradov and then by van der Corput and Pisot [16] 1939, the final result was given by Cassels 1953 [17]. I generalized [18] the result to more dimension and we have for any natural number q with $1 \leqq q \leqq s$

$$D_N(\omega) \leqq C(s)\sqrt{B}\left(1+\log^s \frac{1}{B}\right) \tag{13}$$

where

$$B = \left(\frac{1}{q} + \frac{1}{s}\right)\left(1+2\sum_{h=1}^{q-1} D_{N-h}(\omega^{(h)})\right) \tag{14}$$

If all $\omega(h)$ are uniformly distributed, then $D_{N-h}(\omega^{(h)}) \to 0$ for $N \to \infty$ therefore

$$\overline{\lim_N}\, D_N(\omega) \ll \frac{1}{\sqrt{q}}\left(l+\log^s \frac{1}{q}\right)$$

and therefore $\lim D_N(\omega) = 0$, if we go with $q \to \infty$.

An important result of H. Weyl is the following: Consider the one dimensional sequences $\omega(h) = (\langle x_k h\rangle)$, h integral vector, then the sequence ω is uniformly distributed if and only if the sequences $\omega(h)$ are for each h uniformly distributed. Now what is the connection between the discrepancies $D_N(\omega(h))$ and $D_N(\omega)$? This question was first considered by W. J. Coles [19], a pupil of Cassels, for the two dimensional case. I proved the following general result: For any natural M, we have

$$D_N(\omega) < C(s)\left(\frac{1}{M} + \sum_{0<\|h\|\leqq M} R^{-1}(h) D_N(\omega^{(h)})\right). \tag{15}$$

We can suppose that the h are primitive, that means that the coordinates of h are relative prime. Let us give an example. We have for the discrepancy of $\omega : \{\alpha k^g\}$ after Vinogradov (α irrational, g natural number $\geqq 1$): For any τ with $0 < \tau < 1$

$$D_N(\omega) = 0(N^{\varepsilon}(N^{-\tau} + m^{-1}(\alpha))^{2/\gamma} \qquad (\gamma = 2^g) \tag{16}$$

where $m(\alpha) = \text{Min } q$, with $\{\alpha q\} < N^{-g+\tau}$.

Now we consider the s-dimensional sequences $\omega : (\alpha_1 k^g, \alpha_2 k^g, \ldots, \alpha_s k^g)$ where the $\alpha_1, \ldots \alpha_s$ are linear independent, then we get for the discrepancy of ω with the help of the above formula and the result of Vinogradov

$$D_N(\omega) = 0\left(N^{\varepsilon}\left(N^{-\tau} + \sum_{l=1}^{s} \sum_{(i_1, \ldots i_l)} m^{1/l}(\alpha_{i_1}, \ldots \alpha_{i_l})\right)\right)^{2/\gamma} \tag{17}$$

where $m(\alpha_{i_1} \ldots \alpha_{i_l}) = \min (q_1 \ldots q_l)$, so that

$$\{\alpha_{i_1} q_1 + \ldots + \alpha_{i_l} q_l\} < N^{-g+\tau}. \tag{18}$$

For $s = 2$ this is an old result of Kusmin and Skopin [20] 1934, who work directly with Weyl-sums. Let us make some remarks about uniform distribution in a compact space. The difficulty is the definition of discrepancy. I would propose to define the discrepancy in the following way:

$$D_N = \sup_K \left| \frac{1}{N} \sum_{k=1}^{N} \chi_K(x_k) - \mu(K) \right|$$

about all convex bodies K. This definition depends on the definition of a convex set in a compact space. This is easy, if in this space exists a Riemann metric. A special case namely the spheres $S^m : p_1^2 + \ldots + p_m^2 = \langle p, p \rangle = 1$ and for the special sequence $\{g/\sqrt{n}\}$, where g are lattice points, with $\langle g, g \rangle = n$ (n natural numbers) is already known. For $m \geqq 4$, this sequence is uniformly distributed on S^m, with

$$\mu(K) = \frac{\text{surface of } K}{\text{surface of } S}$$

(K a Jordan measurable set on S^m), and we have for the discrepancy in the above sense

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

($N = r(n)$ the number of solutions of $\langle g, g \rangle = n$). This was shown by Malyshew [21] but natural expressed in other terms.

NOTES

[1] Compare for these definitions the report of J. Cigler and G. Helmberg "Neuere Entwicklungen der Theorie der Gleichverteilung", Jahresberichte der DMV 64(1961) S. 1–50. (short C & H).
There is a continuous analogue studied especially by L. Kuipers and B. Meulenbeld, the so called C-uniform distribution: Let $x(t)$ be a real-valued measurable vector function in s dimension, defined on the interval $0 \leqq t < \infty$, then $x(t)$ mod 1 is C uniformly distributed, if for any Q in E^s:

$$\lim_{T\to\infty} \frac{1}{T}\int_0^T \chi_Q(\{x(t)\})dt = \mu(Q)$$

(compare § 11 of the article cited above).

[2] See § 6 of C. & H. and J. F. Koksma Diophantische Approximationen Ergebn. Math. u. Grenzgeb. Bd. 4, Heft 4, Berlin 1936.
If a is a vector in E^s and if there exist an η and a $C > 0$, that $R(h)^\eta \{\langle ha\rangle\} \geqq C$ for all lattice points $h \neq 0$, then $D_N = O(N^{-1/\eta} \log^s N)$. If $a = (\alpha_1, \dots \alpha_s)$, where $1, \alpha_1 \dots \alpha_s$ linear independent algebraic numbers in a field K of degree $s+1$, then $\eta = s$. It is unknown if there exist a with $\eta = 1$ (Littlewood problem). For almost all a $\eta = 1+\varepsilon(\varepsilon > 0)$. This result can be improved (W. Schmidt).

[3] I. M. Vinogradow,
The method of trigonometric sums in the theory of numbers.
See also the report of Loo-Keng Hua in Enzykl. math. Wiss. 12, Heft 13, I 1950.

[4] Easier are to find metrical theorems:
We have sequences (x_n) depending on a parameter θ and it is asked for an estimate for the Discrepancy D for all θ, except a set of measure zero. There are important results of Koksma, Cassels, Erdös compare C & H § 9. There exist metrical Theorems where instead the Lebesguemeasure is considered Hausdorffmeasure or the Capacity of the set (R. Salem). In the proof of the theorems is of great use a theorem of J. F. Koksma and Gál.
For the C-uniform distribution metrical theorems seems not to be published. In this case it seems interesting to study vector functions $x(t)$ depending not only on a parameter a but on a continuous function and to consider the Discrepancy for all θ except a set Wienermeasure zero. Some results in this direction have been found by the author (unpublished).

[5] Compare J. W. S. Cassels, An introduction to Diophantine Approximation, Cambridge Univ. Press. 1957 and the paper of J. G. van der Corput, et Ch. Pisot, Proc. Kon. Ned. Akad. v. Wet. 42, 476–486 (1939) and for the continuous case E. Hlawka, Ann. Math. pure and appl. IV. Ser. 49, 311–326 (1960).

[6] Mathematica, Zutphen B 11, 7–11 (1942).

[7] Ann. Mat. pure appl. IV, Ser. 54, 325–333 (1961).

[8] Compare the report of Koksma 2.

[9] Mathem. Z. 28, 171–199 (1928).

[10] Monatshefte für Math. 65 (1961).

[11] Compare for instance L. Schmetterer.
Einführung in die Mathematische Statistik, Springer-Verlag, Wien (1956) Kap. 7. For (8) see E. Hlawka, Math. Ann. 150, 259–267 (1963).

[12] A proof was given at the Jahresversammlung der DMV, Bonn 1962.
If we define $\Delta_i f = f(\ldots, p_i + h_i, \ldots) - f(\ldots, p_i, \ldots)$ (other coordinates fixed) and if we suppose that there exists a constant C and α with $0 < \alpha < 1$ such that $|\Delta_{i_1, \ldots i_r} f| \leqq (\alpha + 1/2)^r C |h_{i_1} \ldots h_{i_r}|^a$ (the so called class H) for all $1 \leqq r \leqq s, i_1 < i_2 < \ldots < i_r$ then $|\lambda_N(f) - \mu(f)| \leqq 6^s C D_N^\alpha(\omega)$. J. F. Koksma told me at this symposium that he has for $s = 1$ and for continuous functions some results of the form (9).
[13] Compare [7].
[14] Medd. mat. Sem. Univ. Lund 2 (1935) 97 S.
Kunigl. Fysiografiska Sätslkapets Lund För handlinger 6 (1936) S. 113–131.
[15] Using methods of the geometry of numbers.
[16] Compare the paper of these authors cited in [5].
[17] Math. Ann. 126, 108–118 (1953).
[18] Math. Zeitschrift 77, 273–84 (1961).
[19] Proc. Cambridge 53, 781–89 (1957).
[20] H 3 B A H CCCP (7) 1934 S. 547–60.
[21] Compare for instance the literature in the paper of Pommerenke in Acta Arithmetica V and A. V. Malysev Dokladi Akad. CCCR 133 (1960) 1294–97.

(Oblatum 29-5-63).