Uniform Distribution Theory **5** (2010), no.1, 65–78

uniform distribution theory

VARIATIONS ON THE KOKSMA-HLAWKA INEQUALITY

Glyn Harman

ABSTRACT. The Koksma-Hlawka inequality gives a bound for the error in approximating the multidimensional integral of a function by the average of the function over a discrete set of points. The error is essentially a product of a measure of the function's variation and the discrepancy of the set of points. For the standard form of the inequality this discrepancy is with respect to boxes aligned to the axes and the measure of variation is then dependent on this choice of axes. This rules out various natural choices of function (or even integration domains) and we here explore a new choice for the measure of variation that enables a wider class of functions to be investigated, expanding comments we stated without proof in the author monograph *Metric Number Theory*, Oxford, 1998, p. 161-2.

Communicated by Oto Strauch

Dedicated to the memory of Professor Edmund Hlawka

1. Introduction

Write $U = [0,1)$ and let f be a function defined on U with bounded variation $V = V(f)$. Koksma's inequality [16] (see also [20, Chapter 2, Theorem 5.1]) states that, for any sequence $x_n \in U$, we have

$$\left| \frac{1}{N} \sum_{n=1}^{N} f(x_n) - \int_0^1 f(x)\, dx \right| \leq \frac{V}{N} D_N^*(x_n), \tag{1}$$

where

$$D_N^*(x_n) = \sup_{x \in U} \left| \sum_{\substack{0 \leq x_n < x \\ n \leq N}} 1 - xN \right|.$$

2010 Mathematics Subject Classification: 11K38, 11K45, 65D30.
Keywords: Numerical integration, discrepancy, convex set, variation.

Hlawka [14] (or see [20, Chapter 2, Theorem 5.5]) obtained a generalization of this result to higher dimensions which requires f to have bounded variation in the sense of Hardy [11] and Krause [17]. This means that f has bounded variation in the sense of Vitali [26], and so does its restriction to each lower dimensional face of U^k. See [20, p.147] for the details. It is the purpose of this note to try to circumvent certain of the difficulties associated with this measure of a function's variation. We shall be particularly concerned with integrating a "nicely behaved" function over a convex subset of U^k. This corresponds to integrating a function (the product of our well-behaved function with the characteristic function of a convex body) which does *not* have the requisite bounded variation. This theme has been anticipated by Zaremba [29], but we shall take an alternative approach and finish with rather different results (although we shall also obtain some of the earlier results as well with new proofs). We note that producing variations on the Koksma-Hlawka inequality is still an active area of research (see, for example [22, 9]), while the inequalities themselves are finding a wide range of applications in very different spheres.

To give some indication of the problems in this area, consider in one dimension the characteristic function of an interval. Then $V(f) = 2$. Taking the supremum over all intervals in U the inequality (1) then reduces to

$$D_N(x_n) \leq 2D_N^*(x_n),$$

where

$$D_N(x_n) = \sup_{0 \leq x \leq y < 1} \left| \sum_{\substack{x \leq x_n < y \\ n \leq N}} 1 - (x-y)N \right|.$$

Of course, this result can be proved directly very quickly, but we note that it is in fact sharp, and we also have $D_N^*(x_n) \leq D_N(x_n)$. If we are working in k dimensions, where $k \geq 2$, then the characteristic function of a convex body in U^k only has bounded variation in the sense of Hardy and Krause if it is a box aligned with the coordinate axes. Thus, for example, the characteristic functions of a ball or a tilted box do not have bounded variation in this sense. Since the Koksma-Hlawka inequality uses discrepancy with respect to boxes aligned with the axes, the fact that such characteristic functions are ruled out is quite natural. To elucidate this phenomenon and to make further progress we must introduce more notation.

Let $\mathbf{x_n}$ be a sequence in U^k and suppose that $f(\mathbf{x})$ is a bounded measurable function on U^k. Given a class $\mathcal{S}$ of measurable subsets (to avoid problems of overlap we assume with diameter less than 1) of U^k we write

$$D_N(\mathcal{S}, \mathbf{x}_n) = \sup_{\mathcal{D}\in\mathcal{S}} \left| \sum_{\substack{n\le N \\ \mathbf{x}_n\in\mathcal{D}}} 1 - N\lambda(\mathcal{D}) \right|,$$

where $\lambda(\cdot)$ denotes k-dimensional Lebesgue measure. Write $\mathcal{B}$ for the class of boxes in U^k with sides parallel to the co-ordinate axes, $\mathcal{T}$ for the class of all boxes with no such restriction, and $\mathcal{C}$ for the class of convex sets. In these definitions everything is to be considered $(\operatorname{mod} \mathbb{Z}^k)$ so that U^k is a torus, and a subset of U^k is convex if and only if there exists $\mathbf{y}$ such that the set $\{\mathbf{x}+\mathbf{y} : \mathbf{x}\in\mathcal{D}\}$, where all the co-ordinates are understood to be taken $(\operatorname{mod} 1)$, is convex in the conventional sense. We use the same convention for boxes.

With the above notation we clearly have

$$D_N(\mathcal{B}, \mathbf{x}_n) \le D_N(\mathcal{T}, \mathbf{x}_n) \le D_N(\mathcal{C}, \mathbf{x}_n).$$

By known results in the theory of irregularities of distribution we are aware [10] that there exist sequences with N terms $(\mathbf{x}_n)$ such that

$$D_N(\mathcal{B}, \mathbf{x}_n) \ll (\log N)^{k-1}, \tag{2}$$

yet, for any sequence $(\mathbf{x}_n)$, we have (see [3])

$$D_N(\mathcal{T}, \mathbf{x}_n) \gg N^{\frac{1}{2}(1-1/k)}, \tag{3}$$

and (see [24])

$$D_N(\mathcal{C}, \mathbf{x}_n) \gg N^{1-2/(k+1)}. \tag{4}$$

For (2) the choice of sequence $(\mathbf{x}_n)$ depends on N since otherwise the best result that can be obtained is $D_N(\mathcal{B}, \mathbf{x}_n) \ll (\log N)^k$. Both (3) and (4) are sharp except possibly for a logarithmic factor as shown in [2, 4, 25] (see also [6]). Thus no result of the form

$$\left| \frac{1}{N}\sum_{n=1}^{N} f(\mathbf{x}_n) - \int_{U^k} f(\mathbf{x})\,d\mathbf{x} \right| \le \frac{V(f)}{N} D_N(\mathcal{B}, \mathbf{x}_n) \tag{5}$$

could hold if $V(f)$ remained finite for the characteristic functions of tilted rectangles, or, more generally, if f has some discontinuity not along a line parallel to the axes. It is this problem we address by replacing $\mathcal{B}$ in (5) with alternative classes. From one perspective the main idea of this paper is to replace the underlying idea of Riemann integration in the Koksma-Hlawka inequality with Lebesgue integration. From another it is this change of class of sets. We should remark that the Koksma-Hlawka inequality is in fact more complicated than (5) as it involves not just one term on the right hand side, but the sum of several products of variations of restrictions of the function multiplied by the corresponding discrepancies of the sequence [20, p. 151 (5.18)]. The Koksma-Hlawka inequality gives a better estimate than our results in the case of bounded Hardy-Krause variation and $N \to \infty$, whereas our results may be superior if the Hardy-Krause variation is unbounded or, if N is fixed and the Vitali variations are not small relative to N.

To take an example, suppose we wished to integrate the function

$$f(\mathbf{x}) = \frac{1}{x_1 \ldots x_k(1 - x_1 - \ldots - x_k)} \tag{6}$$

over a convex set (or, more generally, a set which is the set difference of two unions of convex sets) contained in the subset of U^k with

$$x_1 \geq x_2 \geq \ldots \geq x_k \geq x_0, 1 - x_1 - \ldots - x_k \geq x_0, \tag{7}$$

where $(k+1)^{-1} \geq x_0 > 0$. This problem arises frequently in various sieve problems: see, for example, Chapter 3 onwards in [13]. The function which arises is often a little more complicated, involving Buchstab's function $\omega(u)$, but in many cases it reduces to the problem we have stated. What is the optimal strategy for obtaining a numerical approximation to the integral? The standard form of the Koksma-Hlawka inequality is of no use here. An apparently naive approach is to treat the integral as k one-dimensional integrals, so that if we have N points in U^k we have $N^{1/k}$ points in U at our disposal for each integral. This immediately greatly reduces the quality of approximation we could expect, but at least we have removed the constraints of the Hardy-Krause variation. Of course, if the convex sets are polytopes (as usually happens for this example), then standard mathematical software packages will evaluate these integrals in terms of logarithms and polylogarithms and then give a numerical approximation to them. We shall come back to this problem after proving our main results.

First we require some further notation. Given f, a bounded measurable function on U^k, we write

$$\begin{aligned} \mathcal{E}(f,\alpha) &= \big\{\mathbf{x} \in U^k : \ f(\mathbf{x}) \geq \alpha\big\}, \\ \ell(f) &= \inf_{\mathbf{x}\in U^k} f(\mathbf{x}), \\ u(f) &= \sup_{\mathbf{x}\in U^k} f(\mathbf{x}), \\ S(f) &= \big\{\mathcal{E}(f,\alpha) : \ \ell(f) \leq \alpha \leq u(f)\big\}. \end{aligned}$$

Since the set $\{\mathbf{x} : f(\mathbf{x}) = \alpha\}$ is commonly called a level set we shall designate $\mathcal{E}(f,\alpha)$ a *super-level set.* Also, we write $K(f,\alpha)$ for the smallest number of convex bodies (if such a number exists) $\mathcal{F}_1, \ldots, \mathcal{F}_r$ such that $\mathcal{E}(f,\alpha)$ can be expressed in the form

$$\left(\bigcup_{j=1}^{m-1} \mathcal{F}_j\right) \Big\backslash \left(\bigcup_{j=m}^{r} \mathcal{F}_j\right),$$

for some m. Here we use the convention that points are counted according to multiplicity, so that if $\mathbf{x} \in \mathcal{E}(f,\alpha)$ then $\mathbf{x}$ belongs to h sets $\mathcal{F}_j, j < m$ and to $h-1$ sets $\mathcal{F}_j$ with $j \geq m$. Thus

$$\lambda(\mathcal{E}(f,\alpha)) = \sum_{j=1}^{m-1} \lambda(\mathcal{F}_j) - \sum_{j=m}^{r} \lambda(\mathcal{F}_j).$$

If $\mathcal{E}(f,\alpha)$ cannot be so expressed then $K(f,\alpha)$ is undefined. If $K(f,\alpha)$ is defined for all α then we write

$$V^*(f) = \int_{\ell(f)}^{u(f)} K(f,\alpha)\, d\alpha, \tag{8}$$

provided that the right hand side of (8) exists as a *Riemann* integral. We shall explain later that the existence of $V^*(f)$ forces f to be Riemann integrable on U^k. This could be proved directly, but it is easier to see this as a consequence of Theorem 1.

We note that in one dimension for a function of bounded variation and at most a finite number of discontinuities we have

$$V^*(f) = \tfrac{1}{2}V(f).$$

For continuous functions this result is due to Banach[1] and can also be found in [21, Chapter 8.5 Theorem 3]. In these works the expression $N(y)$ occurs (called the *Banach indicatrix*) which is the number of solutions to $f(x) = y$. So, for continuous functions, this corresponds to $2K(f,y)$ in our notation. The passage to functions with a finite number of discontinuities is straightforward by taking a sequence f_n of continuous functions converging pointwise to f.

This new measure $V^*(f)$ (which we briefly introduced in [12, p. 162]) of the variation of a function has a number of advantages. For example, it does not require f to be very regular in behaviour, and is independent of rotations and translations of the variables. As such it matches the *isotropic* properties of $D_N(\mathcal{C}, \mathbf{x}_n)$ (see [29]). We note that other authors have produced different generalizations of Banach's formula in higher dimensions, for example using the number of simply connected components of the super-level sets (see [19, 18, 27]). Of course, the class of simply connected sets will be of no use to us since the discrepancy with respect to this class is always N.

We observe that the measure $V^*(f)$ naturally assigns the value n to the sum of the characteristic functions of n convex bodies. Of course, just by taking a continuous approximation to the characteristic function of a convex body we can illustrate the contrast in behaviour with the Vitali variation of a function. For example, let $f_n(\mathbf{x})$ be defined by:

$$f_n(\mathbf{x}) = \begin{cases} 1 & \text{if } |\mathbf{x}| \le \frac{1}{4}, \\ 4n - 4(4n-1)|\mathbf{x}| & \text{if } \frac{1}{4} < |\mathbf{x}| < (4 - 1/n)^{-1}, \\ 0 & \text{if } \mathbf{x} \ge (4 - 1/n)^{-1}. \end{cases}$$

Then, for each n, $f_n(\mathbf{x})$ is a continuous function taking values in $[0,1]$ for all $\mathbf{x}$ yet its variation in the sense of Vitali tends to infinity with n. On the other hand, $V^*(f_n) = 1$ for all n.

Theorem 1. *Let $f(\mathbf{x})$, $\mathbf{x} \in U^k$ be a bounded measurable function, and suppose that $(\mathbf{x}_n)$ is a sequence in U^k. Write*

$$E(f, N, \mathbf{x}_n) = \left| \frac{1}{N} \sum_{n=1}^{N} f(\mathbf{x}_n) - \int_{U^k} f(\mathbf{x})\, d\mathbf{x} \right|.$$

Then

$$E(f, N, \mathbf{x}_n) \le \frac{u(f) - \ell(f)}{N} D_N\big(\mathcal{S}(f), \mathbf{x}_n\big). \tag{9}$$

Moreover, if $V^*(f)$ exists, then

$$E(f, N, \mathbf{x}_n) \le \frac{1}{N} V^*(f) D_N(\mathcal{C}, \mathbf{x}_n), \tag{10}$$

and

$$E(f, N, \mathbf{x}_n) \le C_k \frac{1}{N} V^*(f) N^{1-1/k} \big(D_N(\mathcal{B}, \mathbf{x}_n)\big)^{1/k}. \tag{11}$$

Here the constant C_k in (11) *can be taken as* $(4k\sqrt{k} + 1)$.

Remark.

Remark. By results of de Bruijn and Post [8] and Binder [7] (see also [20, p. 6]), for every non-Riemann integrable function there exists a uniformly distributed sequence such that $E(f, N, \mathbf{x}_n) \not\to 0$. Hence (11) cannot be true unless f is Riemann-integrable. This justifies our earlier comment that the existence of $V^*(f)$ implies f is Riemann-integrable.

The main new result here is (10). However, it should be noted that (9), although unlikely to be useful in applications because of the need to find discrepancy with respect to the super-level sets of f, is applicable to functions for which no measure of variation is possible. As an extreme example, consider in one dimension the characteristic function of the set of irrationals: so it does not have bounded variation! In this case $\mathcal{S}(f) = \{U, U \setminus \mathbb{Q}\}$. Let $x_n = n\sqrt{2}$. Then $D_N(\mathcal{S}(f), x_n) = 0$ and, of course, the left hand side of (9), which is precisely this discrepancy (times $1/N$), is indeed 0.

To compare our result with Hlawka's second generalization of the Koksma inequality [15] we need yet more notation. Let $\mathcal{P}$ be a partition of U^k and write $s(f; \mathcal{P})$ for the difference between the upper and lower Darboux sums for f corresponding to $\mathcal{P}$. Let $\nu(\mathcal{P})$ be the maximum length of all the edges of the boxes in $\mathcal{P}$. For a positive real number α set

$$S(f; \alpha) = \sup_{\nu(\mathcal{P}) \leq \alpha} s(f; \mathcal{P}).$$

Then Hlawka's result shows that

$$E(f, N, \mathbf{x}_n) \ll S(f; \delta),$$

where

$$\delta = \left(N^{-1} D_N(\mathcal{B}, \mathbf{x}_n)\right)^{1/k}.$$

This quickly leads to the result

$$D_N(\mathcal{C}, \mathbf{x}_n) \ll N^{1-1/k} \left(D_N(\mathcal{B}, \mathbf{x}_n)\right)^{1/k} \tag{12}$$

which shows that (11) follows from (10). The inequality (12) appears in different notation as (11) in [15] or [20, Chapter 2, Theorem 1.6]. We have taken the constant in (11) from the latter result. Our result (10) is a different quite general $k-$dimensional form of Koksma's inequality. The results (11) and (12) may appear to be rather weak, but Zaremba [28], who coined the term *isotropic discrepancy* for $D_N(\mathcal{C}, \mathbf{x}_n)$, showed that (12) is best possible. His example was for sequences $n\mathbf{x} \pmod 1$ for certain rational $\mathbf{x} \in \mathbb{R}^k$. We gave our own proof of this for the sequence $n\mathbf{x} \pmod 1$ for *any* $\mathbf{x} \in \mathbb{R}^k$ in [12, Theorem 5.24A]. To be precise we showed the following.

Theorem 2. *For any* $\mathbf{x} \in \mathbb{R}^k$ *write* $\mathbf{x}_n = n\mathbf{x} \pmod 1$. *Then we have*

$$D_N(\mathcal{T}, \mathbf{x}_n) \gg_k N^{1-1/k}. \tag{13}$$

We note that, for almost all $\mathbf{x} \in \mathbb{R}^k$ we have

$$D_N(\mathcal{B}, \mathbf{x}_n) \ll (\log N)^{k+\epsilon} \tag{14}$$

for any $\epsilon > 0$ as was shown by Beck [5]. In [12, Theorem 5.25] it was shown that adapting Schmidt's method [23] leads to

$$D_N(\mathcal{T}, \mathbf{x}_n) \ll N^{1-1/k} (\log N)^{1+1/k+\epsilon}$$

for almost all $\mathbf{x}$. In fact, the author should have observed that just using (12) with (14) leads to the slightly better and more general bound:

$$D_N(\mathcal{C}, \mathbf{x}_n) \ll N^{1-1/k}(\log N)^{1+\epsilon}.$$

In either case we see that (13) is sharp except possibly for a logarithmic factor.

2. Proof of Theorem 1

Let M be a (large) positive integer and write $\delta = (u(f) - \ell(f))/M$. Put

$$g(h) = (h+1)\delta + \ell, \qquad \mathcal{K}(h) = \mathcal{E}\big(f, g(h-1)\big) \setminus \mathcal{E}\big(f, g(h)\big),$$

$$\mu(h) = \lambda\big(\mathcal{K}(h)\big) = \lambda\big(\mathcal{E}\big(f, g(h-1)\big)\big) - \lambda\big(\mathcal{E}\big(f, g(h)\big)\big),$$

and

$$E_N(\mathcal{D}) = \frac{1}{N} \sum_{\substack{n \le N \\ \mathbf{x}_n \in \mathcal{D}}} 1 - \lambda(\mathcal{D}),$$

for any measurable subset $\mathcal{D}$ of U^k. Then

$$\begin{aligned} \frac{1}{N} \sum_{n=1}^{N} f(\mathbf{x}_n) &\le \frac{1}{N} \sum_{h=0}^{M} g(h) \sum_{\mathbf{x}_n \in \mathcal{K}(h)} 1 \\ &= \sum_{h=0}^{M} g(h)\mu(h) + \sum_{h=0}^{M} g(h) E_N\big(\mathcal{K}(h)\big) \\ &= S_1 + S_2, \text{ say.} \end{aligned}$$

Since f is a bounded measurable function we have

$$S_1 \longrightarrow \int_{U^k} f(\mathbf{x})\, d\mathbf{x} \quad \text{as} \quad M \longrightarrow \infty.$$

Also, by partial summation,

$$\begin{aligned} S_2 &= \sum_{h=0}^{M} \big(g(h) - g(h+1)\big) \sum_{r=0}^{h} E_N\big(\mathcal{K}(r)\big) \\ &\qquad + g(M+1) \sum_{r=0}^{M} E_N\big(\mathcal{K}(r)\big) \\ &= \sum_{h=0}^{M} \delta E_N\big(\mathcal{E}(f, g(h))\big) \end{aligned}$$

since

$$E_N\big(\mathcal{K}(h)\big) = E_N\big(\mathcal{E}\big(f, g(h-1)\big)\big) - E_N\big(\mathcal{E}\big(f, g(h)\big)\big),$$

$$E_N\big(\mathcal{E}(f,\ell)\big) = E_N\big(U^k\big) = 0, \quad \text{and} \quad E_N\big(\mathcal{E}(f, u+\delta)\big) = E_N(\emptyset) = 0,$$

so, in particular,

$$\sum_{r=0}^{h} E_N\big(\mathcal{K}(r)\big) = -E_N\big(\mathcal{E}(f, g(h))\big).$$

Hence

$$|S_2| \le M\delta \sup_{\ell \le \alpha \le u} \big|E_N\big(\mathcal{E}(f,\alpha)\big)\big| = \frac{u(f) - \ell(f)}{N} D_N\big(\mathcal{S}(f), \mathbf{x}_n\big).$$

Since one can obtain a lower bound similarly for

$$\frac{1}{N}\sum_{n=1}^{N} f(\mathbf{x}_n) \tag{15}$$

with the same bound for the error, this establishes (9).

To obtain (10) we note that we have shown that

$$|S_2| \le \sum_{h=0}^{M-1} \delta \left|E_N\big(\mathcal{E}\big(f, g(h)\big)\big)\right|$$

and this last sum can be expressed as

$$\sum_{h=0}^{M-1} \delta \left| \sum_{r=1}^{K(g(h))} \sigma(r) E_N\big(\mathcal{F}\big(r, g(h)\big)\big) \right|,$$

where $\mathcal{F}(r, g(h))$ are convex sets, and $\sigma(r) = \pm 1$, arising from the decomposition of $\mathcal{E}(f, g(h))$ into convex sets. Hence

$$|S_2| \le \sup_{\mathcal{D} \in \mathcal{C}} E_N(\mathcal{D}) \sum_{h=0}^{M-1} \delta K\big(g(h)\big).$$

Since $K(\alpha)$ is Riemann integrable, we have

$$\sum_{h=0}^{M-1} \delta K\big(g(h)\big) \longrightarrow \int_{\ell(f)}^{u(f)} K(\alpha)\, d\alpha \quad \text{as} \quad M \longrightarrow \infty.$$

As in the proof of (9) we can obtain the same estimate for the error when producing a lower bound for (15), and so this completes the proof of (10).

3. An example

We first state an elementary lemma that relates $V^*(\chi f)$ to $V^*(f)$. Here χf means *product* and not composition, of course.

Lemma 1. *If $\chi(\mathbf{x})$ is the characteristic function of a convex set $\mathfrak{C}$ in U^k, then*

$$V^*(\chi f) \le \max\bigl(0, \ell(f), -u(f)\bigr) + V^*(f).$$

P r o o f. We give the proof only for the case $\ell(f) \ge 0$, the other cases ($u(f) \le 0$, $\ell(f) < 0 < u(f)$) follow similarly. Clearly (since the intersection of two convex sets is a convex set)

$$K(\chi f, \alpha) \le K(f, \alpha) \quad \text{if} \ \ \alpha \ge \ell(f).$$

On the other hand,

$$K(\chi f, \alpha) = 1 \qquad\qquad \text{if} \ \ 0 \le \alpha < \ell(f).$$

Since

$$\ell(\chi f) = 0,$$
$$u(\chi f) \le u(f),$$

we have

$$V^*(\chi f) \le \int_0^{\ell(f)} 1\, d\alpha + \int_{\ell(f)}^{u(f)} K(f, \alpha)\, d\alpha = \ell(f) + V^*(f).$$

This completes the proof. □

We now consider the function (6) integrated over a convex subset of U^k, where (7) holds. However, similar results can be obtained for a wide range of well-behaved functions integrated over a convex set. In our situation we can suppose that $f(\mathbf{x}) = 0$ outside the region given by (7). It follows from the above that

$$V^*(f) = F(k, x_0),$$

where $F(k, x_0)$ is a function that can be bounded in terms of k and x_0. To be precise, we have the following elementary result.

Lemma 2. *In the above terminology we have*

$$F(k, x_0) \le \frac{2}{x_0^k(1 - kx_0)}.$$

P r o o f. It is clear that our function attains its maximum value at the point with all co-ordinates equal to x_0. It then only remains to note that the complement of any super-level set for (6) is convex. To prove this it suffices to show that

$$\begin{gathered} x_1 \dots x_k(1 - x_1 - \dots - x_k) > \beta, \\ y_1 \dots y_k(1 - y_1 - \dots - y_k) > \beta \\ \Longrightarrow \\ z_1 \dots z_k(1 - z_1 - \dots - z_k) > \beta, \end{gathered}$$

where $z_j = (x_j + y_j)/2, j = 1, \dots, k$. This follows simply from the Arithmetic--Mean/Geometric-Mean inequality which gives

$$z_j \geq \sqrt{x_j y_j}, \quad (1 - z_1 - \dots - z_k) \geq \sqrt{(1 - x_1 - \dots x_k)(1 - y_1 - \dots - y_k)}.$$

□

By [4, 25], given N, there exists a set of points $\mathbf{x}_1, \dots, \mathbf{x}_N$ with

$$D_N(\mathcal{C}, \mathbf{x}_n) \ll N^{1-2/(k+1)}(\log N)^{g_k},$$

where

$$g_k = \begin{cases} 4 & \text{if } k = 2, \\ \frac{3}{2} & \text{if } k = 3, \\ 2/(k+1) & \text{if } k \geq 4. \end{cases}$$

Combining this with (10) then gives

$$\left| \int_{\mathfrak{C}} f(\mathbf{x})\, d\mathbf{x} - \frac{1}{N} \sum_{n=1}^{n} f(\mathbf{x}_n) \right| \ll F(k, x_0) N^{-2/(k+1)} (\log N)^{g_k}. \tag{16}$$

This should be compared with the bound obtained by repeated application of Koksma's inequality which leads - in the best case of equally spaced points - to

$$\left| \int_{\mathfrak{C}} f(\mathbf{x})\, d\mathbf{x} - \frac{1}{N} \sum_{n=1}^{n} f(\mathbf{x}_n) \right| \ll V^{\dagger} N^{-1/k}. \tag{17}$$

Here $V^{\dagger}$ can be bounded explicitly in terms of k and x_0. Of course, as we have already indicated, we have no hope of applying the standard Koksma-Hlawka inequality here, and (16) will be best possible for some $\mathfrak{C}$ apart possibly from the logarithmic factor in its N-dependence.

We can now see the advantages and disadvantages of our approach. We have nearly doubled the exponent in N for the saving from $1/k$ to $2/(k+1)$. We have lost only a small logarithmic factor in doing this. However, our measure of variation here - $F(k, x_0)$ - may not be easy to calculate in general, and we have no algorithm for determining the set of points to use (since the arguments in [4, 25] are probabilistic). Nevertheless, this method gives better theoretical bounds for large N for the approximate evaluation of an important class of integrals. Should an explicit construction be discovered for finding good sets of points relative to the isotropic discrepancy, this could lead to a feasible numerical method.

4. Further Questions

The first question we raise is the relation between the Vitali variation $V(f)$ and $V^*(f)$. Is there a general inequality which gives an upper bound for $V^*(f)$ in terms of $V(f)$? Or can one find such a relation for a restricted family of functions f? We have shown that there can be no relation in the other direction. At first sight it appears that what would make $V^*(f)$ large - the super-level sets being far from convex - would make the Vitali variation large too.

The second question is whether one can make any connection between $V^*(f)$ and the Fourier coefficients of f. Again, what would make $V^*(f)$ large should make the Fourier coefficients behave badly, and good behaviour of the Fourier coefficients should make $V^*(f)$ small.

The final question is whether the new Koksma-Hlawka inequality we have derived (10) has any practical significance or whether it is just of a theoretical interest.

Acknowledgement. The author would like to thank the referee for pointing out the implications of papers [7, 8] and other useful remarks.

REFERENCES

[1] BANACH, S.: *Sur les lignes rectifiables et les surfaces dont l'aire est finie*, Fund. Math. **7** (1925), 225–236.

[2] BECK, J.: *Some upper bounds in the theory of irregularities of distribution*, Acta Arith. **43** (1984), no. 2, 115–130.

[3] BECK, J.: *Irregularities of distribution I*, Acta Math. **159** (1987), no. 1–2, 1–49.

[4] BECK, J.: *On the discrepancy of convex plane sets*, Monatsh. Math. **105** (1988), no. 2, 91–106.

[5] BECK, J.: *Probabilistic diophantine approximation I, Kronecker sequences*, Ann of Math. **140** (1994), no. 1, 109–160.

[6] BECK, J. – CHEN, W.W.L: *Irregularities of Distribution*, Cambridge Tracts in Math. Vol. **89**. Cambridge University Press, Cambridge, 1987.

[7] BINDER, C.: *Über einen Satz von de Bruijn und Post*, Sitzgsber. Österr. Akad. Wiss. Math.-naturw. Kl., Abt. II **179** (1970), 233–251.

[8] DE BRUIJN, N.D. – POST, K.A.: *A remark on uniformly distributed sequences and Riemann integrability*, Indag. Math. **30** (1968), 149–150.

[9] DICK, J.: *Koksma-Hlawka inequalities of fractional order*, Ann. Mat. Pura Appl. (4) **187** (2008), no. 3, 385-403.

[10] HALTON, J.H.: *On the efficiency of certain quasi-random sequence of points in evaluating multi-dimensional integrals*, Numer. Math. **2** (1960), 84–90.

[11] HARDY, G.H.: *On double Fourier series, and especially those which represent the double zeta-function with real and incommensurable parameters*, Quart. J. Mech. Appl. Math. Quart. J. Math. **37** (1906), 53–79.

[12] HARMAN, G.: *Metric Number Theory*, London Math. Soc. Monographs New Series Vol. **18**, The Clarendon Press, Oxford University Press, New York, 1998.

[13] HARMAN, G.: *Prime-detecting Sieves*, London Math. Soc. Monographs New Series **33**, Priceton University Press, Princenton, NJ, 2007.

[14] HLAWKA, E.: *Funktionen von beschränker Variation in der Theorie der Gleichverteillung*, Ann. Mat. Pura Appl. **54** (1961), 325-333.

[15] HLAWKA, E.: *Discrepancy and Riemann integration*, in: *Studies in Pure Mathematics.* Academic Press, London, 1971, pp. 121–129.

[16] KOKSMA, J.F.: *Een algemeene stelling inuit de theorie der gelijkmatige verdeeling modulo 1*, Mathematica (Zutphen B) **11** (1942-43), 7–11.

[17] KRAUSE, J.M.: *Fouriersche Reihen mit zwei veranderlichen Grössen*, Ber. Verh. Sachs Akad. Wiss. Leipzig, Math.-naturw. Kl. **55** (1903), 164–197.

[18] KRONROD, A.S.: *On functions of two variables*, Uspekhi Mat. Nauk **5** (1950), 24–134. (In Russian)

[19] MULHOLLAND, H.P.: *The total variation of a function of two variables*, Proc. London Math. Soc. (2) **46** (1940). 290–311. Correction: Proc. London Math. Soc. (2) **50** (1949), 559–560.

[20] KUIPERS, L. – NIEDERREITER, H.: *Uniform Distribution of Sequences*, *Pure and Applied Mathematics.* Wiley-Interscience [John Wiley & Sons], New York, 1974. Reprinted: Dover Publications, 2006.

[21] NATANSON, I.P.: *Theory of Functions of a Real Variable* (translated by L.F. Boronwith wisth the collaboration of Edwin Hewitt), Frederick Ungar Publishing Co., New York, 1955.

[22] 22 NIEDERREITER, H.: *Error bounds for quasi-Monte Carlo integration with uniform point sets*, J. Comput. Appl. Math. **150** (2003), no. 2, 283–292.

[23] 23 SCHMIDT, W.M.: *Metrical theorems on fractional parts of real sequences*, Trans. Amer. Math. Soc. **110** (1964), 493–518.

[24] SCHMIDT, W.M.: *Irregularities of distribution IX*, Acta Arith. **27** (1975), 385–396.

[25] STUTE, W.: *Convergence rates for the isotrope discrepancy*, Ann. Probab. **5** (1977), no. 5, 707–723.

[26] VITALI, G.: *Sui gruppi di punti e sulle funzioni di variabili reali*, Atti Acad. Sci. Torino **43** (1908), 75-92.

[27] VITUSHKIN, A.G.: *On Higher-dimensional Variations* (in Russian), Gosudarstv. Izdat. Tehn.-Teor. Lit., Moscow, 1955.
[28] ZAREMBA, S.K.: *Good lattice points in the sense of Hlawka and Monte Carlo integration*, Monatsch. Math. **72** (1968), 264–269.
[29] ZAREMBA, S.K.: *La discrépance isotrope et l'intégration numérique*, Ann. Math. Pura Appl. **87** (1970), 125–135.

Received September 22, 2009
Accepted February 23, 2010

Glyn Harman
Department of Mathematics
Royal Holloway University of London
Egham, Surrey TW20 0EX
UNITED KINGDOM
E-mail: G.Harman@rhul.ac.uk