Quantitative versions of a result of Hecke in the theory of uniform distribution mod 1

by

H. NIEDERREITER* (Princeton, N.J.)

1. Introduction. Let $a$ be an irrational number. Then the sequence $(na)$, $n = 0, 1, \ldots$, is uniformly distributed mod 1, and so we have

$$\lim_{N \to \infty} \frac{1}{N} \sum_{n=0}^{N-1} f(na) = \int_0^1 f(t) \, dt$$

for every Riemann-integrable function $f$ on $[0, 1)$, where $(x)$ denotes the fractional part of the real number $x$. Since the Abel summation method includes the summation method of arithmetic means, it follows that

$$\lim_{r \to 1-} \sum_{n=0}^{\infty} f(na)r^n = \int_0^1 f(t) \, dt.$$  

From this observation, Hecke [3] deduced easily that the power series $\sum_{n=0}^{\infty} (na)x^n$ cannot be continued analytically across the unit circle. More generally, one can show by Hecke's method that the power series $\sum_{n=0}^{\infty} g((na))x^n$ has the unit circle as its natural boundary whenever $g$ is a Riemann-integrable function for which all but finitely many of the integrals $\int_{\mathbb{Z}} g(t)z^{m+1} \, dt$, $m \in \mathbb{Z}$, are nonzero [see [6], Ch. 1, Theorem 2.4].

For other results on noncontinuable power series of the above type, see [6], Ch. 1, Sect. 2, and the survey article of Schwarz [17].

We remark that in the argument leading to (1), the sequence $(na)$ may, of course, be replaced by any sequence $(x_n)$, $n = 0, 1, \ldots$, of real numbers that is uniformly distributed mod 1. Evidently, an analogous

* This research was initiated while the author was a participant of the 1973 Summer Research Institute in Number Theory at the University of Michigan and was also supported by NSF Grant GP-38418X1.
Our first estimate is an analogue of the so-called Koksma–Hlawka inequality (see [6], Ch. 2, Theorem 3.5) and will be valid for any function \( f \) which is of bounded variation on \( I^r \) in the sense of Hardy and Krause. For the definition of this concept of variation, see [6], Ch. 2, Definition 5.2.

For technical reasons, we have to introduce one more class of discrepancies, which may be called truncated Abel discrepancies.

**Definition 2.** Let \( \omega = (\omega_n), n = 0, 1, \ldots \), be a sequence in \( R^r \). Then, for \( 0 < r < 1 \) and for a positive integer \( N \), we set

\[
D_{r,N}(\omega) = \sup_j \left( \frac{1}{1-\epsilon^N} \sum_{n=0}^{N-1} \alpha_j(\omega_n) \right) - \lambda(J),
\]

where the supremum is extended over all subintervals \( J \) of \( I^r \) of the form \( J = [0, \epsilon^0) \times \cdots \times [0, \epsilon^{r-1}) \), and where \( \alpha_j \) and \( \lambda(J) \) stand for the characteristic function and for the \( r \)-dimensional Lebesgue measure of \( J \), respectively.

The method of Hlawka depends on reducing questions concerning the rate of convergence in (2) or concerning the Abel discrepancy to corresponding questions in the quantitative theory of uniform distribution mod 1 with respect to the summation method of arithmetic means. However, this involves a certain loss of precision. In the present paper, we shall improve and complement several results of Hlawka by using a different (and more direct) method. The gist of our method is to estimate the rate of convergence in (2) in terms of the Abel discrepancy, and then to estimate the Abel discrepancy directly by using an analogue of the Erdős–Turán–Koksma inequality. In the last section, we prove some results on irregularities of distribution for the Abel discrepancy. It should be noted that the principal results of this paper can be extended to other types of summation methods. The author intends to treat this subject in detail on another occasion.

2. Integration errors. We show first how to estimate the rate of convergence in (2) in terms of the Abel discrepancy of the sequence involved. Let \( \omega = (\omega_n), n = 0, 1, \ldots \), be an arbitrary sequence in \( R^r \). Then, for a given Riemann-integrable function \( f \) on \( I^r \) and for \( 0 < r < 1 \), we introduce the "integration error"

\[
\delta_r(f, \omega) = \left( 1 - r \right) \sum_{n=0}^{\infty} f(\omega_n) - \int_0^1 \cdots \int_0^1 f(\omega^0) \, d\omega^0 \cdots d\omega^r,
\]

Our first estimate is an analogue of the so-called Koksma–Hlawka inequality (see [6], Ch. 2, Theorem 3.5) and will be valid for any function \( f \) which is of bounded variation on \( I^r \) in the sense of Hardy and Krause. For the definition of this concept of variation, see [6], Ch. 2, Definition 5.2.

For technical reasons, we have to introduce one more class of discrepancies, which may be called truncated Abel discrepancies.

**Definition 2.** Let \( \omega = (\omega_n), n = 0, 1, \ldots \), be a sequence in \( R^r \). Then, for \( 0 < r < 1 \) and for a positive integer \( N \), we set

\[
D_{r,N}(\omega) = \sup_j \left( \frac{1}{1-\epsilon^N} \sum_{n=0}^{N-1} \alpha_j(\omega_n) \right) - \lambda(J),
\]

where the supremum is extended over all subintervals \( J \) of \( I^r \) of the form \( J = [0, \epsilon^0) \times \cdots \times [0, \epsilon^{r-1}) \), and where \( \alpha_j \) and \( \lambda(J) \) stand for the characteristic function and for the \( r \)-dimensional Lebesgue measure of \( J \), respectively.

The method of Hlawka depends on reducing questions concerning the rate of convergence in (2) or concerning the Abel discrepancy to corresponding questions in the quantitative theory of uniform distribution mod 1 with respect to the summation method of arithmetic means. However, this involves a certain loss of precision. In the present paper, we shall improve and complement several results of Hlawka by using a different (and more direct) method. The gist of our method is to estimate the rate of convergence in (2) in terms of the Abel discrepancy, and then to estimate the Abel discrepancy directly by using an analogue of the Erdős–Turán–Koksma inequality. In the last section, we prove some results on irregularities of distribution for the Abel discrepancy. It should be noted that the principal results of this paper can be extended to other types of summation methods. The author intends to treat this subject in detail on another occasion.

2. Integration errors. We show first how to estimate the rate of convergence in (2) in terms of the Abel discrepancy of the sequence involved. Let \( \omega = (\omega_n), n = 0, 1, \ldots \), be an arbitrary sequence in \( R^r \). Then, for a given Riemann-integrable function \( f \) on \( I^r \) and for \( 0 < r < 1 \), we introduce the "integration error"
But
\[ 0 \leq \sum_{n=N}^{\infty} e_n(x_n)r^n \leq \sum_{n=-\infty}^{\infty} r^n = \frac{r^N}{1-r} \]
and
\[ 0 \leq \frac{r^N}{1-r} \sum_{n=0}^{N-1} e_n(x_n)r^n \leq \sum_{n=0}^{\infty} r^n = \frac{r^N}{1-r} \]
so that
\[ |g(x_0, \ldots, x_{n-1}; x_n) - g(x_0, \ldots, x_{n-1})| \leq r^N. \]
Together with (3), we arrive at the desired inequality.

In the following discussion, we use essentially the same notation as in [6], Ch. 2, Sect. 5, and [18]. In particular, the difference operators \( A_{1}, \ldots, p \) and \( \Delta_{p+1}, \ldots, p \) and the summation symbols \( \sum_{j_0}^{j_1} \) are defined in exactly the same way as in these references. By a partition \( P \) of \( I^s \), we mean a set of finite sequences \( \eta_{j_0}^{(0)}, \eta_{j_1}^{(1)}, \ldots, \eta_{j_m}^{(m)} \) with \( 0 = \eta_{j_0}^{(0)} \leq \eta_{j_1}^{(1)} \leq \cdots \leq \eta_{j_m}^{(m)} = 1 \) for \( j = 1, 2, \ldots, s \). The following basic lemma can be found in [6], Ch. 2, Lemma 5.2, and [18].

Lemma 2. Let \( P \) be a partition of \( I^s \), consisting of the \( s \) sequences \( \eta_{j_0}^{(0)}, \eta_{j_1}^{(1)}, \ldots, \eta_{j_m}^{(m)} \) \( (j = 1, 2, \ldots, s) \), and let \( Q \) be a second partition of \( I^s \), consisting of the \( s \) sequences \( \xi_{j_0}^{(0)}, \xi_{j_1}^{(1)}, \ldots, \xi_{j_m}^{(m)} \) \( (j = 1, 2, \ldots, s) \). Furthermore, let \( f \) and \( g \) be two functions on \( I^s \). Then we have the identity
\[ \sum_{j_0=0}^{m_0} \cdots \sum_{j_s=0}^{m_s} f(\xi_{j_0}^{(0)}, \ldots, \xi_{j_s}^{(s-1)}) A_{j_0} \cdots A_{j_s} g(\eta_{j_0}^{(0)}, \ldots, \eta_{j_s}^{(s)}) \]
\[ = \sum_{n=0}^{s} (-1)^n \sum_{j_0=0}^{m_0} \cdots \sum_{j_n=0}^{m_n} f(\xi_{j_0}^{(0)}, \ldots, \xi_{j_n}^{(n-1)}) A_{j_0} \cdots A_{j_n} g(\eta_{j_0}^{(0)}, \ldots, \eta_{j_n}^{(n)}) \times \]
\[ \quad \times \Delta_{j_{n+1}} \cdots \frac{r^n}{1-r} \sum_{j_{n+1}=0}^{m_{n+1}} \cdots \sum_{j_r=0}^{m_r} f(\xi_{j_{n+1}}^{(n+1)}, \ldots, \xi_{j_r}^{(r-1)}) A_{j_{n+1}} \cdots A_{j_r} g(\eta_{j_{n+1}}^{(n+1)}, \ldots, \eta_{j_r}^{(r)}) \].

On the right-hand side, when \( p = 0 \), the summation symbols referring to \( \eta_{j_0}^{(0)}, \ldots, \eta_{j_s}^{(s)} \), as well as \( A_{j_0}, \ldots, p \), are understood to disappear, and similarly, when \( p = s \), then \( \Delta_{j_{s+1}}, \ldots, p \) should be disregarded, the variables \( \eta_{j_{s+1}}, \ldots, \eta_{j_r}^{(r)} \) disappearing altogether.

We note that if \( s = 1 \), then Lemma 2 reduces to the familiar formula for summation by parts. The analogue of the Koksma–Hlawka inequality has the analogous form.

Theorem 1. Let \( \omega = (\omega_n) \), \( n = 0, 1, \ldots \), be a sequence in \( \mathbb{R}^s \), and let \( f \) be a function that is of bounded variation \( V(f) \) on \( I^s \) in the sense of Hardy and Krause. Then, for any \( 0 < r < 1 \) we have
\[ \Delta_r(f, \omega) \leq V(f) D_r(\omega). \]

Proof. Without loss of generality, we may suppose that \( \omega \) is a sequence in \( I^s \). Let us put \( \omega_n = (\omega_n^{(0)}, \ldots, \omega_n^{(s)}) \) for \( n \geq 0 \). We shall first establish a "truncated" version of the inequality (5). For this purpose, we fix a positive integer \( N \).

By an admissible double partition of \( I^s \), we shall mean an ordered pair \((P, Q)\) of partitions \( P \) and \( Q \) of \( I^s \) satisfying the following conditions. First of all, \( P \) consists of the \( s \) sequences \( \eta_{j_0}^{(0)}, \eta_{j_1}^{(1)}, \ldots, \eta_{j_m}^{(m)} \) \( (j = 1, 2, \ldots, s) \), and \( Q \) consists of the \( s \) sequences \( \xi_{j_0}^{(0)}, \xi_{j_1}^{(1)}, \ldots, \xi_{j_m}^{(m)} \) \( (j = 1, 2, \ldots, s) \), and these are related by
\[ 0 = \xi_{j_0}^{(0)} \leq \xi_{j_1}^{(1)} \leq \cdots \leq \xi_{j_m}^{(m)} \leq \cdots \leq \xi_{j_0}^{(m)} = 1 \]
for \( j = 1, 2, \ldots, s \). Moreover, for each \( j = 1, 2, \ldots, s \), the sequence \( \xi_{j_0}^{(0)} \), \ldots, \( \xi_{j_m}^{(m)} \) should at least contain the numbers \( \omega_0^{(0)}, \ldots, \omega_{j_m}^{(m-1)} \).

With such an admissible double partition being chosen, we apply Lemma 2 with the given function \( f \) and with the function \( g \) defined by
\[ g(x_0, \ldots, x_{n-1}; x_n) = \frac{1-r}{1-r} \sum_{n=0}^{N-1} \sum_{j_0=0}^{m_0} \cdots \sum_{j_s=0}^{m_s} f(\eta_{j_0}^{(0)}, \ldots, \eta_{j_s}^{(s-1)}; x_n) \]
for \( (x_0, \ldots, x_{n-1}) \in I^s \) where \( c(\omega_0^{(0)}; \ldots; \omega_{j_m}^{(m-1)}; x_n) \) denotes the characteristic function of \( [\omega_0^{(0)}, \omega_1^{(0)}] \times \cdots \times [\omega_{j_m}^{(m-1)}, x_n] \). Then the left-hand side of (4) attains the following form:
\[ \frac{1-r}{1-r} \sum_{n=0}^{N-1} \sum_{j_0=0}^{m_0} \cdots \sum_{j_s=0}^{m_s} f(\xi_{j_0}^{(0)}, \ldots, \xi_{j_s}^{(s-1)}; x_n) A_{j_0} \cdots A_{j_s} g(\eta_{j_0}^{(0)}, \ldots, \eta_{j_s}^{(s)}) \]
\[ = \frac{1-r}{1-r} \sum_{n=0}^{N-1} \sum_{j_0=0}^{m_0} \cdots \sum_{j_s=0}^{m_s} f(\xi_{j_0}^{(0)}, \ldots, \xi_{j_s}^{(s-1)}; x_n) A_{j_0} \cdots A_{j_s} c(\eta_{j_0}^{(0)}, \ldots, \eta_{j_s}^{(s)}; \omega_n) \]
\[ - \sum_{n=0}^{N-1} \sum_{j_0=0}^{m_0} \cdots \sum_{j_s=0}^{m_s} f(\xi_{j_0}^{(0)}, \ldots, \xi_{j_s}^{(s-1)}; x_n) A_{j_0} \cdots A_{j_s} c(\eta_{j_0}^{(0)}, \ldots, \eta_{j_s}^{(s)}; \omega_n). \]

Now \( A_{j_0} \cdots A_{j_s} c(\eta_{j_0}^{(0)}, \ldots, \eta_{j_s}^{(s)}; \omega_n) \equiv c(\eta_{j_0}^{(0)}, \eta_{j_1}^{(1)}; \ldots; \eta_{j_s}^{(s)}; \omega_n; \omega_n) \), so that the first term on the right-hand side of (6) is equal to
\[ \frac{1-r}{1-r} \sum_{n=0}^{N-1} \sum_{j_0=0}^{m_0} \cdots \sum_{j_s=0}^{m_s} f(\xi_{j_0}^{(0)}, \ldots, \xi_{j_s}^{(s-1)}; x_n) \]
\[ \times c(\eta_{j_0}^{(0)}, \eta_{j_1}^{(1)}; \ldots; \eta_{j_s}^{(s)}; \omega_n; \omega_n). \]
For fixed \( a, 0 \leq a \leq N - 1 \), consider the inner sum in (7) over \( i_1, \ldots, i_4 \). There is a unique \( s \)-tuple \((i_1, \ldots, i_4)\) for which
\[
\chi_a \in [\eta^{(1)}_{i_1}, \eta^{(1)}_{i_1+1}] \times \cdots \times [\eta^{(a)}_{i_a}, \eta^{(a)}_{i_a+1}].
\]
From the definition of the admissible double partition \((P, Q)\) it follows then that
\[
\chi_a = (\eta^{(1)}_{i_1+1}, \ldots, \eta^{(a)}_{i_a+1}).
\]
Therefore the expression in (7) is nothing else but
\[
\frac{1 - \rho}{1 - \rho N} \sum_{a=0}^{N-1} f(\chi_a)\rho^a.
\]
Altogether, we have shown that the left-hand side of (4) reduces to
\[
\frac{1 - \rho}{1 - \rho N} \sum_{a=0}^{N-1} f(\chi_a)\rho^a - \sum_{i_0=0}^{m_0-1} \cdots \sum_{i_{l_1}=0}^{m_{l_1}-1} f(\eta^{(1)}_{i_1+1}, \ldots, \eta^{(1)}_{i_1+1}) A_{i_2, \ldots, i_{l_1+1}} \eta^{(1)}_{i_1+1} \cdots \eta^{(l_1+1)}_{i_{l_1+1}},
\]
Now consider the right-hand side of (4). It is important to note that \( g(t^{(1)}, \ldots, t^{(a)}) \equiv 0 \) whenever at least one of the \( t^{(b)} \) is zero, and that \( g(1, \ldots, 1) \equiv 0 \). Therefore, the term on the right-hand side of (4) corresponding to \( p = 0 \), namely,
\[
\frac{1 - \rho}{1 - \rho N} \sum_{a=0}^{N-1} f(\chi_a)\rho^a - \sum_{i_0=0}^{m_0-1} \cdots \sum_{i_{l_1}=0}^{m_{l_1}-1} f(\eta^{(1)}_{i_1+1}, \ldots, \eta^{(1)}_{i_1+1}) A_{i_2, \ldots, i_{l_1+1}} \eta^{(1)}_{i_1+1} \cdots \eta^{(l_1+1)}_{i_{l_1+1}},
\]
is equal to zero. For example, for \( 1 \leq p \leq s \), only those terms are left where all the variables \( t^{(p+1)}, \ldots, t^{(a)} \) are replaced by 0. Hence, the right-hand side of (4) reduces to
\[
\sum_{p=1}^{s} (-1)^p \sum_{i_0=0}^{m_0-1} \cdots \sum_{i_p=0}^{m_p-1} g(\eta^{(1)}_{i_1}, \ldots, \eta^{(1)}_{i_1}) \cdots g(\eta^{(a)}_{i_a}, 1, \ldots, 1) \times A_{i_2, \ldots, i_{l_1+1}} \eta^{(1)}_{i_1} \cdots \eta^{(l_1+1)}_{i_{l_1+1}}.
\]
Using the definition of the \( p \)-dimensional variation \( \mathcal{V}^{(p)} \) on \( \mathbb{P}^p \), we obtain that the above expression is bounded in absolute value by
\[
D_{r,N}(\mathcal{V}) \sum_{p=1}^{s} \sum_{i_0=0}^{m_0-1} \cdots \sum_{i_p=0}^{m_p-1} \mathcal{V}^{(p)}(f(t^{(1)}, \ldots, t^{(a)}, 1, \ldots, 1)) = D_{r,N}(\mathcal{V}) V(f).
\]
Thus we arrive at the inequality
\[
\frac{1 - \rho}{1 - \rho N} \sum_{a=0}^{N-1} f(\chi_a)\rho^a - \sum_{i_0=0}^{m_0-1} \cdots \sum_{i_{l_1}=0}^{m_{l_1}-1} f(\eta^{(1)}_{i_1+1}, \ldots, \eta^{(1)}_{i_1+1}) A_{i_2, \ldots, i_{l_1+1}} \eta^{(1)}_{i_1+1} \cdots \eta^{(l_1+1)}_{i_{l_1+1}} \leq V(f) D_{r,N}(\mathcal{V}).
\]
We note that
\[
\frac{1 - \rho}{1 - \rho N} \sum_{a=0}^{N-1} f(\chi_a)\rho^a - \sum_{i_0=0}^{m_0-1} \cdots \sum_{i_{l_1}=0}^{m_{l_1}-1} f(\eta^{(1)}_{i_1+1}, \ldots, \eta^{(1)}_{i_1+1}) A_{i_2, \ldots, i_{l_1+1}} \eta^{(1)}_{i_1+1} \cdots \eta^{(l_1+1)}_{i_{l_1+1}} \leq V(f) D_{r,N}(\mathcal{V}).
\]
and the definition of an admissible double partition implies that the sum over \( i_1, \ldots, i_4 \) on the left of (9) is a Riemann sum for
\[
\int_0^1 \cdots \int_0^1 f(t^{(1)}, \ldots, t^{(a)})\rho^a dt^{(1)} \cdots dt^{(a)}.
\]
The other terms in (9) are independent of the chosen admissible double partition \((P, Q)\). By letting \((P, Q)\) run through a sequence of admissible double partitions with
\[
\max_{1 \leq j \leq s, 1 \leq a \leq s} (\eta^{(j)}_{i_j+1} - \eta^{(j)}_{i_j}) \to 0,
\]
we will therefore obtain the inequality
\[
\frac{1 - \rho}{1 - \rho N} \sum_{a=0}^{N-1} f(\chi_a)\rho^a - \sum_{i_0=0}^{m_0-1} \cdots \sum_{i_{l_1}=0}^{m_{l_1}-1} f(\eta^{(1)}_{i_1+1}, \ldots, \eta^{(1)}_{i_1+1}) A_{i_2, \ldots, i_{l_1+1}} \eta^{(1)}_{i_1+1} \cdots \eta^{(l_1+1)}_{i_{l_1+1}} \leq V(f) D_{r,N}(\mathcal{V}).
\]
This is the truncated version of (3). By letting \( N \to \infty \) in (10) and taking into account Lemma 1, we deduce (ii) itself. We remark that the method yields even a somewhat sharper inequality, in the same sense as in [6], Ch. 2, Theorem 5.5.

In the case \( s = 1 \), one can establish an estimate for \( D_{r}(f, \omega) \) that is valid for all continuous functions \( f \) on \([0, 1]\). We first need an alternative representation for the truncated Abel discrepancy \( D_{r,N}(\mathcal{V}) \) that is analogous to [8], Theorem 1, and [13], Lemma 2.4. We note that the definition of \( D_{r,N}(\mathcal{V}) \) also makes sense for a finite sequence of \( N \) numbers; in this case, we simply write \( D_{r,N}(\mathcal{V}) \) for the truncated Abel discrepancy.

\textbf{Lemma 3.} Let \( x_0, x_1, \ldots, x_N \) be \( N \) points in \([0, 1]\), and let \( a_0 \leq a_1 \leq \ldots \leq a_{N-1} \) be an arrangement of these points in nonincreasing order, so that \( a_0, a_1, \ldots, a_{N-1} \) is a permutation of \( 0, 1, \ldots, N - 1 \). Then, for any \( 0 < \tau < 1 \) we have
\[
D_{r,N}^{\tau} \max_{j=0,1,\ldots,N-1} \max \left\{ \left| \frac{1 - \rho}{1 - \rho \tau} \sum_{a=0}^{a_j} \rho^a \right|, \left| \frac{1 - \rho}{1 - \rho \tau} \sum_{a=0}^{a_j-1} \rho^a \right| \right\},
\]
where, as usual, an empty sum is meant to be zero.

Proof. For notational convenience, we write \( x_{N-1} = 0 \) and \( x_N = 1 \). Moreover, let \( \xi_j \) denote the characteristic function of the interval \([0, \tau]\). 

Then,
\[
D_{r,N} = \max_{f=1, \ldots, N} \sup_{x_{j-1} < x_j} \left| \frac{1-r}{1-rN} \sum_{n=0}^{N-1} q_n(x_{n+1})^r - t \right|
\]
\[
= \max_{f=1, \ldots, N} \sup_{x_{j-1} < x_j} \left| \frac{1-r}{1-rN} \sum_{n=0}^{j-1} p_n - t \right|
\]
\[
= \max \left\{ \left| \frac{1-r}{1-rN} \sum_{n=0}^{j-1} p_n - x_{j-1} \right|, \left| \frac{1-r}{1-rN} \sum_{n=0}^{j-1} p_n - x_j \right| \right\}.
\]
In the same way as in the proof of [8], Theorem 1, one shows that one may drop the restriction \(x_{j-1} < x_j\) in the first maximum. Therefore,
\[
D_{r,N} = \max_{f=1, \ldots, N} \left\{ \left| x_{j-1} - \frac{1-r}{1-rN} \sum_{n=0}^{j-1} p_n \right|, \left| x_j - \frac{1-r}{1-rN} \sum_{n=0}^{j-1} p_n \right| \right\}
\]
\[
= \max \left\{ \left| x_{j-1} - \frac{1-r}{1-rN} \sum_{n=0}^{j-1} p_n \right|, \left| x_{j} - \frac{1-r}{1-rN} \sum_{n=0}^{j-1} p_n \right| \right\}.
\]
The last step is valid because we only dropped the terms
\[
|x_{j-1} - 0| \quad \text{and} \quad |x_j - \frac{1-r}{1-rN} \sum_{n=0}^{N-1} p_n|,
\]
both of which are zero.

We recall that the modulus of continuity of a continuous function \(f\) on \([0, 1]\) is defined by
\[
M(h) = \sup_{x,y \in [0,1]} |f(x) - f(y)| \quad \text{for} \quad h \geq 0.
\]
The following is an analogue of an inequality of the author in [10], Theorem 3, [11].

**Theorem 2.** Let \(\omega = (x_n), n = 0, 1, \ldots,\) be a sequence of real numbers, and let \(f\) be a continuous function on \([0, 1]\) with modulus of continuity \(M\). Then, for any \(0 < r < 1\) we have
\[
\delta_r(f, \omega) \leq M(D_r(\omega)).
\]

**Proof.** Without loss of generality, we may assume that \(\omega\) is a sequence in \([0, 1]\). Fix a positive integer \(N\). As in Lemma 3, let \(x_0 = x_1 = \cdots = x_{2N-1}\) be an arrangement of the points \(x_0, \ldots, x_{N-1}\) in nondecreasing order. Put
\[
s_j = \frac{1-r}{1-rN} \sum_{n=0}^{j-1} p_n^r \quad \text{for} \quad 1 \leq j \leq N, \quad \text{and} \quad s_0 = 0.
\]

Then,
\[
\frac{1-r}{1-rN} \sum_{n=0}^{N-1} f(x_n) r^n - \int_0^1 f(t) dt = \frac{1-r}{1-rN} \sum_{j=1}^{N-1} f(x_j) r^j - \frac{1-r}{1-rN} \sum_{j=1}^{N-1} s_j^j f(t) dt
\]
\[
= \frac{1-r}{1-rN} \sum_{j=0}^{N-1} f(x_j) r^j - \sum_{j=1}^{N-1} (s_{j+1} - s_j) f(\xi_j)
\]
with \(s_j < \mathcal{E}_j < \mathcal{E}_{j+1}\) for \(0 \leq j \leq N-1\), by the mean-value theorem of integral calculus. It follows that
\[
\frac{1-r}{1-rN} \sum_{n=0}^{N-1} f(x_n) r^n - \int_0^1 f(t) dt \leq \frac{1-r}{1-rN} \sum_{j=0}^{N-1} (f(x_j) - f(\xi_j)) r^j.
\]

From (11) we deduce that
\[
|s_j - s| \leq D_{r,N}(\omega) \quad \text{for} \quad 0 \leq j \leq N-1,
\]
and so (12) implies
\[
\left| \frac{1-r}{1-rN} \sum_{n=0}^{N-1} f(x_n) r^n - \int_0^1 f(t) dt \right| \leq \frac{1-r}{1-rN} M(D_{r,N}(\omega)) \sum_{j=0}^{N-1} r^j = M(D_{r,N}(\omega)).
\]

Letting \(N \rightarrow \infty\) in (13) and using Lemma 1 as well as the fact that \(M(h)\) is a continuous function of \(h\) because of the uniform continuity of \(f\) on \([0, 1]\), we arrive at the desired inequality.

**3. Abel discrepancy and exponential sums.** In the same way as in [6], Ch. 2, Corollary 5.1, one deduces from Theorem 1 that for any sequence \(\omega = (x_n), n = 0, 1, \ldots,\) in \(R\) and any integer \(h \neq 0\) the inequality
\[
\left| \frac{1-r}{1-rN} \sum_{n=0}^{N-1} \epsilon_m^{|m|} r^m \right| \leq |h| D_r(\omega)
\]
holds for \(0 < r < 1\). Conversely, the Abel discrepancy can be estimated in terms of the weighted exponential sums occurring in (14).

**Theorem 3.** Let \(\omega = (x_n), n = 0, 1, \ldots,\) be a sequence of real numbers. Then, for any \(0 < r < 1\) and for any positive integer \(m\),
\[
D_r(\omega) \leq \frac{4}{m^2 + 1} \frac{4(1-r)}{r} \sum_{m=1}^{\infty} \left( \frac{1}{h} - \frac{1}{m+1} \right) \left| \sum_{n=0}^{\infty} \epsilon_m^{|m|} r^m \right|.
\]

**Proof.** The following theorem was established by the author and W. Philipp ([14], Theorem 1). Let \(\Phi\) be nondecreasing on \([0, 1]\) with \(\Phi(0) = 0\) and \(\Phi(1) = 1\), let \(\Theta\) satisfy a Lipschitz condition on \([0, 1]\),
with constant $K$, and suppose $G(0) = 0$ and $G(1) = 1$. Then for any positive integer $m$ we have
\[
\sup_{t \in [0, 1]} |F(t) - G(t)| \leq \frac{4K}{m+1} + \frac{4}{\pi} \sum_{n=1}^{m} \left( \frac{1}{n} - \frac{1}{m+1} \right) |\hat{F}(h) - \hat{G}(h)|,
\]
where $F$ and $\hat{G}$ denote the Fourier-Stieltjes transforms of $F$ and $G$, respectively. We apply this result with
\[
F(t) = (1-t) \sum_{n=0}^{\infty} \alpha_n(t) e^{\pi i n t} \quad \text{and} \quad G(t) = t \quad \text{for} \quad 0 < t < 1,
\]
where $\alpha_n$ is the characteristic function of $[0, t)$. We have $\hat{G}(h) = 0$ for all integers $h \neq 0$, as well as
\[
\hat{F}(h) = \int_0^1 e^{\pi i n t} dF(t) = (1-t) \int_0^1 e^{\pi i n t} d\left( \sum_{n=0}^{\infty} \alpha_n(t) e^{\pi i n t} \right).
\]
Using a well-known theorem on Stieltjes integrals (see [15], pp. 120–121), we can write
\[
\hat{F}(h) = (1-t) \sum_{n=0}^{\infty} \int_0^1 e^{\pi i n t} dt \alpha_n(t) e^{\pi i n h} = (1-t) \sum_{n=0}^{\infty} e^{\pi i n h} \alpha_n(t) e^{\pi i n h}
\]
for all integers $h$. This completes the proof of Theorem 3.

There is also a multidimensional analogue of Theorem 3. First we need some notation. For a lattice point $h = (h^1, \ldots, h^m) \in \mathbb{Z}^m$, define
\[
A(h) = \max_{1 \leq i \leq m} |h^i|, \quad \text{and} \quad B(h) = \int_0^1 \max(\min(|h^i|, 1), 1).
\]
Also, let $\langle x, y \rangle$ denote the standard inner product of $x, y \in \mathbb{R}^m$.

**Theorem 4.** Let $t = (t^1, \ldots, t^m)$, $0 < t < 1$, be a sequence in $\mathbb{R}^m$. Then, for any $0 < r < 1$ and for any positive integer $m$ we have
\[
D_r(t) \leq C_r \left( \frac{1}{m+1} + (1-r) \sum_{n=0}^{m} \frac{1}{\max(\min(|h^i|, 1), 1)} \right)
\]
with a constant $C_r$ only depending on $r$.

**Proof.** The result can be deduced from [14], Theorem 2, in the same way as Theorem 3 was deduced from [14], Theorem 1. We note that we have used a slightly modified definition of $B(h)$, but that the $B(h)$ from [14] is at least as large as the $B(h)$ employed in the present paper.

By analyzing the proof of [14], Theorem 2, one could, in fact, give an explicit value for the constant $C_r$. However, due to the way in which the inductive argument in [14] operates, this value would be fairly large.
Because of (16) and (17), the intervals
\[ [a, a + \frac{1}{K} H(x)] \] and \[ [y_1 + \frac{1}{K} H(y_1), y_1] \]
can have at most one point in common. Therefore,
\[ \int_0^{\frac{2}{K} H(x)} H^2(t) dt = \int_a^{a + \frac{1}{K} H(x)} H^2(t) dt + \int_{a + \frac{1}{K} H(x)}^a H^2(t) dt + \int_{y_1 + \frac{1}{K} H(y_1)}^{y_1} H^2(t) dt + \int_{y_1}^{y_1 + \frac{1}{K} H(y_1)} H^2(t) dt \]
\[ \geq \int_a^{a + \frac{1}{K} H(x)} H(a) + H(x)(a - \frac{1}{K}) dt + \int_{y_1 + \frac{1}{K} H(y_1)}^{y_1} H(y_1) + H(y_1)(y_1 - \frac{1}{K}) dt \]
\[ = \frac{1}{3K} H^3(x) + \frac{1}{3K} (H(y_1))^3 = \frac{1}{3K} H^3(x) + \frac{1}{3K} (H(y_1))^3. \]
From the inequality \( a^3 + b^3 \geq \frac{1}{2} (a + b)^3 \) for nonnegative real numbers \( a \) and \( b \), we infer
(18) \[ \int_0^{\frac{2}{K} H(x)} H^2(t) dt \geq \frac{1}{12K} (H(x) - H(y))^3. \]
It follows easily that (18) holds, in fact, for all \( a, y \in \mathbb{R} \), and so
(19) \[ |H(x) - H(y)|^3 \leq 12K \int_0^{\frac{2}{K} H(x)} H^2(t) dt \]
for all \( a, y \in \mathbb{R} \).

By the same reasoning as in the proof of [2], Theorem 1, we obtain
(20) \[ \int_0^{\frac{2}{K} H(x)} H^2(t) dt = \frac{1}{2\pi^2} \sum_{h=1}^{\infty} \frac{1}{h^3} (\hat{F}(h) - \hat{G}(h))^2. \]

(We have corrected a typographical error in that paper.) By combining (19) and (20) and noting that the left-hand side of (15) is equal to \( \sup_{x \in \mathbb{R}} |H(x) - H(y)| \), one completes the proof of the lemma.

**Theorem 5.** Let \( \omega = (x_n), n = 0, 1, \ldots \), be a sequence of real numbers. Then, for any \( 0 < r < 1 \) we have
(21) \[ D_r(\omega) \leq \left( \frac{6}{\pi^3} \sum_{h=1}^{\infty} \frac{1}{h^3} \left( 1 - r \right)^2 \sum_{n=0}^{\infty} \sin^2 n \pi h \right)^{1/3}. \]

**Proof.** For \( F \) and \( G \) in Lemma 4, one chooses the same functions as in the proof of Theorem 3.

We remark that the constant \( \frac{6}{\pi^3} \) in (21) is best possible, as one verifies by choosing for \( \omega \) the sequence \((x_n)\) with \( x_n = 0 \) for \( n = 0, 1, \ldots \).
with the constant implied by the Landau symbol only depending on $a$. This completes the proof of the theorem.

Given an explicit function $\psi$, one will, of course, establish the final estimate for $D_1(\omega)$ by choosing $m$ so as to minimize the expression on the right-hand side of (22). We mention briefly the important special case where $a$ is of finite type in the sense of Koksma [5], p. 28.

Corollary. Let $a$ be an irrational of finite type $\eta$. Then the Abel discrepancy of the sequence $\omega = (na)$, $n = 0, 1, \ldots$, satisfies

$$D_\varepsilon(\omega) = O((1-r)^{1/2} \eta^{-2})$$

for every $\varepsilon > 0$.

Proof. According to [6], Ch. 2, Lemma 3.1, the irrational $a$ is of finite type $\eta$ if and only if $\eta$ is the infimum of all real numbers $\tau$ for which there exists a positive constant $c = c(\tau, a)$ such that $a$ is of type $< \psi$, where $\psi(q) = cq^{\tau-1}$. Thus, given an $\varepsilon > 0$, we can apply Theorem 6 with the function $\psi(q) = c(\eta + \varepsilon, a)q^{\eta-1/2}$. Then, since $\eta > 1$, we obtain

$$D_{\varepsilon}(\omega) \leq C\left(\frac{1}{m+1} + (1-r)\left(\log^2 m + m^{\eta-1/2} + \sum_{k=1}^{m} h^{-1/2} k\right)\right)$$

$$\leq C\left(\frac{1}{m+1} + (1-r)m^{\eta-1/2}\right)$$

with constants $C, C'$ independent of $r$ and $m$. Now choose $m = \left\lfloor(1-r)^{-1/2}\eta\right\rfloor$ and we arrive at the desired estimate.

On the basis of Theorem 4, one can also establish estimates for the Abel discrepancy of sequences $\omega$ in $\mathbb{R}^s$ of the form $\omega = (na)$, $n = 0, 1, \ldots$, where $a = (a^{(1)}, \ldots, a^{(s)}) \in \mathbb{R}^s$ with $1, a^{(1)}, \ldots, a^{(s)}$ linearly independent over the rationals. For such $a$, we introduce a notion of type as follows (compare with [17], Definition 6.1).

**Definition.** Let $a = (a^{(1)}, \ldots, a^{(s)}) \in \mathbb{R}^s$ with $1, a^{(1)}, \ldots, a^{(s)}$ linearly independent over the rationals. For a real number $\eta$, the $s$-tuple $a$ is said to be of finite type $\eta$ if $\eta$ is the infimum of all numbers $\sigma$ for which there exists a positive constant $c = c(\sigma, a)$ such that

$$R^s(\langle h, a \rangle) \geq c$$

holds for all lattice points $h \in \mathbb{Z}^s$ with $h \neq 0$.

It follows easily from the Minkowski linear forms theorem that we always have $\eta \geq 1$. For $s = 1$, the above concept reduces, of course, to Koksma's notion of finite type. For certain $a$ the inequality (23) need not hold for any finite $\sigma$. Such an $a$ could be called of infinite type. On the other hand, explicit examples of $s$-tuples of finite type $\eta = 1$ are known. Schmidt [10] has shown that $a = (a^{(1)}, \ldots, a^{(s)})$ is of finite type $\eta = 1$ whenever the $a^{(i)}$, $1 \leq i \leq s$, are real algebraic numbers for which $1, a^{(1)}, \ldots, a^{(s)}$ are linearly independent over the rationals. Also, it follows from a result of Baker [1] that $a = (a^{(1)}, \ldots, a^{(s)})$, with distinct nonzero rationals $a_1, \ldots, a_s$, is of finite type $\eta = 1$.

**Theorem 7.** Let $a \in \mathbb{R}^s$ be of finite type $\eta = 1$. Then the Abel discrepancy of the sequence $\omega = (na)$, $n = 0, 1, \ldots$, satisfies

$$D_{\varepsilon}(\omega) = O((1-r)^{-\varepsilon})$$

for every $\varepsilon > 0$.

Proof. For a lattice point $h \in \mathbb{Z}^s$ with $h \neq 0$, we consider the weighted exponential sum

$$S_h = \left| \sum_{n=0}^{\infty} e^{2\pi i (h, na)} n^s \right| = \left| \sum_{n=0}^{\infty} e^{2\pi i (h, na)} n^s \right|.$$

In the same way as in the proof of Theorem 6, one shows that

$$S_h \leq \frac{2}{3\|h, a\|^s}.$$

Therefore, for any $0 < r < 1$ and for any positive integer $m$ we obtain from Theorem 4 that

$$D_{\varepsilon}(\omega) < C\left(\frac{1}{m+1} + (1-r)\sum_{0 \leq d < m \epsilon} R^{s-1}(h) \|\langle h, a \rangle\|^{-1}\right).$$

Now it was shown in the proof of [13], Theorem 6.1, (see also [9]) that under the given condition on $a$ we have

$$\sum_{0 < d < m} R^{s-1}(h) \|\langle h, a \rangle\|^{-1} = O(m^s)$$

for every $\varepsilon > 0$. We conclude that

$$D_{\varepsilon}(\omega) < C\left(\frac{1}{m+1} + (1-r)m^s\right),$$

with a constant $C'$ independent of $r$ and $m$. The proof is completed by choosing $m = \left\lfloor(1-r)^{-1}\right\rfloor$.

Theorem 7 can be extended to include any $a \in \mathbb{R}^s$ of finite type. This is achieved by generalizing the crucial estimate (24). It follows, in fact, from [6], Ch. 2, Exercises 3.15 and 3.16, that for $a \in \mathbb{R}^s$ of finite type $\eta$ one has

$$\sum_{0 < d < m} R^{s-1}(h) \|\langle h, a \rangle\|^{-1} = O(m^{s(\eta-1/2)}),$$

for every $\varepsilon > 0$. If $\omega = (na)$, $n = 0, 1, \ldots$, this yields then the estimate

$$D_{\varepsilon}(\omega) = O((1-r)^{-\varepsilon})$$

for every $\varepsilon > 0$.\[8 -- Acta Arithmetica XXVIII]
We note that Hawka [4] has shown the following result. Let \( \mathbf{a} = (a^0, \ldots, a^n) \) with \( a^0, 1 \leq i \leq n \), being integers of a fixed real algebraic number field of degree \( s+1 \) such that \( 1, a^1, \ldots, a^n \) are linearly independent over the rationals, and set \( \omega = (ma), \; n = 0, 1, \ldots \). Then, for a function \( f \) that is of bounded variation \( V(f) \) on \( I^2 \) in the sense of Hardy and Krause, we have
\[
\delta_r(f, \omega) \leq c(\alpha, \varepsilon) V(f)(1-r)^{|\omega| - s}
\]
for \( 0 < r < 1 \) and any \( \varepsilon > 0 \), where \( c(\alpha, \varepsilon) \) only depends on \( \alpha \) and \( \varepsilon \). This result can be improved considerably. Namely, by combining Theorems 1 and 7, we see that the exponent \( (1/s) - s \) in (25) can be replaced by \( 1 - s \). This is an indication of the strength of our method.

5. Irregularities of distribution. In this section, we prove some results concerning lower bounds for \( D_r(\omega) \) in the one-dimensional case. For a sequence \( \omega = (x_n), \; n = 0, 1, \ldots \), of real numbers, Hlawka [4] has introduced the (modified) Abel discrepancy
\[
D^*_r(\omega) = \sup_J \left( 1 - r \right) \sum_{n=0}^{\infty} \varphi_{J}(x_n)r^{n} - \lambda(J)
\]
for \( 0 < r < 1 \), where the supremum is now extended over all subintervals \( J \) of \( (0, 1) \) of the form \( J = (t_1, t_2) \). By the usual argument, one shows that
\[
D_r(\omega) \leq D^*_r(\omega) \leq 2D_r(\omega)
\]
for any \( 0 < r < 1 \) and any \( \omega \). On the other hand, by a result of Hlawka [4], we always have
\[
D^*_r(\omega) \geq 1 - r.
\]
Therefore, for any \( 0 < r < 1 \) and any \( \omega \), the inequality
\[
D_r(\omega) \geq \frac{1-r}{2}
\]
holds. This lower bound is best possible, as the following theorem shows.

THEOREM 8. For any \( 0 < r < 1 \), there exists a sequence \( \omega = (x_n), \; n = 0, 1, \ldots \), of real numbers depending on \( r \) such that
\[
D_r(\omega) = \frac{1-r}{2}.
\]

Proof. For given \( r, \; 0 < r < 1 \), consider the sequence \( \omega = (x_n), \; n = 0, 1, \ldots \), given by
\[
x_n = \frac{2 - r^n - r^{n+1}}{2} \quad \text{for} \quad n = 0, 1, \ldots
\]
Evidently, this is an increasing sequence contained in \( [0, 1] \) with \( \lim_n x_n = 1 \).

For notational convenience, we set \( x_{-1} = 0 \). Choose \( \varepsilon \) with \( 0 < \varepsilon < 1 \); then there is a unique \( j \geq 0 \) such that \( x_{j+1} - 1 < x_j \). With \( \rho_j \) denoting the characteristic function of \( [0, t] \), we get
\[
\left| (1-r) \sum_{n=0}^{\infty} \rho_j(x_n)r^n - t \right| = \left| (1-r) \sum_{n=0}^{j-1} r^n - t \right|
\]
\[
= \left| 1 - r^j - t \right| \leq \max \{ 1 - r^{j+1}, |1-r^j - x_j| \}.
\]
Now for \( j \geq 1 \), we have
\[
|1-r^{j+1} - x_j| \leq \frac{r^j - r^{j+1}}{2} \leq \frac{1 - r}{2},
\]
and this holds as well for \( j = 0 \). Similarly, for \( j \geq 0 \) we get
\[
|1-r^j - x_j| \leq \frac{r^j - r^{j+1}}{2} \leq \frac{1 - r}{2},
\]
so that
\[
\left| (1-r) \sum_{n=0}^{\infty} \rho_j(x_n)r^n - t \right| \leq \frac{1-r}{2}
\]
for \( 0 < t < 1 \). Since the inequality is trivial for \( t = 0, 1 \), we obtain \( D_r(\omega) \leq (1-r)/2 \), and together with (26) the formula for \( D_r(\omega) \) is established.

The above theorem shows also, of course, that Hlawka's lower bound for \( D^*_r(\omega) \) is best possible.

The estimate in the Corollary of Theorem 6 is best possible in the following sense.

THEOREM 9. Let \( a \) be an irrational of finite type \( \eta \). Then the Abel discrepancy of the sequence \( \omega = (ma), \; n = 0, 1, \ldots, \) satisfies
\[
D_r(\omega) = \Omega((1-r)^{(2\eta + 1)/r})
\]
for every \( \varepsilon > 0 \).

Proof. Since the result is trivial for \( \eta = 1 \), we assume \( \eta > 1 \) in the rest of the proof. For given \( \varepsilon > 0 \), choose a real number \( \delta \) with
\[
0 < \delta < \min \{ \eta - 1, \frac{\epsilon a^2}{\eta^2 + 1} \},
\]
and then determine \( \gamma > 0 \) from the equation
\[
\frac{1 + \gamma}{\eta - \delta} = \frac{1}{\eta + \varepsilon}.
\]
By a well-known characterization of rationals of finite type \( \eta \) (see [6], Ch. 2, Definition 3.4), we have \( \lim_{q \to \infty} q^{-\omega(q)} \theta(q) = 0 \), where \( \omega \) runs through the positive integers. In particular, there exist infinitely many positive integers \( q \) and corresponding integers \( \omega \) such that

\[
\left| a - \frac{p}{q} \right| < q^{-1+\omega(q)}.
\]

Choose one such \( q \), and set \( N = \lfloor q^n \rfloor \) and \( r = 1 - N^{-1+\omega(q)} \). By writing

\[
a = \frac{p}{q} + \theta q^{-1-\omega(q)} \quad \text{with} \quad |\theta| < 1,
\]

we get for \( 1 \leq n \leq N \),

\[
n \alpha = \frac{n p}{q} + \theta_n \quad \text{with} \quad |\theta_n| < q^{-1-\omega(q)}.
\]

It follows that none of the numbers \( \theta, \{\theta\}, \{2\theta\}, \ldots, \{N\theta\} \) lies in the interval

\[
J = [q^{-1-\omega(q)}, q^{-1-\omega(q)} - (1-r)^{-1+\omega(q)}].
\]

Therefore,

\[
D<sub>r</sub>(\omega) \geq \lambda(J) - (1-r) \sum_{n=1}^{N} \nu_{n}(\{n\theta\}) r^{n} \geq \lambda(J) - (1-r) \sum_{n=1}^{N} r^{n} = \lambda(J) - r^{-1+\omega(q)}.
\]

For sufficiently large \( q \), we have \( \lambda(J) \geq 1/2q \). Moreover,

\[
r^{-1} \geq \frac{2N}{2(1-r)^{-1+\omega(q)}},
\]

and so

\[
q \leq \frac{1}{2}(1-r)^{-(1+\omega(q))} < 2(1-r)^{-(1+\omega(q))}.
\]

Therefore,

\[
D<sub>r</sub>(\omega) \geq \frac{1}{2}(1-r)^{-(1+\omega(q))} \geq (1-r)^{-1+\omega(q)}
\]

for an infinite sequence of values of \( r \) tending to 1. The proof will be complete once we show that

\[
\lim_{r \to 1} (1-r)^{-1+\omega(q)} = 0.
\]

By the substitution \( u = 1/(1-r) \), the above limit relation is equivalent to

\[
\lim_{v \to \infty} \left( 1 - \frac{1}{u} \right) u^{1+\omega(q)} = 0.
\]

However, the latter limit relation follows from \( 1 - \frac{1}{u} \leq e^{-v} \) and

\[
\lim_{v \to \infty} e^{-v} e^{-v} = 0.
\]