TN rl.

FORSCHUNG - AUSBILDUNG - WEITERBILDUNG

Bericht Nr. 115

PARTICLE METHODS

Helmut Neunzert

UNIVERSITÄT KAISERSLAUTERN
Fachbereich Mathematik
Postfach 3049

D -67653 Kaiserslautern

PARTICLE METHODS

H. Neunzert, Kaiserslautern

Introduction:

Particle Methods (also called "Finite Pointset Methods" = FPM in pointing to some fundamental similarity with FEM or FDM, or Simulation Methods or Monte Carlo Methods) are numerical methods for solving evolution equations for functions f(t,P)

$$\frac{\partial f}{\partial t} + \mathrm{div}_P(V[f]f) = Q[f] \quad , \tag{1}$$

where V and Q may depend on f in quite different kinds. P is from a domain Ω in $\mathbb{R}^k$ where k is in general quite big ($k \geq 3$, often $k=6$) - this is the proper application field for particle methods.

For conserving boundary conditions like

$$\int_{\partial\Omega} f\langle V,n\rangle d\omega = 0$$

and for $\int_\Omega Q dP = 0$ the equation (1) is a conservation law, i.e.

$$\int_\Omega f(t,P)dP = \int_\Omega f(0,P)dP = 1 \; .$$

If moreover the nonnegativity of f(0,P) is conserved during the evolution, the solution $f(t,\cdot)$ is a normalized density, i.e.

$$f \geq 0 \quad \text{and} \quad \int_\Omega f(t,P)dP = 1 \; .$$

For numerically solving (1) we approximate these densities by particle ensembles. Our first task will be to clarify this concept: In which sense do particles approximate functions? This question can be answered in the frame of 4 different settings: There is a measure theoretic, a number theoretic, a statistical and a functionalanalytical approach. All of them contribute to the subject and supply techniques which can be used. This will be the first part of the lectures. In the second part we will describe applications for typical equations of type (1): Kinetic questions occurring in plasmaphysics and stellardynamics, in semi conductor technology and space flight; other applications in fluid dynamics etc. will only be touched.

PART I: THE APPROXIMATION OF FUNCTIONS BY PARTICLESETS - DIFFERENT ASPECTS

We consider for a domain Ω in $\mathbb{R}^k$ the class

$$M_{ac}(\Omega) = \left\{f: \Omega \to \mathbb{R}^+ / \int_\Omega f \, dP = 1\right\}$$

of normalized densities. We want to approximate f by an ensemble of N "particles" in Ω which is given by a family $\left[(P_i,\alpha_i)_{i=1,\ldots,N}\right]$ of points $P_1,\ldots,P_N$ and weights $(\alpha_1,\ldots,\alpha_N)$. Two such families are considered to be equal or equivalent, if they differ only by enumeration. (An ensemble is not a set, since we allow two elements (P_i,α_i) and (P_j,α_j), $i \neq j$, to be equal.) We normalize by assuming $\alpha_i \geq 0$ for $i=1,\ldots,N$ and $\sum_{i=1}^{N} \alpha_i = 1$. An ensemble is given by N(k+1) data, which can be used for approximation. Sometimes, the points are fixed (for example by forming a regular grid) and the weights vary, sometimes the weights are fixed (for example $\alpha_i = \frac{1}{N}$, $i=1,\ldots,N$) and the points are chosen freely. We denote the ensemble by $\omega_N^\alpha = \left[(P_i,\alpha_i)_{i=1,\ldots,N}\right]$. What do we mean by saying that "ω_N^α approximates f"? We describe 4 possible concepts.

§ 1 <u>The measuretheoretic interpretation</u>

Let $M(\Omega)$ be the set of all nonnegative Borel measures on Ω with $\mu(\Omega)=1$. Each density $f \in M_{ac}$ defines a Borel measure μ_f by

$$\mu_f(A) = \int_A f \, dP \quad \text{for all Borel measurable sets A ,}$$

where the integral is taken in the sense of Lebesgue.
Clearly $\mu_f \in M(\Omega)$ and μ_f is absolutely continuous with respect to the k-dimensional Lebesgue measure. This justifies the notations $M_{ac}(\Omega)$.
The ensemble ω_N^α may also be interpreted as a measure by defining

$$\delta_{\omega_N^\alpha} := \sum_{j=1}^{N} \alpha_j \delta_{P_j} \; ,$$

$\delta_{\omega_N^\alpha}$ is a discrete measure in $M(\Omega)$.

Having identified densities and particle ensembles as mathematical objects of the same kind, we have the possibility to define "convergence of ensembles to densities" and distances between them.

Convergence:

(See for example P. Billingsley, Convergence of Probability Measures, Wiley 1968.)

Definition 1.1:

A sequence $(\mu_N)_{N\in\mathbb{N}}$ in M is said to "converge weakly" to $\mu \in M$, if

$\int_\Omega \phi d\mu_N \rightarrow \int_\Omega \phi d\mu$ for $N\to\infty$ and all bounded, continuous real functions ϕ on Ω, i.e. for all $\phi \in C^b(\Omega)$.

Especially, if we have a sequence of particle ensembles with growing particle number N, i.e. if we have

$$\left(\delta_{\omega_N^\alpha}\right)_{N\in\mathbb{N}} = \left(\sum_{j=1}^{N} \alpha_j^N \delta_{P_j^N}\right)_{N\in\mathbb{N}} ;$$

we say that ω_N^α (or $\delta_{\omega_N^\alpha}$) converges weakly to $f \in M_{ac}$, if for $N\to\infty$

$$\int_\Omega \phi d\omega_N^\alpha = \sum_{j=1}^{N} \alpha_j^N \phi(P_j^N) \longrightarrow \int_\Omega \phi d\mu_f = \int_\Omega \phi f dP$$

for all $\phi \in C^b(\Omega)$.

One realizes that the problem to approximate an integral $\int_\Omega \phi f dP$ by a sum like $\sum_{j=1}^{N} \alpha_j^N \phi(P_j^N)$ is the classical problem of numerical quadrature. Then P_j^N are the knots and α_j^N the integration weights and one can play with the space of test functions ϕ, which in our case is $C^b(\Omega)$. We come back to this aspect in § 4.

We collect the most important facts about weak convergence in $M(\Omega)$.

A Borel measurable set M is called a μ-continuity set, if $\mu(\partial M) = 0$ (∂M is the topological boundary of M).

Theorem 1.2 (Portmanteau Theorem)

The statement (1) that μ_n converges weakly to μ is equivalent to each of the following statements

(2) $\limsup \mu_n(F) \le \mu(F)$ for all (relatively) closed sets F in Ω

(3) $\lim_{n\to\infty} \mu_n(M) = \mu(M)$ for all μ-continuity sets M

Proof: (See Billingsley page 12)

(1) => (2). Let F be closed and $\delta > 0$. For $\varepsilon > 0$ define

$$G_\varepsilon = \left\{ P \in \Omega \;/\; d(P,F) < \varepsilon \right\}, \text{ where } d(P,F) = \inf\left\{ \|P-Q\| / Q \in F \right\}$$

and $\|\cdot\|$ denotes the Euclidean distance in $\mathbb{R}^k$.

G_ε decreases to F for $\varepsilon \downarrow 0$, and therefore $\mu(G_\varepsilon) \downarrow \mu(F)$. For sufficiently small ε we therefore get $\mu(G_\varepsilon) < \mu(F)+\delta$.

$$\text{Define } \eta(t) := \begin{cases} 1 & \text{if } t \le 0 \\ 1-t & \text{if } 0 \le t \le 1 \\ 0 & \text{if } 1 \le t \end{cases} \text{ and } \Phi_F(P) := \eta\left(\tfrac{1}{\varepsilon} d(P,F)\right)$$

Φ_F is bounded and continuous,

$\Phi_F(P)=1$ for $P \in F$, $\Phi_F(P)=0$

for $P \notin G_\varepsilon$ and $0 \le \Phi_F \le 1$.

Especially, $\Phi_F \in C^b(\Omega)$. From (1) follows $\lim\limits_{n\to\infty} \int_\Omega \Phi_F d\mu_n = \int_\Omega \Phi_F d\mu$.

By construction

$$\mu_n(F) = \int_F \Phi_F d\mu_n \le \int_\Omega \Phi_F d\mu_n$$

and

$$\int_\Omega \Phi_F d\mu = \int_{G_\varepsilon} \Phi_F d\mu \le \mu(G_\varepsilon) \le \mu(F)+\delta \;.$$

This implies

$$\limsup \mu_n(F) \le \lim_{n\to\infty} \int_\Omega \Phi_F d\mu_n = \int_\Omega \Phi_F d\mu \le \mu(F)+\delta \;.$$

Since δ is arbitrary, we get (2)

(2) => (1). Assume (2) and take $\Phi \in C^b(\Omega)$. We show first

$$(1') \quad \limsup_{n\to\infty} \int_\Omega \Phi d\mu_n \le \int_\Omega \Phi d\mu \;.$$

By linear transformation, we may reduce this problem to the case, where $0 < \Phi(P) < 1$ (since $\int (\Phi(P)+\alpha) d\mu_n = \int \Phi d\mu_n + \alpha$).

Choose k arbitrary but fixed and define

$$F_i = \left\{ P \;/\; \tfrac{i}{k} \le \Phi(P) \right\}, \; i=0,\dots,k \;.$$

F_i is closed and since $0<\Phi<1$, we have $\Omega = \bigcup\limits_{i=0}^{k-1} F_i$. Therefore

$$\sum_{i=1}^{k} \frac{i-1}{k}\, \mu\left(\left\{P \,/\, \frac{i-1}{k} \le \phi(P) < \frac{i}{k}\right\}\right) \le \int \phi d\mu \le \sum_{i=1}^{k} \frac{i}{k}\, \mu\left(\left\{P/\frac{i-1}{k} \le \phi(P) < \frac{i}{k}\right\}\right).$$

The right hand side gives

$$\sum_{i=1}^{k} \frac{i}{k}[\mu(F_{i-1})-\mu(F_i)]$$

$$= \frac{1}{k}[\mu(F_o)-\mu(F_1)] + \frac{2}{k}[\mu(F_1)-\mu(F_2)] + \frac{3}{k}[\mu(F_2)-\mu(F_3)] + \ldots + \frac{k}{k}\mu(F_{k-1})$$

$$= \frac{1}{k}\underbrace{\mu(F_o)}_{=1} + \frac{1}{k}\mu(F_1) + \frac{1}{k}\mu(F_2) + \ldots + \frac{1}{k}\mu(F_k) = \frac{1}{k} + \sum_{i=1}^{k} \frac{1}{k}\mu(F_i) .$$

The left hand side gives similarly $\frac{1}{k}\sum_{i=1}^{k} \mu(F_i)$, so that

$$\frac{1}{k}\sum_{i=1}^{k} \mu(F_i) \le \int \phi d\mu \le \frac{1}{k} + \frac{1}{k}\sum_{i=1}^{k} \mu(F_i) .$$

(2) means $\limsup \mu_n(F_i) \le \mu(F_i)$; therefore

$$\int \phi d\mu_n \le \frac{1}{k} + \frac{1}{k}\sum_{i=1}^{k} \mu_n(F_i) \Longrightarrow \limsup_{n\to\infty} \int \phi d\mu_n \le \frac{1}{k} + \frac{1}{k}\sum_{i=1}^{k} \mu(F_i)$$

$$\le \frac{1}{k} + \int \phi d\mu .$$

k was arbitrary, fixed. Let now k tend to ∞; we obtain (1'). If we do the same with $(-\phi)$ we get $\liminf \int \phi d\mu_n \ge \int \phi d\mu$. This implies

$$\int \phi d\mu \le \liminf \int \phi d\mu_n \le \limsup \int \phi d\mu_n \le \int \phi d\mu$$

and therefore the weak convergence.

(2) => (3). By complementation we see immediately that (2) is equivalent to $\liminf \mu_n(G) \ge \mu(G)$ for all open sets G. If we denote the interior of a set A by $\overset{o}{A}$, its closure by $\overline{A}$, (2) implies

$$\mu(\overline{A}) \ge \limsup \mu_n(\overline{A}) \ge \limsup \mu_n(A) \ge \liminf \mu_n(A)$$

$$\ge \liminf \mu_n(\overset{o}{A}) \ge \mu(\overset{o}{A}) .$$

If A is a μ-continuity set, $\mu(\partial A)=0$ and therefore $\mu(\overline{A}) = \mu(\overset{o}{A})$ and we get (3).

(3) => (2). Clearly $\partial\left\{P \in \Omega/d(P,F) \leq \varepsilon\right\} \subset \left\{P \in \Omega/d(P,F) = \varepsilon\right\}$,

therefore these boundaries are distinct for different ε and only countably many of them may have positive measure. Therefore exists a sequence $(\delta_k)_{k\in\mathbb{N}}$, $\delta_k \downarrow 0$, such that $G_k := \left\{P \in \Omega/d(P,F) < \delta_k\right\}$ are μ-continuity sets. (3) implies $\limsup \mu_n(F) \leq \lim \mu_n(G_k) = \mu(G_k)$ for each k. If F is closed, then $G_k \downarrow F$ and $\mu(G_k) \downarrow \mu(F)$, such that (2) holds. □

Besides the concept of weak convergence there is the concept of "vague" convergence. (Since $\Omega \subset \mathbb{R}^k$ is a metric space, Borel measures are identical with Baire measures - see [Bauer, 40.4, Korollar 2].):

$$\mu_n \to \mu \text{ vaguely if } \int \Phi d\mu_n \to \int \Phi d\mu \text{ for all } \Phi \in C_o(\Omega) ,$$

i.e. for all continuous functions Φ with compact support.
Since we have $\mu_n(\Omega) = \mu(\Omega) = 1$, if follows from [Bauer, 45.7 Satz] that

$$\mu_n \to \mu \text{ weakly if and only if it converges vaguely.}$$

We may therefore add to theorem 1.2 a fourth equivalent statement

(4) $\lim_{n\to\infty} \mu_n = \mu$ in the vague sense .

(4) tells us that testing a sequence for weak convergence it is enough to use C_o as class of test functions. Such a class of functions may be called a "convergence determining class". This notion is used for classes of measurable sets too: A class of measurable sets D is called "convergence determining", if for all sequences μ_n the weak convergence $\mu_n \to \mu$ is equivalent to

$$\mu_n(M) \to \mu(M) \quad \text{for all } \mu\text{-continuity sets } M \in D.$$

In order to construct a simple and useful convergence determining class, we define k-dimensional intervals, for example

$$R_{(\overset{o}{Q},\overset{o}{P}]} := \left\{P \in \Omega/\overset{o}{Q}_i < P_i \leq \overset{o}{P}_i,\ i=1,\ldots,k\right\}$$

or $$R_{\overset{o}{Q}} := \left\{P \in \Omega/P_i \leq \overset{o}{Q}_i,\ i=1,\ldots,k\right\}$$

and $R_{(\overset{o}{Q},\overset{o}{P})}$ ect. correspondingly.

Theorem 1.5: (Billingsley page 17)

The class

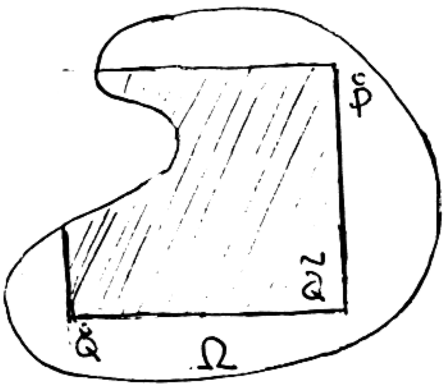
$$\mathcal{R} := \left\{ R_{(\mathring{Q},\mathring{P}]} / \mathring{Q},\mathring{P} \in \Omega \right\}$$

is a convergence determining class.

Proof:

Theorem 1.2(3) tells us that $\mu_n \to \mu$ implies $\mu_n\left(R_{(\mathring{Q},\mathring{P}]}\right) \to \mu\left(R_{(\mathring{Q},\mathring{P}]}\right)$ for all μ-continuity sets $R_{(\mathring{Q},\mathring{P}]}$. Therefore we only have to prove the opposite direction.

Take now an interval $R_{(\mathring{Q},\mathring{P}]}$; its boundary (with respect to Ω) is made of at most 2k hyperplanes, each of dimension k-1. We denote by U the class of those intervals, for which these hyperplanes have μ-measure 0



$R_{(\mathring{Q},\mathring{P}]}$ has altogether 2^k vertices, some of which may lay outside of Ω. For $R_{(\mathring{Q},\mathring{P}]} \in U$, clearly

$$\mu_n\left(R_{(\mathring{Q},\mathring{P}]}\right) \longrightarrow \mu\left(R_{(\mathring{Q},\mathring{P}]}\right)$$

since the boundary is included in these hyperplanes.

Now, U is closed under finite intersections. Therefore, if $R_1,\dots,R_m \in U$, then

$$\mu_n\left(\bigcup_{i=1}^m R_i\right) = \sum_{i=1}^m \mu_n(R_i) - \sum_{i,j=1}^m \mu_n(R_i \cap R_j) + \dots$$

$$\longrightarrow \sum_{i=1}^m \mu(R_i) - \sum_{i,j=1}^m \mu(R_i \cap R_j) + \dots = \mu\left(\bigcup_{i=1}^m R_i\right).$$

In this way, we come from U to the class of all finite union of elements in U. We try to do the jump to open sets G and claim, that every open set is a countable union of elements in U. Take a $\mathring{P} \in \Omega$ and $\varepsilon > 0$ sufficiently small such that the ball $K_\varepsilon(\mathring{P})$ around $\mathring{P}$ with radius $\varepsilon > 0$ is still in Ω. In K_ε there exists an interval $R_{(\mathring{P}-\delta E,\mathring{P}+\delta E]}$, $E=(1,\dots,1)$, which belongs to U; this follows from the fact that considering parallel hyperplanes belonging to different $\delta < \frac{\varepsilon}{\sqrt{k}}$, only countably many can

have positive meassure; we may therefore select a δ such that all these hyperplanes have μ-measure 0 and therefore belong to U.
Now, if G is an open set and $\mathring{P} \in G$ arbitrary, there exists an $R \in U$ such that $\mathring{P} \in \mathring{R} \subset R \subset G$. These $\mathring{R}$ form an open cover of the open set G. In $\mathbb{R}^k$ (as in any separable space) one can select a finite or countably infinite sequence $(R_i)_{i \in \mathbb{N}}$, such that $\bigcup_{i=1}^{\infty} \mathring{R}_i \supset G$; since all $R_i \subset G$, this implies

$$G = \bigcup_{i=1}^{\infty} \mathring{R}_i .$$

Now, given $\varepsilon > 0$, choose m such that $\mu\left(\bigcup_{i=1}^{m} R_i\right) > \mu(G)-\varepsilon$. Then

$$\mu(G)-\varepsilon < \mu\left(\bigcup_{i=1}^{m} R_i\right) = \lim_{n\to\infty} \mu_n\left(\bigcup_{i=1}^{m} R_i\right) \leq \liminf_{n\to\infty} \mu_n(G) .$$

Since $\varepsilon > 0$ was arbitrary, we get

$$\mu(G) \leq \liminf_{n\to\infty} \mu_n(G) \quad \text{for all open sets } G$$

which is, as we mentioned in the proof, equivalent to condition (2) in theorem 1.3.
We conclude that $\mu_n \to \mu$ weakly, i.e. U is a convergence determining in $\mathcal{R}$, therefore $\mathcal{R}$ is itself convergence determining. □

Remarks:

Since any μ-continuity set $R_{\mathring{Q}}$ can be constructed by finite unions or differences of elements of U, even $\tilde{\mathcal{R}} = \left\{R_{\mathring{Q}} \,/\, \mathring{Q} \in \Omega\right\}$ is convergence determining. One calls

$$F(\mathring{Q}) = \mu\left(R_{\mathring{Q}}\right)$$

the "distribution function of μ". We know now that $\mu_n \to \mu$ iff $F_n(\mathring{Q}) \to F(\mathring{Q})$ for all $\mathring{Q}$, for which $R_{\mathring{Q}}$ is a μ-continuity set. We may transform this condition:

Corollary:

$\mu_n \to \mu$ weakly iff $F_n(\mathring{Q}) \to F(\mathring{Q})$ for all continuity points $\mathring{Q}$ of F.

Proof:

We have to show that F is continuous at $\mathring{Q}$ iff $R_{\mathring{Q}}$ is a μ-continuity set.

Now: F continuous at $\overset{o}{Q} \Longleftrightarrow \forall \varepsilon > 0 \ \exists$ open δ-cube $R_{(\overset{o}{Q}-\delta E, \overset{o}{Q}-\delta E)}$ around $\overset{o}{Q}$ such that

$$|F(P)-F(\overset{o}{Q})| < \varepsilon \quad \text{for all } P \in R_{(\overset{o}{Q}-\delta E, \overset{o}{Q}+\delta E)} .$$

F is continuous from above (respectively from below), if this is true for all $P \in R_{[\overset{o}{Q}, \overset{o}{Q}+\delta E)}$ (respectively $P \in R_{(\overset{o}{Q}-\delta E, \overset{o}{Q}]}$).

Clearly, F is nondecreasing (with respect to the semiorder in $\mathbb{R}^k$) and therefore continuous from above if

$$F(\overset{o}{Q}) = \inf\left\{F(\overset{o}{Q}+\delta E)/\delta > 0\right\} .$$

The right hand side gives

$$\inf\left\{\mu\left(\left\{P/P \le \overset{o}{Q}+\delta E\right\}\right)/\delta > 0\right\}$$

$$= \mu\left(\bigcap_{\delta > 0} \left\{P/P \le \overset{o}{Q}+\delta E\right\}\right) = \mu\left(\left\{P/P \le \overset{o}{Q}\right\}\right) ,$$

which is $F(\overset{o}{Q})$; therefore F is always continuous from above.

Again, since F is nondecreasing, it is enough to show that additionally F is continuous from below.

Continuity from below is equivalent to $F(\overset{o}{Q}) = \sup\left\{F(\overset{o}{Q}-\delta E)/\delta > 0\right\}$.

Now

$$\sup\left\{F(\overset{o}{Q}-\delta E)/\delta > 0\right\} = \mu\left(\bigcup_{\delta > 0} \left\{P/P \le \overset{o}{Q}-\delta E\right\}\right) = \mu\left(\left\{P/P < \overset{o}{Q}\right\}\right) .$$

This last expression differs from $F(\overset{o}{Q})$ exactly by $\mu(\delta R_{\overset{o}{Q}})$, which is zero iff $R_{\overset{o}{Q}}$ is a μ-continuity set. □

Now the question arises, which densities can be approximated by ensembles or more general, which measures can be approximated by discrete measures.

Theorem 1.3 (see [Bauer, 45.4]):
The discrete measures are dense in $M(\Omega)$.

Proof:

We may use the concept of vague convergence and have to show that for any $\overset{o}{\mu} \in \mathcal{M}$ and for an arbitrary finite set $\{\phi_1,\ldots,\phi_n\} \subset C_o(\Omega)$ and for each $\varepsilon>0$ there exists

$$\delta^{\alpha}_{\omega_N} = \sum_{j=1}^{N} \alpha_j \delta_{P_j}$$

such that

$$\left| \int \phi_i d\overset{o}{\mu} - \int \phi_i d\delta^{\alpha}_{\omega_N} \right| < \varepsilon \ , \quad i=1,\ldots,n \ .$$

Take a compact set K such that $\bigcup_{i=1}^{n}$ supp $\phi_i \subset K$ and choose $\eta>0$ so that $\eta\overset{o}{\mu}(K)<\varepsilon$. For each $Q \in K$ there exists a neighbourhood $U(Q)$, such that

$$|\phi_i(Q')-\phi_i(Q'')| \leq \eta \quad \text{for } Q',Q'' \in \overline{U(Q)} \ , \quad i=1,\ldots,n \ .$$

$\{U(Q) | Q \in K\}$ is an open covering of K; K is compact and therefore

$$K \subset U(Q_1) \subset \ldots \subset U(Q_k) \quad \text{for certain } Q_1,\ldots,Q_k \ .$$

Put

$$A_1 = K \cap \overline{U(Q_1)}, A_2 = K \cap \overline{U(Q_2)} \setminus A_1,\ldots,A_k := K \cap \overline{U(Q_k)} \setminus A_1\cup\ldots\cup A_{k-1} \ ;$$

then $A_1,\ldots,A_k$ are pairwise disjoint Borel sets with

$$K = A_1 \cup \ldots \cup A_k$$

and $|\phi_i(Q')-\phi_i(Q'')|\leq\eta$ if $Q',Q'' \in A_j$ for arbitrary $j=1,\ldots,k$, $i=1,\ldots,n$. We choose now $P_j \in A_j$ arbitrary and consider

$$\delta' = \sum_{j=1}^{k} \overset{o}{\mu}(A_j)\delta_{P_j} \ .$$

δ' is a discrete but not normalized measure since

$$\delta'(\Omega) = \sum_{j=1}^{k} \overset{o}{\mu}(A_j) = \overset{o}{\mu}(K)$$

which may be less than $\overset{o}{\mu}(\Omega) = 1$, if $K\neq\Omega$. In this (generic) case we choose an arbitrary $P_{k+1} \in \Omega\setminus K$, $\alpha_j=\overset{o}{\mu}(A_j)$, $j=1,\ldots,k$ and $\alpha_{k+1} = 1-(\alpha_1+ \ldots +\alpha_k)$. With $k+1 = N$ the discrete measure

$$\delta^{\alpha}_{\omega_N} = \sum_{j=1}^{N} \alpha_j \delta_{P_j}$$

has the property we want:

Since $P_{k+1} \notin K$ and $\bigcup_{i=1}^{n} \operatorname{supp} \Phi_i \subset K$ we have $\Phi_i(P_{k+1}) = 0$; therefore

$$\left|\int \Phi_i d\mathring{\mu} - \int \Phi_i d\delta^{\alpha}_{\omega_N}\right| = \left|\int \Phi_i d\mathring{\mu} - \int \Phi_i d\delta'\right|$$

$$= \left| \sum_{j=1}^{k} \int_{A_j} \Phi_i d\mathring{\mu} - \sum_{j=1}^{k} \mathring{\mu}(A_j)\Phi_i(P_j)\right| = \left| \sum_{j=1}^{k} \int_{A_j} (\Phi_i(P) - \Phi_i(P_j))d\mathring{\mu}\right|$$

$$\leq \sum_{j=1}^{k} \int_{A_j} \left|\Phi_i(P) - \Phi_i(P_j)\right| d\mathring{\mu}(P) \leq \eta \sum_{j=1}^{k} \mathring{\mu}(A_j) = \eta\mathring{\mu}(K) < \varepsilon .$$

□

How much do we loose, if we wish all weights to be equal: Which measures μ can be approximated by discrete measures $\delta_{\omega_N} = \frac{1}{N} \sum_{j=1}^{N} \delta_{P_j}$? A theorem of Niederreiter gives the answer (H. Niederreiter, Compositio Matematico 25, p. 93-99, 1972 - he proves it under the assumption that Ω is compact, but we don't need this assumption).

Lemma:

Let $\mu = \sum_{j=1}^{M} \alpha_j \delta_{Q_j}$ be a discrete measure. Then there exists a sequence $(P_j)_{j\in\mathbb{N}}$ such that for all $N \in \mathbb{N}$ and all measurable sets $A \subset \Omega$ we get

$$\left|\frac{1}{N} \sum_{j=1}^{N} \delta_{P_j}(A) - \mu(A)\right| \leq \frac{C(\mu)}{N} ,$$

where for example $C(\mu) = (M-1)[\frac{M}{2}]$.

Proof:

We show first that (P_j) exists such that

$$\left|\delta_{\omega_N}(\{Q_k\}) - \mu(\{Q_k\})\right| \leq \frac{M-1}{N}$$

for all k, $1 \leq k \leq M$ and all N. We do it by induction with respect to M. If M=1, $\mu = \delta_Q$ and if we put $P_j = Q$, we get a zero difference.

Assume we have proved the lemma for all discrete measures whose support consists of M points. Take now $\mu = \sum_{j=1}^{M+1} \lambda_j \delta_{Q_j}$ $(\lambda_j > 0, \sum_{j=1}^{M+1} \lambda_j = 1)$.

Define

$$\nu = \sum_{j=1}^{M} \frac{\lambda_j}{\lambda_1 + \ldots + \lambda_M} \delta_{Q_j} ;$$

by assumption we have a sequence $(R_n)_{n \in \mathbb{N}}$, such that

$$\left| \frac{1}{N} \sum_{n=1}^{N} \delta_{R_n}(Q_j) - \nu(Q_j) \right| \leq \frac{M-1}{N} \quad \forall N,\ j=1,\ldots,M .$$

Since $\lambda_1 + \ldots + \lambda_M < 1$, we get $\beta := \frac{1}{\lambda_1 + \ldots + \lambda_M} > 1$ and the equation $n = [m\beta]$ for given n has at most one solution m. If there is such m, we put $P_n = R_m$ - if not, we put $P_n = Q_{M+1}$. This sequence $(P_n)_{n \in \mathbb{N}}$ fulfills $|\delta_{\omega_N}(\{Q_k\}) - \mu(\{Q_k\})| \leq \frac{M}{N}$ for $k=1,\ldots,M+1$:

Consider first Q_i for some i, $1 \leq i \leq M$. Then

$\sum_{n=1}^{N} \delta_{P_n}(\{Q_i\})$ is equal to the number of natural numbers m

with $[m\beta] \leq N$ and $R_m = Q_i$; but this number is equal to

$$\sum_{k=1}^{L} \delta_{R_k}(\{Q_i\})$$

with $L = \max\{m / [m\beta] \leq N\}$. If $\frac{N+1}{\beta} \in \mathbb{N}$, then $L = \frac{N+1}{\beta} - 1 = [\frac{N+1}{\beta}] - 1$ (since $[\frac{N+1}{\beta}\beta] = N+1$ and not $\leq N$); if $\frac{N+1}{\beta} \notin \mathbb{N}$, we get $L = [\frac{N+1}{\beta}]$.

Now

$$\left| \frac{1}{N} \sum_{n=1}^{N} \delta_{P_n}(\{Q_i\}) - \underbrace{\sum_{j=1}^{M+1} \lambda_j \delta_{Q_j}(\{Q_i\})}_{\lambda_i = (\lambda_1 + \ldots + \lambda_M)\nu(\{Q_i\})} \right|$$

$$= \left| \frac{1}{L} \sum_{k=1}^{L} \delta_{R_k}(\{Q_i\}) \frac{L}{N} - (\lambda_1 + \ldots + \lambda_M)\nu(\{Q_i\}) \right|$$

$$\leq \frac{L}{N}\left|\frac{1}{L}\sum_{k=1}^{L}\delta_{R_k}(\{Q_i\}) - \nu(\{Q_i\})\right| + \nu(\{Q_i\})\left|\frac{L}{N} - (\lambda_1+\ldots+\lambda_M)\right|$$

$$\underbrace{\qquad\qquad\qquad}_{\leq \frac{M-1}{L}}$$

$$\leq \frac{M-1}{N} + \nu(\{Q_i\})\left|\frac{[(N+1)(\lambda_1+\ldots+\lambda_M)]-\varepsilon}{N} - (\lambda_1+\ldots+\lambda_M)\right|$$

where $\varepsilon=0$ or 1.

We calculate that, if $\frac{N+1}{\beta} \in \mathbb{N}$, the last expression gives

$$\left|\frac{(N+1)(\lambda_1+\ldots+\lambda_M)-1}{N} - \frac{N(\lambda_1+\ldots+\lambda_M)}{N}\right| = \left|\frac{(\lambda_1+\ldots+\lambda_M)-1}{N}\right| \leq \frac{1}{N}$$

and if $\frac{N+1}{\beta} \notin N$, we get again

$$\left|\frac{[(N+1)(\lambda_1+\ldots+\lambda_M)] - N(\lambda_1+\ldots+\lambda_M)}{N}\right| \leq \frac{1}{N}$$

(since $[(N+1)(\lambda_1+\ldots+\lambda_M)] = [N(\lambda_1+\ldots+\lambda_M)]$ or $[N(\lambda_1+\ldots+\lambda_M)]+1$);

we end up with

$$\left|\frac{1}{N}\sum_{n=1}^{N}\delta_{P_n}(\{Q_i\})-\mu(\{Q_i\})\right| \leq \frac{M-1}{N} + \frac{1}{N} = \frac{M}{N}\ .$$

It remains to show the same estimate for Q_{M+1}. We have $P_n = Q_{M+1}$ for N-L points P_n, i.e.

$$\frac{1}{N}\sum_{n=1}^{N}\delta_{P_n}(\{Q_{M+1}\}) = \frac{N-L}{N} = 1-\frac{L}{N}$$

and therefore

$$\left|\frac{1}{N}\sum_{n=1}^{N}\delta_{P_n}(\{Q_{M+1}\}) - \mu(\{Q_{M+1}\})\right| = \left|1-\frac{L}{N}-(1-(\lambda_1+\ldots+\lambda_M))\right|$$

$$= \left|-\frac{L}{N}+(\lambda_1+\ldots+\lambda_M)\right| \leq \frac{1}{N}$$

as above.

So we have proved the lemma for $A = \{Q_i\}$, where instead of $C(\mu)$ we had M-1.

Take now an arbitrary measurable set $A \subset \Omega$, then for all discrete measures μ we have $\mu(A) = \mu(A \cap \mathrm{supp}(\mu))$. We can restrict ourselves to sets $A \subset \mathrm{supp}\,\mu$. Since $\mu(\mathrm{supp}\,\mu \setminus A) = 1-\mu(A)$, we can restrict ourselves even to those sets A with card $A \leq [\frac{M}{2}]$. Therefore

$$\left|\frac{1}{N}\sum_{n=1}^{N} \delta_{P_n}(A)-\mu(A)\right| \leq \frac{M-1}{N} \text{ card } A \leq [\frac{M}{2}]\frac{M-1}{N} .$$

□

Theorem 1.4:

To each $\mu \in \mathcal{M}$ there exists a sequence $(P_n)_{n\in\mathbb{N}}$ such that

$$\frac{1}{N}\sum_{n=1}^{N} \delta_{P_n}$$

converges weakly to μ.

Proof:

According to theorem (1.2) we have to show that

$$\frac{1}{N}\sum_{n=1}^{N} \delta_{P_n}(M) \longrightarrow \mu(M)$$

for all μ-continuity sets M.

From theorem 1.3 we know the existence of a sequence $(\mu_j)_{j\in\mathbb{N}}$ of discrete measures, which converges to μ:

$$\mu_j \to \mu \quad \text{weakly}$$

The lemma tells us that for each μ_j we have a sequence $(P^j_n)_{n\in\mathbb{N}}$ such that

$$\left|\frac{1}{N}\sum_{n=1}^{N} \delta_{P^j_n}(A)-\mu_j(A)\right| \leq \frac{C(\mu_j)}{N} .$$

Set $C_j := C(\mu_j)$ and choose a natural number $r_j \geq j(C_1+\ldots+C_{j+1})$, $r_o=0$.

To construct (P_n) we represent each n by

$$n = r_o+r_1+\ldots+r_{j-1}+s \quad \text{with} \quad 0<s\leq r_j$$

(j and s is uniquely defined). Put $P_n := P^j_s$, i.e.

$$P_1 = P^1_1,\ldots,P_{r_1} = P^1_{r_1},\ P_{r_1+1} = P^2_1,\ldots,P_{r_1+r_2} = P^2_{r_2},\ldots .$$

For a μ-continuity set M we have to calculate $\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \delta_{P_n}(M)$. We take $N \geq r_1$ and decompose

$$N = r_1+r_2+\ldots+r_k+s, \quad 0<s\leq r_{k+1} .$$

Then

$$\sum_{n=1}^{N} \delta_{P_n}(M) = \delta_{P_1^1}(M)+\ldots+\delta_{P_{r_1}^1}(M)+\ldots+\delta_{P_1^k}(M)+\ldots+\delta_{P_{r_k}^k}(M)+$$

$$\ldots+\delta_{P_1^{k+1}}(M)+\ldots+\delta_{P_s^{k+1}}(M)$$

$$= \sum_{j=1}^{k} \left(\sum_{i=1}^{r_j} \delta_{P_i^j}(M) \right) + \sum_{i=1}^{s} \delta_{P_i^{k+1}}(M)$$

and we get

$$\left| \frac{1}{N} \sum_{n=1}^{N} \delta_{P_n}(M) - \mu(M) \right| = \left| \sum_{j=1}^{k} \frac{r_j}{N} \left(\frac{1}{r_j} \sum_{i=1}^{r_j} \delta_{P_i^j}(M) - \mu_j(M) \right) \right.$$

$$\left. + \frac{s}{N} \left(\frac{1}{s} \sum_{i=1}^{s} \delta_{P_i^{k+1}}(M) - \mu_{k+1}(M) \right) + \sum_{j=1}^{k} \frac{r_j}{N} \mu_j(M) + \frac{s}{N} \mu_{k+1}(M) - \mu(M) \right|$$

$$\leq \sum_{j=1}^{k} \frac{r_j}{N} \frac{C_j}{r_j} + \frac{s}{N} \frac{C_{k+1}}{s} + \left| \frac{1}{N} \left(\sum_{j=1}^{k} r_j \mu_j(M) + s\mu_{k+1}(M) \right) - \mu(M) \right|$$

$$\leq \underbrace{\frac{1}{r_k} \sum_{j=1}^{k+1} C_j}_{\leq \frac{1}{r_k} \cdot \frac{r_k}{k} \to 0} + \left| \frac{1}{N} \left[\sum_{j=1}^{k} r_j \mu_j(M) + s\mu_{k+1}(M) \right] - \mu(M) \right|$$

with $N\to\infty$, since then $k\to\infty$ and $\frac{1}{k} \to 0$.

In order to show that $\lim_{N\to\infty} \left| \frac{1}{N} \left(\sum_{j=1}^{k} r_j \mu_j(M) + s\mu_{k+1}(M) \right) - \mu(M) \right| = 0$

we realize the following: We restrict to $N=r_1+\ldots+r_k+r_{k+1}$, then

$$\frac{1}{N}\left(\sum_{j=1}^{k} r_j\mu_j(M) + s\mu_{k+1}(M)\right) = \frac{1}{r_1+\ldots+r_{k+1}} \sum_{j=1}^{k+1} r_j\mu_j(M)$$

$$= \sum_{j=1}^{k+1} C_j^{k+1}\mu_j(M) \quad \text{with} \quad C_j^{k+1} = \frac{r_j}{r_1+\ldots+r_{k+1}} .$$

We have to prove: If $\mu_j(M) \to \mu(M)$ and $\sum_{j=1}^{k} C_j^k = 1$ for all k, then $\sum_{j=1}^{k} C_j^k\mu_j(M) \to \mu(M)$.

This is a generalization of the theorem of Cauchy about the convergence of arithmetic means: If $\sum_{j=1}^{k} C_j^k = C_k \to 1$, $C_j^k \geq 0$ and $C_j^k \downarrow 0$ for $k\to\infty$, then $\sum_{j=1}^{k} C_j^k x_j \to x$ follows from $x_j \to x$. (K.Knopp, Theorie und Anwendung der unendlichen Reihe, 43, Satz 5).

This is, what we need and everything is done. □

We have defined, what we mean by (weak or vague) convergence, we have given convergence criteria and we have discussed the density of (special) discrete measures in $M(\Omega)$. What we need for numerical purposes are distances of measures, especially distances of particle ensembles to densities. Probability theory supplies us with some different concepts of distances, who generate weak convergence. A metric $\rho(\mu,\nu)$ in $M(\Omega)$ generates weak convergence, if

$$\mu_n \longrightarrow \mu \quad \text{iff} \quad \rho(\mu_n,\mu) \longrightarrow 0 .$$

The first metric of $M(\Omega)$ was introduced by Prohorov (see "Convergence of Random Processes and Limit Theorems in Probability Theory", Theory of Probability and its Applications 1 (1956), pp. 157-214). It generates the convergence, is however not easy to handle and was therefore never used (as far as we know) for numerical purposes. It may seem that

$$d_{TV}(\mu,\nu) := \sup\left\{|\mu(A)-\nu(A)| \;/\; A \text{ measurable set in } \Omega\right\} ,$$

which is called distance of total variation, is a simple choice. But this distance doesn't generate the weak convergence, which is easily seen by choosing $\mu=\delta_P$, $\nu=\delta_Q$, $P\neq Q$, which gives trivially

$$d_{TV}(\delta_P,\delta_Q) = 1 ,$$

i.e. it doesn't tend to zero when $P\to Q$.

Other ideas are needed and the tool for defining useful distances is the notion of a uniformity class: A subset U of $C^b(\Omega)$ is called uniformity class, if $\mu_n \to \mu$ implies

$$\sup_{\phi \in U} \left| \int \phi d\mu_n - \int \phi d\mu \right| \longrightarrow 0 .$$

If U is large enough such that it is convergence determining (as a class of functions, which means that

$$\int \phi d\mu_n \longrightarrow \int \phi d\mu \quad \forall \phi \in U \quad \text{implies } \mu_n \to \mu) ,$$

then

$$d(\mu,\nu) = \sup_{\phi \in U} \left| \int \phi d\mu - \int \phi d\nu \right|$$

is a possible choice as a convergence generating metric.

We cite a result by Ranga Rao ("Some Theorems on Weak Convergence of Measures and Applications", Ann. of Math. Statistics 33 (1962), pp. 659-680). If U is an equicontinuous, uniformly bounded class of functions in $C^b(\Omega)$, then U is a uniformity class. (We use that our basic set Ω with the Euclidean metric is a separable metric space.) For example, the class of all Lipschitz continuous functions with a uniform Lipschitz constant, i.e.

$$\left\{\phi \in C^b(\Omega) \;/\; |\phi(P)-\phi(Q)| \leq L\|P-Q\|\right\} , \text{ L fixed}$$

is an equicontinuous set.
We define

$$\|\phi\|_L := \sup_{P \neq Q} \frac{|\phi(P)-\phi(Q)|}{\|P-Q\|}$$

to be the infimum of all possible Lipschitz constants of ϕ and the Lipschitz norm

$$\|\phi\|_{BL} := \|\phi\|_\infty + \|\phi\|_L$$

(where $\| \ \|_\infty$ denotes the usual $\mathbf{L}^\infty$-norm).

Then $D := \left\{\phi \in C^b \;/\; \|\phi\|_{BL} \leq 1\right\}$ is a uniformity class and we may define

$$d_{BL}(\mu,\nu) = \sup_{\phi \in D} \left| \int \phi d\mu - \int \phi d\nu \right| .$$

Very similar are definitions using

$$D_1 := \left\{\phi \in C^b \ / \ \max\left\{\|\phi\|_\infty, \|\phi\|_L\right\} \le 1\right\}$$

or

$$D_2 := \left\{\phi \in C^b \ / \ 0 \le \phi(P) \le 1 \quad \text{and} \quad \|\phi\|_L \le 1\right\} ,$$

which create $d_{BL}^{(1)}$ and $d_{BL}^{(2)}$. It is easy to prove that d_{BL}, $d_{BL}^{(1)}$ and $d_{BL}^{(2)}$ are equivalent metrics. For example,

$$d_{BL} \le d_{BL}^{(1)} \le 2d_{BL} \ :$$

$\phi \in D$ means $\|\phi\|_\infty + \|\phi\|_L \le 1$; it follows that $\|\phi\|_\infty \le 1$ and $\|\phi\|_L \le 1$, i.e. $\max\left\{\|\phi\|_\infty, \|\phi\|_L\right\} \le 1$ and $\phi \in D^{(1)}$. Therefore $D \subset D^{(1)}$ and

$$d_{BL}(\mu,\nu) = \sup_{\phi \in D} \left|\int \phi d\mu - \int \phi d\nu\right| \le \sup_{\phi \in D^{(1)}} \left|\int \phi d\mu - \int \phi d\nu\right| = d_{BL}^{(1)}(\mu,\nu) \ .$$

If $\phi \in D^{(1)}$, we get $\|\phi\|_\infty \le 1$ and $\|\phi\|_L \le 1$ such that $\|\phi\|_{BL} \le 2$; $\phi' = \frac{\phi}{2}$ is therefore in D and

$$d_{BL}(\mu,\nu) = \sup_{\Psi \in D} \left|\int \Psi d\mu - \int \Psi d\nu\right| \ge \left|\int \phi' d\mu - \int \phi' d\nu\right| = \frac{1}{2}\left|\int \phi d\mu - \int \phi d\nu\right| \ .$$

It follows that $\frac{1}{2} \sup_{\phi \in D^{(1)}} \left|\int \phi d\mu - \int \phi d\nu\right| \le d_{BL}(\mu,\nu)$.

Similar estimates hold for all distances. We are therefore free to use the one of the three distances, which is most appropriate for our purposes and we will denote it always by d_{BL}.
We see for example that

$$d_{BL}(\delta_P, \delta_Q) = \|P-Q\|, \text{ if } \|P-Q\| \text{ is small enough}$$

- we have some hope that d_{BL} generates the weak convergence.

Theorem 1.6:

$$d_{BL}(\mu,\nu) = \sup\left\{\left|\int \phi d\mu - \int \phi d\nu\right| \ / \ \phi \in C^b \text{ with } \|\phi\|_\infty + \|\phi\|_L \le 1\right\}$$

is a metric in $\mathcal{M}(\Omega)$, which generates the weak convergence.

Proof:

The usual proof uses the Prohorov metric, but we want to ... . We follow therefore H.G. Kellerer ("Markov-Komposition und eine Anwendung

auf Martingale", Math. Ann. 198 (1972), pp. 99-122), which is based on $d_{BL}^{(2)}$. We use (without proof) a lemma:
If $\phi \in C^b$, $0 \le \phi \le 1$, then there exists an increasing sequence $(u_n)_{n\in\mathbb{N}}$ and a decreasing sequence $(o_n)_{n\in\mathbb{N}}$ with

$$\frac{u_n}{n}, \frac{o_n}{n} \in D^{(2)} \quad \text{and} \quad u_n \uparrow \phi,\ o_n \downarrow \phi \ .$$

We have now to show that

$$\int \phi d\mu_n \to \int \phi d\mu \ \forall \phi \in D^{(2)} \quad \text{implies} \quad \int \phi d\mu_n \to \int \phi d\mu \ \forall \phi \in C^b \ .$$

If $\phi \in C^b$ we may assume that $|\phi| \le 1$ and decompose

$$\Phi = \Phi_+ - \Phi_- \quad \text{with} \quad 0 \le \Phi_\pm \le 1,\ \Phi_\pm \in C^b \ .$$

It is enough to show the convergence for Φ_+. Φ_+ may be approximated from below and from above by (u_m) and (o_m):

$$\begin{aligned}\int \Phi_+ d\mu_n - \int \Phi_+ d\mu &\le \int o_m d\mu_n - \int u_m d\mu \\ &= \left(\int o_m d\mu_n - \int o_m d\mu\right) + \left(\int o_m d\mu - \int u_m d\mu\right)\end{aligned}$$

The theorem of monotoneous convergence tells us that

$$0 \le \int 0_m d\mu - \int u_m d\mu < \frac{\varepsilon}{2} \quad \text{for } m \ge M.$$

Since $\frac{o_M}{M} \in D^{(2)}$ and $d_{BL}^{(2)}(\mu_n,\mu) \to 0$, we get

$$\left|\int \frac{o_M}{M}\, d\mu_n - \int \frac{o_M}{M}\, d\mu\right| < \frac{\varepsilon}{2M} \quad \text{for } n \ge N_1 \ .$$

This shows that

$$\int \Phi_+ d\mu_n - \int \Phi_+ d\mu \le \varepsilon \quad \text{for } n \ge N_1$$

and similarly we get

$$\int \Phi_+ d\mu - \int \Phi_+ d\mu_n \le \varepsilon \quad \text{for } n \ge N_2 \ .$$

□

We want to mention a completely different approach due to Wasserstein, which was used for particle methods in some papers by Dobrushin.
The definition runs as follows:
We consider Borel measure on $\Omega \times \Omega$ and their projection onto $M(\Omega)$:
For $P \in M(\Omega\times\Omega)$ define $\mu_P \in M(\Omega)$ by $\mu_P(A) = P(A,\Omega)$ and $\nu_P \in M(\Omega)$ by

For $P \in M(\Omega\times\Omega)$ define $\mu_P \in M(\Omega)$ by $\mu_P(A) = P(A,\Omega)$ and $\nu_P \in M(\Omega)$ by $\nu_P(B) = P(\Omega,B)$ for all measurable $A,B \subset \Omega$. Given $\mu,\nu \in M(\Omega)$ we denote by $P[\mu,\nu]$ the set of all measures in $M(\Omega\times\Omega)$, whose projections are the given μ,ν:

$$P[\mu,\nu] := \left\{P \in M(\Omega\times\Omega) \;/\; \mu_P=\mu, \nu_P=\nu\right\}$$

$P[\mu,\nu]$ contains the product measure $\mu \otimes \nu$, but certainly more.
The Wasserstein metric is now defined by

$$d_W(\mu,\nu) := \inf\left\{\int \|x-y\| dP(x,y) \;/\; P \in P[\mu,\nu]\right\} .$$

This definition seems quite complicated; it is however quite easy to use in convergence proofs as mentioned - this was done by Dobrushin (...). d_W was also considered by Kantorovich and Rubinstein - it generates the weak convergence in $M(\Omega)$. The relation between d_{BL} and d_W is investigated in a book by Dudley.

One may ask about the possibility of really computing these distances numerically - our aim is to develop numerical methods, and we should therefore be able to compute the error. There are in fact some steps in this direction concerning d_{BL}; since we will - for practical purposes - mainly use another distance concept originally from number theory we omit a further discussion of d_{BL} and d_W here.

§ 2 Numbertheoretic aspects

Numbertheory supplies a completely different concept for the relation between densities and particle ensembles: It is the concept of the asymptotic distribution of a sequence and originated from a paper by Herman Weyl in 1916: "Über die Gleichverteilung von Zahlen modulo Eins" (Math. Annalen 77, p. 313-352), which was concerned with Kronecker's approximation theorem, i.e. with diophantine approximation. Our main sources are: L. Kuipers, H. Niederreiter: Uniform Distribution of Sequences, Wiley 1974; I.M. Sobol: Punkte, die einen mehrdimensionalen Würfel gleichmäßig ausfüllen, MIR 1985, Moskau (Russian).

We describe the concept first in the <u>most simple case</u> - k=1 and $f(x)=\chi_{[0,1]}(x)$, where χ is again the characteristic function (μ is here the Lebesgue measure restricted to [0,1]). The generalization to higher dimenions and to more general f will then be relatively easy. We take a sequence $(\omega_N)_{N\in\mathbb{N}}$

$$\omega_N = (x_1^N,\dots,x_N^N),\ N \in \mathbb{N}\ \text{(the weights are here always } \tfrac{1}{N})$$

and want to define the convergence of ω_N to $\chi_{[0,1]}$.

We take an interval [a,b), where $0\leq a<b\leq 1$ (if b=1, we take [a,b]) and count the number of points of ω_N in [a,b)

$$\Lambda([a,b);\omega_N) := \sum_{i=1}^{N} \chi_{[a,b)}(x_i^N) \ ;$$

the relative frequency of the points in [a,b) is given by

$$\frac{1}{N}\Lambda([a,b);\omega_N) \ ,$$

and we call the sequence $(\omega_N)_{N\in\mathbb{N}}$ "uniformly distributed" (u.d.), if

$$\lim_{N\to\infty} \frac{1}{N}\Lambda([a,b);\omega_N) = b-a = \int_0^1 \chi_{[a,b)}(x)dx$$

for all $0\leq a<b\leq 1$.

If $(x_i)_{i\in\mathbb{N}}$ is a sequence of points and $\omega_N=(x_1,\dots,x_N)$, we say that (ω_N) is generated by (x_i) and we call $(x_i)_{i\in\mathbb{N}}$ u.d., if $(\omega_N)_{N\in\mathbb{N}}$ has this property.

<u>Examples:</u>

1) $(\frac{1}{2N},\frac{3}{2N},\dots,\frac{2N-1}{2N})$ is a sequence of particle ensembles not generated just by a sequence of points; we will soon see that it is u.d.

2) Let $(x_i)_{i\in\mathbb{N}}$ be the sequence $0,\frac{1}{2}, \frac{1}{4},\frac{3}{4}, \frac{1}{8},\frac{3}{8},\frac{5}{8},\frac{7}{8}, \frac{1}{16},\dots$ We consider $N = 2^{k-1}+2^k$, $k=1,2,\dots$ and $[a,b) = [0,\frac{1}{2})$.

Then

$$\Lambda\left([0,\tfrac{1}{2});\omega_3\right) = 2, \quad \Lambda\left([0,\tfrac{1}{2}),\omega_6\right) = 4,$$

$$\Lambda\left([0,\tfrac{1}{2}),\omega_{2^{k-1}+2^k}\right) = \frac{2^k}{2}+2^{k-1} = 2\cdot 2^{k-1}$$

and therefore

$$\frac{1}{2^{k-1}+2^k}\,\Lambda\left([0,\tfrac{1}{2}),\omega_{2^{k-1}+2^k}\right) = \frac{2\cdot 2^{k-1}}{3\cdot 2^{k-1}} = \frac{2}{3} \not\to \frac{1}{2}\ .$$

$(x_i)_{i\in\mathbb{N}}$ is therefore not u.d. One realizes that the uniform distribution is an asymptotic property - our sequence looks uniformly distributed for $N=2^k$, but is asymmetric inbetween, for $N=2^{k-1}+2^k$.

3) If we define a sequence $(x_i)_{i\in\mathbb{N}}$ by using the dual representation of i:

$$i = \ell_1+\ell_2 2^1+ \dots +\ell_m 2^{m-1}, \quad \ell_k = 0,1$$

and put

$$x_i = p(i) = \ell_1 2^{-1}+\ell_2 2^{-2}+ \dots +\ell_m 2^{-m}\ ,$$

then p(i) is the so-called "van der Corput" series and is u.d. as we shall soon see.

Remark:

Theorem 1.5 tells us that $\left\{(a,b],-\infty<a<b<\infty\right\}$ is a convergence determining class. If we put $\mu_N = \frac{1}{N}\sum_{i=1}^{N}\delta_{x_i^N}$ and $\mu = \lambda|_{[0,1]}$, we see that $\mu_N \to \mu$ iff

$$\frac{\Lambda((a,b],\omega_N)}{N} \longrightarrow \lambda_{[0,1]}((a,b])\ .$$

Since outside [0,1] μ_N vanishes as does μ, we may restrict our consideration to $0\le a<b\le 1$; moreover, it doesn't make a difference whether we consider intervals (a,b] or [a,b). We can therefore easily connect both concepts: $(\omega_N)_{N\in\mathbb{N}}$ is u.d. iff δ_{ω_N} converges weakly to $\lambda|_{[0,1]}$.

Consequence: $(\omega_N)_{N\in\mathbb{N}}$ is u.d. iff

$$\lim_{N\to\infty} \frac{1}{N}\sum_{i=1}^{N} \Phi(x_i^N) = \int_0^1 \Phi(x)dx$$

for every real valued functions, continuous on [0,1].
It is not complicated to extend the last convergence to all Riemann-integrable functions. However, if we try to do the same for $\Phi \in \mathbb{L}^1[0,1]$ we would fail. Extensions in this direction will be given in §3, where we treat the functionalanalytic aspects.
If we consider Φ as test function for convergence, we want to test as few functions as possible. In this direction the Weyl criterion is famous:

$$(\omega_N)_{N\in\mathbb{N}} \text{ is u.d. iff } \lim_{N\to\infty} \frac{1}{N}\sum_{j=1}^{N} e^{2\pi i k x_j^N} = 0 \quad \forall\, k \in \mathbb{Z},\ k \neq 0 .$$

__Proof:__
Clearly, the u.d. of $(\omega_N)_{N\in\mathbb{N}}$ implies that

$$\frac{1}{N}\sum_{j=1}^{N} \cos 2\pi k x_j^N \to \int_0^1 \cos 2\pi kx dx = 0$$

and the same relation for $\sin 2\pi kx$, which shows that the condition is necessary.
To prove sufficiency, we take first a complex valued continuous function on $\mathbb{R}$ with period 1. For $\varepsilon > 0$ arbitrary, the Weierstraß approximation theorem tells us that there exists a trigonometric polynomial

$$\Phi(x) = \sum_{|h|\le H} \Phi_h e^{2\pi i h x}$$

such that

$$\sup_{0\le x\le 1} |f(x)-\Phi(x)| \le \varepsilon ;$$

then

$$|\int_0^1 f(x)dx - \frac{1}{N}\sum_{j=1}^{N} f(x_j^N)|$$

$$\le \int_0^1 \left|f(x) - \Phi(x)\right|dx + \left|\int_0^1 \Phi(x)dx - \frac{1}{N}\sum_{j=1}^{N}\Phi(x_j^N)\right| + \frac{1}{N}\sum_{j=1}^{N}\left|\Phi(x_j^N) - f(x_j^N)\right|$$

$$\le 2\varepsilon + \left|\int_0^1 \Phi dx - \frac{1}{N}\sum_{k=1}^{N} \Phi(x_j^N)\right| \le 3\varepsilon \quad \text{for N sufficient large.}$$

To get rid of the periodicity condition $f(0) = f(1)$, we have to approximate an arbitrary Φ by a periodic f in $\mathbb{L}^\infty[0,1]$ and do another estimation of the above kind.
The Weyl criterion tells us that u.d. is equivalent to cancellation, if we interpret x_j^N as frequencies of signals with amplitude 1.

It is now important that numbertheory gives a definition for another distance - the discrepancy. We call

$$\sup_{0\le a<b\le 1} \left|\frac{\Lambda([a,b);\omega_N)}{N} - (b-a)\right| = D(\omega_N)$$

the "discrepancy of ω_N". An important result tells us that $(\omega_N)_{N\in\mathbb{N}}$ is u.d. iff $\lim D(\omega_N) = 0$, i.e. the convergence of $\dfrac{\Lambda([a,b);\omega_N)}{N} \longrightarrow (b-a)$ is automatically uniform.

<u>**Proof:**</u>
It is obvious that (ω_N) is u.d. if $D(\omega_N) \to 0$. Vice versa, assume that $\dfrac{\Lambda([a,b);\omega_N)}{N} \to (b-a)$; we want to show that this convergence is uniform with respect to a,b. Choose $m\ge 2$ and $I_k := [\frac{k}{m},\frac{k+1}{m})$, $k=0,\dots,m-1$.
By our assumption, there exists a $N_o=N_o(m)$, such that

$$\frac{1}{m}(1-\frac{1}{m}) \le \frac{\Lambda(I_k;\omega_N)}{N} \le \frac{1}{m}(1+\frac{1}{m}),\ k=0,\dots,m-1,\ N\ge N_o .$$

Take $0\le a<b\le 1$ arbitrary, $I=[a,b)$. We can certainly construct finite unions of intervals I_k, namely J_1 and J_2 such that

$$J_1 \subset I \subset J_2 \quad \text{and} \quad \lambda(I)-\lambda(J_1) < \frac{2}{m},\ \lambda(J_2)-\lambda(I) < \frac{2}{m} .$$

For $N\ge N_o$ we get

$$\lambda(J_1)(1-\frac{1}{m}) \le \frac{\Lambda(J_1;\omega_N)}{N} \le \frac{\Lambda(I;\omega_N)}{N} \le \frac{\Lambda(J_2;\omega_N)}{N} \le \lambda(J_2)(1+\frac{1}{m})$$

and therefore

$$(\lambda(I)-\frac{2}{m})(1-\frac{1}{m}) \le \frac{\Lambda(I;\omega_N)}{N} \le (\lambda(I)+\frac{2}{m})(1+\frac{1}{m}) \ ;$$

since $\lambda(I) \le 1$, we get finally

$$-\frac{3}{m} + \frac{2}{m^2} \le \frac{\Lambda(I;\omega_N)}{N} - \lambda(I) \le \frac{3}{m} + \frac{2}{m^2}$$

and therefore

$$D(\omega_N) \leq \frac{3}{m} + \frac{2}{m^2} ,$$

which can be made arbitrarily small. □

Remarks:

1) $D(\omega_N)$ is a distance between δ_{ω_N} and $\lambda|_{[0,1]}$; we have $\delta_{\omega_N} \to \lambda|_{[0,1]}$ iff $D(\omega_N) \to 0$, i.e. $D(\omega_N)$ generates the convergence - if the limit function is $\lambda|_{[0,1]}$.
We will later see that the discrepancy may be extended to a distance between measures, for example by

$$D(\mu,\nu) = \sup_{Q<P} \left| \mu(R_{[Q,P)}) - \nu(R_{[Q,P)}) \right| .$$

We know from theorem 1.5 that $D(\mu_n,\mu) \to 0$ implies $\mu_n \to \mu$; the opposite direction is however not true, the convergence is in general not uniform. If however μ is absolutely continuous with respect to λ, i.e. if it has a density $f \in \mathbb{L}^1$ as in our cases, then

$$\mu_n \to \mu \quad \text{implies} \quad D(\mu_n,\mu) \to 0 .$$

2) The definition may be changed a bit, for example by

$$D^*(\omega_N) := \sup_{0 \leq x \leq 1} \left| \frac{A([0,x);\omega_N)}{N} - x \right| ,$$

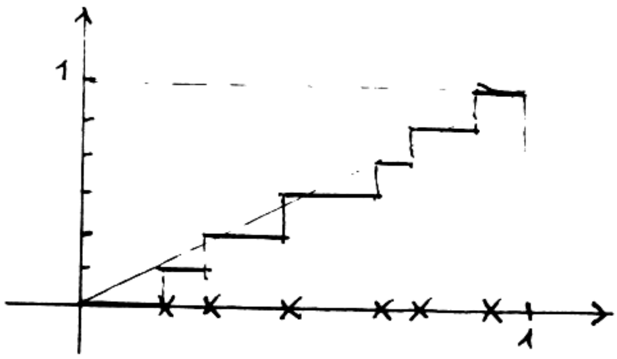
i.e. by considering only intervals of the form $[0,x)$.
It is trivial to see that

$$D^* \leq D \leq 2D^*$$

such that $D^*(\omega_N) \to 0$ is equivalent to $D(\omega_N) \to 0$ and therefore to the fact that ω_N is u.d.
D^* is easy to compute: Consider

$$S_N(x) = \frac{A([0,x);\omega_N)}{N} ;$$

it jumps by $\frac{1}{N}$ at any point x_j of ω_N



The difference is maximal at $x=x_j$, such that

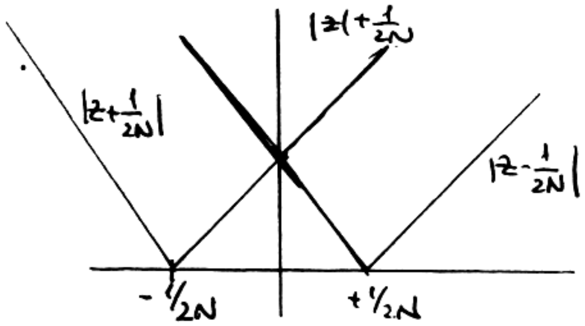
$$D^* = \max_{1\le j\le N} \max\left\{|S_N(x_j+0)-x_j|,|S_N(x_j-0)-x_j|\right\} .$$

We therefore need only to take the maximum of 2N numbers. If we put $x_1,\dots,x_N$ into order, i.e. if $x_1\le x_2\le \dots \le x_N$, then (assuming first that they are all different)

$$S_N(x_j+0) = \frac{j}{N} \quad\text{and}\quad S_N(x_j-0) = \frac{j-1}{N}$$

and

$$D^* = \max_{1\le j\le N} \max\left\{|\frac{j-1}{N}-x_j|,|\frac{j}{N}-x_j|\right\} .$$



Now

$$\frac{j}{N} -x_j = \frac{j-\frac{1}{2}}{N} -x_j + \frac{1}{2N} = z + \frac{1}{2N}$$

$$\frac{j-1}{N} - x_j = \frac{j-\frac{1}{2}}{N} - x_j - \frac{1}{2N} = z - \frac{1}{2N} .$$

Since $\max\left\{|z+\frac{1}{2N}|,|z-\frac{1}{2N}|\right\} = |z| + \frac{1}{2N}$, we get

$$D^* = \frac{1}{2N} + \max_{1\le j\le N}\left|\frac{j-\frac{1}{2}}{N} - x_j\right|.$$

This is a quite important formula, which we will often use. It is easy to see that it is true even if some of the x_j are equal.
In order to compute D^*, if $\omega_N = \{x_1,\dots,x_N\}$ is ordered (which may create especially in higher dimensions a lot of computational effort), then we have just to compute the largest of the distances

$$\left|\frac{2j-1}{2N} - x_j\right|, \quad j=1,\dots,N .$$

We see immediately that our first example

$$\omega_N = \left\{\frac{1}{2N}, \frac{3}{2N}, \ldots, \frac{2N-1}{2N}\right\}$$

gives the optimal possible discrepancy $D^*(\omega_N) = \frac{1}{2N}$ and is therefore u.d.

Remarks:

1) We see that sequences $(\omega_N)_{N\in\mathbb{N}}$ exist with

$$D(\omega_N) \sim \frac{C}{N}$$

and that this convergence speed is the optimal one. The situation is different, if $(\omega_N)_{N\in\mathbb{N}}$ is generated by a sequence $(x_i)_{i\in\mathbb{N}}$ of points: In this case there exists an absolute constant C such that

$$D^*(\omega_N) \geq C \cdot \frac{\ln N}{N} \quad \text{for infinitely many } N \in \mathbb{N}$$

(Theorem 2.3 in [K,N], p. 109). The proof is quite technical. This seems to be a low convergence rate; however, one has to keep in mind in discussing convergence rates for particle methods that using discrepancy as a distance, $0(\frac{\ln N}{N})$ is the best order one can possibly achieve. There are also upper bounds for D or D^*, which are not so easy to handle; it is however clear that for some concrete examples $0(\frac{\ln N}{N})$ can be reached so that this convergence order is really optimal.
The best example is the van der Corput series p(i). It was shown (Haber, S. "On a sequence of points of interest for numerical quadrature, J. Res. Nat. Bur. Standards, Sect. B70, pages 127-136, 1966) that for this sequence

$$D^*(\omega_N) \leq \frac{1}{3\ln 2} \frac{\ln N}{N} + 0(\frac{1}{N})$$

and that the constant $\frac{1}{3\ln 2}$ cannot be improved.

2) We have defined $D^*(\omega_N)$ as $\|S_N - id\|_{\mathbb{L}^\infty[0,1]}$; we may therefore call D^* the $\mathbb{L}^\infty$-discrepancy; there are also $\mathbb{L}^p$-discrepancies defined by $(1 \leq p < \infty)$

$$D^*_{(p)}(\omega_N) = \|S_N - id\|_{\mathbb{L}^p[0,1]} = \left(\int_0^1 |S_N(x) - x|^p dx\right)^{1/p} .$$

There are connections between $D^*_{(p)}$ and the bounded Lipschitz distance d_{BL} we mentioned in chapter 1.

We shall now show how u.d. sequences (ω_N) may be used for numerical integration. We mentioned already that u.d. is equivalent to

$$\lim \frac{1}{N} \sum_{i=1}^{N} f(x_i^N) = \int_0^1 f(x)dx$$

for Riemann integrable functions f. How fast can this convergence be in dependence on the smoothness of f?
We consider functions f of bounded variation V[f], which can be represented as the difference of 2 monotoneous functions and can be used as a distribution in Riemann-Stieltjes integrals.

Lemma (KN, p. 143)
Let $\omega_N = (x_1,\dots,x_N)$ be ordered: $x_1 \le x_2 \le \dots \le x_N$ and let f be of bounded variation V[f]. Then (using $x_o=0$, $x_{N+1}=1$)

$$\frac{1}{N} \sum_{n=1}^{N} f(x_n) - \int_0^1 f(x)dx = \sum_{n=0}^{N} \int_{x_n}^{x_{n+1}} (t-\frac{n}{N})df(t) \ .$$

Proof:

$$\sum_{n=0}^{N} \int_{x_n}^{x_{n+1}} (t-\frac{n}{N})df(t) = \int_0^1 tdf(t) - \sum_{n=0}^{N} \frac{n}{N}[f(x_{n+1})-f(x_n)]$$

$$= tf(t)\Big|_0^1 - \int_0^1 f(t)dt + \frac{1}{N} \sum_{n=0}^{N} f(x_n)-f(1) \quad (\text{using } \int_0^1 gdf = gf\Big|_0^1 - \int_0^1 fdg)$$

$$= \frac{1}{N} \sum_{n=0}^{N} f(x_n) - \int_0^1 f(t)dt$$

□

With this lemma we get easily the so-called Koksma inequality.

Theorem 2.1:
If f is of bounded variation V[f] in [0,1] and $\omega_N = (x_1,\dots,x_N)$, then

$$\left|\frac{1}{N} \sum_{n=1}^{N} f(x_n) - \int_0^1 f(t)dt\right| \le V[f]D^*(\omega_N) \ .$$

Proof:

We may assume $x_1 \le x_2 \le \ldots \le x_N$; for fixed n we have

$$\left|t-\frac{n}{N}\right| \le \max\left(\left|x_n-\frac{n}{N}\right|, \left|x_{n+1}-\frac{n}{N}\right|\right) \le D^*(\omega_N) \quad \text{if } x_n \le t \le x_{n+1} .$$

Therefore

$$\left|\frac{1}{N}\sum_{n=1}^{N} f(x_n) - \int_0^1 f(t)dt\right| \le \sum_{n=0}^{N} \int_{x_n}^{x_{n+1}} \left|t-\frac{n}{N}\right| \left|df(t)\right|$$

$$\le D^*(\omega_N) \sum_{n=0}^{N} \int_{x_n}^{x_{n+1}} |df(t)| = D^*(\omega_N)V[f] .$$

□

Remark:

If one allows f to be differentiable with $f' \in \mathbf{L}^1[0,1]$, then $V[f] = \int_0^1 |f'(t)|dt$ and the proof is a bit easier but uses essentially the same ideas. For continuous functions there are other estimates: If $f \in C[0,1]$ and

$$M_f(h) = \sup\left\{|f(x)-f(y)| \,/\, x,y \in [0,1],\ |x-y| \le h\right\}$$

denotes the "modulus of continuity", then

$$\left|\frac{1}{N}\sum_{n=1}^{N} f(x_n) - \int_0^1 f(t)dt\right| \le M_f(D^*(\omega_N)) .$$

Proof:

Again $\omega_N = (x_1,\ldots,x_N)$, $x_1 \le \ldots \le x_N$. The mean value theorem gives us the existence of ξ_n, $\frac{n-1}{N} < \xi_n < \frac{n}{N}$ such that

$$\int_0^1 f(t)dt = \sum_{n=1}^{N} \int_{\frac{n-1}{N}}^{n/N} f(t)dt = \frac{1}{N}\sum_{n=1}^{N} f(\xi_n)$$

and

$$\left|\int_0^1 f(t)dt - \frac{1}{N}\sum_{n=1}^{N} f(x_n)\right| \le \frac{1}{N}\sum_{n=1}^{N} \left|f(x_n)-f(\xi_n)\right| .$$

If $x_n \ge \xi_n$, then $|x_n-\xi_n| < |x_n-\frac{n-1}{N}| \le D^*(\omega_N)$ and similarly, if $x_n < \xi_n$, then $|x_n-\xi_n| < |x_n-\frac{n}{N}| \le D^*(\omega_N)$. The statement follows from the definition of M_f.

□

Things are a bit different, if f is an unbounded function in [0,1], say that $f(x) \to \infty$ for $x \to 0$ but $\int_0^1 f(x)dx$ exists. Then $\frac{1}{N}\sum_{i=1}^{N} f(x_i)$ doesn't converge necessarily to $\int_0^1 f\,dx$ for an u.d. sequence $(x_i)_{i\in\mathbb{N}}$. Sobol has shown that with $a_N := \min_{0\leq i\leq N-1} x_i$ the following statement is true:

If

$$D^*(\omega_N) \int_{a_N}^{1} |f'(x)|dx \to 0 \quad \text{for } N\to\infty ,$$

then

$$\frac{1}{N}\sum_{i=1}^{N} f(x_i) \longrightarrow \int_0^1 f(x)dx .$$

Example:

Take $p(i) = \ell_1 2^{-1}+\ldots+\ell_m 2^{-m}$ for $i=\ell_1+2\ell_2+\ldots+2^{m-1}\ell_m$, then $(\ell_m \neq 0)$

$$(N+1)a_N = (N+1)2^{-m} \geq \frac{2^{m-1}}{2^m} = \frac{1}{2} \quad \text{and}$$

$$\leq \frac{2^m}{2^m} = 1 \text{ , such that } a_N \sim \frac{c}{N} .$$

Since $D(\omega_N) \sim \frac{\ln N}{N}$, we have to check, whether

$$\frac{\ln N}{N} \int_{c/N}^{1} |f'(x)|dx \longrightarrow 0 ;$$

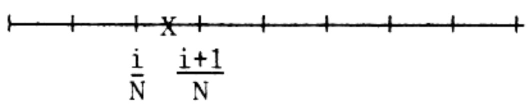
for $f(x) = x^{-\lambda}$, $f'(x) = -\lambda x^{-\lambda-1}$ we get $\int_{c/N}^{1} |f'(x)|dx = \lambda(1-(\frac{c}{N})^{-\lambda})$

so that $\frac{\ln N}{N}(1-(\frac{N}{c})^{\lambda}) \to 0$, if $\lambda<1$, i.e. if $\int_0^1 f\,dx$ is convergent.

The construction of sequences with good discrepancy

We will first follow Sobol:

We take $N=2^{\nu}$ and consider $\omega_N = (x_1,\ldots,x_N)$. We divide [0,1] into N segments of length $\frac{1}{N}$ and call ω_N a P_o-grid, if each segment contains exactly one point of ω_N.



Ordering the points x_i, we see that $\frac{i-1}{N} \leq x_i < \frac{i}{N}$, $i=1,\dots,N$. Therefore $\left|\frac{i-\frac{1}{2}}{N} - x_i\right| \leq \frac{1}{2N}$ and $D^* \leq \frac{1}{N}$ (remember that only sequences ω_N generated by point sequences had always an order worse than $\frac{\ln N}{N}$, meanwhile $O(\frac{1}{N})$ was possible in the general case). P_o-grids in one dimension are quite trivial constructions, but will become more interesting in higher dimensions.

In one dimension one may use the concept of LP_o-sequences for u.d. point sequences $(x_n)_{n\in\mathbb{N}}$.

In each sequence a section $x_{k2^s}, x_{k2^s+1}, \dots, x_{(k+1)2^s-1}$ with $s=1,2,\dots$ and $k=0,1,2,\dots$ is called a binary section. Choose for example $h=2^s$ and divide (x_n) in sections of length h

$$[x_o,\dots,x_{h-1}],\ [x_h,\dots,x_{2h-1}].$$

Now $(x_i)_{i\in\mathbb{N}}$ is called a LP_o-sequence, if any binary section is a P_o-grid.

If $2^s-1 \leq N < 2^{s+1}-1$, then the dual representation of N contains at most s times the digit 1. If the number of 1's is t, we have $t \leq s = [\log_2(N+1)]$, where $[z] = \max\left\{n \in \mathbb{N}/n\leq z\right\}$ is the Gauss bracket. Now $D^*(\omega_N) \leq \frac{t}{N}$ for a LP_o-sequence and therefore

$$D(\omega_N) \leq \frac{[\log_2(N+1)]}{N} .$$

It follows that LP_o-sequences have an optimal behaviour; especially they are all u.d.

Sequences of a binary rational type

Let V be an infinite matrix

$$V = (v_{sj})_{sj\geq 1} = \begin{pmatrix} v_{11} & v_{12} & \cdots \\ v_{21} & v_{22} & \cdots \\ v_{31} & v_{32} & \cdots \\ \cdots\cdots & & \end{pmatrix}$$

with $v_{sj} \in \{0,1\}$. Each column contains a finite non-empty set of ones; then each column defines a rational number in its dual representation

$$V_s := 0,v_{s1}v_{s2}\cdots$$

V is called direction-matrix and the numbers $V_1, V_2, \dots$ are called direction numbers. A sequence $(r(i))_{i\in\mathbb{N}_o}$ is called of binary rational type (shortly BR-sequence), if

(i) $r(0) = 0$

(ii) $r(2^s) = V_{s+1}$, $s=0,1,2,\ldots$

(iii) for $2^s < i < 2^{s+1}$ we define $r(i) = r(2^s)*r(i-2^s)$, where * denotes the bitwise addition modulo 2 (exclusive or)

For example $5/16*7/8 = 0.0101*0.1110 = 0.1011 = 11/16$

$13/16*19/32 = 0.11010*0.10011 = 0.01001 = 9/32$.

It is easy to see that r(i) may also be expressed by

$$r(i) = e_1V_1*e_2V_2*\ldots*e_mV_m \quad \text{for } i=e_m\ldots e_1 \text{ (dual)};$$

for example $r(25) = r(11001) = V_1*V_4*V_5$.

Sobol has shown that a BR-sequence is a LP_o-sequence if V has only ones in the main diagonal and above it only zeroes.

Example 1:

The van der Corput sequence p(i) is the most simple BR-sequence:

$$V = \begin{pmatrix} 1 & 0 & \ldots & \\ 0 & 1 & 0 & \ldots \\ \vdots & & & \end{pmatrix}$$

Then $V_s=2^{-s}$, therefore $V_s*V_r = V_s+V_r$ and

$$r(i) = p(i) = e_1 2^{-1}+e_2 2^{-2}+\ldots+e_m 2^{-m} .$$

Therefore (p(i)) is a BR-sequence and consequently an LP_o-sequence with

$$D^* \leq \frac{t}{N} ,$$

this is for some N better than $\frac{1}{3\ln 2}\frac{\ln N}{N} + O(\frac{1}{N})$.

The fact that $(p(i))_{i\in\mathbb{N}}$ is an LP_o-sequence implies also that every section consisting of 2^m points, i.e. every ensemble ω_{2^m} is a P_o-grid and has therefore optimal discrepancy.

Example 2:

The matrix

$$C = (c_{sj}) \quad \text{with} \quad c_{sj} = \binom{s-1}{j-1} \quad \text{for } s \geq j \text{ and } c_{sj}=0 \text{ for } j>s$$

is called Pascal matrix (for quite obvious reasons).

We take now $V = (v_{sj})$ with $v_{sj}=c_{sj}$ mod 2 and get

$$V = \begin{pmatrix} 1 & 0 & . & . & . & 0 & \\ 1 & 1 & 0 & . & . & . & \\ 1 & 0 & 1 & 0 & . & . & . \\ 1 & 1 & 1 & 1 & 0 & . & . & . \\ 1 & 0 & 0 & 0 & 1 & & \end{pmatrix}, \text{ which fulfils Sobol's condition.}$$

The sequence $(q(i))_{i\in\mathbb{N}}$ given by this V, i.e.

$$q(i) = 0.1e_1 * 0.11e_2 * 0.101e_3 * \ldots * V_n e_m \quad \text{for } i=e_m \ldots e_1$$

is therefore a LP_o-sequence.

p(i) and q(i) are connected but at the same time independent. For example: If p(i)=q(k) for some i,k, the p(k)=q(i). However, $(p(i),q(i))_{i\in\mathbb{N}}$ defines a pointsequence in $[0,1]^2$, which has good properties as a 2-dimensional sequence.

The more dimensional case

We consider now the case of $k>1$, but choose as density still the uniform distribution in the k-dimensional unit cube, i.e. $f(P) = \chi_{[0,1]^k}(P)$, $P \in \mathbb{R}^k$. We consider k-dimensional intervals $[\overset{o}{P},\overset{o}{Q}) \subset [0,1]^k$ as in § 1, define $\Lambda([\overset{o}{P},\overset{o}{Q});\omega_N) = \sum_{i=1}^{N} \chi_{[\overset{o}{P},\overset{o}{Q}]}(P_i^N)$, when

$\omega_N = (P_1^N,\ldots,P_N^N)$ and call $(\omega_N)_{n\in\mathbb{N}}$ u.d., if

$$\lim_{N\to\infty} \frac{\Lambda([\overset{o}{P},\overset{o}{Q});\omega_N)}{N} = \lambda_k([\overset{o}{P},\overset{o}{Q})) \quad \text{for all intervals } [\overset{o}{P},\overset{o}{Q}).$$

Again $(\omega_N)_{N\in\mathbb{N}}$ is u.d. iff $\lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} \phi(P_i^N) \to \int_{[0,1]^k} \phi dx$ for all Riemann integrable functions f and the Weyl criterion gives a sufficient and necessary condition

$$\lim_{N\to\infty} \frac{1}{N} \sum_{j=1}^{N} e^{2\pi i \langle h,P_j^N\rangle} = 0 \quad \text{for every lattice point } h \in \mathbb{Z}^k,\ h\neq 0 .$$

The consequence is that $(\omega_N)_{N\in\mathbb{N}}$ is u.d. iff

$$(\langle h, P_j^N \rangle)_{j=1,\dots,N}$$

are one-dimensional u.d. sequences.
We get now the generalization of the concept of discrepancy to higher dimension

$$D(\omega_N) := \sup_{0 \leq \overset{o}{P} < \overset{o}{Q} \leq E} \left| \frac{\Lambda([\overset{o}{P},\overset{o}{Q});\omega_N)}{N} - \lambda_k([\overset{o}{P},\overset{o}{Q}))\right|$$

and

$$D^*(\omega_N) := \sup_{\overset{o}{Q} \in [0,1]^k} \left| \frac{\Lambda([0,\overset{o}{Q});\omega_N)}{N} - \lambda_k([0,\overset{o}{Q}))\right| .$$

Again, it is easy to see that

$$D^* \leq D \leq 2^k D^*$$

and the proof that the convergence is uniform, such that

$$(\omega_N)_{N \in \mathbb{N}} \text{ u.d. } \Longleftrightarrow D(\omega_N) \to 0 \quad \text{for } N \to \infty$$

is similar to the one-dimensional case.
However, for $k > 1$ one may define other discrepancies, for example the so-called "isotropic discrepancy"

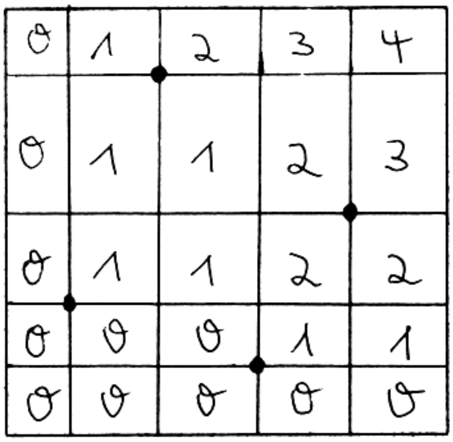
$$J(\omega_N) := \sup_{c \in C} \left| \frac{\Lambda(c;\omega_N)}{N} - \lambda_k(c)\right| ,$$

where C is the family of all convex subsets of $[0,1]^k$. One may (see [K.N.], p. 94) restrict the consideration to all open and closed convex polytops in $[0,1]^k$ and gets in using this fact

$$D \leq J \leq (4k\sqrt{k}+1)D^{1/k}$$

such that (ω_N) is u.d. iff $J(\omega_N) \to 0$.

The computation of D or D^* is not as easy as in the one-dimensional case: One cannot restrict the attention to the points themselves but has also to consider the intersection points of all planes passing through the N ensemble points; this set contains already N^k points - see 2.1.



ω_4 together with 4^2 intersection points and the the values of $\Lambda([0,P);\omega_N)$ for $P \in [0,1]^2$.

It is easy to realize that Λ jumps also in the intersection points which do not belong to ω_N.
Niederreiter gives a formula, which generalizes the explicit expression to higher dimensions.
What are the best bounds for $D^*(\omega_N)$? As for k=1 we have to distinguish between general sequences of point sets and those, who are generated by point sequences. Here are still open questions; the conjectures are, that

$$D^*(\omega_N) = O\left(\frac{\ln^{k-1} N}{N}\right) \quad \text{is optimal for general sequences } (\omega_N)$$

and

$$D^*(\omega_N) = O\left(\frac{\ln^{k} N}{N}\right) \quad \text{is optimal for sequences, which are generated by point sequences.}$$

One can construct sequences with this rate, however one can prove only for k=1 that they are optimal.
What is shown (see [K.N], page 105) is that

$$D^*(\omega_N) > C_k \frac{\ln^{\frac{k-1}{2}} N}{N} \quad \text{in the general case}$$

and

$$D^*(\omega_N) > C'_k \frac{\ln^{k/2} N}{N} \quad \text{for infinitely many } N \text{ in the case of point sequences,}$$

where C_k, C'_k are constants only depending on the dimension. The upper bounds are as in one dimension not very helpful.

We shall give special constructions with corresponding convergence rates quite soon; but first we shall present a generalization of the Koksma inequality, which is now called "Koksma-Hlawka inequality". We need a careful definition of the notion of "bounded variation".
We are still in $[0,1]^k$, consider $k \geq 2$ and a function $f(P) = f(p_1,\dots,p_k)$ defined there. A partition Z of $[0,1]^k$ is given by partions $Z^{(j)}$ on the coordinate axes

$$Z^{(j)} = \left(z_i^{(j)}\right)_{0 \leq i \leq m_j} \quad \text{with } 0 = z_o^{(j)} \leq z_1^{(j)} \leq \dots \leq z_{m_j}^{(j)} = 1 .$$

The difference operator with respect to the j-th coordinate is

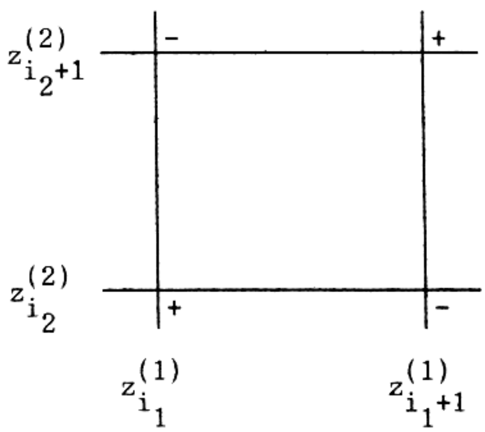
$$\Delta_j f(p_1,\dots,p_{j-1},z_i^{(j)},p_{j+1},\dots,p_k)$$
$$:= f(p_1,\dots,p_{j-1},z_{i+1}^{(j)},p_{j+1},\dots,p_k)$$
$$- f(p_1,\dots,p_{j-1},z_i^{(j)},p_{j+1},\dots,p_k) .$$

We denote by $\Delta_{j_1,\dots,j_p}$ just the composition $\Delta_{j_1} \circ \dots \circ \Delta_{j_p}$ (the operators commute). A straightforward generalization of the one-dimensional variation is the variation "in the sense of Vitali"

$$V^{(k)}[f] = \sup_Z \sum_{i_1=0}^{m_1-1} \dots \sum_{i_k=0}^{m_k-1} \left| \Delta_{1,\dots,k} f\left(z_{i_1}^{(1)},\dots,z_{i_k}^{(k)}\right) \right| .$$

Realize that, since for example for k=2 we sum up terms like

$$\Delta_{1,2} f\left(z_{i_1}^{(1)},z_{i_2}^{(2)}\right) =$$



the variation is zero if f is not depending on one of the variables. $V^{(k)}[f] < \infty$ doesn't mean high regularity. In order to get more smoothness, one has to restrict the variation on the faces of $[0,1]^k$, defined by the fact that one or more coordinates have the value 0 or 1.

We say that $f: [0,1]^k \to \mathbb{R}$ is of bounded variation in the sense of Hardy and Krause, if f itself as well as its restrictions to all lower dimensional faces of $[0,1]^k$ are of bounded variation in the sense of Vitali. In 2 dimensions this means that

$$V^{(2)}[f] + V^{(1)}[f(0,\cdot)] + V^{(1)}[f(1,\cdot)] + V^{(1)}[f(\cdot,0)] + V^{(1)}[f(\cdot,1)] < \infty$$

In order to formulate the k-dimensional version of the Koksma-Hlavka inequality we need some notations.
Given $\ell \le k$ and an ℓ-tupel $(i_1,\dots,i_\ell)$ with $1 \le i_1 < i_2 < \dots < i_\ell \le k$, we denote by $\omega_N[i_1,\dots,i_\ell]$ the ℓ-dimensional pointset given as projection of ω_N to the ℓ-dimensional face of $[0,1]^k$ defined by $p_j = 1$ for $j \notin \{i_1,\dots,i_\ell\}$. Similarly we denote by $f[i_1,\dots,i_\ell]$ the restriction of a function f on $[0,1]^k$ to the same face.

Theorem 2.1': (k-dimensional Koksma-Hlawka inequality)
If f is of bounded variation in the sense of Hardy and Krause and $\omega_N = (P_1,\dots,P_N)$, then

$$\left|\frac{1}{N}\sum_{n=1}^{N} f(P_n) - \int_{[0,1]^k} f(P)dP\right| \le \sum_{\ell=1}^{k} \sum_{1\le i_1<\dots<i_\ell\le k} V^{(\ell)}[f[i_1,\dots,i_\ell]] \cdot \\ \cdot D^*(\omega_N[i_1,\dots,i_\ell]) .$$

We get a simplified version, if we define

$$V(f) := \sum_{\ell=1}^{k} \sum_{1\le i_1<\dots<i_\ell\le k} V^{(\ell)}[f[i_1,\dots,i_\ell]]$$

as Hardy-Krause variation. Since it is obvious that

$$D^*(\omega_N[i_1,\dots,i_\ell]) \le D^*(\omega_N) ,$$

we get

$$\left|\frac{1}{N}\sum_{n=1}^{N} f(P_n) - \int_{[0,1]^k} f(P)dP\right| \le V(f)D^*(\omega_N) .$$

We shall prove a more general version, where we compare ω_N with a general density f instead of just $\lambda_k|_{[0,1]^k}$.
We want to mention that there are estimates, if we restrict the integration to a Jordan measurable subset E of $[0,1]^k$, i.e. for

$$\left|\frac{1}{N}\sum_{n=1}^{N} f(P_n) - \int_E f(P)dP\right|$$

(see H. Niederreiter, Application of diophantine approximations to numerical integration, Diophantine Approximation and its Application (C.F. Osgood, ed.) Ac. Press, 1973, p. 129-199).
There is also an estimate using the modulus of continuity:

$$\left|\frac{1}{N}\sum_{n=1}^{N} f(P_n) - \int_{[0,1]^k} f\,dP\right| \le A_k M_f\left(D^*(\omega_N)^{\frac{1}{k}}\right) .$$

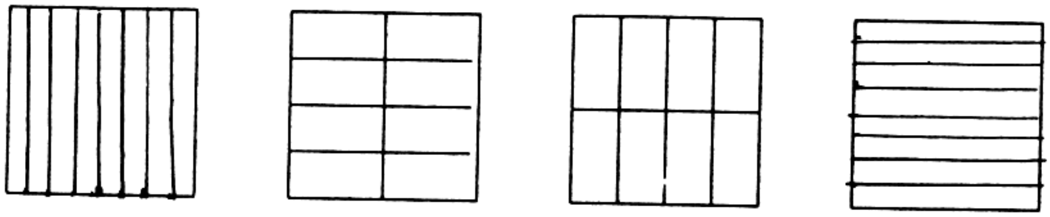
With $A_1=1$ we get the one-dimensional result. For $k\ge 2$ it is shown by Proinov ("Discrepancy and Integration of Continuous Functions", J. of Appr. Theory) that $A_k=4$ is always possible.

We shall now try to construct particle ensembles with low discrepancy. This is much harder than in the one-dimensional case.
Let's begin with the "trivial" solution for $N=M^k$:

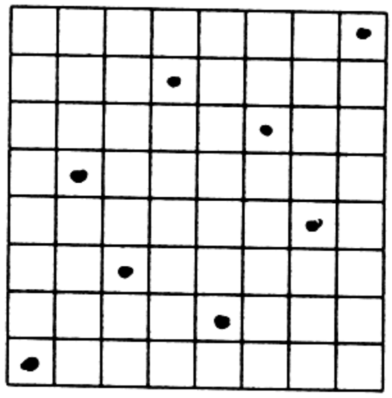
$$\omega_N = \left(\frac{i_1-\frac{1}{2}}{M},\frac{i_2-\frac{1}{2}}{M},\dots,\frac{i_k-\frac{1}{2}}{M}\right),\ 1\le i_1,\dots,i_k\le M .$$

One can check that $\left|\frac{A([0,P))}{N} - \lambda_k([0,P))\right|$ is maximal for example for $P=P_1=(1,\dots,1,\frac{1}{2M})$, namely $\left|0-\frac{1}{2M}\right| = \frac{1}{2M} = \frac{1}{2}N^{-1/k}$. (For example for $P=P_2=(1,\dots,1,\frac{3}{2M})$ one gets $\left|\frac{M}{N}-\frac{3}{2M}\right| = \frac{1}{2M}+(\frac{1}{M}-\frac{M}{N}) = \frac{1}{2M}+(\frac{M^{k-1}-M}{N}) > \frac{1}{2M}$.)
Therefore $D^*(\omega_N) = \frac{1}{2}N^{-1/k}$: This set is optimal for k=1, but becomes worse with growing k (already for k=2 we get only $\frac{1}{\sqrt{N}}$) and approaches for $k\to\infty$ even the worst case. (For k=3 the result is worse than real random effects! $k\ge 3$ are therefore sometimes called "damned dimensions".)

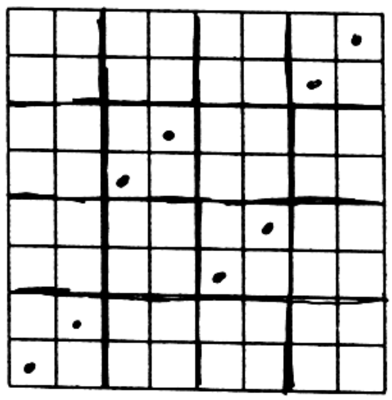
Again we may define P_o-grids as in the onedimensional case. We define binary intervals as $\ell_1\times\dots\times\ell_k$ where each ℓ_i is a onedimensional interval of the form $[(j-1)2^{-m},j2^{-m})$, $1\le j\le 2^m$, m=0,1,2,... Binary intervals in 2 dimensions with volume $|\ell_1|\cdot|\ell_2| = \frac{1}{8}$ are shown in the following 4 pictures:



We call ω_N for $N=2^\nu$ a P_o-grid, if each binary interval I of volume $\frac{1}{N}$ contains just one point of ω_N. However, as one can realize by the case $N=2^3$, k=2, there exist not only N different binary intervals but 32; a P_o-grid would have one point in each of these 32 intervals. This is possible:



One can construct P_o grids in 2 and 3 dimensions, as we shall soon see, but not in higher dimensions: For k=4 there exists no P_o-grid with more than 4 points. Since higher dimensions are important for particle methods, we have to weaken the conditions: We call ω_N with $N=2^\nu$ a P_τ-grid ($\tau=0,1,2,\dots,\nu-1$), if each binary interval of volume $\frac{2^\tau}{N}$ contains 2^τ points.



P_1-grid with 8 points

It is easier to construct P_τ-grids and it is possible for arbitrary dimensions if one increases τ with k, i.e. if $\tau=\tau(k)$. One knows only that $\tau(1) = \tau(2) = \tau(3) = 0$, $\tau(4) = 1$ and $\tau(k) = 0(k \cdot \ln k)$ is possible. For P_τ-grids one can show that

$$D^*(\omega_N) = 0\left(\frac{\ln^{k-1} N}{N}\right) ,$$

which is presumably the optimal rate. Here only the constants depend on τ. We may now generalize the concept of LP_0-sequences: We call a sequence of points $(P_n)_{n\in\mathbb{N}}$ an LP_τ-sequence, if every binary section with at least $2^{\tau+1}$ points is a P_τ-grid.

This gives a method to construct higher dimensional low discrepancy sequences starting with lower dimensions:

If $(P_n)_{n\in\mathbb{N}}$ is an LP_τ-sequence in $[0,1]^{k-1}$ $(k \geq 2)$, then for any given N the set

$$\omega_N = \left\{(0,P_1),(\tfrac{1}{N},P_2),\ldots,(\tfrac{N-1}{N},P_N)\right\}$$

is a P_τ-grid in $[0,1]^k$ (Sobol, p. 37).

A similar, but more general result is due to Niederreiter (Point sets and Sequences with small Discrepancy, Monatshefte Math. 104 (1987), p. 273-337, Lemma 8.9):

If $(P_n)_{n\in\mathbb{N}}$ is a sequence in $[0,1]^{k-1}$ and $\tilde{\omega}_M = (P_1,\ldots,P_M)$ then for ω_N constructed as above

$$ND^*(\omega_N) \leq \max_{1\leq M\leq N} MD^*(\tilde{\omega}_M) + 1 .$$

Therefore, if $(\tilde{\omega}_M)$ is a low discrepancy sequence, ω_N is a low discrepancy set. This can be used: We start with van der Corput $(p(n))_{n\in\mathbb{N}}$ and construct $\omega_N = \left\{(\frac{0}{N},p(1)),\ldots,(\frac{N-1}{N},p(N))\right\}$ - we get a 2-dimensional P_0-grid. Moreover: Use p(i) and the sequence q(i) constructed in using the Pascal matrix. Then (p(i),q(i)) defines an LP_0-sequence in 2 dimensions (we shall see that later) and

$$\omega_N = \left\{(\tfrac{0}{N},p(1),q(1)),\ldots,(\tfrac{N-1}{N},p(N),q(N))\right\}$$

is a P_0-grid in 3 dimensions.

<u>Constructions of LP_τ-sequences</u> are quite often based on the idea that each coordinate forms a one-dimensional LP_0-sequence; however, these LP_0-sequences must be independent.

A method of construction is provided by the so-called monocyclic

operators. All our calculations are now done modulo 2, i.e. we are working in the Galois field GF(2).
We consider infinite sequences $u = (\dots,u_{-2},u_{-1},u_0,u_1,u_2,\dots) \in GF(2)^{\mathbb{Z}}$ and linear operators $L: GF(2)^{\mathbb{Z}} \longrightarrow GF(2)^{\mathbb{Z}}$ defined by

$$(Lu)_i := u_{i+m}+a_{m-1}u_{i+m-1}+ \dots +a_1u_{i+1}+u_i$$

where $a_j \in GF(2)$, $j=1,\dots,m-1$. m is called the order of L.
Denoting by $\underline{0}$ the sequence of zeroes in $GF(2)^{\mathbb{Z}}$, we may ask for solutions of

$$Lu = \underline{0} \ .$$

This equation is easy to solve: Prescribe a section of length m of u arbitrary - then the rest of u is uniquely defined by the equation. For example prescribe $(u_1,\dots,u_m)$, then

$$u_{i+m} = a_{m-1}u_{i+m-1}+ \dots + u_i$$

can be recursively solved for i=1, then i=2,3 etc. (remember that we are in GF(2), where + equals -) and similarly
$u_i = u_{i+m}+a_{m-1}u_{i+mm-1}+ \dots +a_1u_{i+1}$ can be solved for i=0,-1 etc.
Proceeding in that way one gets a unique solution of $Lu = \underline{0}$ with m "free parameters" $u_1,\dots,u_m$. In GF(2) we have 2^m possibilities to choose $u_1,\dots,u_m$ differently, i.e. we get 2^m different solutions of $Lu = \underline{0}$. We consider sections of length m of a solution of $Lu=\underline{0}$: $(u_1,\dots,u_m),(u_2,\dots,u_{m+1}),\dots$. If two of these sections are identical, say $(u_1,\dots,u_m)$ and $(u_\tau,\dots u_{\tau+m-1})$, then the sequence u is periodic with period length $(\tau-1)$. At most 2^m of these sections can be different, therefore τ is at most equal to 2^m and the longest period possible is 2^m-1.
We call L monocyclic, if $Lu = \underline{0}$ has at least one solution of period 2^m-1.

Examples:

$(Lu)_i = u_i+u_{i+1}$ (m=1, maximal period 1; $Lu=\underline{0}$ has 2 solutions: $u=\underline{0}$ and $u=\underline{1}$)

$(Lu)_i = u_{i+3}+u_{i+2}+u_i$ (m=3, maximal period $2^3-1=7$ and $Lu=\underline{0}$ has a solution $\dots,\underset{u_1}{1},\ \underset{u_2}{0},\ \underset{u_3}{0},1,0,1,1,\ \underset{u_7}{1},\ \underset{u_1=u_8}{0},0,\dots$, i.e. a solution of maximal period.)

It is easy to see that for monocyclic operators every nontrivial solution has period 2^m-1 and they all differ from each other only by a shift. Therefore $(Lu)_i = u_{i+3}+u_{i+2}+u_{i+1}+u_i$ doesn't define a monocyclic operator, since we get different period lengths, for example through

$$u = (\ldots,1,0,0,1,1,0,0,1,\ldots) \quad \text{or} \quad (\ldots,1,0,1,0,1,0,\ldots) .$$

Further monocyclic operators are given by $u_{i+2}+u_{i+1}+u_i$ and $u_{i+3}+u_{i+2}+u_i$.

Given a monocyclic operator we can construct LP_τ-sequences: First we assign to L a direction matrix $V = (v_{s,j})$ by

(i) the first m columns of V, i.e. $(v_{s,1})_{s\geq 1},\ldots,(v_{s,m})_{s\geq 1}$ are solutions of $Lu=\underline{0}$ (more precisely: solve $(Lu)_i = 0$ for $i=1,2,\ldots$).

(ii) each of the sequent columns of $(v_{s,k})_{s\geq 1}$, $k>m$ solves

$$Lv_{\cdot,k} = v_{\cdot,k-m} .$$

(iii) V fulfills the condition sufficient to give an LP_o-sequence: There are ones in the main diagonal and only zeroes above.

((iii) follows from (ii), if it is true for the first m columns.) Therefore such a V defines an LP_o-sequence "connected with the operator L".

Example:

We consider the Pascal matrix which gave rise to $(q(i))_{i\in\mathbb{N}}$ and check that it is connected with $(Lu)_i = u_i+u_{i+1}$. This operator has as we know order 1 and one nontrivial solution $u=\underline{1}$. Therefore the first column of V consists of 1's. Now we have to solve

$$v_{i+1,k}+v_{i,k} = v_{i,k-1}$$

recursively with respect to k. One checks easily that this is the recurrence relation for binomial coefficients (modulo 2)

$$\binom{i+1}{k} = \binom{i}{k} + \binom{i}{k-1} .$$

Sobol has shown that the following is true:

(1) Let $L_1,\ldots,L_{k-1}$ be different monocyclic operators with orders $m_1,\ldots,m_{k-1}$. If we construct for each j a LP_o-sequence $(P_j(i))_{i\in\mathbb{N}}$ connected with L_j, then the sequence $(\tilde{Q}_i)_{i\in\mathbb{N}}$ with

$$\tilde{Q}_i := (p_1(i),\ldots,p_{k-1}(i))$$

defines an LP_τ-sequence in $[0,1]^{k-1}$ with $\tau = \sum_{j=1}^{k-1} (m_j-1)$.

(2) If one adds one coordinate defined by the van der Corput sequence $(p(i))_{i\in\mathbb{N}}$ we get that

$$Q_i = (p(i),\tilde{Q}(i))$$

is an LP_τ sequence in $[0,1]^k$ with the same τ.

Generalization of the Sobol constructions of P_0-grid etc. to bases b different from 2

We fix a base $b \in \mathbb{N}$, $b\geq 2$. We define (instead of binary intervals) elementary intervals in base b as $\ell_1\times\ldots\times\ell_k$, where each ℓ_i is a one-dimensional interval of the form $[(j-1)b^{-m},jb^{-m})$ with $1\leq j\leq b^m$ and arbitrary $m=0,1,2,\ldots$.
We consider again point sets $\omega_N \subset [0,1]^k$ with $N=b^m$ and call it P_o-grid in base b, if each elementary interval in base b of volume 1/N contains exactly one point. The generalization to P_τ-grids is obvious - then each such interval of volume $b^\tau/N = b^{\tau-m}$ contains exactly b^τ points. (Niederreiter denotes such a P_τ-grid in base b as a (τ,m,k)-net in base b - see H. Niederreiter, Point Sets and Sequences with Small Discrepancy, Monatshefte für Mathematik 104, pages 273-337, 1987.)

In order to understand the consequence of changing the base b we consider the question how many elementary intervals of volume $1/N=b^{-m}$ exist - if we fix N and change b. We take an example: $k=4$, $N=b^m$ and construct intervals π of volume b^{-m}; if π has edges of length $b^{-m_1},\ldots,b^{-m_4}$, then the volume of π is

$$b^{-m_1} \ldots b^{-m_4} = b^{-m}$$

or $m_1+\ldots+m_4 = m$. We need all nonnegative integer solutions of this equation; elementary number theory tells us that we get $(m+1)(m+2)(m+3)/6$ such solutions. Each solution defines a certain "type" of elementary interval - and there are N intervals of each type. Therefore, for $N=2^7=128$ ($b=2$, $m=7$), we have $720/6 = 120$ types and $120\times128 = 15.360$ intervals. For a P_o-grid this means 15.360 conditions for ω_N. If we choose $N=5^3=125$, we get 20 types and 2500 conditions, for $N=11^2$ we get 10 types and 1210 conditions.

We realize: If we fix N (at least approximately) and enlarge b, we get less conditions for ω_N to be a P_o-grid in base b; it is therefore indeed possible to construct P_o-grids in base b in all dimensions, if $b\geq 3$.

We generalize also the concept of LP_τ-sequences: A sequence of points $(P_i)_{i\in\mathbb{N}}$ in$[0,1]^k$ is called LP_τ-sequence (or - in Niederreiter's notation a (τ,k) sequence in base b) if for all integers $j\geq 0$ and $m\geq\tau$ the pointset $\left\{P_{jb^m+1},\ldots,P_{(j+1)b^m}\right\}$ is a P_τ-grid (or (τ,m,k)-net) in base b.

Remark:

Sobol's original work was concerned with b=2; H. Faure (Discrépance de suites associées à un système de numération (en dimension s), Acta Arithm 41, 337-351, 1982) had considered the case $\tau=0$, i.e. LP_o-sequences in arbitrary bases. Niederreiter has an upper bound for (τ,m,k)-nets in base b:

If ω_N is a (τ,m,k)-net in base $b\geq 3$, then

$$ND(\omega_N) \leq b^\tau \sum_{i=0}^{q} \binom{k-1}{i}\binom{m-\tau}{i}\left[\frac{b}{2}\right]^i \text{ , when } q=\min(m-\tau,k-1) \text{ .}$$

This gives, if we consider the behaviour with respect to $m-\tau$

$$ND(\omega_N) \leq \frac{1}{(k-1)!}b^\tau\left[\frac{b}{2}\right]^{k-1}(m-\tau)^{k-1} + O((m-\tau)^{k-2}) \text{ .}$$

Since $N=b^m$ or $m=\log^b N$ this gives

$$D(\omega_N) \leq \frac{1}{(k-1)!}b^\tau\left(\frac{\left[\frac{b}{2}\right]}{\ln b}\right)^{k-1} \frac{(\ln N)^{k-1}}{N} + O\left(\frac{(\ln N)^{k-2}}{N}\right)$$

which is an optimal rate.

The first construction of P_o-grids and LP_o-sequences in base b are due to Faure (see above), who showed that for $b\geq k$ one can construct this kind of LP_o-sequences. He used a generalization of the (p(i),q(i))-idea in 2 dimensions. (p(i)) and (q(i)) correspond to V=E and V=Pascal matrix C(mod2). It is easy to realize that $C^2=E(\text{mod}2)$ - even more general $C^b=E(\text{mod}b)$. This gives us Faure's idea: Choose k matrices out of $E,C(\text{mod}b),\ldots,C^{b-1}(\text{mod}b)$ (if $b\geq k$), use these matrices as direction matrices, which define now "sequences of rational type in base b" (instead of binary rational sequences). Use now these sequences as coordinates of point sequences - this defines an LP_o-sequence in base b in $[0,1]^k$. Faure has shown that for this sequence $(P_n)_{n\in\mathbb{N}}$ and $\omega_N=(P_1,\ldots,P_N)$ we have

$$D^*(\omega_N) = 0\left(\frac{\ln^k N}{N}\right) \quad \text{and even} \quad D^*(\omega_N) = 0\left(\frac{\ln^{k-1} N}{N}\right) \quad \text{for } N=b^m .$$

Remark:

We are not aware of an investigation telling us, which way - along LP_τ-sequences or other bases - is the better one. It may depend on the size of N: One should realize that not the asymptotic behaviour but the smaller $D^*(\omega_N)$ for given N is the most important fact.

There is a third idea to produce low discrepancy sequences in higher dimensions, which is connected with the names of Hammersley & Halton: Hammersley generalized van der Corput's idea to other bases b - if $i=e_m b^{m-1}+\ldots+e_2 b+e_1$ ($e_j \in \{0,\ldots,b-1\}$), then $\Phi_b(i) = e_1 b^{-1}+\ldots+e_m b^{-m}$ (or simple $\Phi_b(e_m \ldots e_1) = 0,e_1 e_2 \ldots e_m$ in base b). Clearly $\Phi_2=p$. Halton proposed now the following: Take prime number $r_1,\ldots,r_k$ and construct

$$P(i) = (\Phi_{r_1}(i),\ldots,\Phi_{r_k}(i)) \ , \ i \in \mathbb{N} ;$$

then

$$D^*(\omega_N) = 0\left(\frac{\ln^k N}{N}\right) ,$$

which is the same asymptotic as for LP_τ or LP_o in base b. Halton's sequences are quite simple and widely used in particle methods.
A last attempt to compare the three possibilities to get low discrepancy sequences: If we write

$$D^*(\omega_N) = A_k \frac{\ln^k N}{N} + o\left(\frac{\ln^k N}{N}\right) ,$$

then

$$\ln A_k = 0(k \ln k) \quad \text{for } k\to\infty$$

for the best choice of $r_1,\ldots,r_k$ in Halton,

$$\ln A_k = 0(k \ln \ln k) \quad \text{for minimal } \tau$$

in LP_τ-sequences and

$$A_k \to 0 \quad \text{for } LP_o\text{-sequences in minimal base b.}$$

This is in favour of Faure's idea in high dimensions. LP_τ and LP_o in base b seem to be favourable since they have sections ($N=b^m$) with the better asymptotic $0\left(\frac{\ln N^{k-1}}{N}\right)$.

Non uniform distributions:

Until now we have compared ω_N with $\chi_{[0,1]^k}$. We take now more general densities - we could even use measures $\mu \in M$ but restrict ourselves to $M_{ac}(\Omega)$. For Ω we may discuss compact and not compact sets in $\mathbb{R}^k$; since $[0,1)^k$ is a paradigm for compact sets and $\mathbb{R}^k$ for non compact we shall concentrate on $\Omega=\mathbb{R}^k$. However, in our applications Ω quite often is the exterior of a compact set; we may treat this case by extending the density of $f \in M_{ac}(\Omega)$ to whole $\mathbb{R}^k$ by putting $f(P) = 0$ for $P \in \Omega^c$. This is certainly not an ideal solution, but easier than to care for boundary terms.

Let therefore be f a density in $\mathbb{R}^k$, i.e. $f \in M_{ac}(\mathbb{R}^k)$, $[P,Q)$ an interval of $\mathbb{R}^k$ and $\Lambda([P,Q);\omega_N)$ = number of points of ω_N in $[P,Q)$. We call $(\omega_N)_{N\in\mathbb{N}}$ "asymptotically f-distributed" or simply "f-distributed", if

$$\lim_{N\to\infty} \frac{\Lambda([P,Q);\omega_N)}{N} = \int_{[P,Q)} f\, dR .$$

Again, (ω_N) is f-distributed iff $\frac{1}{N}\sum_{i=1}^{N} \Phi(P_i) \to \int_{\mathbb{R}^k} \Phi f\, dP$

for all bounded R-integrable functions Φ on $\mathbb{R}^k$.

We define

$$D(\omega_N;f) := \sup_{P\leq Q} \left|\frac{\Lambda([P,Q);\omega_N)}{N} - \int_{[P,Q)} f\, dR\right|$$

and call it discrepancy of ω_N with respect to f (or simply f-discrepancy); similarly $D^*(\omega_N;f) = \sup_{Q\in\mathbb{R}^k}\left|\frac{\Lambda(R_Q;\omega_N)}{N} - \int_{R_Q} f\, dP\right|$ with $R_Q=\{P/P\leq Q\}$.

We shall now prove the most important theorems in this most general setting.

Theorem 2.2:

$(\omega_N)_{N\in\mathbb{N}}$ is f-distributed iff $D^*(\omega_N;f) \to 0$.

Proof:

Sufficiency is trivial - we show necessity. Choose $m \in \mathbb{N}$ and a partition of the p_1-axis by

$$-\infty = p_1^{(0)} < p_1^{(1)} < \dots < p_1^{(m-1)} < p_1^{(m)} = \infty$$

such that

$$\int_{p_1^{(i)}}^{p_1^{(i+1)}} \int_{\mathbb{R}^{k-1}} f\, dP = \frac{1}{m} \quad \text{for } i=0,1,\dots,m-1$$

(this is possible since $f \in \mathbb{L}^1$ and $\int f\, dP = 1$.)

Then choose a partition of the p_2-axis such that

$$\int_{p_1^{(i)}}^{p_1^{(i+1)}} \int_{p_2^{(j)}}^{p_2^{(j+1)}} f\, dP = \frac{1}{m^2} \quad \text{for } i,j=0,1,\dots,m-1 .$$

Continuing in this way we get altogether m^k rectangles R_α in $\mathbb{R}^k$ such that

$$\int_{R_\alpha} f\, dP = \frac{1}{m^k}, \quad \alpha=1,\dots,m^k .$$

We know that $\frac{\Lambda(R_\alpha;\omega_N)}{N} \longrightarrow \int_{R_\alpha} f\, dP = \frac{1}{m^k}$, such that

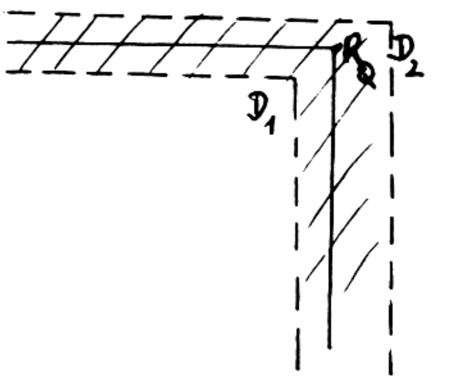
$$\frac{1}{m^k}(1-\frac{1}{m^2}) \leq \frac{\Lambda(R_\alpha;\omega_N)}{N} \leq \frac{1}{m^k}(1+\frac{1}{m^2}) \quad \text{for } N \geq N_o .$$

Choose $Q \in \mathbb{R}^k$ arbitrary but fixed and let $I_1 = \{\alpha/R_\alpha \subset R_Q\}$, I_2 be the smallest set of indices α such that $R_Q \subset \bigcup_{\alpha\in I_2} R_\alpha = D_2$. With D_1 we denote $\bigcup_{\alpha\in I_1} R_\alpha$, such that $D_1 \subset R_Q \subset D_2$.

Clearly

$$\int_{D_1} f\, dP \leq \int_{R_Q} f\, dP \leq \int_{D_2} f\, dP .$$

What about the difference $\int_{D_2/D_1} f\, dP$? The layer between D_1 and D_2 is



essentially $(k-1)$-dimensional, i.e. we need $C\cdot m^{k-1}$ intervals R_α to fill it up (C depends only on dimension); therefore

$$\int_{D_2/D_1} f\, dP \leq C\,\frac{m^{k-1}}{m^k} = \frac{C}{m} .$$

Since also $\Lambda(D_1;\omega_N) \le \Lambda(R_Q;\omega_N) \le \Lambda(D_2;\omega_N)$ we get

$$\int_{D_1} f\, dP - \frac{\Lambda(D_2;\omega_N)}{N} \le \int_{R_Q} f\, dP - \frac{\Lambda(D_2;\omega_N)}{N} \le \int_{D_2} f\, dP - \frac{\Lambda(D_2;\omega_N)}{N} .$$

The left hand side

$$\int_{D_1} f\, dP - \frac{\Lambda(D_2;\omega_N)}{N} = \int_{D_1} f\, dP - \frac{\Lambda(D_2;\omega_N)}{N} - \frac{\Lambda(D_2;\omega_N)}{N}$$

$$\ge \#I_1 \frac{1}{m^k} - \#I_1 \frac{1}{m^k}(1+\frac{1}{m^2}) - Cm^{k-1} \frac{1}{m^k}(1+\frac{1}{m^2})$$

$$= -\#I_1 \frac{1}{m^{k+2}} - \frac{C}{m}(1+\frac{1}{m^2}) \ge -\frac{1}{m^2} - \frac{C}{m}(1+\frac{1}{m^2})$$

(here $\#I_1$ is the cardinality of I_1, which is less than m^k).
Similarly the right hand side may be estimated

$$\int_{D_2} f\, dP - \frac{\Lambda(D_1;\omega_N)}{N} \le \frac{C}{m} + \frac{1}{m^2} .$$

We arrive at

$$-\frac{1}{m^2} - \frac{C}{m}(1+\frac{1}{m^2}) \le \int_{R_Q} f\, dP - \frac{\Lambda(R_Q;\omega_N)}{N} \le \frac{1}{m^2} + \frac{C}{m} \qquad \text{for } N \ge N_o.$$

For arbitrary $\varepsilon > 0$ we can find an m_o such that $\frac{1}{m_o^2} + \frac{C}{m_o}(1+\frac{1}{m_o^2}) < \varepsilon$ and choose now N_o corresponding to m_o; then

$$\left|\int_{R_Q} f\, dP - \frac{\Lambda(R_Q;\omega_N)}{N}\right| < \varepsilon \quad \text{for } N \ge N_o \text{ and all } Q.$$

This proves the uniformity of the convergence with respect to Q.

□

We may continue as in the case of uniform distributions:
For k=1 we get an explicit expression for $D(\omega_N;f)$, if $\omega_N = (x_1,\dots,x_N)$ is ordered $x_1 \le x_2 \le \dots \le x_N$:

$$D^*(\omega_N;f) = \max_{1\le j\le N} \max\left\{\left|\frac{j-1}{N} - \int_{-\infty}^{x_j} f\, dx\right|, \left|\frac{j}{N} - \int_{-\infty}^{x_j} f\, dx\right|\right\}$$

$$= \frac{1}{2N} + \max_{1\le j\le N} \left|\frac{j-\frac{1}{2}}{N} - \int_{-\infty}^{x_j} f\, dx\right|$$

Since we cannot generalize this result to $k>1$, we omit the proof; however, we conclude that the smallest D^*, which is again $\frac{1}{2N}$, is given for x_j, which are defined by

$$\int_{-\infty}^{x_j} f\,dx = \frac{j-\frac{1}{2}}{N} .$$

We may rewrite this by introducing the distribution $F(y) = \int_{-\infty}^{y} f\,dx$, which is increasing and continuous: $F(x_j) = y_j$, where $(y_1,\ldots,y_N) = \tilde{\omega}_N$ has the optimal discrepancy with respect to the uniform distribution. This is a more general result, which is easily proved, if $f>0$ everywhere, i.e. F is strictly increasing such that F^{-1} exists. Then $D^*(\omega_N;f) = D^*(F\omega_N)$ with $F\omega_N = (F(x_1),\ldots,F(x_N))$:

$$\begin{aligned} D^*(F\omega_N) &= \sup_{0\le y\le 1} \left|\frac{1}{N}\,\#\left\{i/F(x_i)<y\right\} - y\right| \\ &= \sup_{0\le y\le 1} \left|\frac{1}{N}\,\#\left\{i/x_i<F^{-1}(y)\right\} - y\right| \\ &= \sup_{-\infty\le x\le +\infty} \left|\frac{1}{N}\,\#\left\{i/x_i<x\right\} - F(x)\right| \\ &= D^*(\omega_N;f) . \end{aligned}$$

This is also true without the assumption that $f>0$ everywhere - we have just to be a bit more careful, since F^{-1} doesn't exist.
We will use this result for generating f-distributed sequences later.

We shall now proceed to the second important result in the theory - the Koksma-Hlawka inequality. Earlier we had defined the Hardy-Krause variation etc. in $[0,1]^k$ - we are now in $\mathbb{R}^k$. We need a bit more: The functions Φ we want to integrate are not only defined on $\mathbb{R}^k$, but even on $\bar{\mathbb{R}}^k$ - $\Phi(x_1,\ldots,x_k) \in \mathbb{R}$ even if some of the coordinates have the value $+\infty$ or $-\infty$.
We consider again an ℓ-tupel $\beta = (i_1,\ldots,i_\ell)$, $1\le i_1<\ldots<i_\ell\le k$ and denote by $\hat{x}_\beta$ the $(k-\ell)$-tupel consisting of all components different from $x_{i_1},\ldots,x_{i_\ell}$; further $\Phi[\beta](x_{i_1},\ldots,x_{i_\ell}) = \Phi(x)/\hat{x}_\beta = (+\infty,\ldots,+\infty)$, $\omega_N[\beta]$ analogously and for $f \in M_{ac}$ we put

$$f\{\beta\}(x_{i_1},\ldots,x_{i_\ell}) = \int_{-\infty}^{+\infty} \ldots \int_{-\infty}^{+\infty} f(x)d\hat{x}_\beta .$$

In the definition of the variations of Φ we need now partitions $Z = (Z^{(1)},\dots,Z^{(k)})$ with $Z^{(j)} = (z_o^{(j)},\dots,z_{m_j}^{(j)})$, $-\infty=z_o^{(j)}<\dots<z_{m_j}^{(j)}=+\infty$, $j=1,\dots,k$, so that $V^{(k)}[\Phi]$ is exactly defined as on page 36.
We get now

Theorem 2.1":
For $f \in M_{ac}(\mathbb{R}^k)$, Φ of bounded Koksma-Hlawka variation in $\mathbb{R}^k$ and $\omega_N = (P_1,\dots,P_N)$ we have

$$\left|\frac{1}{N}\sum_{n=1}^{N} \Phi(P_n) - \int_{[0,1]^k} \Phi(P)f(P)dP\right|$$

$$\le \sum_{\ell=1}^{k} \sum_{1\le i_1<\dots<i_\ell\le k} V^{(\ell)}[\Phi[i_1,\dots,i_\ell]]D^*(\omega_N[i_1,\dots,i_\ell],f\{i_1,\dots,i_\ell\})$$

$$\le V[\Phi]D^*(\omega_N,f) .$$

__Proof for k=2__ (for general k the proof is more technical, but principally the same) and for the simpler version $\le V[\Phi]D_N^*(\omega_N,f)$:
We take a grid Z in $\mathbb{R}^2$, defined by $Z^{(1)}\times Z^{(2)}$, where

$$Z^{(1)} : -\infty = x_o < \dots < x_I = \infty , \quad Z^{(2)} : -\infty = y_o < \dots < y_J = +\infty$$

so that the grid knots are $Q_{i,j} = (x_i,y_j)$, $0\le i\le I$, $0\le j\le J$.
We had already $\Delta_{1,2}\Phi(Q_{ij}) = \Phi(Q_{i+1,j+1})-\Phi(Q_{i+1,j})-\Phi(Q_{i,j+1})+\Phi(Q_{i,j})$, which we now denote also by $\Delta^{ij}\Phi$. In this notation

$$V^{(2)}[f] = \sup_Z \sum_{i=0}^{I-1} \sum_{j=0}^{J-1} |\Delta^{ij}\Phi| .$$

We need the following

__Lemma 1:__
Let Φ, h be defined in $\bar{\mathbb{R}}^2$, then

$$\sum_{i=0}^{I-1} \sum_{j=0}^{J-1} \Phi(Q_{i+1,j+1})\Delta^{ij}h$$

$$= \sum_{i=0}^{I-1} \sum_{j=0}^{J-1} h(Q_{ij})\Delta^{ij}\Phi + \sum_{i=0}^{I-1} h(Q_{i,0})[\Phi(Q_{i+1,0})-\Phi(Q_{i,0})] -$$

$$- \sum_{i=0}^{I-1} h(Q_{i,J})(\phi(Q_{i+1,J})-\phi(Q_{i,J})) + \sum_{j=0}^{J-1} h(Q_{0,j})[\phi(Q_{0,j+1})-\phi(Q_{0,j})]$$

$$- \sum_{j=0}^{J-1} h(Q_{I,j})[\phi(Q_{I,j+1})-\phi(Q_{I,j})] + \phi(Q_{I,J})h(Q_{I,J})$$

$$+ \phi(Q_{0,0})h(Q_{0,0}) - \phi(Q_{I,0})h(Q_{I,0}) - \phi(Q_{0,J})h(Q_{0,J}) .$$

The proof of the Lemma is straightforward, applying twice Abel summation

$$\sum_{\nu=1}^{k} a_\nu b_\nu = \sum_{\nu=1}^{k} A_\nu(b_\nu - b_{\nu+1}) - A_o b_1 + A_k b_{k+1}, \quad \text{where } A_\nu = a_o + \ldots + a_\nu .$$

□

If ϕ is of bounded Koksma-Hlawka variation, we may write it as

$$\phi = \phi_+ - \phi_- - \phi(-\infty,-\infty) ,$$

where ϕ_+, ϕ_- are monotonously increasing, bounded functions. We consider therefore such functions ψ and $f \in M_{ac}(\mathbb{R}^2)$. We need a

Lemma 2:

For each $\varepsilon > 0$ there exists a partition Z, such that for

$$R_{ij} = R_{(Q_{ij}, Q_{i+1,j+1}]}$$

the inequality

$$0 \le \sum_{i=0}^{I-1} \sum_{j=0}^{J-1} [\psi(Q_{i+1,j+1}) - \psi(Q_{i,j})] \int_{R_{i,j}} f \, dP \le \varepsilon .$$

holds.

Remark:

The expressions

$$O(Z) = \sum_{i=0}^{I-1} \sum_{j=0}^{J-1} \psi(Q_{i+1,j+1}) \int_{R_{i,j}} f \, dP$$

and

$$U(Z) = \sum_{i=0}^{I-1} \sum_{j=0}^{J-1} \psi(Q_{i,j}) \int_{R_{i,j}} f \, dP$$

are a kind of upper and lower Riemann-Stieltjes sums for $\int_{\mathbb{R}^2} \psi f\, dP$ and the lemma shows that $0 \leq \int \psi f\, dP - U(Z) \leq \varepsilon$, $0 \leq O(Z) - \int \psi f\, dP \leq \varepsilon$.

Proof of Lemma 2:

Let $F(Q) = \int_{R_Q} f\, dP$. Since $f \in \mathbb{L}^1(\mathbb{R}^2)$, for any $\delta > 0$ there exists a partition Z such that

$$|F(P)-F(Q)| \leq \delta \quad \text{for } P,Q \in R_{i,j}, \text{ where } 0 \leq i \leq I-1,\ 0 \leq j \leq J-1\ ;$$

this follows from

$$|F(P)-F(Q)| \leq |F(Q_{i+1,j+1})-F(Q_{i,j})|$$

$$= \int_{R_{Q_{i+1,j+1}}-R_{Q_{i,j}}} f\, dP$$

$$\leq \int_{R_\delta \cap [R_{Q_{i+1,j+1}}-R_{Q_{i,j}}]} f\, dP + \frac{\delta}{2} < \frac{\delta}{2}+\frac{\delta}{2} = \delta\ ,$$

where we have chosen a square R_δ such that $\int_{R^2-R_\delta} f\, dP < \frac{\delta}{2}$ and then the partition fine enough, such that

$$R_\delta \cap [R_{Q_{i+1,j+1}}-R_{Q_{i,j}}]$$

has sufficiently small Lebesgue measure.

Now

$$\int_{R_{i,j}} f\, dP = \Delta^{ij}F\ , \quad 0 \leq i \leq I-1,\ 0 \leq j \leq J-1\ ;$$

(realize that F is defined in $\bar{\mathbb{R}}^2$, for example by $F(Q_{i,0}) = F(Q_{0,j})=0$).
We use Lemma 1 to get

$$0 \leq O(Z)-U(Z) = \sum_{i=0}^{I-1} \sum_{j=0}^{J-1} [\psi(Q_{i+1,j+1})-\psi(Q_{i,j})]\Delta_{i,j}F$$

$$= - \sum_{i=0}^{I-1} \sum_{j=0}^{J-1} [F(Q_{i+1,j+1})-F(Q_{i,j})]\Delta_{i,j}\psi + \sum_{i=0}^{I-1} [F(Q_{i+1,J})-F(Q_{i,J})] \cdot$$

$$\cdot\ [\psi(Q_{i+1,J})-\psi(Q_{i,J})] + \sum_{j=0}^{J-1} [F(Q_{I,j+1})-F(Q_{I,j})][\psi(Q_{I,j+1})-\psi(Q_{I,j})]\ .$$

We remember that the F-differences are less than δ and the ψ-expression may be estimated by the one- and two-dimensional Vitali variations $V^{(1)}[\psi]$ and $V^{(2)}[\psi]$.

In this way we arrive at

$$0 \le O(Z)-U(Z) \le \delta[V^{(2)}[\psi]+V^{(1)}[\psi(\cdot,\infty)] + V^{(1)}[\psi(\infty,\cdot)]] = \delta V[\psi] .$$

For $\delta = \frac{\varepsilon}{V[\psi]}$ we get the result of Lemma 2. □

We can now prove the theorem:

Assume first that Φ is monotonously increasing and bounded and choose for given $\varepsilon>0$ according to Lemma 2 a partition Z such that

$$0 \le \sum_{i=0}^{I-1} \sum_{j=0}^{J-1} [\Phi(Q_{i+1,j+1})-\Phi(Q_{i,j})] \int_{R_{ij}} f \, dP \le \varepsilon .$$

We mention that R_{ij} contains even unbounded intervals including points at infinity (for i or j=0 or I-1, J-1) so that

$$\bigcup_{i,j=0}^{I-1,J-1} R_{i,j} = \bar{\mathbb{R}}^2 ;$$

moreover $R_{i,j} \cap R_{i',j'} = \emptyset$ for $(i,j) \neq (i',j')$. Therefore

$$\sum_{i=0}^{I-1} \sum_{j=0}^{J-1} \chi_{R_{ij}}(P) = 1 \quad \forall P \in \bar{\mathbb{R}}^2 .$$

Now consider the stepfunctions

$$\Phi_1(Q) = \sum_{i,j} \Phi(Q_{ij})\chi_{R_{ij}}(Q), \quad \Phi_2(Q) = \sum_{i,j} \Phi(Q_{i+1,j+1})\chi_{R_{ij}}(Q) ;$$

since Φ is monotonously increasing, we have $\Phi_1 \le \Phi \le \Phi_2$ and

$$\int \Phi_1 f \, dP \le \int \Phi f \, dP \le \int \Phi_2 f \, dP .$$

Now consider $\omega_N = (P_1,\dots,P_N)$ (for the first time in the proof); we get

$$\frac{1}{N} \sum_{\ell=1}^{N} \Phi_1(P_\ell) \le \frac{1}{N} \sum_{\ell=1}^{N} \Phi(P_\ell) \le \frac{1}{N} \sum_{\ell=1}^{N} \Phi_2(P_\ell) .$$

Put

$$L(\Phi) := \frac{1}{N} \sum_{\ell=1}^{N} \Phi(P_\ell) - \int \Phi f \, dP$$

and realize that

$$\int(\phi_1-\phi_2)f\,dP + L(\phi_1) \le L(\phi) \le L(\phi_2) + \int(\phi_2-\phi_1)f\,dP .$$

This yields the estimation

$$|L(\phi)| \le \max_{r=1,2} |L(\phi_r)| + \int(\phi_2-\phi_1)f\,dP$$

$$= \max_{r=1,2} |L(\phi_r)| + \sum_{i=0}^{I-1} \sum_{j=0}^{J-1} [\phi(Q_{i+1,j+1})-\phi(Q_{i,j})] \int_{R_{i,j}} f\,dP .$$

We had constructed Z such that the last term is less or equal ε and arrive at

$$|L(\phi)| \le \max |L(\phi_r)| + \varepsilon .$$

It remains to estimate $L(\phi_r)$:
Since $Q_{i+1,j+1} \in R_{ij}$ we have $\phi_r(Q) = \phi_r(Q_{i+1,j+1})$ for $Q \in R_{ij}$, r=1,2 and

$$L(\phi_r) = \sum_{i=0}^{I-1} \sum_{j=0}^{J-1} \phi_r(Q_{i+1,j+1})L(\chi_{R_{ij}}) .$$

We want to apply Lemma 1 to this expression and have to express

$$L(\chi_{R_{ij}}) \text{ as } \Delta_{i,j}H \text{ for some H ;}$$

$$\text{with } h(P;Q) := \begin{cases} \chi_{R(P)}(Q) & \text{for } P \succ (-\infty,-\infty) \\ 0 & \text{if } P=(-\infty,y) \text{ or } P=(x,-\infty) \end{cases} ,$$

we get

$$\Delta_{i,j}h(\cdot,Q) = \chi_{R_{i,j}}(Q) \text{ and } L(\chi_{R_{i,j}}) = L(\Delta_{i,j}h(\cdot,Q)) = \Delta_{ij}^{(P)}(L_Q h(P;Q))$$

with a selfexplaining notation.
With

$$H(P) := L_Q h(P;Q)$$

we get

$$L(\phi_r) = \sum_{i,j} \phi_r(Q_{i+1,j+1})\Delta_{i,j}H .$$

Before applying Lemma 1, we realize that

$$H(Q) = L(h(Q;\cdot)) = \frac{1}{N}\sum_{\ell=1}^{N} h(Q;P_\ell) - \int_{\mathbb{R}^2} h(Q;P)f(P)dP$$

$$= \frac{1}{N}\sum_{\ell=1}^{N} \chi_{R(Q)}(P_\ell) - \int_{R(Q)} f\, dP$$

for $Q>(-\infty,-\infty)$ and zero else, especially

$$H(Q_{I,J}) = \frac{1}{N}\sum_{\ell=1}^{N} \chi_{\mathbb{R}^2}(P_\ell) - \int_{\mathbb{R}^2} f\, dP = 0 .$$

Now apply Lemma 1:

$$L(\phi_r) = \sum_{i=0}^{I-1}\sum_{j=0}^{J-1} H(Q_{i,j})\Delta_{i,j}\phi_r - \sum_{i=0}^{I-1} H(Q_{i,J})[\phi_r(Q_{i+1,J})-\phi_r(Q_{i,J})]$$

$$- \sum_{j=0}^{J-1} H(Q_{I,j})[\phi_r(Q_{I,j+1})-\phi_r(Q_{I,j})] .$$

Since $|H(Q)| \leq D^*(\omega_N;f)$ we conclude

$$|L(\phi_r)| \leq D^*(\omega_N;f)[V^{(2)}[\phi_r]+V^{(1)}[\phi_r(\cdot,\infty)]+V^{(1)}[\phi_r(\infty,\cdot)]] .$$

Clearly the variations of ϕ_1, ϕ_2 are less than those of ϕ, therefore

$$|L(\phi_r)| \leq V[\phi]\cdot D^*(\omega_N;f)$$

and

$$|L(\phi)| \leq V[\phi]D^*(\omega_N;f)+\varepsilon$$

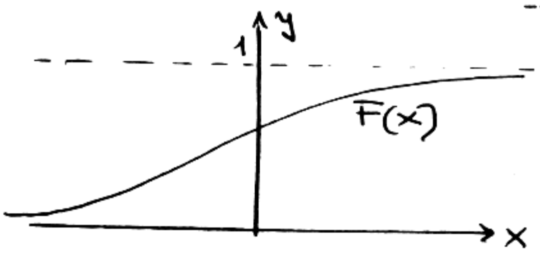
for arbitrary $\varepsilon>0$.
We arrive at the final result, at least for monotonously increasing bounded ϕ; but by decomposition of an arbitrary ϕ of bounded Hardy-Krause variation and by realizing $L(\phi(-\infty,-\infty)) = 0$, we get

$$|L(\phi)| \leq D^*(\omega_N;f)[V[\phi_+]+V[\phi_-]] \leq D^*(\omega_N;f)V[\phi]$$

for all ϕ of the considered class. □

We can now turn to the question how to construct ω_N with small $D(\omega_N;f)$ for given f. We know the answer for $f=\chi_{[0,1]^k}$ - how can we transform uniformly distributed sequences (ω_N) into f-distributed ones? Or do there exist direct methods for construction?

The one-dimensional case is relatively simple: Take $f \in \mathcal{M}_{ac}(\mathbb{R})$ and define as we did before $F(x) = \int_{-\infty}^{x} f(t)dt$; F maps $\mathbb{R}$ into $(0,1)$.



Take $\omega_N = (y_1, \dots, y_N) \subset (0,1)$ u.d. and define

$$x_j = F^{-1}(y_j) \ .$$

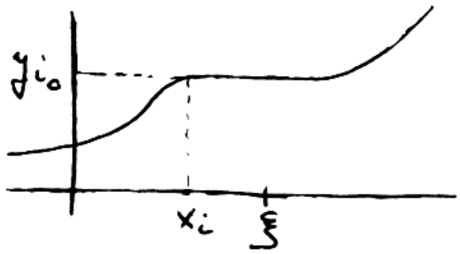
F is monotonously increasing, but not necessarily invertible; if it is not, we define

$$F^{-1}(y) := \inf\{x / F(x) \geq y\} \ .$$

For $\tilde{\omega}_N := F^{-1}\omega_N$ we get

$$D^*(\tilde{\omega}_N; f) = D^*(\omega_N)$$

- we have seen this already above at least if F^{-1} exists in the classical sense



$$\#\{i/x_i \leq \xi\} = \#\{i/F^{-1}(y_i) \leq \xi\}$$

$$= \#\{i/y_i \leq F(\xi)\}$$

$$D^*(\tilde{\omega}_N; f) = \sup_{\xi \in \mathbb{R}} \left| \frac{1}{N}\#\{i/x_i \leq \xi\} - F(\xi) \right| = \sup_{\xi \in \mathbb{R}} \left| \frac{1}{N}\#\{i/y_i \leq F(\xi)\} - F(\xi) \right|$$

$$= \sup_{0 \leq \eta \leq 1} \left| \frac{1}{N}\#\{i(y_i < \eta\} - \eta \right| = D^*(\omega_N) \ .$$

<u>**Example:**</u>

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{for } x \geq 0 \\ 0 & \text{else} \end{cases} \implies F(x) = (1 - e^{-\lambda x})$$

$$F^{-1}(y) = -\frac{1}{\lambda}\ln(1-y), \text{ i.e.}$$

$$\left\{\left(-\frac{1}{\lambda}\ln(1-y_1), \dots, -\frac{1}{\lambda}\ln(1-y_N)\right)\right\}_{N \in \mathbb{N}}$$

is f-distributed, if $(y_1, \dots, y_N)$ is uniformly distributed.

(Since with $(y_1, \dots, y_N)$ also $(1-y_1, \dots, 1-y_N)$ is uniformly distributed, we may also use $(-\frac{1}{\lambda}\ln y_1, \dots, -\frac{1}{\lambda}\ln y_N)$ as f-distributed sequence.)

The problem is that the inversion of F may create problems - as it does for example for $f(x) = \frac{1}{\sqrt{\pi}} e^{-x^2/2}$, an important special case.

Hlawka & Mück (A Transformation of Equidistributed Sequences - in Zaremba, S.K. (Hrsg.) Appl. of Number Theory to Num. Analysis, Acad. Press 1972) made a proposal which is useful only in higher dimensions. Assume that $\mathrm{supp}(f) \subset [0,1]$, so that $F(x) = 0$ for $x \le 0$, $F(x) = 1$ for $x \ge 1$ and assume moreover that f is strictly increasing such that F^{-1} exists.

For given $y \in (0,1)$ we have to find $F^{-1}(y)$. To get it we take a uniformly distributed set $\omega_N = (z_1, \ldots, z_N) \subset [0,1)$ and realize that

$$\left| \frac{1}{N} \sum_{j=1}^{N} \chi_{[0,\alpha)}(z_j) - \alpha \right| \le D^*(\omega_N) .$$

Take $\frac{1}{N} \sum \chi_{[0,\alpha)}(z_j)$ as an approximation for α, if $\alpha = F^{-1}(y)$:

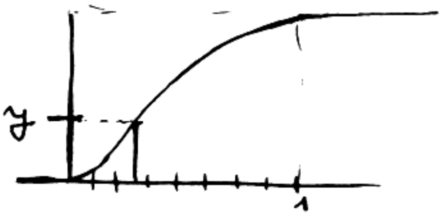
$$\frac{1}{N} \sum_{j=1}^{N} \chi_{[0,F^{-1}(y))}(z_j) = \frac{1}{N} \# \left\{ i / z_i \le F^{-1}(y) \right\} = \frac{1}{N} \# \left\{ i / F(z_i) \le y \right\}$$

$$= \frac{1}{N} \# \left\{ i / 0 \le y - F(z_i) \right\} .$$

Generally $0 \le 1 + y - F(z_i) < 2$ (since $0 \le y < 1$ and $0 \le F(z_i) \le 1$) and the Gauss bracket $[1+y-F(z_i)]$ has the values 0 or 1, 1 just when $y - F(z_i) \ge 0$. This leads to $\frac{1}{N} \# \left\{ i \ / \ 0 \le y - F(z_i) \right\} = \frac{1}{N} \sum_{i=1}^{N} [1+y-F(z_i)]$ as approximation for $F^{-1}(y)$. If we use for y and z the same uniformly distributed sequence $(y_1, \ldots, y_N)$ we get

$$x_j = F^{-1}(y_j) \approx \frac{1}{N} \sum_{i=1}^{N} [1+y_j-F(y_i)] , \quad 1 \le j \le N$$

as an f-distributed sequence. If one looks a bit nearer to this idea and take $z_i = \frac{2i-1}{2N}$, one realizes that the determination of $F^{-1}(y)$ is done just by testing whether $F(\frac{1}{2N}) < y$, $F(\frac{3}{2N}) < y$ ... until we find the first j with $F(\frac{2j+1}{2N}) \geq y$ - then $\frac{2j-1}{2N}$ is a $\frac{1}{N}$-approximation for $F^{-1}(y)$.

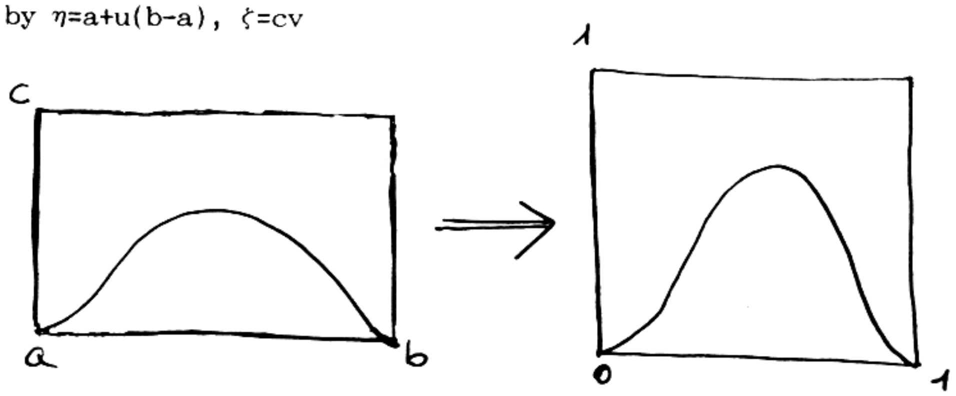


The regula falsi is certainly a better choice - in one dimension!

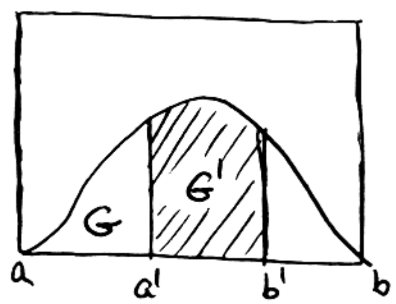
Another idea is given by the rejection method: Assume that $\text{supp}(f) \subset [a,b]$ and $|f(x)| \leq C$. We define a mapping

$$[0,1]^2 \longrightarrow [a,b]\times[0,c] \ , \ (u,v) \longrightarrow (\eta,\zeta)$$

by $\eta=a+u(b-a)$, $\zeta=cv$



Now choose $\left((u_1,v_1),\dots,(u_N,v_N)\right)_{N\in\mathbb{N}}$, a uniformly distributed sequence in $[0,1]^2$, compute $(\eta_i,\zeta_i) = \Gamma_i$. $\omega_N = (\Gamma_1,\dots,\Gamma_N)$ is uniformly distributed in $[a,b]\times[0,c]$. If $\zeta_i \leq f(\eta_i)$ choose $x_i = \eta_i$ - if $\zeta_i > f(\eta_i)$, we forget Γ_i. We get a subset $\tilde{\omega}_M = (x_{i_1},\dots,x_{i_M})_{M\in\mathbb{N}}$ of $(\eta_1,\dots,\eta_N)$ which is f-distributed:



$$\frac{\Lambda(\omega_N;G')}{N} \longrightarrow \frac{\lambda(G')}{c(b-a)} = \frac{\int_{a'}^{b'} f(t)dt}{c(b-a)}$$

$$\frac{\Lambda(\omega_N;G)}{N} \longrightarrow \frac{\lambda(G)}{c(b-a)} = \frac{1}{c(b-a)}$$

The number of points Γ_i with $\zeta_i \leq f(\eta_i)$ is $\Lambda(\omega_N;G)$; since $M = \Lambda(\omega_N;G)$, we get

$$\frac{1}{M}\Lambda(\tilde{\omega}_M;[a',b']) = \frac{\Lambda(\omega_N;G')}{\Lambda(\omega_N;G)} = \frac{\Lambda(\omega_N;G')}{N}\ \frac{N}{\Lambda(\omega_N;G)} \longrightarrow \int_{a'}^{b'} f(t)dt \ .$$

Next we mention the Box-Muller method for generating f-distributed sequences where $f(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$.

The idea is to use

$$\frac{1}{\sqrt{2\pi}} e^{-x^2/2} \cdot \frac{1}{\sqrt{2\pi}} e^{-y^2/2} = \frac{1}{2\pi} e^{-(x^2+y^2)/2}$$

and to transform it into polarcoordinates. Since (K_R is the circle with radius R)

$$\frac{1}{2\pi} \iint_{K_R} e^{(-x^2+y^2)/2} \, dxdy = \frac{1}{2\pi} \int_0^R \int_0^{2\pi} e^{-r^2/2} r \, dr d\phi$$

it is equivalent to the density $re^{-r^2/2} \frac{1}{2\pi} \chi_{[0,2\pi]}(\phi)$, i.e. the uniform distribution with respect to the angle and the $re^{-r^2/2}$-distribution with respect to r. $re^{-r^2/2}$ has a primitive $\int_0^r se^{-s^2/2} \, ds = 1-e^{-r^2/2}$, $r \geq 0$ and, starting with a uniformly distributed set $(u_1,\ldots,u_N)$, we get by $F(r_i) = u_i$ a $re^{-r^2/2}$-distributed set $(r_1,\ldots,r_N)$; now $r_i^2 = -2\ln(1-u_i)$ (or $-2\ln(\tilde{u}_i)$ with an as well uniformly distributed sequence $\tilde{u}_i$).
If we use another uniformly distributed sequence v_i, we get a $\frac{1}{2\pi} \chi_{[0,2\pi]}$ distributed sequence $(\phi_1,\ldots,\phi_N)$ through $\phi_i = 2\pi v_i$.
Therefore

$$x_i^{(1)} = (-2 \ln \tilde{u}_i)^{1/2} \cos 2\pi v_i \quad \text{and}$$

$$x_i^{(2)} = (-2 \ln \tilde{u}_i)^{1/2} \sin 2\pi v_i$$

are normally distributed sequences.

In more than one dimension there are not so many ideas on the market - if the density doesn't factorize as the Maxwell distribution does.
We use here the idea of Hlawka & Mück ("Über eine Transformation von gleichverteilten Folgen II", Computing 9, 127-138, 1972) and consider a nonuniform distribution in $[0,1]^k$, more precisely one with a positive continuous density $f \in M_{ac}([0,1]^k)$.

We define now T: $[0,1]^k \to [0,1]^k$ by

$$T_j(P) = T_j(p_1,\ldots,p_j)$$

$$= \frac{\int_0^{p_j} (\int_0^1 \ldots \int_0^1 f(p_1,\ldots,p_{j-1},\tau_j,\ldots,\tau_k)d\tau_{j+1}\ldots d\tau_k)d\tau_j}{\int_0^1 (\int_0^1 \ldots \int_0^1 f(p_1,\ldots,p_{j-1},\tau_j,\ldots,\tau_k)d\tau_{j+1}\ldots d\tau_k)d\tau_j}$$

for $1 \leq j \leq k$. We realize that $p_j \to T_j(p_1,\ldots,p_{j-1},p_j)$ is strongly increasing (since $f > 0$!). It follows that T is injective: If $P \neq Q$, say $p_{\mathfrak{j}} \neq q_{\mathfrak{j}}$ and $p_j = q_j$ for $1 \leq j < \mathfrak{j}$, then $T_{\mathfrak{j}}(P) \neq T_{\mathfrak{j}}(Q)$. T is also surjective:
$T_k(p_1,\ldots,p_{k-1},0) = 0$ and $T_k(p_1,\ldots,p_{k-1},1) = 1$.
If $f \in C^1([0,1]^k)$, then T is a regular transformation, i.e. a bijective differentiable mapping with nonvanishing Jacobian

$$J_T = \frac{\partial(T_1,\ldots,T_k)}{\partial(p_1,\ldots,p_k)} = \frac{\partial T_1}{\partial p_1} \ldots \frac{\partial T_k}{\partial p_k} \quad (\text{since } \frac{\partial T_j}{\partial p_i} = 0 \quad \text{for } i > j) \ ,$$

$$\frac{\partial T_j}{\partial p_j} = \frac{I_j}{I_{j-1}} \ , \text{ where } I_j = \int_0^1 \ldots \int_0^1 f(p_1,\ldots,p_j,\tau_{j+1},\ldots,\tau_k)d\tau_{j+1},\ldots,d\tau_k$$

and $I_o = \int_{[0,1]^k} f(P)dP = 1$ and $I_k = f$. We get

$$J_T = I_k = f > 0 \ .$$

This is the essential idea: T is a transformation, whose Jacobian is the given f. It follows that T^{-1}: $[0,1]^k \to [0,1]^k$ is a regular transformation too, whose Jacobian $J_{T^{-1}}$ is $\frac{1}{f}$. We therefore get

$$J_{T^{-1}}(P) = \frac{1}{f(T^{-1}(P))}$$

and

$$\int_{[0,1]^k} \Phi(Q)f(Q)dQ = \int_{[0,1]^k} \Phi(T^{-1}(P))f(T^{-1}(P))J_{T^{-1}}(P)dP$$

$$= \int_{[0,1]^k} \Phi(T^{-1}(P))dP$$

and realize that T^{-1} transforms uniform distribution into f-distribution: Let $\tilde{\omega}_N = (Q_1,\dots,Q_N) \subset (0,1)^k$ and $P_j = T(Q_j)$, $Q_j = T^{-1}(P_j)$, $j=1,\dots,N$ and Φ R-integrable, then

$$\left|\frac{1}{N}\sum_{j=1}^{N}\Phi(Q_j) - \int_{[0,1]^k}\Phi f\,dQ\right| = \left|\frac{1}{N}\sum_{j=1}^{N}\Phi(T^{-1}(P_j)) - \int_{[0,1]^k}\Phi\circ T^{-1}dP\right| ;$$

we see: If $\omega_N = (P_1,\dots,P_N)_{N\in\mathbb{N}}$ is uniformly distributed, then

$$\frac{1}{N}\sum_{j=1}^{N}\Phi\circ T^{-1}(P_j) \longrightarrow \int_{[0,1]^k}\Phi\circ T^{-1}dP \quad \text{for all } \Phi\circ T^{-1}$$

and therefore

$$\frac{1}{N}\sum_{j=1}^{N}\Phi(Q_j) \longrightarrow \int_{[0,1]^k}\Phi f\,dQ \quad \text{for all } \Phi$$

In order to get some information about $D^*(T^{-1}\omega_N;f)$ if $D^*(\omega_N)$ is given, we use the Koksma-Hlawka inequality once more.
We know that

$$\left|\frac{1}{N}\sum_{j=1}^{N}\Phi(Q_j) - \int\Phi f\,dP\right| \le D^*(\omega_N)V(\Phi\circ T^{-1}) .$$

If $\Phi\circ T^{-1}$ is smooth enough, one may substitute the Hardy-Krause variation by

$$\sum_{\substack{1\le \ell\le k\\(j_1,\dots,j_\ell)}}\int_{[0,1]^k}\left|\frac{\partial^\ell(\Phi\circ T^{-1})}{\partial p_{j_1}\dots\partial p_{j_\ell}}\right|dP .$$

One can manipulate this expression: Assume that

$$\left|\frac{\partial^\ell T_i}{\partial q_{j_1}\dots\partial q_{j_\ell}}\right| \le M$$

for all derivations up to order k and all components i.
Using the fact that $T^{-1}=S$ fulfills

$$T_i(S_1(P),\dots,S_k(P)) = P_i$$

or

$$\sum_{j=1}^{k} \frac{\partial T_i}{\partial q_j} \frac{\partial S_j}{\partial p_r} = \delta_{ir} ,$$

we get by Cramer's rule

$$\left|\frac{\partial S_j}{\partial p_r}\right| \le \frac{(k-1)!M^{k-1}}{f} ,$$

where as before f is the density, which we assumed to be positive. If we now care for that $f(P) \ge m \ \forall P \in [0,1]^k$, then

$$\left|\frac{\partial S_j}{\partial p_r}\right| \le C_1(k)\frac{M^k}{m} .$$

From this one gets by induction

$$\left|\frac{\partial^{\ell} S_j}{\partial p_{r_1} \dots \partial p_{r_\ell}}\right| \le C_\ell(k)\left(\frac{M^k}{m}\right)^{2\ell-1}$$

and finally an estimate for the difference:

$$\left|\frac{1}{N} \sum_{j=1}^{N} \Phi(Q_j) - \int \Phi f dP\right| \le C^{(k)} \cdot K\left(\frac{M^k}{m}\right)^{2k-1} D^*(\omega_N) ,$$

if

$$\left|\frac{\partial^{\ell} \Phi}{\partial q_{\nu_1} \dots \partial q_{\nu_\ell}}\right| \le K \quad (1 \le \nu_1, \dots, \nu_\ell \le k) .$$

What about $D^*(S\omega_N;f)$ compared to $D^*(\omega_N)$?
It is not as easy as in one dimension, where $D^*(S\omega_N;f) = D^*(\omega_N)$.
We mention that for k=1 from

$$\left|\frac{1}{N} \sum_{j=1}^{N} \Phi(S(Q_j)) - \int \Phi f \, dP\right| \le D^*(\omega_N)V(\Phi \circ S)$$

follows that $D^*(S\omega_N;f) \le 2D^*(\omega_N)$, since for characteristic functions Φ even $\Phi \circ S$ is characteristic and $V(\Phi \circ S) = 2$.
The estimate for $k>1$ is worse:

$$D^*(S\omega_N;f) \le C(D^*(\omega_N))^{1/k} ;$$

for the proof we refer to Hlawka/Mück.
The disadvantage of the method lies again in the need of computing $S=T^{-1}$. We had already a method in one dimension which we now transfer

to $k\geq 1$.

We start with a (uniformly distributed) sequence $(\omega_N)_{N\in\mathbb{N}}$, where $\omega_N = (P^{(1)},\dots,P^{(N)})$, $P^{(j)} = (p_1^{(j)},\dots,p_k^{(j)})$, denote $\hat{\omega}_N = S\omega_N$ and define

$$\tilde{\omega}_N = (\tilde{Q}^{(1)},\dots,\tilde{Q}^{(N)}),\ \tilde{Q}^{(j)} = (\tilde{q}_1^{(j)},\dots,\tilde{q}_k^{(j)}),\ j=1,\dots,N$$

by

$$\tilde{q}_i^{(j)} = \frac{1}{N}\sum_{\ell=1}^{N}[1+p_i^{(j)}-T_i(\tilde{q}_1^{(j)},\dots,\tilde{q}_{i-1}^{(j)},\tilde{p}_i^{(\ell)})]$$

$1\leq i\leq k,\ 1\leq j\leq N$.

$\tilde{\omega}_N$ may be used instead of $\hat{\omega}_N$ as f-distributed sequence since one gets the estimate

$$(*)\qquad \left|\frac{1}{N}\sum_{\ell=1}^{N}\Phi(Q^{(\ell)}) - \int \Phi f\, dP\right| \leq C^{(k)}K\left(\frac{M^k}{m}\right)^{2k-1}(1+2M)^k D^*(\omega_N)\ .$$

The idea is similar to the onedimensional case: We know that $0<p_i^{(j)}<1$. Consider $p_i \to F(p_1,\dots,p_{i-1},p_i)$ for fixed $p_1,\dots,p_{i-1}$: This function is strongly monotonously increasing and takes the values 0 and 1 for $p_i=0$ and $p_i=1$ respectively. For each $p_i^{(j)}$ there exists exactly one $\tilde{\tilde{q}}_i^{(j)} \in (0,1)$ such that $T_i(\tilde{q}_1^{(j)},\dots,q_{i-1}^{(j)},\tilde{\tilde{q}}_i^{(j)}) = p_i^{(j)}$; the construction of $\tilde{Q}^{(j)}$ reads now as

$$\tilde{q}_i^{(j)} = \frac{1}{N}\sum_{i=1}^{N}\underbrace{[1+T_i(\tilde{q}_1^{(j)},\dots,q_{i-1}^{(j)},\tilde{\tilde{q}}_i^{(j)}) - T_i(\tilde{q}_1^{(j)},\dots,\tilde{q}_{i-1}^{(j)},p_i^{(\ell)})]}\ .$$

The terms in the brackets are between 0 and 2, so that the summands are 0 or 1; 1, if $T_i(\dots,\tilde{\tilde{q}}_i^{(j)})-T_i(\dots,p_i^{(\ell)}) \geq 0$, i.e. if $\tilde{\tilde{q}}_i^{(j)} \geq p_i^{(\ell)}$. Therefore

$$\tilde{q}_i^{(j)} = \frac{1}{N}\sum_{\ell=1}^{N}\chi_{[0,\tilde{\tilde{q}}_i^{(j)}]}(p_i^{(\ell)})$$

and we arrive at

$$\left|\tilde{q}_i^{(j)}-\tilde{\tilde{q}}_i^{(j)}\right| = \left|\frac{1}{N}\sum_{\ell=1}^{N}\chi_{[0,\tilde{\tilde{q}}_i^{(j)}]}(p_i^{(\ell)}) - \tilde{\tilde{q}}_i^{(j)}\right| \leq D^*(\omega_N^{(i)})$$

where $\omega_N^{(i)}$ is the one-dimensional set $(p_i^{(1)},\dots,p_i^{(N)})$.

It is now easy to show the estimate - if one uses a little result about the continuity of the discrepancy:

If $\tilde{\Omega}_N = T\tilde{\omega}_N = (\tilde{P}^{(1)},\dots,\tilde{P}^{(N)})$, i.e.

$$\tilde{P}_i^{(j)} = T_i(\tilde{q}_1^{(j)},\dots,\tilde{q}_i^{(j)}) \ , \quad 1\le i\le k,\ 1\le j\le N$$

then

$$\left|\tilde{P}_i^{(j)}-P_i^{(j)}\right| = \left|T_i(\tilde{q}_1^{(j)},\dots,\tilde{q}_i^{(j)})-T_i(\tilde{q}_1^{(j)},\dots,\tilde{q}_{i-1}^{(j)},\tilde{\tilde{q}}_i^{(j)})\right|$$

$$\le \left|\frac{\partial T_i}{\partial q_i}\right|\left|\tilde{\tilde{q}}_i^{(j)}-\tilde{q}_i^{(j)}\right| \le M\cdot D^*(\omega_N^{(i)}) \le M\cdot D^*(\omega_N) \ .$$

The continuity lemma runs as follows: Let $\omega_N^{(1)}$, $\omega_N^{(2)}$ be two sets $(Q^{(1)},\dots,Q^{(N)})$, $(P^{(1)},\dots,P^{(N)})$ respectively and

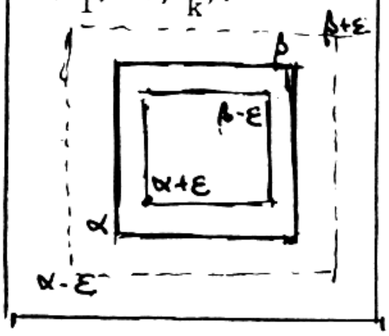
$$\left|Q_j^{(i)}-P_j^{(i)}\right| \le \varepsilon_j \ , \ i=1,\dots,N,\ j=1,\dots,k \ ,$$

then

$$|D^*(\omega_N^{(1)})-D^*(\omega_N^{(2)})| \le \prod_{i=1}^{k}(1+2\varepsilon_i)-1 \ .$$

__Proof of the Lemma:__

Let $\alpha,\beta \in [0,1]^k$, $R = R_{[\alpha,\beta)}$ and $\Lambda(R,\omega_N^{(i)})$ defined as before, $\varepsilon=(\varepsilon_1,\dots,\varepsilon_k)$.



$R^+ := [\alpha-\varepsilon,\beta+\varepsilon) \cap [0,1]^k$,

$R^- = [\alpha+\varepsilon,\beta-\varepsilon)$ ($\varepsilon=\emptyset$ if $\alpha_i+\varepsilon>\beta_i-\varepsilon_i$ for some i).

Since $|q_j^{(i)}-p_j^{(i)}| < \varepsilon_j \ \forall i$,

we learn that

$$\Lambda(R^-,\omega_N^{(2)}) \le \Lambda(R,\omega_N^{(1)}) \le \Lambda(R^+,\omega_N^{(2)}) \ .$$

Since

$$\left|\frac{\Lambda(R^+,\omega_N^{(2)})}{N} - \lambda_k(R^+)\right| \le D^*(\omega_N^{(2)}), \quad \left|\frac{\Lambda(R^-,\omega_N^{(2)})}{N} - \lambda_k(R^-)\right| \le D^*(\omega_N^{(2)})$$

we get

$$-(\lambda_k(R)-\lambda_k(R^-))-D^*(\omega_N^{(2)}) \le \frac{A(R,\omega_N^{(1)})}{N} - \lambda_k(R) \le (\lambda_k(R^+)-\lambda_k(R))+D^*(\omega_N^{(2)})$$

But

$$\lambda_k(R^+)-\lambda_k(R) = \Pi(\beta_i-\alpha_i+2\varepsilon_i)-\Pi(\beta_i-\alpha_i) \le \Pi(1+2\varepsilon_i)-1$$

since the maximal distance is obtained for maximal $\beta_i-\alpha_i$ i.e. for $\beta_i-\alpha_i=1$.

Since $\lambda_k(R)-\lambda_k(R^-) \le \lambda_k(R^+)-\lambda_k(R)$ for the same reason we arrive at

$$D^*(\omega_N^{(1)}) \le D^*(\omega_N^{(2)}) + \Pi(1+2\varepsilon_i)-1 \;;$$

changing the roles of $\omega_N^{(1)}$, $\omega_N^{(2)}$ one gets the result.

For $\varepsilon_i = MD^*(\omega_N)$ we can compare $\tilde{\Omega}_N$ and ω_N:

$$|D^*(\omega_N)-D^*(\tilde{\Omega}_N)| \le (1+2MD^*(\omega_N))^k-1$$

$$\Rightarrow D^*(\tilde{\Omega}_N) \le (1+2M)^k D^*(\omega_N) \qquad (*)$$

Remembering that $\tilde{\Omega}_N = T\tilde{\omega}_N$ we get the estimate (*).

A similar but simpler method (compared to the inversion of T) is used by H. Moock (PhD-thesis). We consider first k=1.

Then we may get a f-distributed sequence in just using as uniformly distributed sequence $\omega_N = (\frac{1}{N},\frac{2}{N},\ldots,\frac{N}{N}) = (y_1,\ldots,y_N)$. Then $\tilde{\omega}_N = F^{-1}\omega_N = (x_1,\ldots,x_N)$ with

$$x_j = F^{-1}(y_j) \quad \text{or by} \quad \int_0^{x_j} f(t)dt = \frac{j}{N},\; j=1,\ldots,N\;.$$

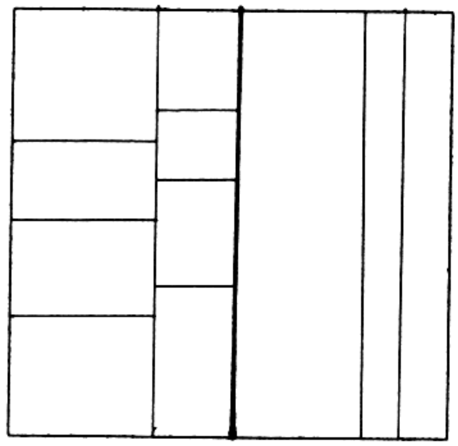
A better idea to construct x_j is to put $F^{-1}(y_j) = \hat{x}_j$; $\hat{x}_j$ is now a partition of [0,1] into segments $[0,\hat{x}_1],[\hat{x}_1,\hat{x}_2],\ldots,[\hat{x}_{N-1},1]$, each carrying the mass $\frac{1}{N}$ and we choose x_j as geometric centre or as centre of gravity of $[\hat{x}_{j+1},\hat{x}_j]$

$$x_j = \frac{\hat{x}_{j-1}+\hat{x}_j}{2} \quad \text{or} \quad x_j = N\int_{\hat{x}_{j-1}}^{\hat{x}_j} f(t)dt\;.$$

In higher dimensions we treat the coordinates successively: We choose $N=N_1\ldots N_k$, divide $[0,1]^k$ first into layers $\hat{q}_1^{i_1-1} \le q_1 \le \hat{q}_1^{i_1}$ through

$$\frac{i_1}{N_1} = \int_0^{\hat{q}^{i_1}} \left(\int_0^1 \ldots \int_0^1 f(q_1,\tau_2,\ldots,\tau_k)d\tau_2 \ldots d\tau_k\right)dq_1 \;;$$

each layer carries a mass of $\frac{1}{N_1}$.



Now we consider the i_1-th layer and divide it into "bars" of mass $(\frac{1}{N})\cdot\frac{1}{N_2}$, given by $\hat{q}^{i_1-1}\leq q_1\leq\hat{q}^{i_1}$, $\hat{q}_2^{i_2-1}\leq q_2\leq\hat{q}_2^{i_2}$ with

$$\frac{i_2}{N_1N_2} = \int_{\hat{q}_1^{i_1-1}}^{\hat{q}_1^{i_1}} \left(\int_0^{\hat{q}_2^{i_2}} \left(\int_0^1 \ldots \int_0^1 f(q_1,q_2,\tau_3,\ldots,\tau_k)d\tau_3 \ldots d\tau_k\right)\right)dq_2dq_1 \;.$$

Continuing in that way we get at the end "such" intervals $P(i_1,\ldots,i_k)$ defined by

$$[\hat{q}_1^{i_1-1},\hat{q}_1^{i_1}] \times [\hat{q}_2^{i_2-1},\hat{q}_2^{i_2}] \times \ldots \times [\hat{q}_k^{i_k-1},\hat{q}_k^{i_k}]$$

each of mass $\frac{1}{N} = \frac{1}{N_1\ldots N_k}$. We may now define $\tilde{\omega}_N$ as centre (of gravity) of these intervals $Q_{(i_1,\ldots,i_k)} = N \int_{P(i_1,\ldots,i_k)} Pf(P)dP$.

The advantage of this construction is that for $\delta_{\omega_N} = \frac{1}{N}\sum \delta_{Q_{(i_1,\ldots,i_k)}}$ the error $\int \Phi \; d\delta_{\omega_N} - \int \Phi f \; dP$ is zero not only for $\Phi=1$ (as always) but also for $\Phi(P)=P$ i.e. the first moment is exact.

§3 The functionanalytic aspect

I follow here an article by P.A. Raviart "An Analysis of Particle Methods" in F. Brezzi (ed.) "Numerical Methods in Fluid Dynamics", Springer Lecture Notes 1127 (1985).

The difference to the previous approach is that one plays with the smoothness of f and on the other hand more with the weights than with points in the approximating discrete measure $\sum_{j=1}^{N} \alpha_j \delta_{P_i}$; moreover one allows countably many instead of finitely many points. We denote by J the index set for the points. So, given $f \in C^0(\mathbb{R}^k)$, we want to approximate it by $\sum_{j\in J} \alpha_j \delta_{P_j}$ and this means to compare

$$\int_{\mathbb{R}^k} f\phi dP \quad \text{with} \quad \sum_{j\in J} \alpha_j \phi(P_j) \quad \text{for some } \phi \in C_0^0(\mathbb{R}^k) \text{ for example.}$$

We interprete this now as the classical problem of numerical quadrature. The construction of $(\alpha_j, P_j)_{j\in J}$ may run as follows: Cover $\mathbb{R}^k$ with a uniform mesh of mesh size $h>0$: For all $j = (j_1,\dots,j_k) \in \mathbb{Z}^k$ denote by B_j the cell

$$B_j = \left\{P \in \mathbb{R}^k;\ (j_i-\tfrac{1}{2})h \leq p_i \leq (j_i+\tfrac{1}{2})h\right\} = R_{[(j-\frac{1}{2})h,(j+\frac{1}{2})h]}$$

and by $P_j = jh$ the centre of B_j.

With $J=\mathbb{Z}^k$, $P_j=jh$ and α_j some approximation for $\int_{B_j} fdP$, for example $\alpha_j = h^k f(P_j)$ (remember that $f \in C^0$) we may try $\sum \alpha_j \phi(P_j)$ as approximation for $\int f\phi dP$.

Now

$$\int f\phi dP - \sum_{j\in J} \alpha_j \phi(P_j) = \sum_{j\in J} \int_{B_j} f\phi dP - \sum_j h^k (f\cdot\phi)(P_j) = \sum_{j\in J} E_j(f\cdot\phi)$$

where

$$E_j(g) = \int_{B_j} g dP - h^k g(P_j) .$$

As a measure for smoothness one may use Sobolev spaces. for example

$$W^{m,p}(\Omega) := \left\{g \in L^p(\Omega),\ D^\alpha g \in L^p \ \text{ for } |\alpha| \leq m\right\}$$

where $1\leq p\leq\infty$ and $m\geq 0$. For p=2 we call $W^{m,2}$ also H^m. The norm in $W^{m,p}(\Omega)$ is defined by

$$\|g\|_{m,p,\Omega} := \left(\sum_{|\alpha|\le m} \|D^\alpha g\|^p_{L^p(\Omega)} \right)^{1/p} ,$$

a seminorm by

$$|g|_{m,p,\Omega} = \left(\sum_{|\alpha|=m} \|D^\alpha g\|^p_{L^p(\Omega)} \right)^{1/p} .$$

A classical result is due to Bramble & Hilbert:
If Ω is bounded and open in $\mathbb{R}^k$ with a Lipschitz-continuous boundary and if L is a linear continuous functional on $W^{m,p}(\Omega)$, $m\ge 1$, $1\le p\le\infty$ with norm $\|L\|$ and

$$L(g) = 0 \quad \text{for all polynomials of degree less or equal } m-1 ,$$

then

$$|L(g)| \le C\|L\| |g|_{m,p,\Omega} \quad \forall g \in W^{m,p}(\Omega)$$

with a fixed constant C.
(Please realize that the right hand side contains the seminorm, not the norm, which would be trivial with C=1.)
As a consequence we get

Lemma 3.1:
There exists a $C>0$ (independent of g,h and j) such that

$$|E_j(g)| \le C \begin{cases} h^{1+\frac{k}{q}} |g|_{1,p,B_j} & \forall g \in W^{1,p}(B_j),\ p>k \\ h^{2+\frac{k}{q}} |g|_{2,p,B_j} & \forall g \in W^{2,p}(B_j),\ p>\frac{k}{2} \end{cases}$$

and $\frac{1}{p}+\frac{1}{q} = 1$.

Proof:
Consider the cube $\hat{B} = [-1,1]^k$ and put

$$\hat{E}(\hat{g}) = \int_{\hat{B}} \hat{g}dP - 2^k\hat{g}(0), \quad \hat{g} \in C^o(\hat{B}) .$$

Clearly, $\hat{E}(g) = 0$ for first order polynomials, therefore $\hat{E}$ is a bounded linear functional on $C^o(\hat{B})$ vanishing on π_1, the set of polynomials of degree ≤ 1.

We now use Sobolev's imbedding theorem telling us that for sufficiently smooth Ω (and $\hat{B}$ is sufficiently smooth) we have

$$W^{m,p}(\Omega) \hookrightarrow C^{o}(\bar{\Omega}) \quad \text{if } m > \frac{k}{p},$$

where the injection of $W^{m,p}(\Omega)$ in $C^{o}(\bar{\Omega})$ is not only continuous but even compact, since Ω is bounded. We use the cases m=1, for which $W^{1,p} \hookrightarrow C^{o}$ if $p>k$, and m=2, for which $W^{2,p} = H^{p} \hookrightarrow C^{o}$ for $p>\frac{k}{2}$. (See for example R. Dautray, J.L. Lions: Math. Analysis and Num. Methods for Science and Technology, Vol. 2, page 139.)
Therefore $g \to \hat{E}(g)$ is continuous even in these Sobolev spaces, and the Bramble-Hilbert lemma tells us that

$$|\hat{E}(\hat{g})| \leq C \begin{cases} |\hat{g}|_{1,p,\hat{B}} & \text{for all } \hat{g} \in W^{1,p}(\hat{B}), \text{ if } p>k \\ |\hat{g}|_{2,p,\hat{B}} & \text{for all } \hat{g} \in W^{2,p}(\hat{B}), \text{ if } p>\frac{k}{2}. \end{cases}$$

We now consider a function g on B_j and transform $B_j = R_{[(j-\frac{1}{2})h,(j+\frac{1}{2})h]}$ onto $[-1,1]^k$ by $x_i = (j_i+\frac{\xi_i}{2})h$:

$$\hat{g}(\xi) := g((j+\tfrac{\xi}{2})h) \quad \text{varies on } \hat{B}$$

and belongs to $W^{m,p}(\hat{B})$, if $g \in W^{m,p}(B_j)$.

Most important, we get (remember that the definition of $|g|_{m,p,\Omega}$ includes only the highest, i.e. the m-th derivative of g)

$$|\hat{g}|_{m,p,\hat{B}} = (\tfrac{h}{2})^{m-k/p}\,|g|_{m,p,B_j}$$

(the exponent m comes from the differentiation, the Jacobian of the transformation $\xi \to x$ gives $(\frac{h}{2})^{-k}$, which, together with the exponent $\frac{1}{p}$ for L^{p} gives $-\frac{k}{p}$). If we realize that

$$E_j(g) = \int_{B_j} g\,dP - h^k g(P_j) = (\tfrac{h}{2})^k\Big(\int_{\hat{B}} \hat{g}\,dP - 2^k\hat{g}(0)\Big) = (\tfrac{h}{2})^k\hat{E}(\hat{g})$$

we get

$$|E_j(g)| = (\tfrac{h}{2})^k|\hat{E}(\hat{g})| \leq C(\tfrac{h}{2})^k|\hat{g}|_{1,p,\hat{B}} = C(\tfrac{h}{2})^k(\tfrac{h}{2})^{1-k/p}|g|_{1,p,B_j}$$

$$= \tilde{C}\, h^{1+k/q}|g|_{1,p,B_j}$$

which is lemma 3.1 for m=1. For m=2 we get the result correspondingly. A stronger result can be obtained in using an asymptotic expansion with respect to h.

Lemma 3.2:

There exist constants $C>0$ and $d_\alpha \in \mathbb{R}$, $|\alpha|\geq 2$, independent of h and j, such that

$$\left|E_j(g) - \sum_{2\leq|\alpha|\leq m-1} d_\alpha h^{|\alpha|} \int_{B_j} D^\alpha g dP\right| \leq C\, h^{(m+\frac{k}{q})}$$

for $g \in W^{m,p}(B_j)$, where $m\geq 3$, $p>\frac{k}{m}$.

Proof:

First we determine of $\hat{d}_\alpha$, $\alpha \in \mathbb{N}^k$, $|\alpha|\geq 2$, such that

$$\hat{L}_m(\hat{g}) := \hat{E}(\hat{g}) - \sum_{2\leq|\alpha|\leq m-1} \hat{d}_\alpha \int_{\hat{B}} D^\alpha \hat{g} dP$$

vanishes for all polynomials of degree $\leq m-1$, i.e. on π_{m-1}.
We know that $\hat{E}$ vanishes on π_1; put therefore $\hat{L}_2 = \hat{E}$. We make now an induction with respect to m. Assume that we found constants $\hat{d}_\alpha$ for $|\alpha|=m$ such that

$$\hat{L}_{m+1}(\hat{g}) = \hat{L}_m(\hat{g}) - \sum_{|\alpha|=m} \hat{d}_\alpha \int_{\hat{B}} D^\alpha \hat{g} dP$$

vanishes on π_m. Clearly, $\hat{L}_{m+1}$ vanishes on π_{m-1}. We therefore have to check only homogeneous polynomials of degree m, i.e. $\hat{g}(P) = P^\beta$ with $|\beta|=m$.

Since

$$\int_{\hat{B}} (D^\alpha P^\beta) dP = \begin{cases} 2^k \alpha! = 2^k \alpha_1! \ldots \alpha_k! & \text{for } \alpha=\beta \\ 0 & \text{else} \end{cases}$$

(remember that $|\alpha| = |\beta| = m$), we have to find $\hat{d}_\alpha$, $|\alpha|=m$ such that

$$\hat{L}_m(P^\beta) = 2^k \beta! \tilde{d}_\beta, \quad |\beta| = m \text{ , i.e. } \tilde{d}_\beta = \frac{\hat{L}_m(P^\beta)}{2^k \cdot \beta!} .$$

We now choose $m\geq 3$, $p>\frac{k}{m}$, such that $W^{m,p}(\hat{B}) \hookrightarrow C^{o}(\hat{B})$; therefore $\hat{L}_m$ is defined on $W^{m,p}$ and vanishes on π_{m-1}. Bramble-Hilbert gives

$$|\hat{L}_m(\hat{g})| \leq C|\hat{g}|_{m,p,\hat{B}} \quad \forall\hat{g} \in W^{m,p}(\hat{B}) \ .$$

The transformation from B_j to $\hat{B}$ as in lemma 3.1 assigns to any $g \in W^{m,p}(B_j)$ a $\hat{g} \in W^{m,p}(\hat{B})$ and since

$$E_j(g) - \sum_{2\leq|\alpha|\leq m-1} \frac{\hat{d}_\alpha}{2^{|\alpha|}} |h|^\alpha \int_{B^j} D^\alpha g dP = (\frac{h}{2})^k \hat{L}_m(\hat{g})$$

we get the result of lemma 3.2 by putting $d_\alpha = \frac{\hat{d}_\alpha}{2^{|\alpha|}}$.

The preceding lemmata now allow to prove the approximation theorem for f by $\sum \alpha_j \delta_{P_j}$ in the sense that we estimate

$$\int f\cdot\phi dP - \sum_{j\in J} \alpha_j \phi(P_j) \ .$$

We choose $J=\mathbb{Z}^k$, P_j=jk and $\alpha_j = h^k f(P_j)$, such that the difference is

$$\int f\phi dP - \sum_{j\in\mathbb{Z}^k} h^k (f\cdot\phi)(P_j) \ .$$

Theorem 3.1:
Let $m\geq 1$, $p>\frac{k}{m}$, $\frac{1}{p}+\frac{1}{q}=1$. Then there exists a constant C (independent of h), such that

$$\left|\int_{\mathbb{R}^k} f\phi dP - h^k \sum_{j\in\mathbb{Z}^k} (f\phi)(P_j)\right| \leq C\cdot h^{m+\frac{k}{q}} \sum_{j\in\mathbb{Z}^k} |f\phi|_{m,p,B_j}$$

for all $f\cdot\phi$ in $W^{m,p}(\mathbb{R}^k) \cap W^{m-1,1}(\mathbb{R}^k)$ if $m\geq 3$, for all $f\cdot\phi$ in $W^{m,p}(\mathbb{R}^k) \cap L^1(\mathbb{R}^k)$ if $m\leq 2$.

Remark 1:
In this result infinitely many points are involved; however, if f or ϕ has compact support, there are only finitely many, say N, where $N\sim h^{-k}$. The estimate is now of order $N^{-m/k}\cdot N^{-1/q}$ $(m>\frac{k}{p})$ which is certainly better than our discrepancy estimates; it depends on the smoothness of $f\cdot\phi$ and is of infinite order for infinitely smooth $f\cdot\phi$. The constants

however include still f which is not the case in Koksma-Hlawka-inequality. It may be that this method has a better asymptotic but need - as a minimum - quite many points to give a reasonable result, if the support is not very small. A careful comparison is still missing.

Proof of theorem 3.1:
Denote $f\cdot\phi$ by g; if $g \in C^o(\mathbb{R}^k) \cap \mathbf{L}^1(\mathbb{R}^k)$, then

$$\int_{\mathbb{R}^k} g\,dP - h^k \sum_{j\in\mathbb{Z}^k} g(P_j) = \sum_{j\in\mathbb{Z}^k} E_j(g) .$$

If $m\leq 2$, lemma 3.1 gives us immediately the result of theorem 3.1 just by summing up. If $m\geq 3$ and $g \in W^{m-1,1}$, i.e. if $D^\alpha g \in \mathbf{L}^1(\mathbb{R}^k)$ for $|\alpha|\leq m-1$ and $D^\alpha g \in \mathbf{L}^p$ for $|\alpha|\leq m$, it is easy to see that

$$\int_{\mathbb{R}^k} D^\alpha g\,dP = 0 \quad \text{for } 2\leq|\alpha|\leq m-1 .$$

We may therefore add

$$\sum_{2\leq|\alpha|\leq m-1} d_\alpha h^{|\alpha|} \int_{\mathbb{R}^k} D^\alpha g\,dP \text{ , i.e.}$$

$$\sum_{j\in\mathbb{Z}^k} E_j(g) = \sum_{j\in\mathbb{Z}^k} \left(E_j(g) - \sum_{2\leq|\alpha|\leq m-1} d_\alpha h^{|\alpha|} \int_{B_j} D^\alpha g\,dP\right) ;$$

Lemma 3.2 now gives the result. □

Remark 2:
As a consequence $h^k \sum_{j\in\mathbb{Z}} f(P_j)\delta_{P_j}$ converges vaguely (i.e. for $\phi \in C_o^o$ - see §1) to f - and therefore also weakly.
So far the approximation theory in the papers of Raviart and his school.
There is another functionalanalytic aspect, which seems appropriate but was not used until now. We need the concept of $W^{s,2} = H^s(\mathbb{R}^k)$ even for negative s and define therefore (see Dautray-Lions):
Let $s \in \mathbb{R}$, then $H^s(\mathbb{R}^k)$ is the subspace of $S'(\mathbb{R}^k)$ (the space of tempered distributions), whose elements u have the property that

$$(1+\|\xi\|^2)^{s/2}\ \hat{u}(\xi) \in \mathbf{L}^2(\mathbb{R}^k) ,$$

where $\hat{u}(\xi)$ is the Fouriertransform of u (in the sense of distribution). If $s=m \in \mathbb{N}$, then $H^s=H^m=W^{m,2}$ defined as above. H^s is even a Hilbert space with scalarproduct $\langle u,v\rangle_s = \int_{\mathbb{R}^k} (1+\|\xi\|^2)^{s/2}\hat{u}(\xi)\overline{\hat{v}(\xi)}d\xi$ and $H^{s1} \subset H^{s2}$ if $s_1 \geq s_2$.

It is now easy to see that $\delta \in H^s$ for $s < -\frac{k}{2}$, since $\hat{\delta}=1$ and

$$\int_{\mathbb{R}^k} (1+\|\xi\|^2)^s |\hat{u}(\xi)|^2 d\xi = \omega_k \int_0^\infty (1+r^2)^s r^{k-1} dr < \infty$$

if $2s+k-1<1$ or $s<-\frac{k}{2}$ (ω_k is the surface measure of the k-dimensional unit sphere). Certainly f belongs to the same space and we can define a distance between f and $\sum \alpha_j \cdot \delta_{P_j} = \delta_{\omega_N}$ by

$$d(f,\delta_{\omega_N}) = \|f-\delta_{\omega_N}\|_s = \left(\int_{\mathbb{R}^k} (1+\|\xi\|^2)^s \left|\hat{f}(\xi) - \sum \alpha_j e^{-i\langle P_j,\xi\rangle}\right|^2 d\xi\right)^{1/2} ,$$

$$s < -\frac{k}{2} .$$

This reminds of the Weyl criterion, where we had already considered $\frac{1}{N}\sum_j e^{2\pi i\langle \xi_j, P_j\rangle}$ (the factor 2π instead of -1 is adopted to the unit intervals).

This concept might be useful since it provides us with a scalarproduct and makes optimal constructions easier. For example if f is given and we look for the best δ_{ω_N} to approximate it, if $N \leq \overset{o}{N}$ we may try to find the orthogonal projection of f on the set of these discrete measures (which is not a linear subspace!). But nothing is yet done in this direction.

We mention that H^s is the dual of H^{-s} and that one may define it in this way; but there has been no use for this fact either.

§4 Statistical aspects

Until now our particle sets were given for each N and we asked what happens for $N\to\infty$. In statistical physics one may start with a probability distribution for ω_N - here in general one assumes that all particles are equal (especially have the same mass $\frac{1}{N}$) and undistinguishable. Then $\omega_N = (P_1,\dots,P_N) \in \mathbb{R}^{Nk}$ and we start with a probability density $F_N(P_1,\dots,P_N)$, such that for measurable $A \subset \mathbb{R}^{Nk}$, we have

$$\int_A F(P_1,\dots,P_N)dP_1,\dots,dP_N = \text{Probability for } \omega_N \in A .$$

Since the particles are undinstinguishable, $F(P_1,\dots,P_N)$ does not change its value if we permute $(P_1,\dots,P_N)$.
There are now several possibilities to connect ω_N with $f \in \mathcal{M}_{ac}$:

1) We may consider the one-particle density

$$f_N(P_1) = \int_{\mathbb{R}^{(N-1)k}} F_N(P_1,P_2,\dots,P_N)dP_2,\dots,dP_N$$

and can ask for the convergence of f_N to f for example in a weak sense i.e.

$$\int f_N\Phi dP \longrightarrow \int f\Phi dP \text{ for all } \Phi \in C^0 .$$

2) For given f, $\omega_N \to D(\omega_N,f)$ is now a stochastic variable and we may consider as convergence the convergence of this variable to zero in probability

$$\text{Prob}\left\{\omega\colon D(\omega_N,f) \geq \varepsilon\right\} = \int_{\{\omega_N | D(\omega_N,f)\geq\varepsilon\}} F_N(\omega_N)d\omega_N \to 0 \text{ for all } \varepsilon ;$$

here the distribution "concentrates" more and more around f.

4) If we assume that the particles $P_1,\dots,P_N$ are independent, identically distributed random variables with density f, then

$$F_N(P_1,\dots,P_N) = f(P_1),\dots,f(P_N) .$$

Normally one takes a sequence $(P_n=P_n(\Omega))_{n\in\mathbb{N}}$ of independent, identically distributed random variables with distribution f; then

$$\text{Probability } \left\{\Omega | (P_n(\Omega))_{n\in\mathbb{N}} \text{ is f-distributed}\right\} = 1.$$

FOLGENDE BERICHTE SIND ERSCHIENEN:

Nr. 31 FORSCHUNG
H.-G. Stark, H. Trinkaus, Ch. Jansson
The Simulation of the Charge Cycle in a Cylinder of a Combustion Engine

Nr. 32 FORSCHUNG
H. Babovsky, F. Gropengießer, H. Neunzert, J. Struckmeier, B. Wiesen
Low Discrepancy Methods for the Boltzmann Equation

Nr. 33 FORSCHUNG
M. Brokate
Some BV properties of the Preisach hysteresis operator

1989

Nr. 34 FORSCHUNG
H. Neunzert
Industrial Mathematics: General Remarks and Some Case Studies

Nr. 35 FORSCHUNG
M. Brokate
On a Characterization of the Preisach Model for Hysteresis

Nr. 36 FORSCHUNG
C.-P. Fritzen, P. Hackh
Optimization of a Spring for Dental Attachments

Nr. 37 FORSCHUNG
U. Helmke, D. Prätzel-Wolters, S. Schmid
Adaptive Synchronization of Interconnected Linear Systems

Nr. 38 FORSCHUNG
U. Helmke, D. Prätzel-Wolters, S. Schmid
Sufficient Conditions for Adaptive Stabilization and Tracking

Nr. 39 FORSCHUNG
U. Helmke, D. Prätzel-Wolters, S. Schmid
Adaptive Tracking for Scalar Minimum Phase Systems

Nr. 40 FORSCHUNG
F.-J. Pfreundt
Nähen als dynamisches System

1990

Nr. 41 FORSCHUNG
H.-G. Stark
Multiscale Analysis, Wavelets and Texture Quality

Nr. 42 FORSCHUNG
I. Einhorn, H. Moock
A Deterministic Particle Method for the Simulation of the Boltzmann Transport Equation of Semiconductors

Nr. 43 FORSCHUNG
F. Gropengießer, H. Neunzert, J. Struckmeier
Computational Methods for the Boltzmann Equation

Nr. 44 FORSCHUNG
S. Nikitin, S. Schmid
Universal Adaptive Stabilizers for One-Dimensional Nonlinear Systems

Nr. 45 FORSCHUNG
P. Hackh
Quality Control of Artificial Fabrics

Nr. 46 FORSCHUNG
S. Körber, B. Wiesen
A Comparison of a Microscopic and a Phenomenological Model for a Polyatomic Gas

Nr. 47 FORSCHUNG
F. Gropengießer, H. Neunzert, J. Struckmeier, B.Wiesen
Several Computer Studies on Boltzmann Flows in Connection with Space Flight Problems

Nr. 48 FORSCHUNG
M. Brokate
Some Remarks on the Discretization of the Preisach Operator

Nr. 49 FORSCHUNG
M. Brokate
On the Moving Preisach Model

1991

Nr. 50 FORSCHUNG
W. Wagner
A Stochastic Particle System Associated with the Spatially Inhomogeneous Boltzmann Equation

Nr. 51 AUSBILDUNG
I.M. Sobol
Punkte, die einen mehrdimensionalen Würfel gleichmäßig ausfüllen

Nr. 52 FORSCHUNG
M. Brokate, A.H. Siddiqi
Sensitivity in the Rigid Punch Problem

Nr. 53 FORSCHUNG
S. Nikitin, D. Prätzel-Wolters
Multiparameter, Polynomial Adaptive Stabilizers

Nr. 54 FORSCHUNG
S. Schmid, D. Prätzel-Wolters
Synchronization through System Interconnections

Nr. 55 FORSCHUNG
D.Prätzel-Wolters, R.D. Reinke
Simple Adaptive Control of a Discrete Almost Strict Positive Real Heat Treatment System

Nr. 56 FORSCHUNG
S. Chen, D. Prätzel-Wolters
Modelling and Controller Design for Heat Treatment Processing of Enamelled Wires

Nr. 57 FORSCHUNG
B. Wiesen
On the Dependence of the Solution of Generalized Boltzmann Equation on the Scattering Cross Section: The Inverse Problem

Nr. 58 FORSCHUNG
G. Engl, R. Rösch
Studien zum Programmsystem PROM (Berechnung des instationären Ladungswechsels von zündenden Mehrzylinder- Verbrennungs-motoren)

Nr. 59 FORSCHUNG
J. Hoffmann, D. Prätzel-Wolters
Controllability Tests for Behaviour Systems in AR-Representations

Nr. 60 FORSCHUNG
M. Bäcker, H. Neunzert, S. Sundar, S. Younis
A 2-D Kaniel Kinetical Scheme for the Isentropic Compressible Flow

Nr. 61 FORSCHUNG
F.-J. Pfreundt, J. Struckmeier
On the Efficiency of Simulation Methods for the Boltzmann Equation on Parallel Computers

Nr. 62 FORSCHUNG
M. Schreiner
Weighted particles in the finite pointset method