Construction of Particlesets to Simulate Rarefied Gases

Michael Hack
AG Technomathematik
University of Kaiserslautern
Germany

Abstract

In this paper a new method is introduced to construct asymptotically f–distributed sequences of points in the $\mathbb{R}^d$. The algorithm is based on a transformation proposed by E.Hlawka and R.Mück.By exploitation of the special structure the method proves to be very efficient. For the numerical tests a new procedure to evaluate the f–discrepancy in two dimensions is proposed.

1 The problem

1.1 The reentry of a space shuttle

For the reentry of a space shuttle into the atmosphere, one is interested in the flow field around the shuttle to be able to adapt the surface in shape and material such that the produced heat can be controlled.

The first phase of the reentry, which takes place in an already rarefied gas is not that crucial in heat production but it determines the initial conditions for the later phases. In this state we can not use normal continuum flow mechanics, i.e. we can not use Navier–Stokes or Euler equations to describe the flow properly.

An attempt to govern this region was done by Bird [5] (starting 1968) by direct simulation methods. From a mathematical point of view it is not quite satisfactory, because there is no underlying equation to be solved and hence no possibility of testing a convergence of the quality of the simulation is given.

On the other hand the equation governing this region, the Boltzmann equation, is quite complicated. Nanbu [16] proposed in 1980 a method directly connected to the Boltzmann equation, starting from that a group in Kaiserslautern is developing numerical schemes for it. These methods are called *"Finite Pointset Methods" FPM* . Convergence could be proven by Babovsky et al. ([3] and [4]).

1.2 The mathematical formulation

We will follow the formulation for gases, consisting of only one kind of molecules, given in [9]. The Boltzmann equation describes the evolution of the density in the position–density space.

We start with an initial density

$$\begin{aligned} f_0 : \Omega \times \mathbb{R}^3 &\longrightarrow \mathbb{R} \\ (x, v) &\longmapsto f_0(x, v) \end{aligned}$$

which develops following the integro–differential equation

$$f_t(t,x,v) + \langle v, f_x(t,x,v)\rangle = J(f,f)$$

where

$$\begin{aligned} J(f,f) &= \int_{\mathbb{R}^3}\int_{\mathbf{S}^2_+} \sigma(\|v-w\|,\eta)\,[f(t,x,v')f(t,x,w') - f(t,x,v)f(t,x,w)]\; d\omega(\eta)\,dw \\ v' &= v - \langle v,\eta\rangle\eta \\ w' &= w + \langle w,\eta\rangle\eta \end{aligned}$$

Here $J(f,f)$ describes the intermolecular forces, and $\sigma(\|v-w\|,\eta)$ corresponds to the differential cross section of the binary collisions.

Ω is bounded by the shuttle surface and an outer boundary to hold the region bounded (for computational reasons). On the inner boundary $\partial\Omega_i$ the molecules are scattered on (resp. in) the top layers of the surface. This is handled by

$$|\langle v,n\rangle|\, f(t,x,v) = \int_{\langle v',n\rangle<0} R(v'\mapsto v;t,x)\,|\langle v',n\rangle|\, f(t,x,v')\,dv'$$

for

$$\begin{aligned} \langle v,n\rangle &> 0 \\ x &\in \partial\Omega_i \\ n &: \text{inner normal on } \partial\Omega_i \text{ in } x \end{aligned}$$

$R(v'\mapsto v;t,x)$ is called scattering kernel. [9] recommend a combination of a specular reflection kernel

$$R_S(v'\mapsto v;t,x) = \delta(v - v' + 2n\langle v',n\rangle)$$

and a diffuse reflection

$$R_D(v'\mapsto v;t,x) = \frac{1}{2\pi k^2 T_0^2(x)}\, e^{-\frac{\|v-u_0\|^2}{2kT_0(x)}}|\langle v,n\rangle|$$

where $T_0(x)$ and $u_0(x)$ denote the temperature and velocity of the surface at x. We see that R_D is completely independent on the past (i.e. v'). The choice of the combination and the influence on the solution is examined in [9].

The outer boundaries are chosen, such that one finds thermodynamic equilibrium and free stream conditions. (see [9]) for details.

1.3 Finite pointset methods

The main idea of Finite Pointset Methods (of particle methods) is somehow to reverse the common process of building macroscopic physical models, namely to describe the discrete nature by continuous models. Here we simulate continuous models (the Boltzmann equation) by discrete particles. The advantage of simulating the continuous model instead of the discrete nature itself, is that using the information of the model one needs much less particles than there are in reality and one is able to compare the simulation with the model, even if there are no direct experimental data of the reality.

The aim of the FPM is to approximate the density f by a finite set of points. The meaning of approximation in our sense is presented in chapters 2 and 3. The usual way to apply a FPM to a time evolution equation like the Boltzmann equation is divided into two steps:

(1) Given an initial value f_0, construct a particle ensemble (see section (2.1)) $\omega_N^0 = (p_i^0, \alpha_i^0)_{i=1}^N$ which approximates f_0.

(2) Change the ensemble in time

$$t \longmapsto \omega_n(t) = (p_i(t), \alpha_i(t))_{i=1}^N$$

where $\omega_N(0) = \omega_N^0$ and $\omega_N(t)$ approximates $f(t, \cdot)$

Remark: Another possibility are refinement methods. Here step (2) is substituted by (time is discretized in steps τ)

(2') Having constructed $\omega_N(k\tau)$, one calculates a density $f^{(k+1)\tau}$ using $\omega_N(k\tau)$ and approximates this by $\omega_N((k+1)\tau)$ like in step (1).

.

1.4 The aim of this theses

The work which is done in this theses is to find a general and effective method for step (1) (resp. for step(2')). For the treatment of step (2) see [9]. Efforts to generalise the model to gas mixtures have been done and are still being done, too.

The initial distributions to be approximated often are Maxwell distributions in the position–velocity space:

$$\begin{aligned} f_M : \mathbb{R}^{2k} &\longrightarrow \mathbb{R}^{2k} \\ (x, v) &\longmapsto \frac{1}{(2\pi T(x))^{\frac{3}{2}}} \rho(x) e^{\frac{\|v-u(x)\|^2}{2T(x)}} \end{aligned}$$

for $k = 2, 3$— i.e. k–dimensional position– and velocity spaces

$\rho(x)$: some density in the position space

$u(x)$: velocity of the shuttle relative to the gas at point x

$T(x)$: Temperature at point x

Here $\rho(x)$ gives the dependency of the density in position (e.g. the gas in higher regions is less dense than in regions closer to the earth surface). These distributions can also approximately be handled in subdividing Ω in cells, where $\rho(x), u(x), T(x)$ is set to be constant in these cells. In this special case one is able to approximate the distribution. But there are also different ideas for the initial distribution, e.g. some which take into considerations some more parameters like pressure and derivatives like $\nabla_x T(x)$, especially if one approaches the fluid–dynamical region.

Those distributions have the property to get zero in some regions which one has to consider in the design of methods to construct approximating pointsets. Some proposals for such distributions are given as correction terms for the Maxwellian.

$$\begin{aligned} & f_{NS} : \mathbb{R}^6 \longrightarrow \mathbb{R}^6 \\ & f_{NS} = f_M \cdot (1 + \varepsilon Q) \\ \text{resp.} \quad & f_{NS}^* : \mathbb{R}^6 \longrightarrow \mathbb{R}^6 \\ & f_{NS}^* = f_M \cdot (1 + \frac{\varepsilon}{2} Q)^2 \end{aligned}$$

where Q depends e.g. on $\rho(x), c(x, v) := v - u(x), \nabla_x u(x), T(x), \nabla_x T(x)$. The advantage of choosing the quadratic form is clearly, that in any way the density stays positive, while in the other form,

one is responsible that ε is chosen appropriately. We will present one twoparametric proposal for such a correction term:

$$\begin{aligned} Q &= \frac{1}{2PT}\sum_{i,j=1}^{3} p_{ij}c_ic_j - \frac{2}{5PT}\left(\frac{5}{2} - \frac{|c|^2}{2T}\right)\sum_{i=1}^{3} q_ic_i \\ q_i &= -\lambda\frac{\partial T}{\partial x_i} \\ p_{ij} &= -\mu\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} - \frac{2}{3}\sum_{\ell=1}^{3}\frac{\partial u_\ell}{\partial x_\ell}\right) \\ P &= \frac{1}{3}\sum_{\ell=1}^{3} p_{\ell\ell} \end{aligned}$$

In this case the averaging of the x–dependencies over cells does not lead to a simple gaussian distribution in the velocity space, such that we have to approximate those distributions by numerical methods like the one, which will be introduced in this thesis.

1.5 The structuring of the thesis

In chapters 2 and 3, we will give a short introduction into the underlying mathematics. At first one has to define what is meant by approximating densities by finite pointsets. To do this, we have to make things comparable, i.e. we have to choose a structure governing both densities and pointsets, measures. This is done in chapter 2. Furthermore it is shown there that our problem of approximation has always a solution in the sense discussed there. In addition we have to choose also some distance between a pointset and a density, such that one can say, whether an approximation is good enough or is not. One kind of such distances is presented in chapter 3.

First steps to solve the problem of constructing pointsets approximating some density f were done by E.Hlawka and R.Mück in 1971 (see [13] and [14]). We will use the principal idea of the algorithm which was proposed there (see chapter 4), but we will develop a numerically more efficient method and estimate the approximation error and asymptotic behaviour of the pointsets constructed in this way. (see chapters 5 and 6). The method introduced here is also proven to converge for any density f, if the number of particles tends to infinity. In the beginning of chapter 5 a catalogue of requirements for construction methods is stated also with respect to step (2'). Hence the method introduced here is not restricted to the framework of the problem shown in the starting sections, but it has also applications in the solution of any transport equation, if one chooses to use FPM's.

In chapter 7 we introduce a new method to calculate f–discrepancies (see section 3.1) in two dimensions. This method was used to test the ability of the constructive method. In more than two dimensions and in the case of the initial distribution for the Boltzmann problem, we are in six dimensions, it is quite complicated to do numerical tests, because the distance we use, the discrepancy (see section 3.1), can not be evaluated for a reasonable number of particles. So we are restricted to qualitative tests (see section 7.3).

One can summarize the structuring:

(1) Mathematical modelling (2 and 3)

(2) Solution (4 to 6 and 7.1)

(3) Tests and application (7.2 and 7.3)

2 Particles, measures and weak convergence

Let $\Omega \subseteq \mathbb{R}^d$, $\lambda = \lambda_d$ denotes the Lebesgue–measure on the $\mathbb{R}^d$ and $\mathcal{B} = \mathcal{B}(\Omega)$ the matching Borel-algebra.

2.1 Particle–ensembles and discrete measures

Definition 2.1 $\omega_n = (\alpha_i, p_i)_{i=1}^n$ *is called* ***particle–ensemble*** $:\Leftrightarrow$

(i) $\alpha_i \in \mathbb{R}^+ \qquad i = 1, \ldots, n$

(ii) $p_i \in \Omega \qquad i = 1, \ldots, n$

Notations:

(1) If p $\in \Omega$, then we denote by δ_p the ***discrete measure***

$$\delta_p(\mathbf{A}) := \begin{cases} 0 & p \notin \mathbf{A} \\ 1 & p \in \mathbf{A} \end{cases} \qquad \forall \mathbf{A} \in \mathcal{B}$$

(2) Let ω_n be a Particle–ensemble. We define the ***discrete measure*** δ_{ω_n} as:

$$\delta_{\omega_n}(\mathbf{A}) := \sum_{i=1}^{n} \alpha_i \delta_{p_i}(\mathbf{A}) \qquad \forall \mathbf{A} \in \mathcal{B}$$

(3) $\mathcal{D}(\Omega) := \{\delta_{\omega_n} \mid \omega_n \text{ particle–ensemble}\}$

(4) $\mathcal{D}_1(\Omega) := \{\delta_{\omega_n} \in \mathcal{D} \mid \delta_{\omega_n}(\Omega) = 1\}$

Remark: $\delta_{\omega_n} \in \mathcal{D}_1(\Omega) \iff \sum_{i=1}^{n} \alpha_i = 1$

2.2 Absolutely continuous measures

Notations:

(1) Let $\mathcal{F}(\Omega) = \{f : \Omega \longrightarrow \mathbb{R}^+ \mid \int_\Omega f \, d\lambda < \infty\}$ be the set of all ***densities*** on Ω.

(2) Given Ω, we denote by $\mathcal{M}_{ac}(\Omega)$ the set of all ***absolutely continuous measures*** (with respect to λ) on Ω, i.e. for any $\mu \in \mathcal{M}_{ac}(\Omega)$, there is a $f \in \mathcal{F}(\Omega)$, such that

$$\mu(\mathbf{A}) = \int_{\mathbf{A}} f \, d\lambda \qquad \forall \mathbf{A} \in \mathcal{B}$$

(3) For any $f \in \mathcal{F}(\Omega)$, let

$$\mu_f(\mathbf{A}) := \int_{\mathbf{A}} f \, d\lambda \qquad \forall \mathbf{A} \in \mathcal{B}$$

Hence $\mu_f \in \mathcal{M}_{ac}(\Omega)$

(4) Analogously we define $\mathcal{F}_1 = \{f \in \mathcal{F} \mid \int_\Omega f \, d\lambda = 1\}$ the set of all ***probability–densities*** and $\mathcal{M}_1$ the set of all ***absolutely continuous probability–densities***.

Remark: If $f \in \mathcal{F}$ and $\kappa = \int_\Omega f \, d\lambda$, then for $\hat{f} := \frac{1}{\kappa} f$ it holds:

$$\hat{f} \in \mathcal{F}_1$$

2.3 Convergence of measures

Definition 2.2 *Let $(\mu_n)_{n\in\mathbb{N}}$ be a sequence of measures and μ a measure. We say $(\mu_n)_{n\in\mathbb{N}}$* ***converges*** *to μ and we write*

$$\begin{aligned} \mu_n &\longrightarrow \mu && :\Leftrightarrow \\ \mu_n(\mathbf{A}) &\longrightarrow \mu(\mathbf{A}) && \forall \mathbf{A} \in \mathcal{B} \end{aligned}$$

But this is not the appropriate kind of convergence for discrete measures as the following example will show.

Example: If $p \in \Omega$ and $(p_n)_{n\in\mathbb{N}} \subset \Omega$ is a sequence of points, where

$$p_n \longrightarrow p \qquad \text{but} \qquad p_i \neq p \quad \forall i \in \mathbb{N}$$

then

$$\delta_{p_n} \not\longrightarrow \delta_p$$

Proof:

Let $\mathbf{A} = \{p\}$. (Then $\mathbf{A} \in \mathcal{B}(\Omega)$)

On one hand we have	$\delta_p(\mathbf{A})$	$=$	$1,$
but on the other hand it holds	$\delta_{p_i}(\mathbf{A})$	$=$	$0 \quad \forall i.$
Hence	$\delta_{p_i}(\mathbf{A})$	$\not\longrightarrow$	$\delta_p(\mathbf{A})$
and therefore	δ_{p_i}	$\not\longrightarrow$	δ_p

□

So we must use another kind of convergence, here the weak convergence of measures.

Definition 2.3 *If $(\mu_n)_{n\in\mathbb{N}}$ is a sequence of measures and μ is a measure, we call $(\mu_n)_{n\in\mathbb{N}}$ to be* ***weakly convergent*** *to μ and we write*

$$\begin{aligned} \mu_n &\xrightarrow{w.} \mu && :\Leftrightarrow \\ \int_\Omega \varphi \, d\mu_n &\longrightarrow \int_\Omega \varphi \, d\mu && \forall \varphi \in \mathcal{C}_b(\Omega) \end{aligned}$$

Remark: Let $\mu_f \in \mathcal{M}_{ac}$, $\hat{f} \in \mathcal{F}_1$ with $\hat{f} = \kappa f$. If $(\hat{\omega}_n = (\alpha_i, p_i)_{i=1}^n)_{n\in\mathbb{N}}$ is a sequence of particle–ensembles with

$$\delta_{\hat{\omega}_n} \xrightarrow{w.} \mu_{\hat{f}},$$

then $(\omega_n)_{n\in\mathbb{N}}$ where $\omega_\nu = \kappa\hat{\omega}_\nu$ (i.e. $\omega_\nu = (\kappa\alpha_i, p_i)_{i=1}^\nu$) converges weakly to μ_f

$$\delta_{\omega_n} \xrightarrow{w.} \mu_f.$$

Therefore if we are able to approximate probability–measures, we can approximate any absolutely continuous measure.

Example: Let us consider again the sequence of points $(p_n)_{n\in\mathbb{N}} \subset \Omega$ and the point $p \in \Omega$ where

$$p_n \longrightarrow p.$$

Now it holds

$$\delta_{p_n} \xrightarrow{w.} \delta_p .$$

Proof:

Let $\varphi \in \mathcal{C}_b$, then it holds

$$\int_\Omega \varphi \, d\delta_{p_i} = \varphi(p_i) \xrightarrow{\varphi \in \mathcal{C}_b} \varphi(p) = \int_\Omega \varphi \, d\delta_p$$

□

For the proof of weak convergence of a sequence of measures it is not necessary to test all $\varphi \in \mathcal{C}_b(\Omega)$. There is a couple of equivalent formulations – see the next theorem. For the special problem, whether ensembles converge weakly see also section 3.1.

Theorem 2.1 *Let $\Omega \subset \mathbb{R}^d$ be compact. $(\mu_n)_{n\in\mathbb{N}}, \mu \subset \mathcal{M}_1(\Omega)$. Then the following statements are equivalent:*

(i) $\mu_n \xrightarrow{w.} \mu$

(ii) $\mu_n(\mathbf{M}) \longrightarrow \mu(\mathbf{M}) \qquad \forall \mathbf{M} \in \mathcal{B}(\Omega) \quad \mu$*–continuous, i.e.* $\mu(\partial \mathbf{M}) = 0.$

(iii) $\mu_n(\mathbf{R}_p) \longrightarrow \mu(\mathbf{R}_p) \qquad \forall \mathbf{R}_p \in \mathcal{R}(\Omega)$

where $\mathcal{R}(\Omega) := \{\mathbf{R}_p \,|\, p \in \Omega,\ \mathbf{R}_p\ \mu$*–continuous*$\}$
and $\mathbf{R}_{(x_1,\ldots,x_d)} := \{(\xi_1,\ldots,\xi_d) \in \Omega \,|\, \xi_i < x_i \ i = 1,\ldots,d\}$

The equivalence of *(i)* and *(ii)* is part of the **Portmanteau – theorem**. For the proof see [17]

2.4 Sequences of points and sequences of ensembles

In this section we want examine special particle–ensembles, induced by sequences of points.

Notations:

(1) Let $(\alpha_n, p_n)_{n\in\mathbb{N}}$, such that $\omega_n := (\alpha_i, p_i)_{i=1}^n$ is a particle–ensemble $\forall n \in \mathbb{N}$. We denote by

$$(\omega_{(\alpha_n, p_n)})_{n\in\mathbb{N}} := (\omega_n)_{n\in\mathbb{N}}$$

the ***sequence of particles*** induced by $(\alpha_n, p_n)_{n\in\mathbb{N}}$.

(2) If $(p_n)_{n\in\mathbb{N}} \subset \Omega$, they induce the ***equiweighted sequence of points***

$$(\omega_{p_n})_{n\in\mathbb{N}} := (\omega_{(\frac{1}{n}, p_n)})_{n\in\mathbb{N}}$$

Especially those equiweighted sequence of points are of interest, because here the least amount of information is needed.

We will show in the next section, that given any absolutely continuous measure μ, it is possible to construct equiweighted sequences of points, such that their induced discrete measures converge weakly to μ.

2.5 Discrete measures are dense

The following theorem — for proof see [2] §45 — states that the discrete measures are dence in $\mathcal{M}_{ac}$:

Theorem 2.2 *Given any $\mu \in \mathcal{M}_{ac}$, there is a sequence of particle–ensembles $(\omega_n)_{n\in\mathbb{N}}$, such that*

$$\delta_{\omega_n} \xrightarrow{w.} \mu$$

Concerning equiweighted sequences of points, we have the following result from [20], which was extended for non compact Ω in [17]:

Theorem 2.3 *Let $\tilde{\omega}_N \in \mathcal{D}_1$ be a particle–ensemble. Then there is a sequence of points $(p_n)_{n\in\mathbb{N}}$ and a constant C, only depending on $\tilde{\omega}_N$, such that*

$$\left|\frac{1}{n}\sum_{i=1}^{n}\delta_{p_i}(\mathbf{A}) - \delta_{\tilde{\omega}_N}(\mathbf{A})\right| \leq \frac{C}{n} \qquad \forall \mathbf{A} \in \mathcal{B}(\Omega) \quad \forall n \in \mathbb{N}$$

Putting these two results together , we gain our required statement (s. [17]):

Theorem 2.4 *Let $\mu \in \mathcal{M}_1$. Then there exists a sequence of points $(p_n)_{n\in\mathbb{N}} \subset \Omega$, such that*

$$\delta_{\omega_{p_n}} \xrightarrow{w.} \mu$$

3 The discrepancy

The terms introduced in this section are originally developed in number theory. We will use the measure–theoretical notation introduced in the section before.

3.1 Asymptotic distributions and the discrepancy

Definition 3.1 *Let $f \in \mathcal{F}_1(\Omega)$. A sequence of points $(p_n)_{n\in\mathbb{N}} \subset \Omega$ is called* ***asymptotically f–distributed*** *:$\Leftrightarrow$*

$$\left|\delta_{\omega_{p_n}}(\mathbf{R}_p) - \mu_f(\mathbf{R}_p)\right| \longrightarrow 0 \qquad \forall \mathbf{R}_p \in \mathcal{R}(\Omega)$$

Remark: $(p_n)_{n\in\mathbb{N}}$ is asymptotically f–distributed $\iff \delta_{\omega_{p_n}} \xrightarrow{w.} \mu_f$

From a practical point of view, it is not only of interest whether a given sequence of points is asymptotically f–distributed, but also how "good" a fixed $\delta_{\omega_{p_n}}$ approximates the given μ_f. This means we need a *distance* in the space of measures. This distance has to induce the weak convergence.

There is a couple of such metrics (see e.g. [17]), but we will only use the *discrepancy*.

Definition 3.2

(i) Let μ and $\nu \in \mathcal{M}_1(\Omega)$. We call

$$\mathbf{D}(\mu,\nu) := \sup_{\mathbf{R}\in\mathcal{R}} |\mu(\mathbf{R}) - \nu(\mathbf{R})|$$

the ***discrepancy*** *of μ and ν*

(ii) If ω_n is a particle–ensemble and $f \in \mathcal{F}_1(\Omega)$, we use the abbreviation:

$$\mathbf{D}_f(\omega_n) := \mathbf{D}(\delta_{\omega_n}, \mu_f)$$

Remarks:

(a) $\mathbf{D}$ is a metric on $\mathcal{M}_1(\Omega)$

(b) There are different kinds of discrepancies, taking the supremum over different sets. The metric introduced above often is called $\mathbf{D}^*$, while $\mathbf{D}$ is used for an equivalent metric.

Theorem 3.1 *If $(\omega_n)_{n\in\mathbb{N}}$ is a sequence of particle–ensembles and $f \in \mathcal{F}_1(\Omega)$, then it holds*

$$\mathbf{D}_f(\omega_n) \longrightarrow 0 \qquad \Longleftrightarrow \qquad \delta_{\omega_n} \xrightarrow{w.} \mu_f$$

Proof:

See [17] (It is shown there, that $\mathbf{D}_f(\omega_n) \longrightarrow 0 \iff (\omega_n)_{n\in\mathbb{N}}$ as.f–distributed — Remember the first remark in this section.) □

The discrepancy is an useful tool to prove weak convergence as well as for error estimates. In one dimension it is quite easy to calculate the discrepancy for a given particle–ensemble, but in higher dimensional cases — and those are the cases we are interested in here — this is a very expensive operation in terms of computational effort. So in this theses a new algorithm for the case of two dimensions is presented in section 7.1.

With the help of the discrepancy, we can state the following result, which shows the connection between particle–ensembles and numerical integration.

3.2 The Koksma–Hlawka–inequality

The goal of the Koksma–Hlawka–inequality is to estimate the error done in calculating

$$\int_\Omega \varphi \, d\delta_{\omega_{p_n}}$$

for a given function φ instead of

$$\int_\Omega \varphi \, d\mu_f$$

This error depends on one hand on $\delta_{\omega_{p_n}}$ — especially on $\mathbf{D}_f(\omega_{p_n})$ but on the other hand also on the choice of φ. The smoother φ is, the better we can approximate the integral by a finite sum. An appropriate measure for this kind of smoothness is the *variation*.

The variation in one dimension is well known, but in higher dimensions there exist a couple of generalizations. We will use the *Hardy– and Krause*–variation, but to introduce it, we need some preparations:

At first a few abbreviations for an easier writing:

Notations:

(1) Let $\mathbf{Z}^{(j)} = \left\{ z_i^{(j)}, \ldots, z_{n(j)}^{(j)} \right\}$ be a subdivision of $[\,0,1\,]$ and let $\varphi : [\,0,1\,]^d \longrightarrow \mathbb{R}$, we write Δ_j for the following operator:

$$\Delta_j \varphi \left(x_1, \ldots, x_{j-1}, z_i^{(j)}, x_{j+1}, \ldots, x_d \right) := \varphi \left(x_1, \ldots, x_{j-1}, z_{i+1}^{(j)}, x_{j+1}, \ldots, x_d \right) - \varphi \left(x_1, \ldots, x_{j-1}, z_i^{(j)}, x_{j+1}, \ldots, x_d \right)$$

(2) The composition of several such operators for different j we abbreviate by:

$$\Delta_{j_1,\ldots,j_m} := \Delta_{j_1} \circ \cdots \circ \Delta_{j_m}$$

$(\Delta_{j_1,\ldots,j_m} = \Delta_{\sigma(j_1),\ldots,\sigma(j_m)}$ for any permutation $\sigma \in \mathbf{S}_m)$

Now we are able to define the *Vitali*–variation:

Definition 3.3 *Let* $\varphi : [0,1]^d \longrightarrow \mathbb{R}$. *We call*

$$\mathbf{V}^{(d)}[\varphi] := \sup_{\substack{\left(\mathbf{Z}^{(j)}=\{z_1^{(j)},\ldots,z_{n_j}^{(j)}\}\right)_{j=1}^{d} \\ \mathbf{Z}^{(j)}:\textit{subdivision}}} \sum_{i_1=1}^{n_1-1} \cdots \sum_{i_d=1}^{n_d-1} \left|\Delta_{1,\ldots,d}\,\varphi\left(z_{i_1}^{(1)},\ldots,z_{i_d}^{(d)}\right)\right|$$

Vitali–variation *of* φ

An easy calculation shows that, if there is one variable, φ does not depend on, then the Vitali–variation of φ is equal to zero. This means, that low Vitali–variation does not imply smoothness.
This is taken into consideration by the Hardy–and Krause–variation:

Notations:

(1) For $[0,1]^d$ the d–dimensional unit–hyper cube, we call $\mathcal{S}_\ell$ with $1 \le \ell \le d$ the set of all ℓ–dimensional boundaries of $[0,1]^d$, i.e. if $\mathbf{S} \in \mathcal{S}_\ell$, then there exists a partitioning

$$\mathbf{I}_0 \uplus \mathbf{I}_1 \uplus \mathbf{J} = \{1,\ldots,d\}$$

$$\text{where} \qquad |\mathbf{I}_0| + |\mathbf{I}_1| = d - \ell \qquad \text{and} \qquad |\mathbf{J}| = \ell,$$

such that for any $(x_1,\ldots,x_d) \in \mathbf{S}$ it holds

$$\begin{array}{lcll} x_\nu & = & 0 & \text{if} \quad \nu \in \mathbf{I}_0 \\ x_\nu & = & 1 & \text{if} \quad \nu \in \mathbf{I}_1 \\ x_\nu & \in & [0,1] & \text{if} \quad \nu \in \mathbf{J}. \end{array}$$

(2) Let $\varphi : [0,1]^d \longrightarrow \mathbb{R}$, $\mathbf{S} \in \mathcal{S}_\ell$. We write $\varphi_{\mathbf{S}}$ for the restriction of φ on $\mathbf{S}$.

Definition 3.4 *Let* $\varphi : [0,1]^d \longrightarrow \mathbb{R}$

(i) $\mathbf{V}[\varphi] := \sum_{\ell=0}^{d} \sum_{\mathbf{S}\in\mathcal{S}_\ell} \mathbf{V}^{(\ell)}[\varphi_{\mathbf{S}}]$

(ii) φ *is said to be of* ***bounded variation in the sense of Hardy and Krause*** $:\Leftrightarrow$

$$\mathbf{V}[\varphi] < \infty$$

Now we are able to state the following theorem:

Theorem 3.2 (The Koksma–Hlawka-inequality)
Let $f \in \mathcal{F}_1([0,1]^d)$, $\{p_1,\ldots,p_N\} \subset [0,1]^d$ *and let* $\varphi : [0,1]^d \longrightarrow \mathbb{R}$ *be of bounded variation in the sense of Hardy and Krause. Then it holds*

$$\left| \int_{[0,1]^d} \varphi \, d\delta_{\omega_{p_N}} - \int_{[0,1]^d} \varphi \, d\mu_f \right| \le \mathbf{V}[\varphi]\, \mathbf{D}_f(\omega_{p_N})$$

For the proof see [17]. (There the inequality is stated for the (more general) case of non compact Ω.)

Remarks:

(a) This inequality is not only true for $\Omega = [\,0,1\,]^d$. There are several generalization, for example for unbounded Ω.

(b) Other formulations of the Koksma–Hlawka–inequality use the modulus of continuity instead of the variation, or the theorem is stated for Lipschitz–continuous functions φ.

3.3 Uniform distributed sequences of points

This section shall give a short introduction into the construction of uniform distributed sequences of points. See e.g.[15] for details of the theory of uniform distribution.

Definition 3.5 *A sequence of points* $(p_n)_{n\in\mathbb{N}} \subset \Omega$ *is called* ***asymptotically uniformly distributed*** $:\Leftrightarrow$
$(p_n)_{n\in\mathbb{N}}$ *is asymptotically* χ_Ω*–distributed where*

$$\chi_\Omega(x) \;:=\; \frac{1}{Vol_\lambda(\Omega)} \qquad \forall x \in \Omega$$

There are several ideas to construct uniformly distributed sequences — and still a lot of work is done in this field. (see e.g. [21]) We are not only interested in the asymptotical behaviour of those sequences, but it is also important — especially for the practical purpose — how the discrepancy of ω_{p_N} develops for relatively small N.

We just mention the following methods:

- I.M.Sobol's $\mathbf{LP}_\tau$ – sequences. (see [26]), which were used for the numerical tests of the methods introduced in this theses.
- H.Faure produces $\mathbf{LP}_0$–sequences with respect to the basis d (see [8]).
- H.Niederreiter developed $\mathbf{LP}_0$–sequences using relationships between polynomials over finite fields (see [19])
- The sequences introduced by Hammersley and Halton (see [11]).

There are some numerical experiments to compare the different methods (see e.g. [22]), but the interpretation of the result is not that easy.

An important application of uniformly distributed sequences follows by theorem 3.2: numerical integration in higher dimensions, because the asymptotics of the sequences above is:

$$\mathbf{D}_{\chi_{[\,0,1\,]^d}}(\omega_{p_N}) = \mathcal{O}\left(\frac{\ln^d N}{N}\right)$$

How uniformly distributed sequences are used to construct f–distributed sequences, we will see in the following chapters.

4 The method by Hlawka and Mück

4.1 The transformation

Notation: For $f \in \mathcal{F}_1([0,1]^d)$ the transformation

$$\mathbf{F} : [0,1]^d \longrightarrow [0,1]^d$$

is given by its coordinate–functions

$$\mathbf{F}_k(x_1,\ldots,x_k) := \frac{\int_0^{x_k}\int_0^1\cdots\int_0^1 f(x_1,\ldots,x_{k-1},\xi_k,\ldots,\xi_d)\,d\xi_d\cdots d\xi_{k+1}\,d\xi_k}{\int_0^1\int_0^1\cdots\int_0^1 f(x_1,\ldots,x_{k-1},\xi_k,\ldots,\xi_d)\,d\xi_d\cdots d\xi_{k+1}\,d\xi_k}$$

Properties 4.1

(1) For fixed $a_1,\ldots,a_{k-1}$ is $\mathbf{F}_k(a_1,\ldots,a_{k-1},x_k)$ is monotonically not descending in x_k.

(2) $\mathbf{F}$ is surjective.

(3) If f is positive (i.e. $f(x) > 0\ \forall x \in [0,1]^d$), Then $\mathbf{F}$ is also injective.

(4) If $f \in \mathcal{C}^1([0,1]^d)$, then $\mathbf{F}$ is a diffeomorphism and it holds:

(a) The Jacobian $\mathbf{J}_F = \left(\frac{\partial \mathbf{F}_i}{\partial x_j}\right)_{i,j=1}^d$ is a lower triangular matrix

(b) The diagonal elements have the structure

$$\frac{\partial \mathbf{F}_k}{\partial x_k} = \frac{\mathbf{I}_k}{\mathbf{I}_{k-1}},$$

$$\text{where } \mathbf{I}_k := \int_0^1\cdots\int_0^1 f(x_1,\ldots,x_k,\xi_{k+1},\ldots,\xi_d)\,d\xi_d\cdots d\xi_k$$

(c) Because of

$$\mathbf{I}_0 = \int_{[0,1]^d} d\mu_f = 1 \text{ and } \mathbf{I}_d = f$$

it holds, that the functional–determinant $|\mathbf{J}_F| = f$

(d) If we define $\mathbf{T} := \mathbf{F}^{-1}$, then it holds

$$|\mathbf{J}_T| = \frac{1}{f}$$

and using the transformation formula for integrals:

$$\int_{[0,1]^d} \varphi(\xi) f(\xi)\,d\xi = \int_{[0,1]^d} \varphi(\mathbf{T}(x)) f(x) |\mathbf{J}_T|\,dx = \int_{[0,1]^d} \varphi(\mathbf{T}(x))\,dx$$

4.2 Transformed sequences of points

Theorem 4.2 *Let $(\omega_{p_n})_{n\in\mathbb{N}}$ be an asymptotically uniformly distributed sequence of points and let f as well as $\mathbf{F}$ and $\mathbf{T}$ be given as in 4.1(4). Then $(\omega_{T(p_n)})_{n\in\mathbb{N}}$ is asymptotically f–distributed*

Proof:

Show $\delta_{\omega_{T(p_n)}} \xrightarrow{w.} \mu_f$

Let $\varphi \in \mathcal{C}_b([0,1]^d)$. Using the definition of μ_f it holds

$$\int\limits_{[0,1]^d} \varphi \, d\mu_f = \int\limits_{[0,1]^d} \varphi f \, d\lambda = \int\limits_{[0,1]^d} \psi \, d\lambda$$

where $\psi := \varphi f$. (Thus $\psi \in \mathcal{C}_b([0,1]^d)$). Because $\omega_{p_n}$is asymptotically uniformly distributed, it holds

$$\int\limits_{[0,1]^d} \psi \, d\delta_{\omega_{p_n}} \longrightarrow \int\limits_{[0,1]^d} \psi \, d\lambda$$

The theorem for transformations of measures (see e.g. [25] Th. 7.26) implies

$$\int\limits_{F([0,1]^d)} \varphi \, d\delta_{\omega_{T(p_n)}} = \int\limits_{[0,1]^d} (\varphi \circ \mathbf{F}) |\mathbf{J}_F| \, d\delta_{\omega_{T(p_n)}} = \int\limits_{[0,1]^d} \varphi f \, d\delta_{\omega_{p_n}} = \int\limits_{[0,1]^d} \psi \, d\delta_{\omega_{p_n}}$$

Putting everything together we get:

$$\int\limits_{[0,1]^d} \varphi \, d\delta_{\omega_{T(p_n)}} = \int\limits_{[0,1]^d} \psi \, d\delta_{\omega_{p_n}} \longrightarrow \int\limits_{[0,1]^d} \psi \, d\lambda = \int\limits_{[0,1]^d} \varphi \, d\mu_f$$

□

In the case, if we do not have the property $f > 0$, the following corollary is true:

Corollary 4.3 *Let $f \in \mathcal{F}_1([0,1]^d) \cap \mathcal{C}^1([0,1]^d)$ and $\mathbf{K} := supp(f)$. Then it holds*

(i) $\mathbf{F} : \mathbf{K} \longrightarrow [0,1]^d$ *defined as above is a diffeomorphism.*

(ii) If we define $\mathbf{T} : [0,1]^d \longrightarrow \mathbf{K}$ *as* $\mathbf{T} := (\mathbf{F}\,|_{\mathbf{K}})^{-1}$ *and $(\omega_{p_n})_{n\in\mathbb{N}}$ is an as. uniformly distributed sequence of points, then $(\omega_{T(p_n)})_{n\in\mathbb{N}}$ is as. f–distributed.*

Proof:

ad (ii): Let $\varphi \in {}_b([0,1]^d)$. Then we get analogously to the proof of 4.2

$$\int\limits_{\mathbf{K}} \varphi \, d\delta_{\omega_{T(p_n)}} \longrightarrow \int\limits_{\mathbf{K}} \varphi \, d\mu_f$$

If $\mathbf{N} := [0,1]^d \setminus \mathbf{K}$, we have

$$\delta_{\omega_{T(p_n)}}(N) = 0 \qquad \text{and} \qquad \mu_f(N) = 0$$

□

Hlawka and Mück can prove the following — not really satisfactory — result concerning the discrepancy of the transformed sequences (see [14])

Theorem 4.4 *Suppose $\omega_{p_n} \subset [0,1]^d$ is a sequence of points and for the transformation $\mathbf{F}$ the Lipschitz–condition*

$$\|\mathbf{F}(x) - \mathbf{F}(y)\|_\infty \leq k\|x - y\|_\infty$$

holds, then there exists a constant c, s.t.

$$\mathbf{D}_f\left(\omega_{T(p_n)}\right) \leq c \cdot \sqrt[d]{\mathbf{D}_{\chi_{[0,1]^d}}\left(\omega_{p_n}\right)}$$

Remarks:

(a) If f is continuously differentiable, then $\mathbf{F}$ is Lipschitz–continuous.

(b) This theorem also implies, that as. uniformly distributed sequences are transformed into f–distributed ones.

(c) If f factorizes $\left(f(x_1,\ldots,x_d) = \prod_{i=1}^{d} f_i(x_i)\right)$, then $\mathbf{F}(\mathbf{R}) \in \mathcal{R}(\Omega)$ for any $\mathbf{R} \in \mathcal{R}(\Omega)$ and therefore

$$\mathbf{D}_f\left(\omega_{T(p_n)}\right) \leq c \cdot \mathbf{D}_{\chi_{[0,1]^d}}\left(\omega_{p_n}\right)$$

4.3 The method

We have seen, that using the transformed sequences of points $\omega_{T(p_n)}$, we have the possibility to reduce the problem of constructing f–distributed sequences to the wellknown and solved problem of constructing uniformly distributed sequences of points.

The difficulty now is hidden in the transformation $\mathbf{T} = \mathbf{F}^{-1}$, which is not given for arbitrary f. We will now briefly introduce the idea of Hlawka and Mück. They use a recurrent formula to approximately find the transformed points.

Notation: Let $\alpha \in \mathbb{R}$, we write

$$[\alpha] := \max_{z \in \mathbb{Z}} \{z \leq \alpha\}$$

Method by Hlawka and Mück

given:	N	number of points to be transformed
	$p^{(1)},\ldots,p^{(N)}$	sequence of points
goal:	$\tilde{q}^{(1)},\ldots,\tilde{q}^{(N)}$	approximation of $\mathbf{T}(p_1)\ldots\mathbf{T}(p_N)$

The $\tilde{q}^{(i)} = (\tilde{q}_1^{(i)},\ldots,\tilde{q}_d^{(i)})$ are calculated by the recurrent formula:

$$\tilde{q}_k^{(i)} := \frac{1}{\mathsf{N}} \sum_{j=1}^{\mathbf{N}} \left[\, 1 + p_k^{(i)} - \mathsf{F}\left(\tilde{q}_1^{(i)},\ldots,\tilde{q}_{i-1}^{(i)}, p_i^{(j)}\right) \right] \qquad \begin{matrix} k = 1,\ldots,d \\ i = 1,\ldots,\mathsf{N} \end{matrix}$$

Hlawka and Mück are able to show (see [14]):

Theorem 4.5 *Suppose f is chosen, such that $\mathbf{F} \in \mathcal{C}^d$ and there is a $\mathbf{M} \in \mathbb{R}$, such that*

$$D^\alpha \mathbf{F} \leq \mathbf{M} \qquad \text{for any multiindex } \alpha \in \mathbb{N}^d \quad \text{where } |\alpha| \leq d$$

If ω_{p_N} is an arbitrary sequence of points and $\omega_{\tilde{q}_N}$ is constructed using the above method, then the discrepancy of $\omega_{\tilde{p}_N} := \omega_{F(\tilde{q}_N)}$ can be estimated by:

$$\mathbf{D}_f\left(\omega_{\tilde{p}_N}\right) \leq (1+2\mathbf{M})^d \, \mathbf{D}_f\left(\omega_{p_N}\right)$$

As asymptotic behaviour, we get:

Corollary 4.6 *If f is chosen like in the theorem above, $(\omega_{p_n})_{n\in\mathbb{N}}$ is as. uniformly distributed and for any $N \in \mathbb{N}$ $\omega_{\tilde{q}_N}$ is constructed by the method of Hlawka and Mück, then*

$$\mathbf{D}_f\left(\omega_{\tilde{q}_N}\right) \longrightarrow 0,$$

i.e. $(\omega_{\tilde{q}_n})_{n\in\mathbb{N}}$ is as. f–distributed.

□

4.4 Some critical remarks

Hlawka and Mück's method provides the ability to find a $\delta_{\omega_{p_N}}$, that approximates μ_f for any given $f \in \mathcal{F}_1$. For the calculation, we only need the knowledge of $\mathbf{F}$ but not of $\mathbf{T} = \mathbf{F}^{-1}$. We get $\mathbf{F}$ by integration of f.

In general we do not have an analytical description of $\mathbf{F}$, but we have to find the value of $\mathbf{F}$ by numerical integration. Therefore the number of evaluations of $\mathbf{F}$ has main influence on the efficiency of the method. If we define

$\Phi(N)$:= Number of evaluations of $\mathbf{F}$ to construct a sequence of N points,

then we get for the Hlawka–/Mück–method

$$\Phi_{\text{H–M}}(N) = d \cdot N^2 \quad \text{i.e.} \quad \Phi_{\text{H–M}} = \mathcal{O}(N^2).$$

For increasing N this leads to enormous computational effort.

An advantage of the method is, that the $\tilde{q}^{(i)}$'s are constructed completely independently. The relationships between the $\tilde{q}^{(i)}$'s are given only by the uniform distribution of the originally $p^{(i)}$'s. Hence we can construct the $\tilde{q}^{(i)}$'s in parallel.

This method constructs sequences of ensembles. For any given N a special set of points $\tilde{q}_N^{(1)}, \ldots, \tilde{q}_N^{(N)}$ is constructed. And for fixed original sequence of uniformly distributed points N is the only parameter to decrease the discrepancy. So if it is necessary to increase the accuracy of the approximation, we have to reconstruct the whole sequence. This is a big disadvantage of the Hlawka/Mück method.

If we want to use a refinement method, then the distributions of two succeeding time steps $f^{(k\tau)}$ and $f^{((k+1)\tau)}$ do not differ much. So it would be good, if we were able to reuse the information gained in step $k\tau$ to construct the sequence of step $(k+1)\tau$. Using the Hlawka/Mück formula this is impossible.

The advantages and disadvantages are written together in the following:

+ The method "works" for arbitrary $f \in \mathcal{F}_1$. (But the error estimates need a more regular f)
+ The asymptotic behaviour $\mathbf{D}_f \longrightarrow 0$.
+ The possibility to construct the $\tilde{q}^{(i)}$'s in parallel.
− $\mathcal{O}(\Phi(N)) = N^2$
− Decreasing the discrepancy leads to complete reconstruction.
− The same problem with refinement.

5 The iterative ansatz

5.1 Iterative methods to construct transformed sequences

The basic idea of the iterative method to construct transformed sequences without explicit evaluation of the transformation $\mathbf{T} = \mathbf{F}^{-1}$ is to replace the direct calculation

$$q^{(i)} = \mathbf{T}p^{(i)} \quad i = 1, \ldots, N$$

by the solution of the nonlinear system of equations

$$\mathbf{F}q^{(i)} = p^{(i)} \quad i = 1, \ldots, N.$$

From the preceding section (4.4) we state the following requirements on the choice of the solver of the system:

(1) The method shall converge for arbitrary $f \in \mathcal{F}_1$, i.e. for given f, $p^{(1)}, \ldots, p^{(N)}$ and $\varepsilon > 0$ the method finds $\tilde{q}^{(1)}, \ldots, \tilde{q}^{(N)}$, such that

$$\left\| \mathbf{F}\tilde{q}^{(i)} - p^{(i)} \right\| \leq \varepsilon \quad i = 1, \ldots, N \tag{1}$$

(2) Given $\varepsilon > 0$ the method shall construct sequences *of points* $(\tilde{q}^{(n)})_{n \in \mathbb{N}}$, which fulfil condition 1.

So we gain the property of extensionalability, but on the other hand we loose the nice asymptotics of the Hlawka/Mück sequences, because we have to introduce the additional parameter ε.

(3) The method shall solve the equations for the separate points independently, such that it is possible to construct the points in parallel.

(4) For given ε the number of evaluations of $\mathbf{F}$ shall only increase linearly with the number of transformed points:

$$\mathcal{O}\big(\Phi(N)\big) = N$$

(5) The method shall be able to reuse the old $\tilde{q}^{(1)}_{[k]}, \ldots, \tilde{q}^{(N)}_{[k]}$ in the case where the new $f^{(k+1)}$ does not differ to much from the old one. This means the method shall support refinement methods.

The requirement (5) especially is fulfilled by iterative methods, i.e. by methods

$$\begin{array}{rcl} \Theta : \mathbb{R}^{d\times N} & \longrightarrow & \mathbb{R}^{d\times N} \\ \left(q^{(1)},\ldots,q^{(N)}\right) & \longmapsto & \left(\theta_1(q^{(1)},\ldots,q^{(N)},),\ldots,\theta_N(q^{(1)},\ldots,q^{(N)})\right) \end{array}$$

$$\text{where} \quad \left\| \mathbf{F}\theta_i(q^{(1)},\ldots,q^{(N)}) - p^{(i)} \right\| \leq \left\| \mathbf{F}q^{(i)} - p^{(i)} \right\| \quad i = 1,\ldots,N$$

Requirement (3) implies, that Θ faktorizes:

$$\theta_i(q^{(1)},\ldots,q^{(N)}) = \theta_i(q^{(i)})$$

If furthermore we require that

$$\theta^{(i)} = \theta^{(1)} =: \theta \quad \forall\, i \in \mathbb{N},$$

then we fulfil requirement (2), too.

This leads to the basic definition of methods using the iterative ansatz:

method using the iterative ansatz

given:	N	Number of the particles to be constructed
	$p^{(1)},\ldots,p^{(N)}$	Particles to be transformed
	$q_{[0]}^{(1)},\ldots,q_{[0]}^{(N)}$	Initial distribution
goal:	$\tilde{q}^{(1)},\ldots,\tilde{q}^{(N)}$	particles fulfilling condition (1)

For any $i = 1,\ldots,$**N**

construct $q_{[0]}^{(i)}, q_{[1]}^{(i)},\ldots,q_{[\nu_i]}^{(i)}$, where $q_{[k+1]}^{(i)} := \theta(q_{[k]}^{(i)})$ and ν_i minimal, such that $q_{[\nu_i]}^{(i)}$ fulfils condition (1).

Set $\tilde{q}^{(i)} := q_{[\nu_i]}^{(i)}$

5.2 Estimates of the discrepancy

Because the method constructs a sequence $(\tilde{q}^{(n)})_{n\in\mathbb{N}}$, such that condition (1) holds, instead of the transformed sequence $(q^{(n)} := \mathbf{T}p^{(n)})_{n\in\mathbb{N}}$, we have to compare the discrepancies.

Theorem 5.1 *Suppose $f \in \mathcal{F}_1([0,1]^d)$ is chosen, such that the coordinate functions $\mathbf{F_1},\ldots,\mathbf{F_d}$ fulfil: For any $(a_1,\ldots,a_{k-1}) \in [0,1]^{k-1}\mathbf{F}_k(a_1,\ldots,a_{k-1},x_k)$ is twice continuously differentiable as a function of x_k and there exists*

$$\mathbf{M}_k := \sup_{(x_1,\ldots,x_k)\in[0,1]^k} \left(\frac{\partial \mathbf{F}_k}{\partial x_k}\Big|_{(x_1,\ldots,x_k)} \right)^{-1}.$$

Let ω_{x_N} and ω_{y_N} be sequences of points fulfilling

$$\|\mathbf{F}x_j - \mathbf{F}y_j\| \leq \varepsilon$$

where $\varepsilon > 0$ is given, then it holds

$$\left|\mathbf{D}_f(\omega_{x_N}) - \mathbf{D}_f(\omega_{y_N})\right| \leq \left(\sum_{k=1}^{d} \mathbf{M}_k \right) \|f\|_\infty\, \varepsilon + \mathcal{O}(\varepsilon^2)$$

Proof:

Let $\mathbf{R} := \mathbf{R}_p = \mathbf{R}_{(p_1,\ldots,p_d)} \in \mathcal{R}([\,0,1\,]^d)$.
We define:

$$\delta_1^+ := \begin{cases} \mathbf{F}_1^{-1}(\mathbf{F}_1(p_1)+\varepsilon) - p_1 & if \quad \mathbf{F}_1(p_1) \le 1-\varepsilon \\ 1-p_1 & else \end{cases}$$

For $k = 2,\ldots,d$ we introduce the following functions:

$$\begin{aligned} \mathbf{F}_k^{p+\delta^+} : \mathbb{R} &\longrightarrow \mathbb{R} \\ p_k &\longmapsto \mathbf{F}_k(p_1+\delta_1^+,\ldots,p_{k-1}+\delta_{k-1}^+,p_k) \end{aligned}$$

and we set

$$\delta_k^+ := \begin{cases} (\mathbf{F}_k^{p+\delta^+})^{-1}(\mathbf{F}_k^{p+\delta^+}(p_k)+\varepsilon) - p_k & \\ & if \quad \mathbf{F}_k^{p+\delta^+}(p_k) \le 1-\varepsilon \\ 1-p_k & else \end{cases}$$

Analogously we define

$$\delta_1^- := \begin{cases} \mathbf{F}_1^{-1}(\mathbf{F}_1(p_1)-\varepsilon) - p_1 & if \quad \mathbf{F}_1(p_1) \ge 0 \\ p_1 & else \end{cases}$$

$$\begin{aligned} \mathbf{F}_k^{p-\delta^-} : \mathbb{R} &\longrightarrow \mathbb{R} \\ p_k &\longmapsto \mathbf{F}_k(p_1-\delta_1^-,\ldots,p_{k-1}-\delta_{k-1}^-,p_k) \end{aligned}$$

$$\delta_k^- := \begin{cases} (\mathbf{F}_k^{p-\delta^-})^{-1}(\mathbf{F}_k^{p-\delta^-}(p_k)+\varepsilon) - p_k & \\ & if \quad \mathbf{F}_k^{p+\delta^+}(p_k) \ge \varepsilon \\ p_k & else \end{cases}$$

Because the first derivatives of any $\mathbf{F}_k^{p\pm\delta^\pm}$ are bounded from below, these $\delta_k^\pm$ are welldefined.

We use the abbreviations:

$$\begin{aligned} \mathbf{R}^+ &:= \mathbf{R}_{(p_1+\delta_1^+,\ldots,p_d+\delta_d^+)} \\ \mathbf{R}^- &:= \mathbf{R}_{(p_1-\delta_1^-,\ldots,p_d-\delta_d^-)} \end{aligned}$$

By definition of $\mathbf{R}^+$ and $\mathbf{R}^-$ it follows

$$\delta_{\omega_{y_N}}(\mathbf{R}^-) \le \delta_{\omega_{x_N}}(\mathbf{R}) \le \delta_{\omega_{y_N}}(\mathbf{R}^+) \tag{2}$$

because if $y_\ell \in \mathbf{R}^-$, then $x_\ell \in \mathbf{R}$, if not there would exist a $k \in \{1,\ldots,d\}$ where $|\mathbf{F}_k(y_\ell) - \mathbf{F}_k(x_\ell)| > \varepsilon$ and this would contradict the assumptions of the theorem. Analogously $x_\ell \in \mathbf{R}$ implies $y_\ell \in \mathbf{R}^+$.

The definition of the discrepancy leads to:

$$\begin{aligned} |\delta_{\omega_{y_N}}(\mathbf{R}^+) - \mu_f(\mathbf{R}^+)| &\le \mathbf{D}_f(\omega_{y_N}) \\ |\delta_{\omega_{y_N}}(\mathbf{R}^-) - \mu_f(\mathbf{R}^-)| &\le \mathbf{D}_f(\omega_{y_N}) \end{aligned}$$

Using the triangular inequality $|a-b| \geq |a| - |b|$ and the positivity of the summands, we can write this as:

$$\begin{aligned} \delta_{\omega_{y_N}}(\mathbf{R}^+) &\geq -\mathbf{D}_f(\omega_{y_N}) + \mu_f(\mathbf{R}^+) \\ \delta_{\omega_{y_N}}(\mathbf{R}^-) &\geq -\mathbf{D}_f(\omega_{y_N}) + \mu_f(\mathbf{R}^-) \end{aligned}$$

Putting this in (2) and subtracting $\mu_f(\mathbf{R})$, we get:

$$\begin{aligned} -(\mu_f(\mathbf{R}) - \mu_f(\mathbf{R}^-)) - \mathbf{D}_f(\omega_{y_N}) \leq & \\ \leq \; & \delta_{\omega_{x_N}}(\mathbf{R}) - \mu_f(\mathbf{R}) \\ \leq \; & \mu_f(\mathbf{R}^+) - \mu_f(\mathbf{R}) + \mathbf{D}_f(\omega_{y_N}) \end{aligned} \tag{3}$$

So we must estimate the differences between the f–Volumes of $\mathbf{R}^+$ and $\mathbf{R}$ (resp. between $\mathbf{R}$ and $\mathbf{R}^-$).

$$\begin{aligned} \mu_f(\mathbf{R}^+) - \mu_f(\mathbf{R}) = & \\ \mu_f\big(\mathbf{R}_{(p_1+\delta_1^+, p_2+\delta_2^+, \ldots, p_d+\delta_d^+)}\big) \; - \; & \mu_f\big(\mathbf{R}_{(p_1, p_2, \ldots, p_d)}\big) \; = \\ \mu_f\big(\mathbf{R}_{(p_1+\delta_1^+, p_2+\delta_2^+, \ldots, p_d+\delta_d^+)}\big) \; - \; & \mu_f\big(\mathbf{R}_{(p_1, p_2+\delta_2^+, \ldots, p_d+\delta_d^+)}\big) \; + \\ \mu_f\big(\mathbf{R}_{(p_1, p_2+\delta_2^+, \ldots, p_d+\delta_d^+)}\big) \; - \; & \mu_f\big(\mathbf{R}_{(p_1, p_2, \ldots, p_d+\delta_d^+)}\big) \; + \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots & \; + \\ \mu_f\big(\mathbf{R}_{(p_1, \ldots, p_{d-1}, p_d+\delta_d^+)}\big) \; - \; & \mu_f\big(\mathbf{R}_{(p_1, \ldots, p_{d-1}, p_d)}\big) \end{aligned} \tag{4}$$

We examine the k'th row:

$$\begin{aligned} \mu_f\big(\mathbf{R}_{(p_1, \ldots, p_{k-1}, p_k+\delta_k^+, \ldots, p_d+\delta_d^+)}\big) - \mu_f\big(\mathbf{R}_{(p_1, \ldots, p_k, p_{k+1}+\delta_{k+1}^+, \ldots, p_d+\delta_d^+)}\big) = & \\ = \int_{p_k}^{p_k+\delta_k^+} \int_0^{p_1} \cdots \int_0^{p_{k-1}} \int_0^{p_{k+1}} \cdots \int_0^{p_d} f(\xi_1, \ldots, \xi_d)\, d\xi_d \cdots d\xi_{k+1} d\xi_{k-1} \cdots d\xi_1\, d\xi_d & \\ \leq \; \delta_k^+ \underbrace{\mathrm{Vol}_{\lambda_{\mathbb{R}^{d-1}}}\left(\mathbf{R}_{(p_1, \ldots, p_{k-1}, p_{k+1}, \ldots, p_d)}\right)}_{\leq 1} \|f\|_\infty & \\ \leq \; \delta_k^+ \, \|f\|_\infty & \end{aligned}$$

Using the definition of δ_k^+, we have:

$$\begin{aligned} \delta_k^+ &\leq \frac{\partial((\mathbf{F}_k^{p+\delta})^{-1})}{\partial x_n}\varepsilon + \mathcal{O}(\varepsilon^2) \\ &\leq \mathbf{M}_k\, \varepsilon + \mathcal{O}(\varepsilon^2) \end{aligned}$$

Hence using (4):

$$\mu_f(\mathbf{R}^+) - \mu_f(\mathbf{R}) \leq \left(\sum_{k=1}^{d} \mathbf{M}_k\right) \|f\|_\infty\, \varepsilon + \mathcal{O}(\varepsilon^2)$$

The analogous examination on the lower bound in (3) leads to:

$$\mu_f(\mathbf{R}) - \mu_f(\mathbf{R}^-) \leq \left(\sum_{k=1}^{d} \mathbf{M}_k\right) \|f\|_\infty\, \varepsilon + \mathcal{O}(\varepsilon^2)$$

R was chosen arbitrarily, hence (3) implies:

$$\mathbf{D}_f(\omega_{x_N}) - \mathbf{D}_f(\omega_{y_N}) \leq \left(\sum_{k=1}^{d} \mathbf{M}_k \right) \|f\|_\infty \, \varepsilon + \mathcal{O}(\varepsilon^2)$$

and interchanging the rolls of $\delta_{\omega_{x_N}}$ and $\delta_{\omega_{y_N}}$ leads to the proof of the statement.

□

Remark: If $f \geq \mu > 0$ and $f \in \mathcal{C}^1$, then f fulfils the assumptions of the statement.

Corollary 5.2 *Suppose f is chosen as in the above theorem and $\omega_{p^{(N)}}$ is a sequence of points. If we denote by $\omega_{q^{(N)}}$ the transformed sequence and by $\omega_{\tilde{q}_\varepsilon^{(N)}}$ a sequence constructed by an iterative method, then*

$$\left| \mathbf{D}_f(\omega_{q^{(N)}}) - \mathbf{D}_f(\omega_{\tilde{q}_\varepsilon^{(N)}}) \right| \leq \left(\sum_{i=1}^{d} \mathbf{M}_i \right) \|f\|_\infty \, \varepsilon + \mathcal{O}(\varepsilon^2)$$

Proof:

Because the constructed sequence has to fulfil condition (1), it holds

$$\left\| \mathbf{F}\tilde{q}_\varepsilon^{(i)} - p^{(i)} \right\| \leq \varepsilon \quad i = 1, \ldots, N$$

By definition of the transformed sequence, we have

$$p^{(i)} = \mathbf{F}q^{(i)} \quad i = 1, \ldots, N$$

hence

$$\left\| \mathbf{F}\tilde{q}_\varepsilon^{(i)} - \mathbf{F}q^{(i)} \right\| \leq \varepsilon \quad i = 1, \ldots, N$$

and the corollary is proved by the latter theorem.

□

If one leaves the sequence $\omega_{q^{(N)}}$ fixed and changes the distribution f — this is important with respect to requirement (5) — one gets

Theorem 5.3 *Let ω_{p_n} be a particle–ensemble and let $f, g \in \mathcal{F}_1(\Omega)$,then*

$$\left| \mathbf{D}_f\left(\omega_{p_n}\right) - \mathbf{D}_g\left(\omega_{p_n}\right) \right| \leq \|f - g\|_{L_1}$$

Proof:

Let $\mathbf{R} \in \mathcal{R}(\Omega)$ be arbitrary but fixed. By the definition of the discrepancy, it holds

$$\left| \delta_{\omega_{p_n}}(\mathbf{R}) - \mu_f(\mathbf{R}) \right| \leq \mathbf{D}_f(\omega_{p_n})$$

On the other hand, we have:

$$\left| \delta_{\omega_{p_n}}(\mathbf{R}) - \mu_g(\mathbf{R}) \right| \leq \left| \delta_{\omega_{p_n}}(\mathbf{R}) - \mu_f(\mathbf{R}) \right| + \left| \mu_f(\mathbf{R}) - \mu_g(\mathbf{R}) \right|$$

and using

$$\begin{aligned}\left|\mu_f(\mathbf{R}) - \mu_g(\mathbf{R})\right| &\leq \int_{\mathbf{R}} |f-g|\, d\lambda \\ &\leq \|f-g\|_{L_1}\end{aligned}$$

we get:

$$\left|\delta_{\omega_{p_n}}(\mathbf{R}) - \mu_g(\mathbf{R})\right| \leq \mathbf{D}_f(\omega_{p_n}) + \|f-g\|_{L_1}$$

and because $\mathbf{R}$ was chosen arbitrarily, we get:

$$\mathbf{D}_g(\omega_{p_n}) - \mathbf{D}_f(\omega_{p_n}) \leq \|f-g\|_{L_1}$$

Interchanging f and g leads to the required result. □

5.3 Asymptotic behaviour of iterative methods

Altogether, using results (4.2) and (5.2), we get:

Corollary 5.4 *Let f be given as in theorem (5.2), furthermore let $\omega_{p^{(n)}}$ be an asymptotically uniformly distributed sequence of points and $\varepsilon_n \longrightarrow 0$. If $\omega_{\tilde{q}^{(n)}_{\varepsilon_n}}$ is chosen, such that for any $n \in \mathbb{N}$ it holds:*

$$\left\| \mathbf{F}\tilde{q}^{(i)}_{\varepsilon_n} - p^{(i)} \right\| \leq \varepsilon_n \quad i = 1, \ldots, n$$

Then $\omega_{\tilde{q}^{(n)}_{\varepsilon_n}}$ is an asymptotically f–distributed sequence of particle ensembles.

Proof:

Let $\omega_{q^{(n)}}$ denote the transformed sequence of points with respect to $\omega_{p^{(n)}}$. Then theorem (4.2) implies:

$$\mathbf{D}_f(\omega_{q^{(n)}}) \longrightarrow 0$$

Using (5.2) and $\varepsilon_n \longrightarrow 0$ we get

$$\left|\mathbf{D}_f(\omega_{q^{(n)}}) - \mathbf{D}_f(\omega_{\tilde{q}^{(n)}_{\varepsilon_n}})\right| \longrightarrow 0$$

Because of the linearity of limits:

$$\mathbf{D}_f(\omega_{\tilde{q}^{(n)}_{\varepsilon_n}}) \longrightarrow 0.$$

□

Hence for the asymptotic behaviour we again need sequences of ensembles and therefore the requirement of extensionability is not completely fulfilled. But we have now two parameters and hence more flexibility. Furthermore:

(1) If $\omega_{\tilde{q}^{(N)}_{\varepsilon_N}}$ is constructed and we need $\omega_{\tilde{q}^{(M)}_{\varepsilon_M}}$ for $M > N$, we get $q^{(1)}_{\varepsilon_M}, \ldots, q^{(N)}_{\varepsilon_M}$ by iterating the $q^{(1)}_{\varepsilon_N}, \ldots, q^{(N)}_{\varepsilon_N}$. So we do not loose gained information as using the Hlawka/Mück–method.

(2) By one additional iteration ε decreases in general so much, that this ε is usable for many N.

5.4 Some remarks on quasi–Newton–methods

Quasi–Newton–methods minimize nonlinear functionals. We choose

$$\Psi(x) \ := \ \left\| \mathbf{F}x - p^{(i)} \right\|^2$$

to be minimized, where the square is taken for regularity aspects. As distance we take the Euclidean one.

Remark: For every step of iteration, it is necessary to calculate the gradient of Ψ. Therefore we need $\Psi \in \mathcal{C}^1$. This is true if $f \in \mathcal{C}^1$.

M.J.D.Powell has shown in [23], that the sequence $(x_{[n]})_{n \in \mathbb{N}}$ constructed by the quasi–Newton–method introduced there — a BFGS–method using inexact line–search — converges for arbitrary $x_{[0]} \in \mathbb{R}^d$ superlinearly to a local minimum of Ψ if the following assumptions are true:

(i) $\Psi \in \mathcal{C}^2$

(ii) Ψ convex

These requirements, especially the convexity, are the main problems if we want to use such methods to transform our sequences of particles, because we can not guarantee them for arbitrary f. (Even smooth f do not lead to convex Ψ). Another objection to those minimization methods in our context is given by the necessity of using the square of the distance, such that the convergence is slow.

An adaptation of a quasi–Newton–method is shown in more detail in my diploma–thesis ([10]).

So quasi–Newton–methods fulfil requirements (2) – (5) given at the beginning of this chapter but they fail concerning requirement (1). So we must choose an other ansatz to the iterative methods.

6 The method

Because minimization–methods lead to the problems shown in the latter section , we have to choose another ansatz to solve the system of nonlinear equations:

6.1 Nonlinear Gauß–Seidel–methods

Those methods extend the Gauß–Seidel–method to solve linear systems.

We want to solve $\Psi = 0$, where Ψ is given by

$$\begin{array}{rccl} \Psi : & \mathbb{R}^d & \longrightarrow & \mathbb{R}^d \\ & x & \longmapsto & \mathbf{F}x - p^{(i)}. \end{array}$$

If we write Ψ as

$$\Psi = (\psi_1, \ldots, \psi_d),$$

then it holds:

$$\begin{array}{rcl} \psi_1 & = & \psi_1(x_1) \\ \psi_2 & = & \psi_2(x_1, x_2) \\ \ldots\ldots\ldots\ldots\ldots\ldots & & \\ \psi_d & = & \psi_d(x_1, \ldots, x_d) \end{array}$$

By definition of the method:

Nonlinear Gauß–Seidel–method

given:	$\Psi = (\psi_1, \ldots, \psi_d)$	system
	$x^{[0]} = (x_1^{[0]}, \ldots, x_d^{[0]})$	starting vector
	ε	stopping condition
goal:	x^*	Solution of $\Psi(x) = 0$

Construct $x^{[0]}, x^{[1]}, \ldots, x^{[\nu]}$ where

$$\begin{array}{lll} x_1^{[k+1]} & \text{solves} & \psi_1(x_1^{[k+1]}, x_2^{[k]}, \ldots, x_d^{[k]}) = 0 \\ x_2^{[k+1]} & \text{solves} & \psi_2(x_1^{[k+1]}, x_2^{[k+1]}, x_3^{[k]}, \ldots, x_d^{[k]}) = 0 \\ \ldots & & \ldots \\ x_d^{[k+1]} & \text{solves} & \psi_d(x_1^{[k+1]}, \ldots, x_d^{[k+1]}) = 0 \end{array}$$

and ν minimal, such that

$$\| \Psi(x^{[\nu]}) \|_\infty < \varepsilon$$

Set $x^* := x_{[\nu]}$

we see, assuming the solution of the one-dimensional equations is found within the bounds of ε, that the stopping condition is fulfilled after one single iteration. The method uses the triangular structure of the transformation $\mathbf{F}$. So the problem is reduced from the solution of a nonlinear system with d unknowns to the solution of d nonlinear equations in one unknown.

6.2 The solution in one dimension

We examine the functions ψ_i by the help of (4.1):

Corollary 6.1 *Let $f \in \mathcal{F}_1([0,1]^d)$ and let the transformation $\mathbf{F}$ and $\Psi = (\psi_1, \ldots, \psi_d)$ be given as above. For given $k \in \{1, \ldots, d\}$ let $(a_1, \ldots, a_{k-1}) \in [0,1]^{k-1}$ be chosen arbitrarily but fixed. Then it holds for the restrictions ψ_k^a where*

$$\begin{array}{rcl} \psi_k^a : \mathbb{R} & \longrightarrow & \mathbb{R} \\ x_k & \longmapsto & \psi_k(a_1, \ldots, a_{k-1}, x_k) : \end{array}$$

(i) $\psi_k^a([0,1]) = [-p^{(i)}, 1 - p^{(i)}]$

(ii) ψ_k^a *is monotonically non decreasing.*

(iii) ψ_k^a *is continuous. Hence* ψ_k^a *has a zero in* $[0,1]$.

(iv) ψ_k^a *is continuously differentiable and*

$$\frac{d\psi_k^a}{dx_k} = \frac{\int_0^1 \cdots \int_0^1 f(a_1, \ldots, a_{k-1}, x_k, \xi_{k+1} \ldots, \xi_d)\, d\xi_d \cdots d\xi_{k+1}}{\int_0^1 \int_0^1 \cdots \int_0^1 f(a_1, \ldots, a_{k-1}, \xi_k, \ldots, \xi_d)\, d\xi_d \cdots d\xi_{k+1}\, d\xi_k}$$

(v) If $f > 0$, then ψ_k^a is strictly monotonically increasing and there is exactly one zero in $[0,1]$

Proof:

These properties are consequences of the properties of the transformation **F** (see 4.1).
□

Especially point (*iv*) is important: The ψ_k^a are differentiable if f is continuous. So we have the possibility to fulfil requirement (1) from section (5.1) in a wide range.

Because the derivatives of the ψ_k^a are given explicitly and the evaluation of those derivatives require an comparable amount of calculation as the evaluation of the functions themselves, a *Newton*–method can be used to find the zeroes. But we have to take into account that the functions are only defined on a bounded set and also the monoticity of the functions should be used.

Therefore we combine a Newton–method with a method of nested intervals.

6.3 The algorithm

Suppose a starting vector $(x_1^{[0]}, \ldots, x_d^{[0]})$ and $\Psi(x) = \mathbf{F}x - p^{(i)}$ are given. We look for the solution $x^* = x_1^*, \ldots, x_d^*$, i.e. x_k^* solves

$$\psi_k^{x^*}(x_k) = 0.$$

To do this, we construct in parallel a sequence $(x_k^{[j]})_{j \in \mathbb{N}}$ containing at least a convergent subsequence

$$x_k^{[j_\iota]} \longrightarrow x_k^*$$

as well as a sequence of intervals $([\,\ell^{[j]}, r^{[j]}\,])_{j \in \mathbb{N}}$ where

$$x_k^* \in [\,\ell^{[j]}, r^{[j]}\,] \; \forall j \in \mathbb{N}. \tag{5}$$

Because of the monoticity and the continuity, condition (5) is fulfilled, if

$$\psi_k^{x^*}(\ell^{[j]}) \leq 0 \text{ and } \psi_k^{x^*}(r^{[j]}) \geq 0.$$

In the beginning, we choose

$$\begin{array}{ll} [\,x_k^{[0]}, 1\,] & \text{if} \quad \psi_k^{x^*}(x_k^{[0]}) < 0 \\ [\,0, x_k^{[0]}\,] & \text{if} \quad \psi_k^{x^*}(x_k^{[0]}) > 0 \end{array}$$

If $\psi_k^{x^*}(x_k^{[0]}) = 0$, then $x_k^* = x_k^{[0]}$ is the solution.
This interval fulfils condition (5).

For each iteration step we calculate to $x_k^{[j]}$ the corresponding Newton iterate:

$$\xi_k^{[j]} := x_k^{[j]} - \frac{\psi_k^{x^*}(x_k^{[j]})}{\psi'^{x^*}_k(x_k^{[j]})}$$

if existing, i.e. if $\psi'^{x^*}_k(x_k^{[j]}) \neq 0$. We know that $\psi'^{x^*}_k \geq 0$. But $\psi'^{x^*}_k(x_k^{[j]})$ may be very small or equal to zero. In this case we have to calculate the new value of $x_k^{[j+1]}$ using another method. Furthermore we have to exclude the case of the sequence $(x_k^{[j]})_{j \in \mathbb{N}}$ getting to cycle. Therefore we introduce the following indicator–functions:

$$\begin{array}{lllll} \delta_j^+ & := & \delta^+(x_k^{[j]}) & := & \left\{ \begin{array}{lll} 1 & \text{if} & \psi_k^{x^*}(x_k^{[j]} > 0) \\ 0 & \text{if} & \psi_k^{x^*}(x_k^{[j]}) < 0 \end{array} \right. \\ \delta_j^- & := & \delta^-(x_k^{[j]}) & := & \left\{ \begin{array}{lll} 0 & \text{if} & \psi_k^{x^*}(x_k^{[j]}) > 0 \\ 1 & \text{if} & \psi_k^{x^*}(x_k^{[j]}) < 0 \end{array} \right. \end{array}$$

In addition we choose constants γ^* and γ_j, such that

$$0 < \gamma^* \leq \gamma_j < 1$$

and we set

$$x_k^{[j+1]} := \begin{cases} \xi_k^{[j]} \\ \text{if } \xi_k^{[j]} \in [\, \ell^{[j]} + \delta_j^+ \gamma_j |r^{[j]} - \ell^{[j]}|,\, r^{[j]} - \delta_j^- \gamma_j |r^{[j]} - \ell^{[j]}| \,] \\ \ell^{[j]} + |\psi_k^{x^*}(\ell^{[j]})| \frac{r^{[j]} - \ell^{[j]}}{\psi_k^{x^*}(r^{[j]}) - \psi_k^{x^*}(\ell^{[j]})} \qquad \text{else} \end{cases}$$

and

$$[\, \ell^{[j+1]}, r^{[j+1]} \,] := \begin{cases} [\, \ell^{[j]}, x_k^{[j+1]} \,] & \text{if} \quad \psi_k^{x^*}(x_k^{[j+1]}) > 0 \\ [\, x_k^{[j+1]}, r^{[j]} \,] & \text{if} \quad \psi_k^{x^*}(x_k^{[j+1]}) < 0 \end{cases}$$

This means if $\xi^{[j]}$ does not lie inside the search interval, we interpolate $x_k^{[j+1]}$ linearly between the old $x_k^{[j]}$ and the other boundary of the search interval (By definition $x_k^{[j]}$ is one of the boundaries.) Using this definition condition (5) holds.

For every $x_k^{[j]}$ we have to evaluate the function $\psi_k^{x^*}$ as well as its derivative. We see

$$\begin{aligned} \psi_k^{x^*}(x_k^{[j]}) &= \frac{\int_0^{x_k^{[j]}} \int_0^1 \cdots \int_0^1 f\left(x_1^*, \ldots, x_{k-1}^*, \xi_k, \ldots, \xi_d\right) d\xi_d \cdots d\xi_k}{\int_0^1 \int_0^1 \cdots \int_0^1 f\left(x_1^*, \ldots, x_{k-1}^*, \xi_k, \ldots, \xi_d\right) d\xi_d \cdots d\xi_{k+1}\, d\xi_k} - p_k^{(i)} \\ \left.\frac{d\psi_k^{x^*}}{dx_k}\right|_{x_k^{[j]}} &= \frac{\int_0^1 \cdots \int_0^1 f\left(x_1^*, \ldots, x_{k-1}^*, x_k^{[j]}, \xi_{k+1} \ldots, \xi_d\right) d\xi_d \cdots d\xi_{k+1}}{\int_0^1 \int_0^1 \cdots \int_0^1 f\left(x_1^*, \ldots, x_{k-1}^*, \xi_k, \ldots, \xi_d\right) d\xi_d \cdots d\xi_{k+1}\, d\xi_k} \end{aligned}$$

that the denominators are the same and they do not depend on $x_k^{[j]}$. Hence we evaluate the denominator in advance and only once. In each step we therefore have only to evaluate the nominators. Is x_k^* found, the nominator of $\left.\frac{d\psi_k^{x^*}}{dx_k}\right|_{x_k^*}$ and the denominator needed to find the zero of $\psi_{k+1}^{x^*}$ are the same, hence we have to calculate a denominator only in the first step. So only two numerical integrations are to be done in one iteration.

Remark: It is obvious that the method automatically normalizes the densities to be approximated: If $\tilde{f} = \kappa f$, then $\tilde{\psi}_k^{x^*} = \psi_k^{x^*}$ as well as $\tilde{\psi'}_k^{x^*} = \psi'^{x^*}_k$.

We will use the abbreviations:

$$\psi_k^{x^*}(x_k^{[j]}) = \frac{Z_k^{[j]}}{N_k} - p^{(i)} \text{ and } \left.\frac{d\psi_k^{x^*}}{dx_k}\right|_{x_k^{[j]}} = \frac{Z'^{[j]}_k}{N_k}$$

So we can sum up the algorithm as shown below:

Transformation using Gauß–Seidel–iteration

given:	$\Psi = (\psi_1, \ldots, \psi_d)$	system
	$x^{[0]} = (x_1^{[0]}, \ldots, x_d^{[0]})$	starting vector
	ε	stopping condition
goal:	x^*	solution of $\Psi(x) = 0$

1 Calculate N_1

2 For $k = 1, \ldots, d$ do

2a Calculate $Z_k^{[0]}$ and set $\ell^{[0]}$ and $r^{[0]}$ correspondingly. Set $j := 0$

2b Calculate $Z'^{[j]}_k$. If $Z'^{[j]} = 0$ go to 2e

2c Calculate $\xi_k^{[j]}$ and check whether

$$\xi_k^{[j]} \in [\, \ell^{[j]} + \delta_j^+ \gamma_j |r^{[j]} - \ell^{[j]}|, r^{[j]} - \delta_j^- \gamma_j |r^{[j]} - \ell^{[j]}| \,]$$

If not go to 2e

2d Set $x_k^{[j+1]} := \xi_k^{[j]}$.
Go to 2f

2e Set $x_k^{[j+1]} := \ell^{[j]} + |\psi_k^{x^*}(\ell^{[j]})| \frac{r^{[j]} - \ell^{[j]}}{\psi_k^{x^*}(r^{[j]}) - \psi_k^{x^*}(\ell^{[j]})}$

2f Set $\ell^{[j+1]}$ and $r^{[j+1]}$ correspondingly.

2g Calculate $Z^{[j+1]}$ and check whether

$$\left| \frac{Z_k^{[j+1]}}{N_k} - p^{(i)} \right| < \varepsilon$$

If not set $j := j + 1$ and go to 2b

3 Set $N_{k+1} := Z^{[j+1]}$
Set $x_k^* := x_k^{[j+1]}$ Increase $k := k + 1$. If $k \leq d$ go back to 2

4 $(x_1^*, \ldots, x_d^*)$ is the solution

6.4 The convergence of the method

Theorem 6.2 *Let $f \in \mathcal{F}_1([0,1]^d)$ and $p^{(i)} \in [0,1]^d$. Suppose $k \in \{1, \ldots, d\}$ and $(x_1^*, \ldots, x_{k-1}^*) \in [0,1]^{k-1}$ are given arbitrarily but fixed. Then the algorithm given in the section above stops for any given $\varepsilon > 0$ after finitely many steps with a solution*

$$|\psi_k^{x^*}(x_k)| = |\mathbf{F}(x_1^*, \ldots, x_{k-1}^*, x_k) - p_k^{(i)}| < \varepsilon$$

Proof:

We show, that the constructed sequence $(x_k^{[j]})_{j \in \mathbb{N}}$ includes a convergent subsequence.

We distinguish two cases:

Case 1 $r^{[j]} - \ell^{[j]} \longrightarrow 0$

It holds for any $j \in \mathbb{N}$:

$$\psi_k^{x^*}(\ell^{[j]}) \leq 0 \leq \psi_k^{x^*}(r^{[j]}),$$

Hence using that $\psi_k^{x^*}$, $\forall j \in \mathbb{N}$ are monotonic:

$$\ell^{[j]} \leq x^* \leq r^{[j]}$$

and therefore

$$\ell^{[j]} \longrightarrow x^* \text{ and } r^{[j]} \longrightarrow x^*.$$

Using

$$\ell^{[j]} \leq x_k^{[j]} \leq r^{[j]} \quad \forall j \in \mathbb{N}$$

we get

$$x_k^{[j]} \longrightarrow x^*$$

Case 2 $r^{[j]} - \ell^{[j]} \geq \zeta > 0 \; \forall j \in \mathbb{N}$

Let us choose ζ to be the limit of the sequence $(r^{[j]} - \ell^{[j]})_{j \in \mathbb{N}}$. Because $(r^{[j]} - \ell^{[j]})_{j \in \mathbb{N}}$ is a monotonically decreasing sequence bounded from below by 0 there exists a limit $\zeta > 0$.

We assume there is no subsequence $(x_k^{[j_i]})_{i \in \mathbb{N}}$, such that $(\psi_k^{x^*}(x_k^{[j_i]}))_{i \in \mathbb{N}} \longrightarrow 0$. Hence

$$|\psi_k^{x^*}(\ell^{[j]})| \not\longrightarrow 0 \text{ and } |\psi_k^{x^*}(r^{[j]})| \not\longrightarrow 0$$

These are monotonically decreasing sequences. Therefore there exists an $\alpha > 0$, which is the lower bound of $|\psi_k^{x^*}(x_k^{[j]})|$. We choose β as an upper bound of $\frac{d\psi_k^{x^*}}{dx_k}$.

We choose a j_0, such that for any $j > j_0$ it holds: $(r^{[j]} - \ell^{[j]}) < \zeta + \varepsilon$ where

$$\varepsilon := \min\left\{\frac{\alpha}{2\beta}, \zeta\gamma^*, \frac{\alpha\zeta}{1+\alpha}, \frac{\zeta\alpha}{2}\right\}$$

We examine the case that $x_k^{[j]} = \ell^{[j]}$ (The other case $x_k^{[j]} = r^{[j]}$ is to be treated analogously.)

We distinguish between two possibilities:
Case a: $x_k^{[j+1]} = \xi_k^{[j]}$: If it holds, that
Case a1: $\psi_k^{x^*}(x_k^{[j+1]}) < 0$
Then we have

$$\begin{aligned} r^{[j+1]} - \ell^{[j+1]} &= r^{[j]} - x^{[j+1]} \\ &= r^{j]} - \left(x[j] - \frac{\psi_k^{x^*}(x_k^{[j]})}{\psi'^{x^*}_k(x_k^{[j]})}\right) \\ &\leq r^{[j]} - \ell^{[j]} + \frac{-\alpha}{\beta} \\ &< \zeta + \varepsilon - \varepsilon \\ &\leq \zeta \end{aligned}$$

But this contradicts the assumption.

Case a2: $\psi_k^{x^*}(x_k^{[j+1]}) > 0$
Using the definition

$$x^{[j+1]} \leq r^{[j]} - \gamma_j(r^{[j]} - \ell^{[j]})$$

we get

$$\begin{aligned} r^{[j+1]} - \ell^{[j+1]} &= x_k^{[j+1]} - \ell^{[j]} \\ &\leq r^{[j]} - \gamma_j (r^{[j]} - \ell^{[j]}) - \ell^{[j]} \\ &< r^{[j]} - \gamma^*(\zeta + \varepsilon) - \ell^{[j]} \\ &\leq \left(1 - \gamma^* + \frac{\varepsilon}{\zeta}\right)\zeta \\ &\leq \left(1 - \gamma^* + \frac{\zeta\gamma^*}{\zeta}\right)\zeta \\ &= \zeta \end{aligned}$$

and we get again a contradiction. There remains

Case b: $x_k^{[j+1]} = \ell^{[j]} + |\psi_k^{x^*}(\ell^{[j]})| \frac{r^{[j]} - \ell^{[j]}}{\psi_k^{x^*}(r^{[j]}) - \psi_k^{x^*}(\ell^{[j]})}$
Case b1 $x^{[j+1]} < 0$:
Again we estimate the new search interval:

$$\begin{aligned} r^{[j+1]} - \ell^{[j+1]} &= r^{[j]} - x^{[j+1]} \\ &= r^{[j]} - \left(\ell^{[j]} + |\psi_k^{x^*}(\ell^{[j]})| \frac{r^{[j]} - \ell^{[j]}}{\psi_k^{x^*}(r^{[j]}) - \psi_k^{x^*}(\ell^{[j]})}\right) \end{aligned}$$

We estimate $|\psi_k^{x^*}(\ell^{[j]})|$ by α and $\psi_k^{x^*}(r^{[j]})$ by 1

$$\begin{aligned} &< \zeta + \varepsilon - \alpha \frac{\zeta}{1+\alpha} \\ &\leq \zeta \end{aligned}$$

So we have still to examine
Case b2: $x^{[j+1]} > 0$:
We denote by $\psi_\ell := |\psi_k^{x^*}(\ell^{[j]})|$

$$\begin{aligned} r^{[j+1]} - \ell^{[j+1]} &= x_k^{[j+1]} - \ell^{[j]} \\ &\leq \ell^{[j]} + \psi_\ell \frac{\zeta + \varepsilon}{\alpha + \psi_\ell} - \ell^{[j]} \\ &\leq \frac{\zeta + \varepsilon}{1 + \alpha} \quad \text{because } \frac{\psi_\ell}{\alpha + \psi_\ell} \leq \frac{1}{\alpha + 1} \\ &\leq \frac{\zeta + \frac{1}{2}\zeta\alpha}{1 + \alpha} \\ &< \zeta \frac{1 + \alpha}{1 + \alpha} = \zeta \end{aligned}$$

Hence the assumption, that there is no convergent subsequence, is false.

□

So it is shown that the method converges. If we choose the initial value $x_k^{[0]}$ close enough to the solution x_k^*, we will obtain a quadratic convergence because of the Newton–iteration. Especially if we do some refinement (i.e. some more iterations) this is of importance since in this case we have a high probability that the condition on the initial value is fulfilled.

Some remarks

Since we used all properties of the transformation **F** shown in Corollary (6.1), this method is able to fulfil the catalogue of requirements given in the beginning of chapter 5 in a wide range:

Theorem (6.2) gives requirement (1). But we have to also take into account the estimations of the discrepancy in chapters 4 and 5.

The other requirements are fulfilled by any iterative method. By the quadratic convergence in the neighbourhood of the solution, guaranteed by this method, it is especially useful to follow slowly changing f. Hence if fulfils requirement (5).

Furthermore we can choose the parameter ε in advance — resp. it can be adapted by some additional iterations on already constructed particles — that we construct sequences of points instead of sequences of ensembles.

The practical tests in the next chapter will show whether those theoretical advantages of the method hold as well if we use the method in practice.

7 Numerical tests

7.1 Numerical evaluation of the discrepancy in two dimensions

To test the method it is necessary to evaluate the discrepancy of the constructed ensembles. The number of num. integrations, necessary to calculate $\mu_f(\mathbf{R}_p)$, determine the efficiency of the method. Here we will introduce a method, which needs fewer integrations than conventional methods.

To calculate the f–discrepancy of $\delta_{\omega_{p_n}}$ it is not necessary to build the supremum

$$\sup_{\mathbf{R}\in\mathcal{R}([0,1]^d)} \left| \delta_{\omega_{p_n}}(\mathbf{R}) - \mu_f(\mathbf{R}) \right|$$

It suffices to take the maximum over a finite set:

Notation: Let $\omega_{p^{(n)}} \in \mathcal{D}_1([0,1]^d)$ be an equiweighted sequence of points where $p^{(i)} = (p_1^{(i)}, \ldots, p_d^{(i)}) i = 1, \ldots, n$.

(1) We call $\mathcal{K}(\omega_{p^{(n)}}) := \{(x_1, \ldots, x_d) \mid x_k \in \{p_k^{(i)} \mid i = 1, \ldots, n\} \text{ or } x_k = 1\}$ the set of all crossing points of $\omega_{p^{(n)}}$.

(2) For any $x = (x_1, \ldots, x_d) \in \mathcal{K}(\omega_{p^{(n)}})$ we denote by $\mathbf{Q}_x$ the closed box:

$$\mathbf{Q}_x := \{(\xi_1, \ldots, \xi_d \mid 0 \le \xi_i \le x_i,\ i = 1, \ldots, d\}$$

Remarks:

(a) The set of all crossing points also encloses the intersection with the boundary $x_i = 1$.

(b) If $f \in \mathcal{F}_1([0,1]^d)$ and $x \in \mathcal{K}(\omega_{p^{(n)}})$, it holds (because of the absolute continuity of μ_f)

$$\mu_f(\mathbf{Q}_x) = \mu_f(\mathbf{R}_x) \text{ but } \delta_{\omega_{p^{(n)}}}(\mathbf{Q}_x) \ge \delta_{\omega_{p^{(n)}}}(\mathbf{R}_x)$$

(c) The following result was stated by Niederreiter in [18] for the discrepancy of uniform distributed sequences.

Theorem 7.1 *Let $f \in \mathcal{F}_1([0,1]^d)$ and $\omega_{p^{(n)}} \in \mathcal{D}_1([0,1]^d)$ equiweighted sequence of points where $p^{(i)} = (p_1^{(i)}, \ldots, p_d^{(i)})$ for $i = 1, \ldots, n$. Then it holds*

$$\mathbf{D}_f(\omega_{p^{(n)}}) = \max_{x \in \mathcal{K}(\omega_{p^{(n)}})} \max \left\{ \left| \delta_{\omega_{p^{(n)}}}(\mathbf{Q}_x) - \mu_f(\mathbf{Q_X}) \right|, \left| \delta_{\omega_{p^{(n)}}}(\mathbf{R}_x) - \mu_f(\mathbf{R_X}) \right| \right\}$$

Proof:

"$\geq$" Let $x \in \mathcal{K}(\omega_{p^{(n)}})$ be given. Since $\mathcal{K}(\omega_{p^{(n)}}) \subseteq [0,1]^d$, we have $\mathbf{D}_f(\omega_{p^{(n)}}) \geq \left| \delta_{\omega_{p^{(n)}}}(\mathbf{R}_x) - \mu_f(\mathbf{R}_x) \right|$.
It remains to show that also $|\delta_{\omega_{p^{(n)}}}(\mathbf{Q}_x) - \mu_f(\mathbf{R}_x)|$ is smaller than $\mathbf{D}_f(\omega_{p^{(n)}})$. Since $\delta_{\omega_{p^{(n)}}}(\mathbf{Q}_x) \geq \delta_{\omega_{p^{(n)}}}(\mathbf{R}_x)$, we have only to consider the case, where $\delta_{\omega_{p^{(n)}}}(\mathbf{Q}_x) > \mu_f(\mathbf{R}_x)$.
We do not have to take care about the intersections with the boundary, because if $x_k = 1$ for some k we look at $\xi \in \mathcal{K}(\omega_{p^{(n)}})$ where $\xi_k = x_k$ if $x_k \neq 1$ and $\xi_k = max\{p_k^{(i)} \mid i = 1, \ldots, n\}$ else. Then $\delta_{\omega_{p^{(n)}}}(\mathbf{Q}_x) = \delta_{\omega_{p^{(n)}}}(\mathbf{Q}_\xi)$ and $\mu_f(\mathbf{R}_\xi) \geq \mu_f(\mathbf{R}_x)$ and hence $|\delta_{\omega_{p^{(n)}}}(\mathbf{Q}_\xi) - \mu_f(\mathbf{R}_\xi)| \geq |\delta_{\omega_{p^{(n)}}}(\mathbf{Q}_x) - \mu_f(\mathbf{R}_x)|$.
So we can choose x such that $x_k \neq 1$ for any k.
Let $((y_1^{(j)}, \ldots, y_d^{(j)}))_{j \in \mathbb{N}} \subseteq [0,1]^d$ be a sequence of points where

$$y_i^{(j)} \geq x_i \ \forall j \in \mathbb{N} \text{ and } y_i^{(j)} \longrightarrow x_i \text{ for } i = 1, \ldots, d$$

In this case it holds:

$$\delta_{\omega_{p^{(n)}}}(\mathbf{R}_{y^{(j)}}) = \delta_{\omega_{p^{(n)}}}(\mathbf{Q}_x) \text{ for almost all } j \in \mathbb{N}.$$

On the other hand we have:

$$\mu_f(\mathbf{R}_{y^{(j)}}) \longrightarrow \mu_f(\mathbf{R}_x) = \mu_f(\mathbf{Q}_x)$$

Altogether we get:

$$\begin{aligned} \mathbf{D}_f(\omega_{p^{(n)}}) &\geq \sup_{j \in \mathbb{N}} \left| \delta_{\omega_{p^{(n)}}}(\mathbf{R}_{y^{(j)}}) - \mu_f(\mathbf{R}_{y^{(j)}}) \right| \\ &\geq \left| \delta_{\omega_{p^{(n)}}}(\mathbf{Q}_x) - \mu_f(\mathbf{Q_X}) \right| \end{aligned}$$

"$\leq$" We assume, there is a $\xi = (\xi_1, \ldots, \xi_d) \in [0,1]^d)$, such that

$$\left| \delta_{\omega_{p^{(n)}}}(\mathbf{R}_\xi) - \mu_f(\mathbf{R}_\xi) \right| > \max_{x \in \mathcal{K}(\omega_{p^{(n)}})} \left| \delta_{\omega_{p^{(n)}}}(\mathbf{R}_x) - \mu_f(\mathbf{R}_x) \right|$$

We distinguish two cases:
Case 1 $\delta_{\omega_{p^{(n)}}}(\mathbf{R}_\xi) > \mu_f(\mathbf{R}_\xi)$

We can choose $\xi = (\xi_1, \ldots, \xi_d)$, such that $\xi_k \geq \min\{p_k^{(i)} \mid i = 1, \ldots, n\}$ $k = 1, \ldots, d$. Else we have $\delta_{\omega_{p^{(n)}}}(\mathbf{R}_\xi) = 0$ and hence $\mu_f(\mathbf{R}_\xi) = 0$. Then we have for any $x \in \mathcal{K}(\omega_{p^{(n)}})$

$$\left| \delta_{\omega_{p^{(n)}}}(\mathbf{R}_x) - \mu_f(\mathbf{R}_x) \right| \geq \left| \delta_{\omega_{p^{(n)}}}(\mathbf{R}_\xi) - \mu_f(\mathbf{R}_\xi) \right|$$

Look at $x = (x_1, \ldots, x_d) \in \mathcal{K}(\omega_{p^{(n)}})$ where

$$x_k = \max\{p_k^{(i)} < \xi_k\} \quad k = 1, \ldots, d$$

It holds

$$\delta_{\omega_{p^{(n)}}}(\mathbf{Q}_x) = \delta_{\omega_{p^{(n)}}}(\mathbf{R}_\xi),$$

since if there were a $p^{(i)} \in \mathbf{R}_\xi \setminus \mathbf{Q}_x$, than there would exist a $k \in \{1, \ldots, d\}$ where $p_k^{(i)} > x_k$ which contradicts the definition of x.
On the other hand we have

$$\mu_f(\mathbf{Q}_x) \leq \mu_f(\mathbf{R}_\xi),$$

since $\mathbf{Q}_x \subseteq \mathbf{R}_\xi$. Hence

$$\delta_{\omega_{p^{(n)}}}(\mathbf{Q}_x) - \mu_f(\mathbf{Q}_x) \geq \delta_{\omega_{p^{(n)}}}(\mathbf{R}_\xi) - \mu_f(\mathbf{R}_\xi)$$

in contradiction to the assumption. Therefore it only remains

Case 2 $\delta_{\omega_{p^{(n)}}}(\mathbf{R}_\xi) < \mu_f(\mathbf{R}_\xi)$

Here we must also take into account the intersections with the boundary.
We define our x now by

$$x_k = \min\{1, p_k^{(i)} > \xi_k\} \quad k = 1, \ldots, d$$

i.e. $x_k = 1$ if $\xi_k > p_k^{(i)} \ \forall i = 1, \ldots, n$.
An analogous deduction, now using $\mathbf{R}_x$, leads to

$$\delta_{\omega_{p^{(n)}}}(\mathbf{R}_x) = \delta_{\omega_{p^{(n)}}}(\mathbf{R}_\xi) \text{ and } \mu_f(\mathbf{R}_x) \geq \mu_f(\mathbf{R}_\xi)$$

and hence

$$\mu_f(\mathbf{R}_x) - \delta_{\omega_{p^{(n)}}}(\mathbf{R}_x) \geq \mu_f(\mathbf{R}_\xi) - \delta_{\omega_{p^{(n)}}}(\mathbf{R}_\xi)$$

again a contradiction to the assumption. Hence the theorem is proven.

□

Notation: We denote by

$$\mathbf{d}_f(x) := \max\left\{\left|\delta_{\omega_{p^{(n)}}}(\mathbf{Q}_x) - \mu_f(\mathbf{Q}_\mathbf{X})\right|, \left|\delta_{\omega_{p^{(n)}}}(\mathbf{R}_\mathbf{X}) - \mu_f(\mathbf{R}_\mathbf{X})\right|\right\}$$

the ***local discrepancy*** of $\delta_{\omega_{p^{(n)}}}$ and μ_f in x.

The theorem above shows, that we have to calculate the f–volume for at most $(n+1)^2$ boxes $\mathbf{R}$. This result is wellknown, but the algorithm which is introduced here shows that it is not necessary to take care about all $x \in \mathcal{K}(\omega_{p^{(n)}})$. Therefore we partition $\mathcal{K}(\omega_{p^{(n)}})$ using the number of $p^{(i)}$ which lie in $\mathbf{R}_x$ resp. in $\mathbf{Q}_x$ and we denote by

$$\begin{aligned}\mathcal{K}_\nu(\omega_{p^{(n)}}) &:= \{x \in \mathcal{K}(\omega_{p^{(n)}}) \mid \delta_{\omega_{p^{(n)}}}(\mathbf{R}_x) = \frac{\nu}{n}\} \\ &\cup \{x \in \mathcal{K}(\omega_{p^{(n)}}) \mid \delta_{\omega_{p^{(n)}}}(\mathbf{Q}_x) = \frac{\nu}{n}\} \quad \nu = 0, \ldots, n\end{aligned}$$

So we can write the discrepancy as:

$$\begin{aligned}\mathbf{D}_f(\omega_{p^{(n)}}) &= \max_{\nu=0,\ldots,n} \max_{x \in \mathcal{K}_\nu(\omega_{p^{(n)}})} \mathbf{d}_f(x) \\ &= \max_{\nu=0,\ldots,n} \max_{x \in \mathcal{K}_\nu(\omega_{p^{(n)}})} \left|\frac{\nu}{n} - \mu_f(\mathbf{R}_\mathbf{X})\right|\end{aligned}$$

and omitting the modulus

$$\mathbf{D}_f(\omega_{p^{(n)}}) = \max_{\nu=0,\ldots,n} \quad \max\{ \frac{\nu}{n} - \mu_f(\mathbf{R}_{x_\nu^{min}}), \mu_f(\mathbf{R}_{x_\nu^{max}}) - \frac{\nu}{n}\}$$

where

$$\begin{aligned} x_\nu^{min} &:= \arg \min_{x \in \mathcal{K}_\nu(\omega_{p^{(n)}})} \mu_f(\mathbf{R}_x) \\ x_\nu^{max} &:= \arg \max_{x \in \mathcal{K}_\nu(\omega_{p^{(n)}})} \mu_f(\mathbf{R}_x) \end{aligned}$$

The $x_\nu^{min/max}$ need not to be unique, it is enough to know one representative.

Remark: We do not need to calculate $x_0^{min} = (0,\ldots,0), x_n^{max} = (1,\ldots,1)$. Furthermore we have in these points:

$$\left| \delta_{\omega_{p^{(n)}}}(\mathbf{R}_{x_{0/n}^{min/max}}) - \mu_f(\mathbf{R}_{x_{0/n}^{min/max}}) \right| = 0$$

In two dimensions it is possible to reduce the number of candidates for $x_\nu^{min/max}$ as the following algorithm will show.

We sort the $(p^{(1)},\ldots,p^{(n)})$ using the first coordinate. The sorted points we call again $(p^{(1)},\ldots,p^{(n)})$. Hence we have $n+1$ stripes

$$[\, p_1^{(i)}, p_1^{(i+1)}\,] \times [\,0,1\,] \qquad i = 0,\ldots,n$$

Theorem 7.2 *Let $\nu \in \{1,\ldots,n\}$ be arbitrary but fixed and let $(p^{(1)},\ldots,p^{(n)})$ be sorted w.r.t. the first coordinate. Then there exist a x_ν^{min}, defined as above, with the following properties:*

(i) $x_\nu^{min} = \left(p_1^{(\xi)}, p_2^{(\eta)}\right)$

(ii) $\left(x_\nu^{min}\right)_1 \geq p_1^{(\nu)}$, *i.e.* $\xi \geq \nu$

(iii) If we sort $(p^{(1)},\ldots,p^{(\xi)})$ w.r.t. the second coordinate, such that

$$p_2^{(\sigma(1))} \leq p_2^{(\sigma(2))} \leq \cdots \leq p_2^{(\sigma(\xi))},$$

it holds

(1) $p_2^{(\eta)} = p_2^{(\sigma(\nu))}$

(2) $p_2^{(\xi)} \leq p_2^{(\sigma(\nu))}$

Proof:

ad (i) Clear, since $x_\nu^{min} \in \mathcal{K}_\nu(\omega_{p^{(n)}}) \subseteq \mathcal{K}(\omega_{p^{(n)}})$

ad (ii) For $(x_1, x_2) \in \mathcal{K}_\nu(\omega_{p^{(n)}})$ where $x_1 < p^{(\nu)}$ it holds:

$$\delta_{\omega_{p^{(n)}}}(\mathbf{R}_{(x_1,x_2)}) \leq \delta_{\omega_{p^{(n)}}}(\mathbf{Q}_{(x_1,x_2)}) < \nu$$

ad (iii) (1) $p_2^{(\eta)} = p_2^{(\sigma(\nu))}$

"$\geq$" If $(x_\nu^{min})_2 < p_2^{(\sigma(\nu))}$, then

$$\delta_{\omega_{p^{(n)}}}(\mathbf{R}_{x_\nu^{min}}) \leq \delta_{\omega_{p^{(n)}}}(\mathbf{Q}_{x_\nu^{min}}) < \nu$$

$$\text{hence} \quad x_\nu^{min} \notin \mathcal{K}_\nu(\omega_{p^{(n)}})$$

"$\leq$" If $(x_\nu^{min})_2 > p_2^{(\sigma(\nu))}$, then

$$\mathbf{R}_{(p_1^{(\xi)}, p_2^{(\sigma(\nu))})} \subseteq \mathbf{R}_{(p_1^{(\xi)}, (x_\nu^{min})_2)}$$

hence

$$\mu_f(\mathbf{R}_{(p_1^{(\xi)}, p_2^{(\sigma(\nu))})}) \leq \mu_f(\mathbf{R}_{(p_1^{(\xi)}, (x_\nu^{min})_2)})$$

Because of $\delta_{\omega_{p^{(n)}}}(\mathbf{Q}_{(p_1^{(\xi)}, p_2^{(\sigma(\nu))})}) = \nu$ we have $(p_1^{(\xi)}, p^{(\sigma(\nu))}) \in \mathcal{K}_\nu(\omega_{p^{(n)}})$

(2) $p_2^{(\xi)} \leq p_2^{(\sigma(\nu))}$
We assume $p_2^{(\xi)}$ to be bigger than $p_2^{(\sigma(\nu))}$. We look at the rectangle $\mathbf{Q}_{(p_1^{(\zeta)}, p_2^{(\eta)})}$ where $\zeta < \xi$ maximal, such that $p_2^{(\zeta)} \leq p_2^{(\sigma(\nu))}$. Using the assumption on $p_2^{(\xi)}$:

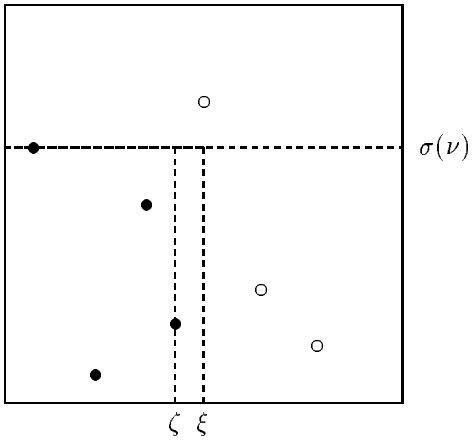
$$\delta_{\omega_{p^{(n)}}}\left(\mathbf{Q}_{(p_1^{(\xi)}, p_2^{(\eta)})} \setminus \mathbf{Q}_{(p_1^{(\zeta)}, p_2^{(\eta)})}\right) = 0$$

Hence $(p_1^{(\zeta)}, p_2^{(\eta)}) \in \mathcal{K}_\nu(\omega_{p^{(n)}})$. Furthermore $(p_1^{(\zeta)}, p_2^{(\eta)})$ is fulfilling the conditions $(i), \ldots, (iii)$. In addition it holds:

$$\mu_f\left(\mathbf{R}_{(p_1^{(\xi)}, p_2^{(\eta)})}\right) \geq \mu_f\left(\mathbf{Q}_{(p_1^{(\zeta)}, p_2^{(\eta)})}\right).$$

Therefore if $(p_1^{(\xi)}, p_2^{(\eta)})$ is a representant of x_ν^{min}, then this also holds for $(p_1^{(\zeta)}, p_2^{(\eta)})$.

Ex: $\nu = 4$, $\xi = 5$ and $p_2^{(5)} > p_2^{(\sigma(\nu))}$



□

If we want to calculate x_ν^{max} we have also to take into account the intersections with the boundary. To get an handier formulation of the following theorem, we introduce the additional points $p^{(0)} := (0, 1)$ and $p^{(n+1)} := (1, 0)$ giving them the weight 0.

Theorem 7.3 *Let $\nu \in \{0, \ldots, n-1\}$ be arbitrary but fixed. Let $(p^{(0)}, \ldots, p^{(n+1)})$ be sorted w.r.t. the first coordinate. Then there exists a x_ν^{max}, defined as above, having the following properties:*

(i) $x_\nu^{max} = \left(p_1^{(\Xi)}, p_2^{(\Upsilon)}\right)$

(ii) $\left(x_\nu^{min}\right)_1 \geq p_1^{(\nu+1)}$, *i.e.* $\Xi \geq \nu + 1$

(iii) If we sort $(p^{(0)}, \ldots, p^{(\Xi)})$ *w.r.t. the second coordinate, such that*

$$p_2^{(\sigma(0))} \leq p_2^{(\sigma(1))} \leq \cdots \leq p_2^{(\sigma(\Xi))},$$

it holds:

(1) $p_2^{(\Upsilon)} = p_2^{(\sigma(\nu+2))}$

(2) $p_2^{(\Xi)} \leq p_2^{(\sigma(\nu+1))}$

Proof:

ad (i) As in the latter theorem.

ad (ii) We know from the proof of the former theorem, that $\Xi \geq \nu$. But if $(p_1^{(\nu)}, p_2^{(i)}) \in \mathcal{K}_\nu(\delta_{\omega_{p^{(n)}}})$, then also $\delta_{\omega_{p^{(n)}}}(\mathbf{R}_{(p_1^{(\nu+1)},1)}) = \nu$. And furthermore we have, that $\mu_f(\mathbf{R}_{(p_1^{(\nu+1)},1)}) \geq \mu_f(\mathbf{R}_{(p_1^{(\nu)},p_2^{(i)})})$.

ad (iii) (1) Similar as in the latter proof:

"$\leq$" If $(x_\nu^{max})_2 > p_2^{(\sigma(\nu+2))}$, then

$$\delta_{\omega_{p^{(n)}}}(\mathbf{Q}_{x_\nu^{max}}) \geq \delta_{\omega_{p^{(n)}}}(\mathbf{R}_{x_\nu^{max}}) > \nu$$

$$\text{hence} \quad x_\nu^{max} \notin \mathcal{K}_\nu(\omega_{p^{(n)}})$$

"$\geq$" If $(x_\nu^{min})_2 < p_2^{(\sigma(\nu+2))}$, then we have

$$\mathbf{R}_{(p_1^{(\Xi)},p_2^{(\sigma(\nu+2))})} \supseteq \mathbf{R}_{(p_1^{(\Xi)},(x_\nu^{max})_2)}$$

hence

$$\mu_f(\mathbf{R}_{(p_1^{(\Xi)},p_2^{(\sigma(\nu+2))})}) \geq \mu_f(\mathbf{R}_{(p_1^{(\Xi)},(x_\nu^{max})_2)})$$

and in addition $\delta_{\omega_{p^{(n)}}}(\mathbf{R}_{(p_1^{(\Xi)},p_2^{(\sigma(\nu+2))})} = \nu$.

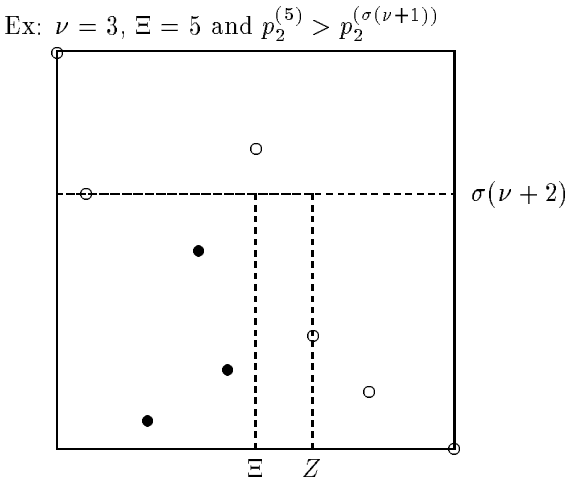
(2) If this case we look at the rectangle $\mathbf{R}_{(p_1^{(Z)},p_2^{(\Upsilon)})}$, where $Z > \Xi$ minimal, such that $p_2^{(Z)} \leq p_2^{(\sigma(\nu+1))}$. Because we introduced the artificial point $p^{(n+1)}$ where $p_2^{(n+1)} = 0$ there exists such a Z in any case. It holds

$$\delta_{\omega_{p^{(n)}}}\left(\mathbf{R}_{(p_1^{(Z)},p_2^{(\Upsilon)})}\right) = \delta_{\omega_{p^{(n)}}}\left(\mathbf{R}_{(p_1^{(\Xi)},p_2^{(\Upsilon)})}\right)$$

but

$$\mu_f\left(\mathbf{R}_{(p_1^{(Z)},p_2^{(\Upsilon)})}\right) \geq \mu_f\left(\mathbf{R}_{(p_1^{(\Xi)},p_2^{(\Upsilon)})}\right)$$

From this point we can follow the argumentation of the proof of the latter theorem.



□

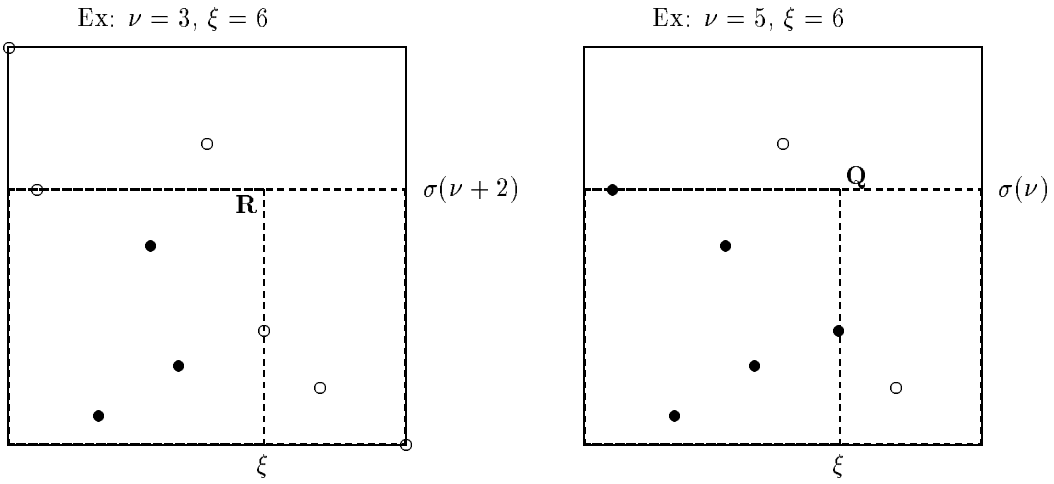
If we take a closer look on the conditions, we see that for the calculation of x_ν^{min} and $x_{\nu-1}^{max}$ the conditions (i),(ii),(iii)(2) to be proven are the same. The only difference appears in condition (iii)(1). Furthermore we need x_ν^{min} for $\nu = 1, \ldots, n$, whereas we need x_ν^{max} for $\nu = 0, \ldots, n-1$. Hence we calculate x_ν^{min} and $x_{\nu-1}^{max}$ in one step. But we hold in mind, that we may use the artificial points $p^{(0)}$ and $p^{(n+1)}$ only to calculate $x_{\nu-1}^{max}$.

Up to now we have only decreased the number of necessary evaluations of local discrepancies separately for x_ν^{min} and x_κ^{max}. But we can see from the following theorem, that many of these evaluations can be used for both using the fitting ν and κ.

Theorem 7.4 *Let ν be arbitrary but fixed. Assume that $\nu < \xi < n+1$ and $p_2^{(\xi)} \le p_2^{(\sigma(\nu))}$. Then the point $(p_1^{(\xi)}, p_2^{(\sigma(\nu+1))})$ is a candidate as well for $x_{\nu-1}^{max}$ as for $x_{\nu+1}^{min}$.*

Proof:

The assumptions are just the conditions for $x_{\nu-1}^{max}$. Furthermore $p_2^{(\xi)} \le p_2^{(\sigma(\nu))}$ implies $p_2^{(\xi)} \le p_2^{(\sigma(\nu+1))}$. Hence the point is a candidate for $x_{\nu+1}^{min}$, too. □



So for the evaluation of $\mathbf{d}_\nu^{min}$, we only have to calculate a μ_f–volume, if $p_2^{(\xi)} = p_2^{(\sigma(\nu))}$. Of course we have to pay for this savings by a slightly bigger amount of administration.

Calculation of the f–discrepancy in two dimensions

given:	μ_f	approximated measure
	$p^{(1)}, \ldots, p^{(n)} \in [0,1]^2$	sequence of points
goal:	$\mathbf{D}_f(\omega_{p^{(n)}})$	discrepancy

Introduce the artificial points $p^{(0)}$ and $p^{(n+1)}$ and resort the sequence, such that $p_1^{(0)} \leq p_1^{(1)} \leq \cdots \leq p_1^{(n+1)}$

For any $\nu = 1, \ldots, n$ do:

Sort $p^{(0)}, \ldots, p^{(\nu)}$ w.r.t the 2^{nd} coordinate into a list holding the ν smallest items: $p_2^{(\sigma(1))} \leq \cdots \leq p_2^{(\sigma(\nu))}$ and save $p^{max} := p_2^{(\sigma(\nu+1))}$

Calculate

$$\mathbf{d}_\nu^{min} := \mathbf{d}_f(p_1^{(\nu)}, p_2^{(\sigma(\nu))})$$
$$\mathbf{d}_{\nu-1}^{max} := \mathbf{d}_f(p_1^{(\nu)}, p^{max})$$

Check for any $\xi = (\nu+1), \ldots, (n+1)$, whether $p_2^{(\xi)} \leq p_2^{(\sigma(\nu))}$.
If this is true

Set $p^{max} := p_2^{(\sigma(\nu))}$ and insert $p_2^{(\xi)}$ into the list.

Set

$$\mathbf{d}_\nu^{min} := \max\{\mathbf{d}_\nu^{min}, \mathbf{d}_f(p_1^{(\nu)}, p_2^{(\sigma(\nu))})\} \text{ if } \xi < n+1$$
$$\mathbf{d}_{\nu-1}^{max} := \max\{\mathbf{d}_{\nu-1}^{max}, \mathbf{d}_f(p_1^{(\nu)}, p^{max})\}$$

Set $\mathbf{D}_f(\omega_{p^{(n)}}) = \max_{\nu=1,\ldots,n} \{\mathbf{d}_\nu^{min}, \mathbf{d}_{\nu-1}^{max}\}$

In order to hold the description of the algorithm easy to read, the consequences of the latter theorem are not implemented in the next table. (But they have been implemented in programs — the implementation is obvious).

Some remarks on the numerical implementation

(a) To handle the artificial point $p^{(0)} = (0, 1)$, we only have to initially set p^{max} to 1.

(b) In every step we only need the maximum of all $\mathbf{d}_\nu^{min}$ and $\mathbf{d}_{\nu-1}^{max}$ calculated so far.

(c) As mentioned above the case of the latter theorem is handled.

Theorem 7.5 *The algorithm needs at most* $\frac{n^2+5n}{2}$ *evaluations of* μ_f*–volumes to calculate* $\mathbf{D}_f(\omega_{p^{(n)}})$.

Proof:

The "worst case" is achieved if the condition on $p_2^{(\xi)}$ is always fulfilled. Providing this case the crossing points, where evaluations are necessary, build an upper triangle. We have

$\frac{n(n-1)}{2}$	used twice	the interior of the triangle
$2n$	used for $\mathbf{d}^{max}$	the boundary points
n	used for $\mathbf{d}^{min}$	the diagonal
$\frac{n^2+5n}{2}$	altogether	

□

It follows that in the worst case the behaviour is still quadratic, but numerical tests can show that for uniformly distributed sequences as input, we get a slightly subquadratic behaviour and we need in the case of 1024 particles less than one forth of integral–evaluations than in the naive evaluation.

7.2 Numerical tests in two dimensions

Using the method introduced in the section above, we are able to calculate and compare the discrepancies of several particle–ensembles.

We choosed μ_f, where

$$\begin{aligned} f : \mathbb{R}^2 &\longrightarrow \mathbb{R} \\ (x, y) &\longmapsto \sin 2\pi x \, 2\pi y + 1, \end{aligned}$$

to be approximated.

This f has two zeroes and furthermore it is quit flat in a neighbourhood of those zeroes ($\nabla f = 0$). So we can test the ability to approximate measures with small f.

We started with an uniform distributed sequence of particles, which was constructed by the method of Sobol (s.[26]), (p_n).

The starting section, consisting of n points had the following discrepancies:

n	$\mathbf{D}_{\chi_{[0,1]^2}}$
128	0,02515
256	0,01146
512	0,00841
1024	0,00430
2048	0,00247

The next table gives the discrepancies of the transformed sequences which we constructed using the method of Hlawka/Mück:

n	$\mathbf{D}_{\mu_f}$
128	0,10938
256	0,08905
512	0,08203
1024	0,07885
2048	0,07406

Because the method introduced in this thesis has the additional parameter ε, we constructed for fixed length n transformed sequences using different values of ε:

n	$\varepsilon = 10^{-1}$	$\varepsilon = 10^{-2}$	$\varepsilon = 10^{-3}$	$\varepsilon = 10^{-4}$	$\varepsilon = 10^{-5}$
128	0,06520	0,03724	0,03928	0,03888	0,03888
256	0,05739	0,02400	0,02289	0,02213	0,02213
512	0,05445	0,01725	0,01609	0,01542	0,01542
1024	0,05417	0,01476	0,01273	0,01273	0,01273

Hence in this example the results of the new method are much better than the results of the Hlawka/–Mück–method which has only the free choice of the number of particles.

In addition we see a value of $\varepsilon = 10^{-2}$ is usable in the full range from $n = 128$ to $n = 1024$. So it is not necessary to refine the transformation for increasing n in this range. Hence the effort in the number of numerical integrations we have to do is of order $\mathcal{O}(n)$ whereas this order is $\mathcal{O}(n^2)$ in the case of the Hlawka–/Mück–method.

7.3 Application to real densities in higher dimensions

Since we were able to reduce the problem of a d–dimensional non-linear system to the problem of d one–dimensional equations, the increase in the effort to solve a $(d+1)$–dimensional system is just the solution of one further one–dimensional equation. However we need in this case for the evaluation of ψ_1 $(d+1)$–dimensional numerical integrations. Even if we use particle methods to calculate them, the effort is increased considerably (see section 3.3). Furthermore we need in higher dimensions more transformed points to get a good approximation.

Another big difficulty is the testing of higher dimensional transformed sequences of points, since we can not evaluate e.g. the discrepancy because the effort would be to high. In our examples, Maxwell–distributions in the velocity resp. phase space

$$\begin{aligned} f &= c\, e^{-\frac{v^2}{2}} \qquad \text{resp.} \\ f &= c\, \rho(x)\, e^{-\frac{v^2}{2}}, \end{aligned}$$

we calculated moments in place of discrepancies. This is for several reasons problematic:

- The Variation in the sense of Hardy & Krause of the functions $\phi = |v|^k$ is not bounded.
- The zeroes of those functions ϕ take place, where f takes its maximum, where hence the most of the transformed particles are placed. Therefore we loose a lot of information if we calculate moments.

So the results can only give a rough estimation of the quality of the constructed sequences of points. We can just prove, whether the distribution "make somehow sense"

We started with uniformly distributed sequences, which were constructed following the method of I.M.Sobol. The first coordinate is the van-der-Corput sequences, the others are the $\mathbf{LP}_\tau$–sequences connected to the monocyclic operators $u_i + u_{i+1}$, $u_{i+2} + u_{i+1} + u_i$, $u_{i+3} + u_{i+1} + u_i$, $u_{i+3} + u_{i+2} + u_i$ and $u_{i+4} + u_{i+1} + u_i$.

First we approximated a three–dimensional Maxwell distribution in the velocity space:

$$\begin{aligned} f : \mathbb{R}^3 &\longrightarrow \mathbb{R}^3 \\ (v_1, v_2, v_3) &\longmapsto e^{\frac{v^2}{2}} \\ &\quad \text{where } v := \|v\|_2 \end{aligned}$$

To get a compact region, we substituted f by $\tilde{f} := f|_{[-2,5\,,\,2,5]^3}$. In this case also this cut–off has influence on the results, but we are only interested in qualitative statements. We distributed 4192 points and calculated the first moment — the expectation value for the modulus of the velocity:

$$\int\limits_{\mathbb{R}^3} v\, e^{\frac{v^2}{2}}\, dv = \frac{\pi}{2}$$

The approximation leaded to:

$\frac{\pi}{2}$			1,571
n	=	512	1,531
n	=	1024	1,535
n	=	2048	1,536
n	=	4192	1,537

The difficulty in approximating moments can especially be seen, if we want to calculate higher moments. For the approximation of

$$\int\limits_{\mathbb{R}^3} v^2\, e^{\frac{v^2}{2}}\, dv = \frac{3\pi\sqrt{\pi}}{2^{5/2}} \approx 2,953$$

using 4096 particles, we got a value of 2,732.

We also approximated Maxwell distributions in the four resp. six dimensional phase space:

$$\begin{aligned} f : \mathbb{R}^{2k} &\longrightarrow \mathbb{R}^{2k} \\ (x, v) &\longmapsto \rho(x) e^{\frac{v^2}{2}} \\ &\quad \text{for } k = 2, 3 \quad \rho(x) = \sum x_i \end{aligned}$$

We transformed in every case 1024 points and calculated again the first moments in velocity. The results are presented in the following table:

	$k = 2$	$k = 3$
Real value	0,98	1,57
Calculated value	1,05	1,63

Even though the calculation can not give that quantitative results (take also in consideration the naiv cut–off), we can state qualitatively, that the method is able to approximate densities also in higher dimensions. The increase in necessary computation time was reasonable. So the time we needed in average to construct one particle on a HP–Workstation A9000/710 was:

$d = 3$	$d = 4$	$d = 6$
0.20s	0.30s	0.58s

8 Conclusions

In this thesis, we developed, based on the transformation introduced by E.Hlawka and R.Mück, a method to construct sequences of points, which approximate some density f. especially we are now able to construct the desired simulations of the initial distribution of the rarefied gas around the shuttle, even if we use the more complicated distributions using the transport parameters λ and μ. Because we used the special properties of the Transformation, it is possible to do the construction in a quite effective manner.

The main–difficulty of all methods, that are based on the Hlawka/Mück–transformation, is caused by the fact that the transformation is a fracture of two integrals. Hence in higher dimensions we have to evaluate in general numerically multidimensional integrals. In our method we can reduce this, because we have to evaluate integrals in the highest dimension only for the solution of the first equation.

The method competes well on the theoretical side by fulfilling the requirements posed in the beginning of section 5.1 as well as on the practical side (see sections 7.2 and 7.3)

Because of its properties the method seems also to be very useful to approximate distributions, that vary in time. Here we have to take into consideration, that we have quadratic convergence in a neighbourhood of the solution. So the rearrangement in every time step only needs few iterations.

References

[1] **I.A.Antonov, V.M.Saleev**, *An Economic Method of Computing LP_τ–sequences* USSR – Computing Mathematics and Math. Physics **19/1**, pp. 252–256, 1979

[2] **H.Bauer**, *Einf. in die Wahrscheinlichkeitstheorie und Maßtheorie*, 3.Aufl., de Gruyter, 1978

[3] **H.Babovsky**, *A Convergence Proof for Nanbu's Boltzmann Simulation Scheme*, European Journal of Mechanics B/Fluids, **8**, no. 1,1989

[4] **H.Babovsky, R. Illner**, *A Convergence Proof for Nanbu's Simulation Method for the full Boltzmann Equation*, SIAM Journal of Numerical Analysis, **26/1**, pp. 45–65, 1989

[5] **G.Bird, J.Moss**, *Direct Simulation of Transitional Flow for Hypersonic Reentry Conditions*, AIAA no **84–0223**, 1984

[6] **R.H.Byrd, R.B.Schnabel, G.A.Shultz**, *Parallel Quasi–Newton Methods for Unconstraint Optimization*, Mathematical Programming **42**, pp. 273–306, 1988

[7] **W.C.Davidon**, *Optimally Conditioned Optimization Algorithms without Line Searches*, Mathematical Programming **9**, pp. 1–30, 1975

[8] **H.Faure**, *Discrèpance de suites associées à une système de numeration (en dimension s)*, Acta Arithmeticae **41**, pp. 337–251, 1982

[9] **F.Gropengießer, H.Neunzert, J.Struckmeier**, *Computational Methods for the Boltzmann Equation*, R.Sigler(ed.) "Applied and Industrial Mathematics", pp. 111–140, 1991

[10] **M.Hack**, *Ein iteratives Verfahren zur Transformation gleichverteilter Teilchen*, Diploma thesis, Kaiserslautern University, 1992

[11] **J.H.Halton**, *On the Efficientcy of Certain Quasi–Random Sequences of Points in Evaluating Multi–Dimensional Integrals*, Num. Math. **2**, pp. 84–90, 1960

[12] **E.Hlawka**, *Funktionen von beschränkter Variation in der Theorie der Gleichverteilung*, Annali di Matematica (Ser.IV) **54**, pp. 325–334, 1961

[13] **E.Hlawka, R.Mück**, *A Transformation of Equidistributed Sequences*, Zeremba (ed.) "Appl. of Numbertheory to Numerical Analysis", 1972

[14] **E.Hlawka, R.Mück**, *Über eine Transformation von gleichverteilten Folgen II*, Computing **9**, pp. 127–138 Springer 1972

[15] **L.Kuipers, H.Niederreiter**, *Uniform Distribution of Sequences*, John Wiley & Sons, 1974

[16] **K.Nanbu**, *Direct Simulation Schemes Derived from the Boltzmann Equation*, J.Phys. Japan **49**, p.2042, 1980

[17] **H.Neunzert**, *Particle Methods*, Script SS 1991, University Kaiserslautern

[18] **H. Niederreiter**, *Discrepancy and Convex Programming*, Ann. Mat.Pura Appl. (serie IV),**93**, pp.89–97, 1972

[19] **H. Niederreiter**, *Low –Discrepency Point–Sets*, Monatshefte für Mathematik **102**, pp. 155–167, 1986

[20] **H. Niederreiter**, *On the Existence of Uniformly Distributed Sequences in Compact Spaces*, Composito Mathematico **25**, pp.93–99, 1972

[21] **H. Niederreiter**, *Random Number Generation and Quasi–Monte Carlo Methods*, CBMS–NSF regional conference series in applied mathematics, **63**, SIAM 1992

[22] **G.Pagès, Y.J.Xiao,** *Sequences with Low Discrepency and Pseudo–Random Numbers: Theoretical Remarks and Numerical Tests*, Prépuplication **91/7**, Labaratoire de mathématiques et modèlisation, 1991

[23] **M.J.D. Powell**, *Some Global Convergence Properties of a Variable Metric Algorithm for Minimization without Exact Line–Search*, SIAM–AMS–Proceedings **9**, pp. 53–72, 1976

[24] **M.J.D. Powell**, *Updating Conjugate Directions by the BFGS–formula*, Mathematical Programming **38**, pp. 29–46, 1987

[25] **W.Rudin**, *Real and Complex Analysis*, 3^{rd} Edition, Mc Graw Hill, 1986

[26] **I.M.Sobol**, *Punkte, die einen mehrdimensionalen Würfel gleichmäßig ausfüllen*, Berichte der AG Technomathematik Nr.**51**, Kaiserslautern University, 1991