



KERNFORSCHUNGSANLAGE JÜLICH
GESELLSCHAFT MIT BESCHRÄNKTER HAFTUNG

Zentralinstitut für Angewandte Mathematik

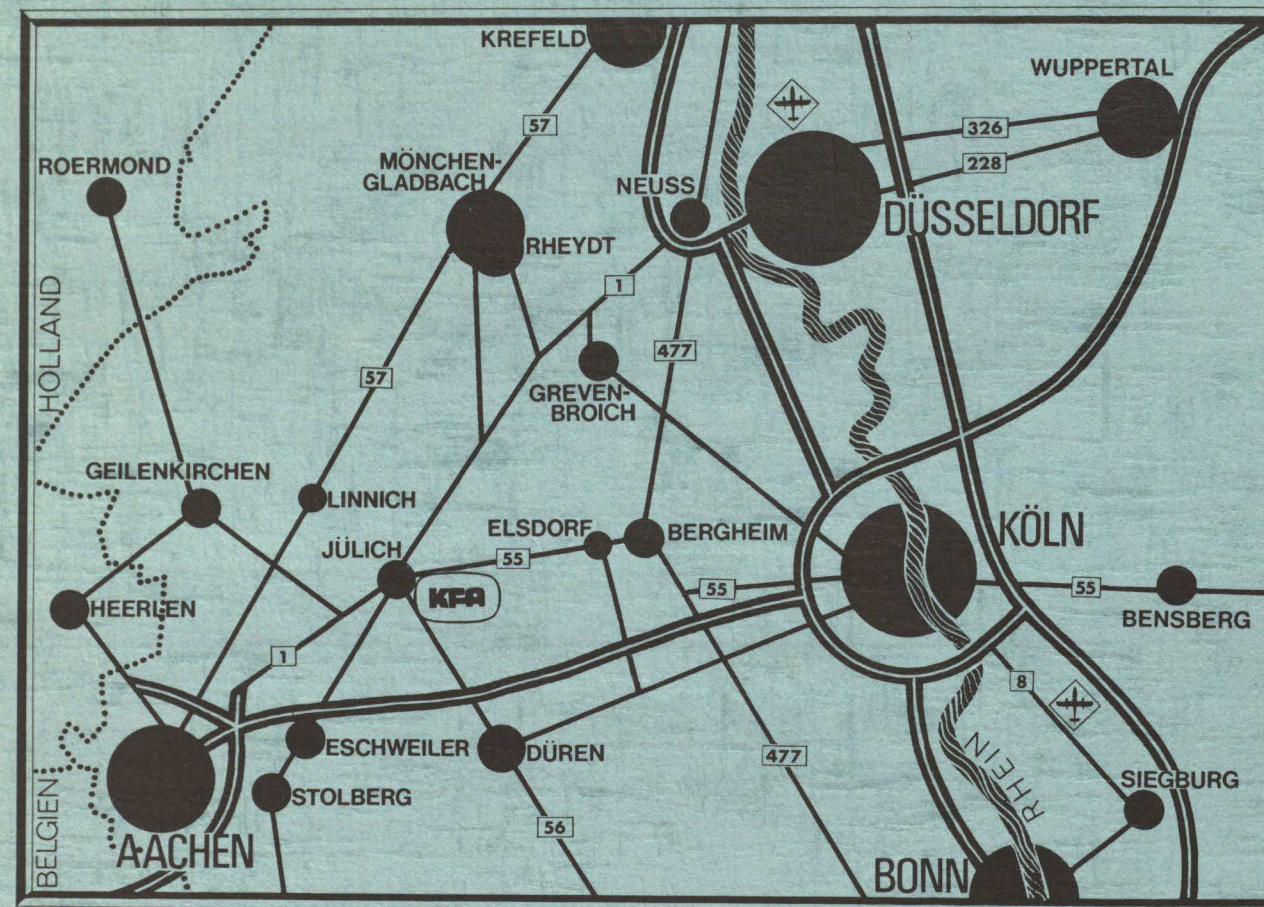
**Zur Anwendung der Approximation
durch endliche Punktmengen auf die Lösung
von Integro-Differentialgleichungen**

von

J. Wick

JÜI - 1124 - MA
Oktober 1974

Als Manuskript gedruckt



Berichte der Kernforschungsanlage Jülich – Nr. 1124
 Zentralinstitut für Angewandte Mathematik Jül – 1124 – MA

Dok.: Integro-Differential Equation - Numerical Solution
 Asymptotic Distribution
 Diophantic Approximation
 Vlasov Equation

Im Tausch zu beziehen durch: ZENTRALBIBLIOTHEK der Kernforschungsanlage Jülich GmbH,
 Jülich, Bundesrepublik Deutschland

**Zur Anwendung der Approximation
durch endliche Punktmengen auf die Lösung
von Integro-Differentialgleichungen**

von
J. Wick

Einleitung

Für die in [10] dargestellte Approximation nichtnegativer integrierbarer Funktionen durch endliche Punktmengen wurde in [12] gezeigt, daß sich damit konvergente Folgen bilden lassen, deren Grenzwerte die Lösungen von Integro-Differentialgleichungen vom Typ der Erhaltungsgleichungen sind. Dabei wird aus einer Punktmenge, die asymptotisch nach der Anfangsbedingung verteilt ist [8], für jeden späteren Zeitpunkt eine gleichmächtige Menge konstruiert, die asymptotisch nach der Lösung verteilt ist.

Da die betrachteten Gleichungen insbesondere in der kinetischen Gastheorie auftreten, läßt sich dort diese Art der Approximation als Darstellung des Gases durch n Teilchen auffassen. Diese Überlegung, Probleme der kinetischen Gastheorie durch das Studium des Verhaltens von Stichproben numerisch zu lösen, findet sich bereits bei Morse [in 1]. Die von Morse und anderen entwickelten Simulationsverfahren unterscheiden sich jedoch in zwei wesentlichen Punkten von der hier angegebenen Approximation. Einmal fehlt jegliche Konvergenzaussage, zum anderen gilt der Einwand, den Hlawka [4] bereits für Monte-Carlo-Integration machte: Eine Vergrößerung der Stichprobe und damit Rechenaufwandes führt nicht notwendig zu einer Verbesserung der Ergebnisse. In unmittelbarem Zusammenhang damit steht die Konstruktion der Anfangsapproximation (§ 2).

Die Approximation ist aber nicht auf Probleme der kinetischen Gastheorie beschränkt, wo sie diese anschauliche Interpretation gestattet und umgekehrt darüber Auskunft liefert, welche Funktionalgleichungen asymptotisch das Verhalten von Massenzuständen beschreiben. Sie läßt sich ebenso z.B. in gewissen Fällen der Wirbeltransportgleichung benutzen. Darüberhinaus ist die Beschränkung auf nichtnegative Funktionen nur aus Darstellungsgründen vorgenommen. Auf die beiden letztgenannten Probleme soll hier jedoch nicht weiter eingegangen werden.

In § 1 werden die theoretischen Grundlagen, wie sie in [10] und [12] dargestellt sind, zusammengestellt. In § 2 wird die Transformation der Anfangsbedingung auf die Gleichverteilung behandelt. In § 3 wird eine möglichst gute Approximation der Gleichverteilung gesucht und in § 4 numerische Ergebnisse am Beispiel der Vlasov-Gleichung diskutiert.

1. Theoretische Grundlagen

Zunächst soll die Approximation nichtnegativer integrierbarer Funktionen über dem R^m durch endliche Punktmengen kurz definiert werden. Eine breitere Darstellung findet sich in [10]. Dazu bezeichnen wir mit X_n den Raum aller n -punktigen Mengen des R^m

$$X_n := \{(P_1, \dots, P_n) : P_i \in R^m, i=1, \dots, n\}.$$

Mit Hilfe des Rechtecks $R(Q)$ aus R^m

$$R(Q) := \{P : P = (p_1, \dots, p_m) \in R^m, p_i < q_i, i=1, \dots, n\}$$

und seiner charakteristischen Funktion $K_{R(Q)}$

$$K_{R(Q)}(P) = \begin{cases} 1 & \text{für } P \in R(Q) \\ 0 & \text{für } P \notin R(Q) \end{cases}$$

definieren wir über X_n die Funktion

$$\lambda(\omega_n, R(Q)) := \frac{1}{n} \sum_{i=1}^n K_{R(Q)}(P_i) \quad \text{mit } \omega_n = (P_1, \dots, P_n) \in X_n.$$

Wir merken hier schon an, daß durch

$$r(\omega_n, \hat{\omega}_n) := \sup_{Q \in R^m} |\lambda(\omega_n, R(Q)) - \lambda(\hat{\omega}_n, R(Q))|$$

für $\omega_n, \hat{\omega}_n \in X_n$ in X_n eine Metrik definiert wird.

Definition: Sei $f(P)$ eine nichtnegative integrierbare Funktion über R^m

$$\text{und } \alpha = \int_{R^m} f(P) dP.$$

Das Paar (α, ω) , wobei $\omega = \{\omega_n \in X_n, n=1, 2, \dots\}$ sei, ist eine Approximation von f , wenn für alle $Q \in R^m$ gilt

$$(1.1) \quad \lim_{n \rightarrow \infty} \alpha \cdot \lambda(\omega_n, R(Q)) = \int_{R(Q)} f(P) dP.$$

Da sich jede integrierbare Funktion als Differenz von nichtnegativen integrierbaren Funktionen schreiben läßt, ist die Erweiterung der Definition durch $(\alpha, \omega^1; \beta, \omega^2)$ offensichtlich. Wir werden uns im weiteren auf den in der Definition genannten Fall beschränken, und dann genügt es, Funktionen zu betrachten, deren Integral Eins ist. Deshalb setzen wir

$$X_0 := \left\{ f(P) : f(P) \geq 0, f(P) \in L_1(R^m), \int_{R^m} f(P) dP = 1 \right\}$$

und schreiben für Folgen $\{\omega_n\}$, die $f \in X_0$ im Sinne der Definition approximieren, kurz

$$\lim_{n \rightarrow \infty} \omega_n = f$$

Die Approximation hat folgende Eigenschaften

a) Zu jedem $f \in X_0$ existiert eine Folge $\{\omega_n\}$ mit $\lim_{n \rightarrow \infty} \omega_n = f$.

b) Gilt für alle $Q \in R^m$ für eine Folge $\{\omega_n\}$ und ein $f \in X_0$ die Relation (1.1), so gilt sie gleichmäßig (Satz von Weyl).

Diese Eigenschaft b) gibt die Möglichkeit, ein Maß für die Güte der Approximation, die Diskrepanz, einzuführen,

$$D(\omega_n, f) = \sup_{Q \in R^m} \left| \lambda(\omega_n, R(Q)) - \int_{R(Q)} f(P) dP \right|.$$

c) Für die Funktion $g(P)$ seien die n -dimensionale Vitali-Variation $V^n(g)$ und alle k -dimensionalen Vitali-Variationen ($k=1, 2, \dots, m-1$) von g beschränkt, wobei $(m-k)$ Komponenten von P gleich ∞ gesetzt sind. Bezeichnen wir die Summe aller dieser Variationen mit $V(g)$, dann gilt (Satz von Koksma-Hlawka)

$$\left| \frac{1}{n} \sum_{i=1}^n g(P_i) - \int_{R^m} f(P) g(P) dP \right| < V(g) D(\omega_n, f).$$

Das Ziel ist es nun, aus einer Approximation $\{\omega_n^0\}$ der Anfangsbedingung $f_0(P)$ für jeden Zeitpunkt t eine Approximation $\{\omega_n(t)\}$ der Lösung einer Gleichung vom Typ

$$\frac{\partial f}{\partial t} + \operatorname{div}_P (f \cdot F[f]) = 0$$

zu konstruieren. Dazu beschreiben wir zunächst die Gleichung genauer.

Es sei ein auf $R^m \times R^m$ definiertes Vektorfeld

$$G(P, Q) := (G_1(P, Q), \dots, G_m(P, Q))$$

gegeben, das folgende Voraussetzungen erfüllt

$G_i(P, Q)$, $i=1, \dots, m$ sind auf $\mathbb{R}^m \times \mathbb{R}^m$ beschränkt und für festes P bzgl. Q Lebesgue - meßbar; darüberhinaus gibt es eine von Q unabhängige Konstante L_G derart, daß

$$(1.2) \quad ||G(P_1, Q) - G(P_2, Q)|| < L_G ||P_1 - P_2|| \quad \text{für alle } P_1, P_2, Q \in \mathbb{R}^m \text{ gilt.}$$

Dabei bezeichnen wir mit $||\cdot||$ die als Maximum der Komponentenbeträge definierte Vektornorm.

Weiterhin sei $T > 0$ und $J := [0, T)$.

Neben G sei ein auf $J \times \mathbb{R}^m$ definiertes, stetiges und beschränktes Vektorfeld $H(t, P)$ gegeben, das ebenfalls einer globalen Lipschitzbedingung mit von t unabhängiger Lipschitzkonstanten L_H genügt.

Wir betrachten nun für eine in $J \times \mathbb{R}^m$ definierte Funktion $f(t, P)$ das Anfangswertproblem

$$(A) \quad \begin{aligned} \frac{\partial f}{\partial t} + \operatorname{div}_P (f \cdot F[f]) &= 0, \\ f(0, P) &= f_0(P), \end{aligned}$$

wobei

$$(1.3) \quad F_i[f](t, P) := \int_{\mathbb{R}^m} G_i(P, Q) f(t, Q) dQ + H_i(t, P), \quad i=1, \dots, m.$$

Wie üblich definieren wir eine Funktion $f(t, P)$ als schwache Lösung des Problems (A), wenn für alle im $J \times \mathbb{R}^m$ beliebig oft differenzierbaren Funktionen, die außerhalb einer Menge $[0, \tau) \times \{P: ||P|| < R\}$ mit $\tau < T$, $R < \infty$ verschwinden, gilt

$$\int_0^T \int_{\mathbb{R}^m} f \{w_t + \langle F[f], \operatorname{grad}_P w \rangle\} dP dt + \int_{\mathbb{R}^m} w(0, P) f_0(P) dP = 0.$$

Für die Anfangsbedingung gelte $f_0 \in X_0$.

Die schwache Lösung f von (A) soll dann in der folgenden, mit X bezeichneten Klasse von auf $J \times \mathbb{R}^m$ definierten Funktionen $f(t, P)$ liegen, die die Eigenschaften besitzen:

- a) für alle $t \in J$ gilt: $f(t, \cdot) \in X_0$,
- b) $f(t, P)$ ist für fast alle P eine in J stetige Funktion von t , und es existiert eine (von f abhängige) Funktion $h \in L^1(\mathbb{R}^m)$ mit

$$f(t, P) < h(P) \quad \text{für alle } t \in J, P \in \mathbb{R}^m.$$

Wir fordern, daß für das Problem (A) eine so definierte schwache Lösung existiert, und werden diese im weiteren kurz als Lösung bezeichnen.

Zur Beschreibung der Approximation dieser Lösung nehmen wir zunächst an, daß für $\omega_n(t) = (Q_1(t), \dots, Q_n(t))$ gilt

$$\lim_{n \rightarrow \infty} \omega_n(t) = f(t, \cdot),$$

und betrachten das System

$$(1.4) \quad \dot{P} = \frac{1}{n} \sum_{k=1}^n G(P, Q_k(t)) + H(t, P)$$

Für alle Anfangswerte $(P_1^0, \dots, P_n^0) = \hat{\omega}_n^0$ mit $\lim_{n \rightarrow \infty} \hat{\omega}_n^0 = f_0$ gilt für die Lösung $\hat{\omega}_n(t)$ von (1.4) [13]

$$\lim_{n \rightarrow \infty} \hat{\omega}_n(t) = f(t, \cdot).$$

Man wird daher erwarten, daß die Lösung des Anfangswertproblems

$$(B) \quad \begin{aligned} \dot{P}_i &= \frac{1}{n} \sum_{k=1}^n G(P_i, P_k) + H(t, P_i) \\ P_i(0) &= P_i^0 \end{aligned} \quad i=1, \dots, n$$

mit $(P_1^0, \dots, P_n^0) = \omega_n^0$ und $\lim_{n \rightarrow \infty} \omega_n^0 = f_0$ eine Approximation der Lösung von (A) ist.

Zunächst formulieren wir die Voraussetzungen, unter denen diese Annahme richtig ist.

Voraussetzung 1: Es existiert ein $h \in L^1(\mathbb{R}^m)$ mit $f(t, P) < h(P)$ für alle $t \in J, P \in \mathbb{R}^m$, für die die Funktion

$$h^*(P) = \int_{R(P)} h(Q) d(Q)$$

global Lipschitzstetig ist.

Voraussetzung 2: Das Anfangswertproblem (B) besitzt mindestens eine Lösung $\omega_n(t)$, $t \in J$.

Voraussetzung 3: Es sei $\omega_n(t)$ eine Lösung von (B) und $\hat{\omega}_n(t) = (\hat{P}_i(t))$ die Lösung von

$$\dot{\hat{P}}_i(t) = F[f](t, \hat{P}_i) , \quad \hat{P}_i(0) = P_i^0 .$$

Dann sollen die Anfangswertprobleme

$$\dot{P} = \frac{1}{n} \sum_{i=1}^n G(P, P_i(t)) + H(t, P)$$

$$\dot{P} = \frac{1}{n} \sum_{i=1}^n G(P, \hat{P}_i(t)) + H(t, P)$$

$$P(t) \Big|_{t=s} = Q$$

für alle $s \in J$, $Q \in \mathbb{R}^m$ in J eindeutige Lösungen besitzen.

Voraussetzung 4: Für festes $P \in \mathbb{R}^m$ sei die Funktion $G(P, \cdot)$ von beschränkter Hardy-Krause-Variation, und die Variation ist für alle P gleichmäßig beschränkt.

Satz 1: Sind die Voraussetzungen 1 - 4 erfüllt und existiert für (A) eine Lösung, so gilt für die Lösungen von (A) und (B)

$$\lim_{n \rightarrow \infty} \omega_n(t) = f(t, \cdot) ,$$

sofern für die Anfangswerte

$$\lim_{n \rightarrow \infty} \omega_n^0 = f_0$$

erfüllt ist.

In [12] ist der Satz für $m = 2$ bewiesen, doch läßt sich der Beweis ohne Schwierigkeiten auf allgemeine Dimensionen übertragen.

In einer Reihe wichtiger Fälle, z.B. der Vlasov-Gleichung sind jedoch die Voraussetzungen an $G(P,Q)$ nicht erfüllt. Um auch hier den Satz anwenden zu können, modifiziert man die Ausgangsgleichung in folgender Weise:

Es sei $\phi(h,r)$ eine für $h > 0$, $r \geq 0$ definierte Funktion mit den Eigenschaften $\phi(h,r) \geq 0$, $\phi(h,r) = 0$ für $r \geq h$ und $\int_{R^m} \phi(h, ||P-Q||) dQ = 1$.

Ersetzt man nun $G(P,Q)$ durch

$$(1.5) \quad \tilde{G}(P,Q) = \int_{R^m} G(P,Q') \phi(h, ||Q-Q'||) dQ' ,$$

so erfüllt $\tilde{G}$ bei geeigneten zusätzlichen Bedingungen (z.B. Glattheit) an ϕ die geforderten Bedingungen.

Es ist interessant, daß ein globaler Existenzsatz für die Vlasovgleichung in mehr als zwei Dimensionen auch diese Modifikation an G erfordert [vgl. 3].

Das System (B) hat dann die Gestalt

$$(B') \quad \dot{P}_i = \frac{1}{n} \sum_{k=1}^n \int_{R^m} G(P_i, Q) \phi(h, |P_k - Q|) dQ + H(t, P_i)$$

2. Transformation der Anfangsbedingung

Zur numerischen Anwendung des Satzes benötigt man zunächst eine geeignete Anfangsapproximation. In [10] ist zum Beweis des Weylschen Satzes zu einer Funktion $f \in X_0$ eine Folge $\{\omega_n\}$ konstruiert worden, für die gilt

$$D(\omega_n, f) < c \cdot (1/n)^{1/m}.$$

Es stellt sich die Frage, ob es zu gegebenem f andere Folgen $\{\omega_n\}$ gibt, deren Diskrepanz eine schneller gegen Null strebende Abschätzung gestattet.

Wir beschränken uns auf den Fall, daß der Träger $\bar{B}$ von f konvex und f im Innern des Trägers (welches wir mit B bezeichnen) stetig und positiv ist. Durch

$$F_k(p_1, \dots, p_k) = \frac{\int_{-\infty}^{p_k} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} f(p_1, \dots, p_{k-1}, x_n, \dots, x_m) dx_m \dots dx_k}{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} f(p_1, \dots, p_{k-1}, x_k, \dots, x_m) dx_m \dots dx_k} \quad k=1, 2, \dots, m$$

wird eine eindeutige Abbildung $F = (F_1, \dots, F_m)$ von B auf den offenen m -dimensionalen Einheitswürfel E definiert, und die Funktionaldeterminante der Abbildung F ist $f(p_1, \dots, p_m)$ [vgl. 6]. Die Umkehrabbildung $F^{-1} = \phi$ von E auf B ist daher ebenfalls eindeutig mit der Funktionaldeterminante $\Delta = 1/f$. Wegen $\phi(E) = B$ gilt daher für beliebige integrierbare Funktionen g

$$\int_B g(P) f(P) dP = \int_E g(\phi(P)) f(\phi(P)) \Delta dP = \int_E g(\phi(P)) dP.$$

Sei nun $\hat{\omega}_n = (P_1, \dots, P_n)$ eine beliebige Punktmenge aus E , so gilt nach dem Satz von Koksma-Hlaska

$$\begin{aligned} (2.1) \quad \left| \frac{1}{n} \sum_{i=1}^n K_{R(Q)}(\phi(P_i)) - \int_{R(Q)} f(P) dP \right| &= \left| \frac{1}{n} \sum_{i=1}^n K_{R(Q)}(\phi(P_i)) - \int_B K_{R(Q)}(P) f(P) dP \right| \\ &= \left| \frac{1}{n} \sum_{i=1}^n K_{R(Q)}(\phi(P_i)) - \int_E K_{R(Q)}(\phi(P)) dP \right| \\ &\leq V(\phi) D(\hat{\omega}_n, K_E(\cdot)) = V(\phi) D(\hat{\omega}_n) \end{aligned}$$

wobei $V(\phi)$ die Summe aus allen k -dimensionalen Vitali-Variationen

$(k=1, \dots, m)$ ist und $(m-k)$ Komponenten des Argumentes von ϕ gleich Eins gesetzt sind.

Aus (2.1) folgt für

$$(2.2) \quad \omega_n = (\phi(P_1), \dots, \phi(P_n))$$

$$(2.3) \quad D(\omega_n, f) \leq V(\phi) D(\hat{\omega}_n) .$$

Damit ist die Frage nach der Konstruktion einer Approximation einer stetigen Funktion $f \in X_0$, deren Träger konvex und im Innern von Null verschieden ist, auf die Aufgabe zurückgeführt, eine modulo Eins gleichverteilte Folge zu finden.

Dieses Problem ist in vielfältiger Gestalt in der Zahlentheorie behandelt worden. Obgleich ich die zugehörige Literatur nicht vollständig kenne, scheinen die theoretischen Untersuchungen dort stets unter der Nebenbedingung geführt worden zu sein, daß ω_{n+1} die Elemente von ω_n enthält. Diese Bedingung ist aber für die hier beabsichtigten Anwendungen nicht nötig.

Für Folgen ω_n mit $\omega_n \subset \omega_{n+1}$ folgt aus einem Satz von Roth [15], daß mit $c > 0$ gilt

$$(2.4) \quad D(\omega_n) \geq c \frac{(\log n)^{\frac{m-1}{2}}}{n} .$$

Ob diese Aussage für die hier betrachteten allgemeineren Folgen ebenfalls gültig ist, kann bisher nicht gesagt werden.

Bevor wir konkrete gleichverteilte Folgen angeben, soll zunächst die Ungleichung (2.3) verbessert werden. Dazu benötigen wir, daß sich die Diskrepanz als Maximum endlich vieler Werte bestimmen läßt [9].

Wir betrachten zu gegebenem $\omega_n = (P_1, \dots, P_n)$ die Menge $G(\omega_n)$ aller Schnittpunkte von je m orthogonalen Hyperebenen

$$H_{ij} := \left\{ Q = (q_1, \dots, q_m) \in R^m : q_j = p_{ij}, P_i = (p_{i1}, \dots, p_{im}) \right\} \begin{matrix} i=1, \dots, n; \\ j=1, \dots, m \end{matrix}$$

$G(\omega_n)$ werde kurz das von ω_n erzeugte Gitter genannt. Bezeichnen wir ferner mit $\overline{R(Q)}$ den Abschluß von $R(Q)$, so gilt

(2.5)

$$D(\omega_n, f) = \max_{Q \in G(\omega_n)} \left(\max(|\lambda(\omega_n, R(Q)) - \int_{R(Q)} f(P) dP|, |\lambda(\omega_n, \overline{R(Q)}) - \int_{R(Q)} f(P) dP|) \right)$$

Nehmen wir für $f(P)$ zusätzlich an, daß gilt

$$f(P) = f_1(p_1) \dots f_m(p_m),$$

so ist

$$F_k(p_1, \dots, p_k) = F_k(p_k) = \frac{\int_{-\infty}^{p_k} f_k(x_k) dx_k}{\int_{-\infty}^{\infty} f_k(x_k) dx_k}.$$

Durch die Umkehrabbildung $\phi = F^{-1}$ wird daher das von $\hat{\omega}_n = (P_1, \dots, P_n)$ in E erzeugte Gitter auf das von $\omega_n = (\bar{\phi}(P_1), \dots, \bar{\phi}(P_n))$ erzeugte Gitter abgebildet. Daraus folgt unmittelbar

$$D(\omega_n, f) = D(\hat{\omega}_n).$$

3. Gleichverteilung modulo Eins

Für reelle a bezeichne $[a]$ die größte ganze Zahl, die kleiner oder gleich a ist.

Dann setzen wir

$$a - [a] = a \bmod 1$$

Hlawka [5] und, unabhängig von ihm, Korobow haben gezeigt, daß für

$\omega_n = (p_1, \dots, p_n)$ mit

$$p_i = \left(\frac{ig_1}{n}, \dots, \frac{ig_m}{n} \right) \bmod 1; \quad i=1, \dots, m$$

gilt

$$D(\omega_n) < c \frac{(\log n)^{\beta n}}{n}.$$

Dabei sind $g_1, \dots, g_m = G$ ganze Zahlen und $\beta = \beta(G) > 0$.

Das Problem ist, G so zu bestimmen, daß β möglichst klein ausfällt, einen sogenannten "guten Gitterpunkt" zu finden. Hierzu schreibt Zaremba in [17]

"In more than two dimensions, no recipe is known for lattice points G and there seems to be no reason to believe that such a recipe could be found. One has to find such lattice points by a laborious procedure of trial and error, or more efficiently, but still quite laboriously, by eliminating "bad" lattice points. There are, however, existence theorems which give us at least the assurance that the search for good lattice points will not be in vain."

Ein anderes Hilfsmittel, ω_n mit kleiner Diskrepanz zu suchen, bietet die Relation (2.5), da zu einer gegebenen Punktmenge die Diskrepanz direkt berechnet werden kann. Zumindest läßt sich auf die Weise klären, ob eine Vergrößerung der Punktmenge eine kleinere Diskrepanz liefert, so daß man entscheiden kann, ob der Aufwand für die eigentliche numerische Rechnung vertretbar ist.

Betrachten wir nun den Fall $m=2$. Sei u_k , $k=1,2,\dots$ die Folge der Fibonacci-Zahlen, definiert durch

$$u_1 = u_2 = 1, \quad u_{k+2} = u_{k+1} + u_k \quad k=1,2,\dots$$

Für $n = u_k$ gilt nach Zaremba [16] für die Diskrepanz von $\omega_n = (p_1, \dots, p_n)$ mit

$$(3.1) \quad p_i = \left(\frac{i}{u_k}, \frac{i u_{k-1}}{u_k} \right) \bmod 1 \quad i=0,1,\dots,u_k-1$$

die Abschätzung

$$D(\omega_n) < \frac{7}{6u_k} \log(6u_k) + 1/u_k.$$

Da in ω_n der Punkt $(0,0)$ enthalten ist, läßt sich diese Punktmenge nicht auf unendlich ausgedehnte Gebiete nach (2.2) transformieren. Wir setzen daher

$$(3.2) \quad p_i = \left(\frac{2i-1}{2u_k}, \frac{2i u_{k-1}-1}{2u_k} \right) \bmod 1 \quad i=1,\dots,u_k.$$

Weil sich bei Fibonacci-Zahlen mit geradem Index zur positiven Winkelhalbierenden symmetrische Punktmenge ergeben, läßt sich die Diskrepanz mit geringem Aufwand nach (2.5) berechnen. Dabei ist

$$D(\omega_{u_k}) = \frac{a}{4u_k^2}.$$

Die Werte sind in der nachfolgenden Tabelle zusammengestellt, in Spalte 5 ist dabei der Quotient aus der unteren Abschätzung nach dem Satz von Roth und der Diskrepanz angegeben

$$q = \frac{(\log u_k)^{1/2}}{u_k} : D(\omega_{u_k})$$

k	u_k	a	$D(\omega_{u_k})$	q
6	8	55	0.2148	0.839
8	21	167	0.09467	0.878
10	55	503	0.04158	0.876
12	144	1503	0.01812	0.854
14	377	4345	0.007643	0.845
16	987	12683	0.003255	0.817
18	2584	35975	0.001347	0.805
20	6765	103179	0.0005636	0.779

Zwar ist die Diskrepanz erheblich kleiner als die Abschätzung von Zaremba angibt, aber die Folge bietet auch keine Hoffnung, daß sich die Schranke aus (2.4) unterschreiten läßt.

Es sei angemerkt, daß die Diskrepanz noch kleiner wird, wenn in (3.1) der Parameter i zwischen 1 und u_k variiert. Dann gehört jedoch (1.1) zu ω_n , und es ergeben sich entsprechende Probleme bei der Transformation.

Mit der Punktmenge (3.2) steht zwar für praktische Rechnungen in zwei Dimensionen eine brauchbare Approximation der Gleichverteilung zur Verfügung, aber aus theoretischer Sicht ist die Situation noch sehr unbefriedigend.

4. Numerische Beispiele

Wir wollen die Anwendung des Satzes an einer Anfangswertaufgabe zur Vlasov-Gleichung in zwei Dimensionen illustrieren. In 1. war schon gesagt worden, daß hier das Vektorfeld $G(P,Q)$ nicht den Bedingungen des Satzes genügt und durch (1.5) modifiziert wird. Es ist klar, daß die Lösung des modifizierten Problems mit h gegen Null gegen die (in zwei Dimensionen existierende) Lösung des ursprünglichen Problems steht.

Trivialerweise ist der Satz auch anwendbar, wenn zusätzliche Randbedingungen derart gestellt werden, daß die Lösung nur in einem gewissen Bereich des R^2 von Null verschieden ist.

In [11], wo der Satz für Gleichungen vom Typ der Vlasovgleichung bei gleichzeitiger Diskretisierung von (B) gezeigt wurde, bereitete gerade die Diskretisierung Schwierigkeiten beim Nachweis der Konvergenz. Deshalb sei hier die Konvergenz bei Benutzung von Differenzenverfahren zur Lösung von (B') nochmals gezeigt.

Zunächst ist [vgl. 10] für beliebige $\omega_n, \hat{\omega}_n \in x_n$

$$D(\hat{\omega}_n, f) \leq r(\omega_n, \hat{\omega}_n) + D(\omega_n, f)$$

Sei nun $(\hat{P}_1(v\Delta t), \dots, \hat{P}_n(v\Delta t))$ die durch ein konvergentes Differenzenverfahren ermittelte Lösung des Systems (B') und $(P_1(t), \dots, P_n(t))$ seine wahre Lösung. Dann existiert für alle $\epsilon > 0$ ein Δt , so daß für alle $v=0, \dots, N; N\Delta t = T$ gilt

$$|P_i(v\Delta t) - \hat{P}_i(v\Delta t)| < \epsilon, \quad i=1, 2, \dots, n$$

Nach Hilfssatz 3 aus [12] ist daher

$$r(\hat{\omega}_n(v\Delta t), \omega_n(v\Delta t)) \leq 2 \left(\sigma_f(\epsilon) + D(\omega_n(v\Delta t), f(v\Delta t, \cdot)) \right)$$

wobei

$$\sigma_f(\epsilon) = \sup_P \left(\max_{Q: |P-Q| < \epsilon} \left| \int_{R(P)} f(Q') dQ' - \int_{R(Q)} f(Q') dQ' \right| \right)$$

ist.

Wir wollen die Vlasov-Gleichung in $0 \leq x \leq b$ mit den Bedingungen

$$f(t, 0, v) = f(t, b, v)$$

$$\int_0^b E(t, x) dx = 0$$

betrachten.

Dann beschreibt sie die Dichteverteilung von Elektronen in einer Orts- und einer Geschwindigkeitskoordinate unter der Annahme, daß die Ionen ruhen und räumlich homogen verteilt sind. Wir benutzen daher die übliche Schreibweise $(x, v) \in \mathbb{R}^2$ an Stelle von $P = (p_1, p_2) \in \mathbb{R}^2$, und zum Vergleich mit anderen Rechnungen die Einheiten (e = Elektronenladung, m = Elektronenmasse, $v_0 \approx$ Elektronentemperatur, N = Ionendichte)

$$\lambda_0 = \left(\frac{m v_0^2}{8\pi N e^2} \right)^{1/2} \quad \text{für den Ort}$$

$$c_p = \left(\frac{\pi m}{N e^2} \right)^{1/2} \quad \text{für die Zeit}$$

$$\text{und } - \left(\frac{N v_0^2 m}{8\pi^3} \right)^{1/2} \quad \text{für die elektrische Feldstärke.}$$

Dann hat die Gleichung die Gestalt

$$\frac{\partial f}{\partial t} + v \frac{\partial f}{\partial x} + E(t, x) \frac{\partial f}{\partial v} = 0$$

$$\frac{\partial E}{\partial x} = 4\pi^2 \left(\int_{-\infty}^{\infty} f(t, x, v) dv - N \right).$$

Die Gleichung für E ist im 2-dimensionalen Fall direkt integrierbar, wir erhalten hier mit $N = 1/b$

$$E(t, x) = 4\pi^2 \left(\int_0^x \int_{-\infty}^{\infty} f(t, \xi, v) dv d\xi - \frac{1}{b} \int_0^b \int_{-\infty}^{\infty} (b-\xi) f(t, \xi, v) dv d\xi + \frac{1}{2} - \frac{x}{b} \right)$$

Das Vektorfeld $G = (g_1(x, v; \xi, n), 4\pi^2 \cdot g_2(x, v; \xi, n))$ hat daher die einfache Gestalt

$$g_1(x, v; \xi, n) = g_1(v) = v$$

$$g_2(x, v; \xi, n) = g_2(x, \xi) = \begin{cases} \frac{\xi - x}{b} + \frac{1}{2} & \xi < x \\ \frac{\xi - x}{b} - \frac{1}{2} & x < \xi \end{cases}$$

Um G so zu modifizieren, daß es den Bedingungen des Satzes genügt, wählen wir die numerisch besonders einfach zu handhabende Funktion

$$\phi(h, r) = \begin{cases} \frac{1}{2h} & |r| \leq h \\ 0 & |r| > h \end{cases}$$

Als $\tilde{G}$ erhält man

$$\begin{aligned} \tilde{g}_1(v) &= v \\ \tilde{g}_2(x, \xi) &= \begin{cases} \frac{\xi - x}{b} - \frac{1}{2} & x - h \leq \xi \\ \frac{\xi - x}{b} + \frac{1}{2} - \frac{x - \xi - h}{2h} & x + h \leq \xi \leq x - h \\ \frac{\xi - x}{b} + \frac{1}{2} & \xi \leq x + h \end{cases} \end{aligned}$$

Wir haben nicht untersucht, ob sich der Aufwand lohnt, wenn man andere Funktionen zur Modifizierung des Problems benutzt, weil bei unseren Rechnungen der Einfluß von h gering war.

Als modifiziertes System (B') ergibt sich

$$\begin{aligned} (4.1) \quad \dot{x}_i &= v_i \\ \dot{v}_i &= \frac{4\pi^2}{n} \sum_{k=1}^n \tilde{g}_2(x_i, x_k) \end{aligned} \quad i=1, \dots, n$$

Wir setzen $x_i(j\Delta t) = x_i^j$, $v_i(j\Delta t) = v_i^j$ und diskretisieren (4.1)

$$v_i^{j+1} = v_i^j + 4\pi^2 \frac{\Delta t}{n} \sum_{k=1}^n \tilde{g}_2(x_i^j, x_k^j)$$

$$x_i^{j+1} = x_i^j + \Delta t v_i^{j+1}$$

Für eine möglichst schnelle Berechnung der Summe $\sum_{k=1}^n \tilde{g}_2(x_i^j, x_k^j)$ bieten sich folgende Wege an:

a) Für jedes j werden die (x_k^j, v_k^j) derart neu indiziert, daß

$$x_1^j \leq x_{1+1}^j \quad l=1, \dots, n-1$$

gilt.

Dann läßt sich $\sum_{k=1}^n \tilde{g}_2(x_{i+1}^j, x_k^j)$ aus $\sum_{k=1}^n \tilde{g}_2(x_i^j, x_k^j)$ durch relativ wenige Additionen berechnen.

b) Zunächst gilt

$$\frac{1}{n} \sum_{k=1}^n \tilde{g}_2(x, x_k^j) = \frac{1}{2} + \frac{1}{n \cdot b} \sum_{k=1}^n x_k^j - \frac{x}{b} - \frac{1}{n} \sum_{k=1}^n \int_x^b \phi(h, |\xi - x_k^j|) d\xi$$

(4.2)

$$= \frac{1}{n \cdot b} \sum_{k=1}^n x_k^j - \frac{1}{2} + \int_0^x \left(\frac{1}{n} \sum_{k=1}^n \phi(h, |\xi - x_k^j|) - 1/b \right) d\xi$$

Das letzte Integral berechnet man mittels numerische Integration an den (für alle j festen) Grenzen $X_1, \dots, X_s$ mit

$$X_{l+1} - X_l = \Delta x \quad l=1, \dots, s-1$$

und interpoliert daraus die Werte $\frac{1}{n} \sum_{k=1}^n \tilde{g}_2(x_i^j, x_k^j)$.

Wir haben fast ausschließlich mit dem zweiten Verfahren gerechnet und dabei $\Delta x = h$ gewählt. Das Integral von X_l bis X_{l+1} wurde aus den Werten des Integranden I in X_{l-1} , X_l , X_{l+1} und X_{l+2} durch

$$\frac{h}{24} (-I(X_{l-1}) + 13 I(X_l) + 13 I(X_{l+1}) - I(X_{l+2}))$$

bestimmt. In dieser Form entspricht die Berechnung des elektrischen Feldes dem "particle-in-cell"-Verfahren von Morse [vgl. in 1], für das damit die Konvergenz gesichert ist.

Wir betrachten das Anfangswertproblem

$$f_s(x, v) = \frac{1}{v_0 \pi^{1/2}} \exp\left(-\frac{v^2}{v_0^2}\right) \left(1 + \kappa \cos \frac{2\pi x}{b}\right) \quad 0 < \kappa < 1, \quad 0 \leq x \leq b,$$

die sogenannte "Landau-Dämpfung".

Für diese Anfangsbedingung gilt

$$\int_0^b \int_{-\infty}^{\infty} f(t, x, v) dv dx = b,$$

wodurch die Ionendichte $N = 1$ wird, was aus physikalischen Gründen vorzuziehen ist. In den beschriebenen Verfahren bedeutet diese Wahl lediglich einen zusätzlichen Faktor b in (4.2).

Für die Lösung dieses Problems gilt bei den angegebenen Randbedingungen

$$f(t, x, v) = f(t, b-x, -v)$$

$$E(t, x) = -E(t, b-x)$$

und daher

$$E(t, 0) = E(t, \frac{b}{2}) = E(t, b) = 0,$$

wie man sofort durch Einsetzen in die Gleichung sieht.

Daher kann man die numerische Rechnung auf das Intervall $0 \leq x \leq b/2$ beschränken, denn ist $\{(x_i, v_i), i=1, \dots, n\}$ eine Approximation von $f(t, x, v)$ in $0 \leq x \leq b/2$, so ist $\{(b-x_i, v_i), i=1, \dots, n\}$ eine von $f(t, x, v)$ in $b/2 \leq x \leq b$. Deshalb und wegen des zusätzlichen Faktors b erhalten wir aus (4.2)

$$\frac{1}{n} \sum_{k=1}^n \tilde{g}_2(x, x_k^j) = \int_0^x \left(\frac{b}{n} \sum_{k=1}^n \phi(h, |\xi - x_k^j|) - 1 \right) d\xi.$$

Nach der linearisierten Theorie [vgl. 14] ist die elektrische Energie $\tilde{\mathcal{E}}(t) = \tilde{E}(t)/4\pi^2$ mit

$$\tilde{E}(t) = \int_0^b E^2(t, x) dx$$

für $b=20$ außer am Anfang nur schwach gedämpft. Für $b=100$ erhält man hier-

nach keine Dämpfung, und für $b=10$ verschwindet $\tilde{E}(t)$ sehr rasch. Wir haben daher den Fall $b=20$ mit dem Ziel, den zeitlichen Verlauf der elektrischen Energie anzugeben, untersucht. Für $t = 0$ gilt

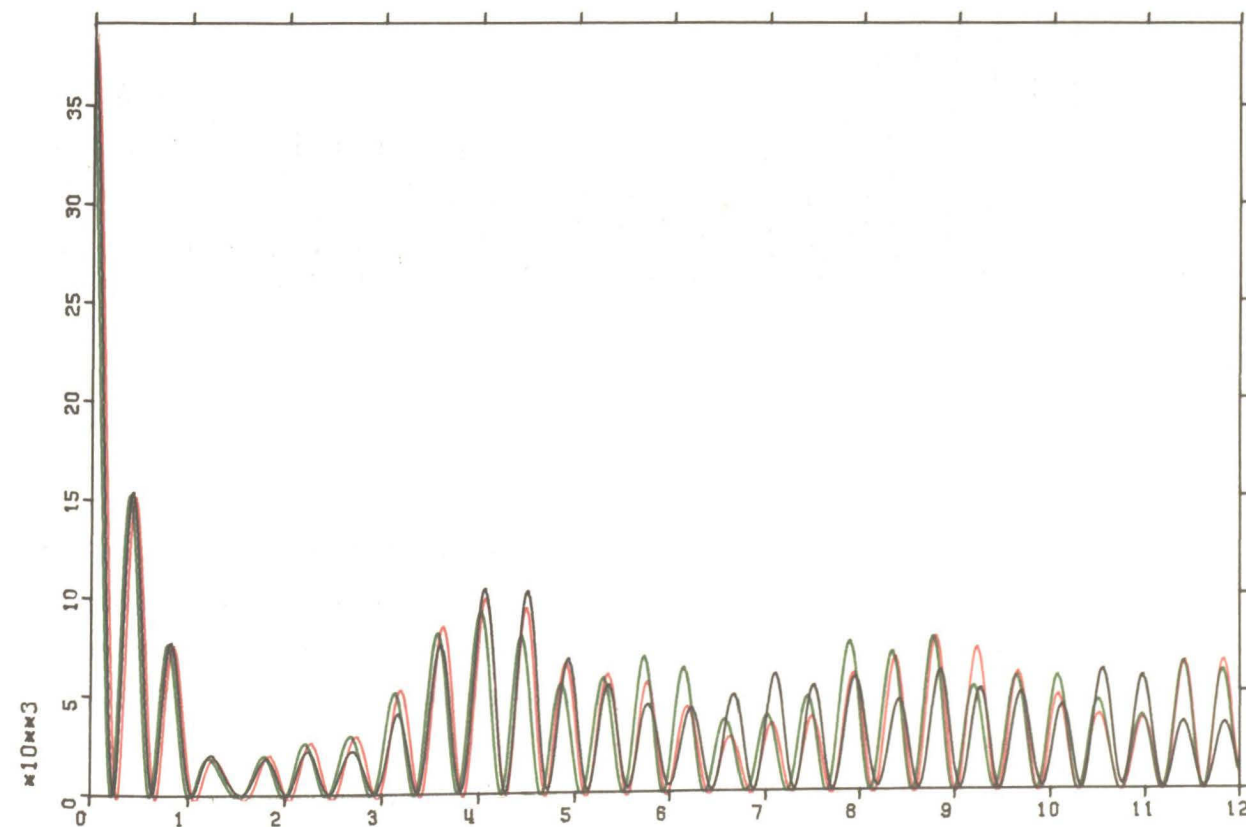
$$\tilde{E}(0) = 2b^3(\pi\kappa)^2,$$

so daß für kleine κ die Anzahl n der Approximationspunkte groß werden muß, damit der Fehler der kinetischen Energie

$$\tilde{E}_{\text{kin}}(t) = \int_0^b \int_{-\infty}^{\infty} v^2 f(t, x, v) dv$$

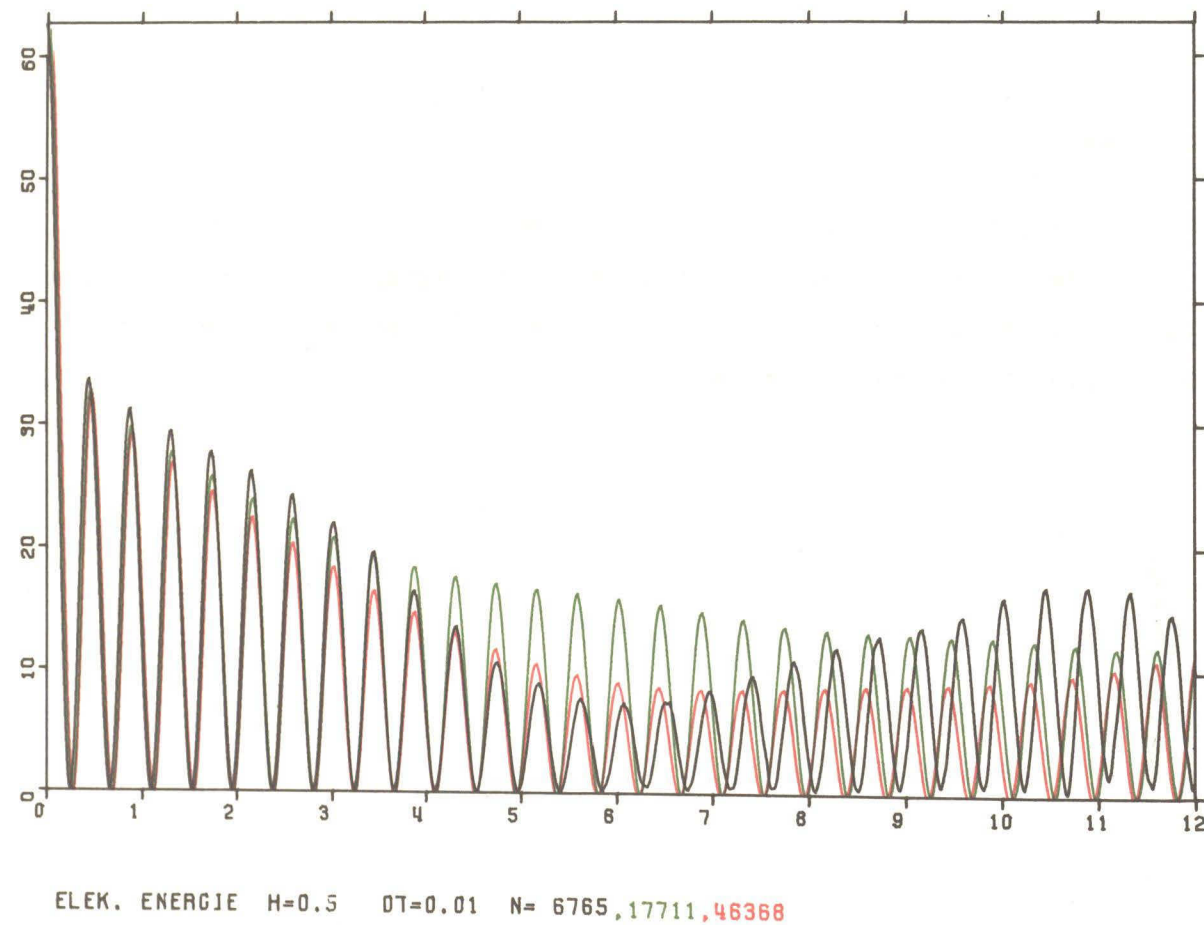
kleiner als die elektrische Energie wird.

In Abb. 1 ist der Verlauf von $\tilde{E}(t)$ bis $t=12$ für $\kappa = 1/2$ für drei verschiedene Werte von n bei festem h und Δt angegeben. Die Rechenzeit für alle drei Fälle betrug auf einer IBM/370-168 insgesamt 5 Minuten, oder pro Approximationspunkt und Zeitschritt 24 μsec .



ELEK. ENERGIE H=0.5 DT=0.01 N= 987, 2584, 6765

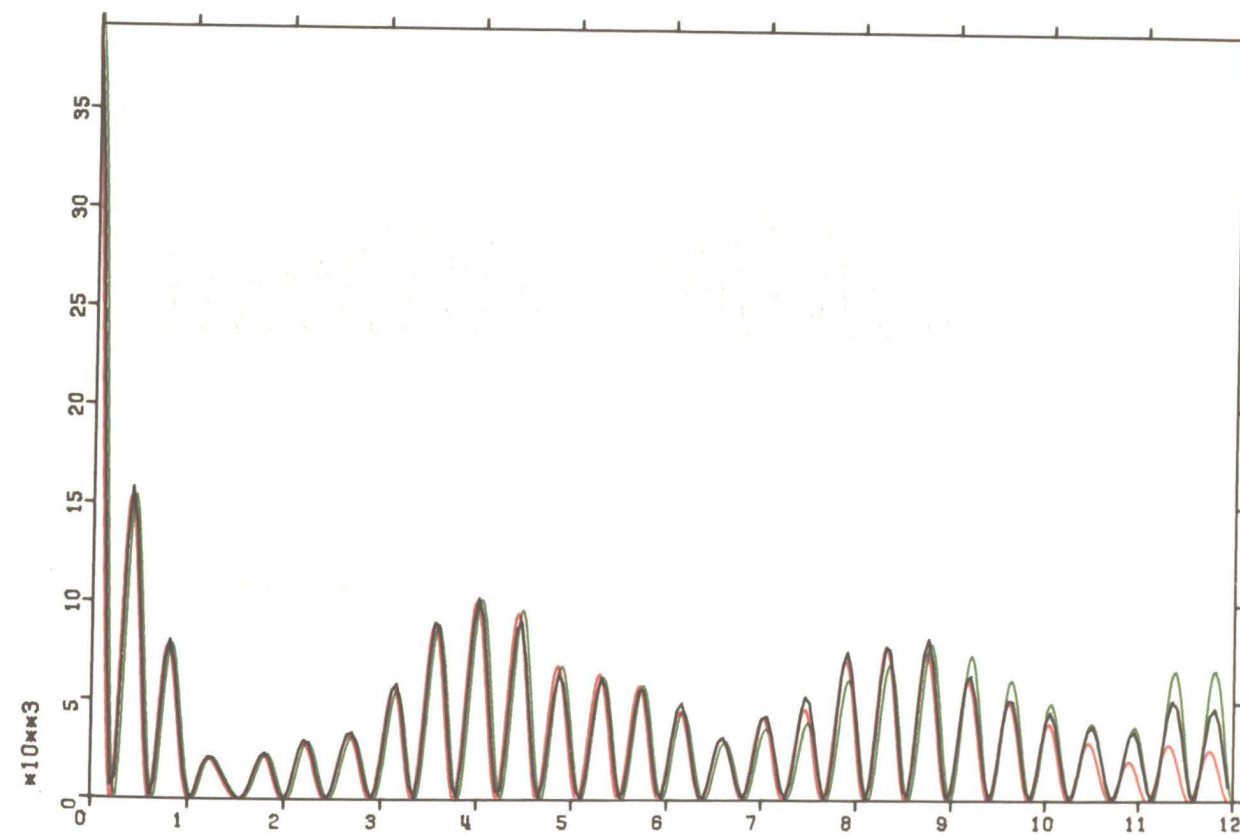
Man sieht, daß man bereits für $\kappa = 1/2$ mit $n = 987$ (d.h. weniger als 30 sec. Rechenzeit) brauchbare Resultate erhält. Doch für $\kappa = 0.02$ hat man nur noch $1/625$ der elektrischen Anfangsenergie $E(0)$ im Vergleich zu $\kappa = 1/2$, zwischen den Diskrepanzen für $n = 987$ und $n = 6765$ aber nur den Faktor 5.78. Die Schwierigkeiten, kleine Störungen des Gleichgewichts numerisch zu beherrschen, zeigt Abb. 2.



Auffallend ist, daß auch bei $\kappa = 0.02$ die elektrische Energie nach $t = 7.5$ wieder ansteigt. Das Auftreten einer solchen "Oberschwingung" hat bereits Kadomzev [7] aus Überlegungen mit der linearisierten Theorie gezeigt und ihre Wellenlänge proportional zu $E^{-1/2}(b/4,0)$ angegeben. Wenn wir bei $\kappa = 1/2$ diese Wellenlänge mit 5 ansetzen, wäre für $\kappa = 0.02$ danach eine Wellenlänge von 25 zu erwarten. Dies würde eine Rechnung bis $t = 40$ erfordern (75 min. Rechenzeit für $n = 46368$ und $\Delta t = 0.01$). Wir haben diese

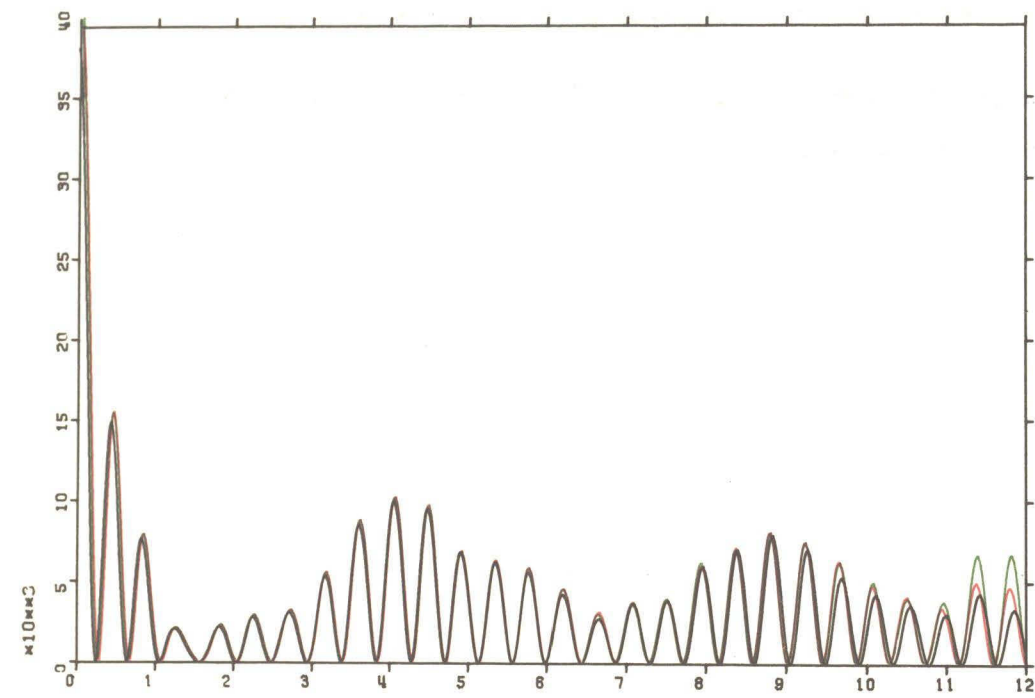
Rechnung nicht durchgeführt; doch wenn man die Zeitpunkte der ersten Minima der "Oberschwingung" für die beiden x -Werte betrachtet, so findet man zwischen diesen näherungsweise ebenfalls den Faktor 5.

Der zweite, die Rechenzeit wesentlich beeinflussende Parameter ist Δt . Die Abhängigkeit zeigt Abb. 3.

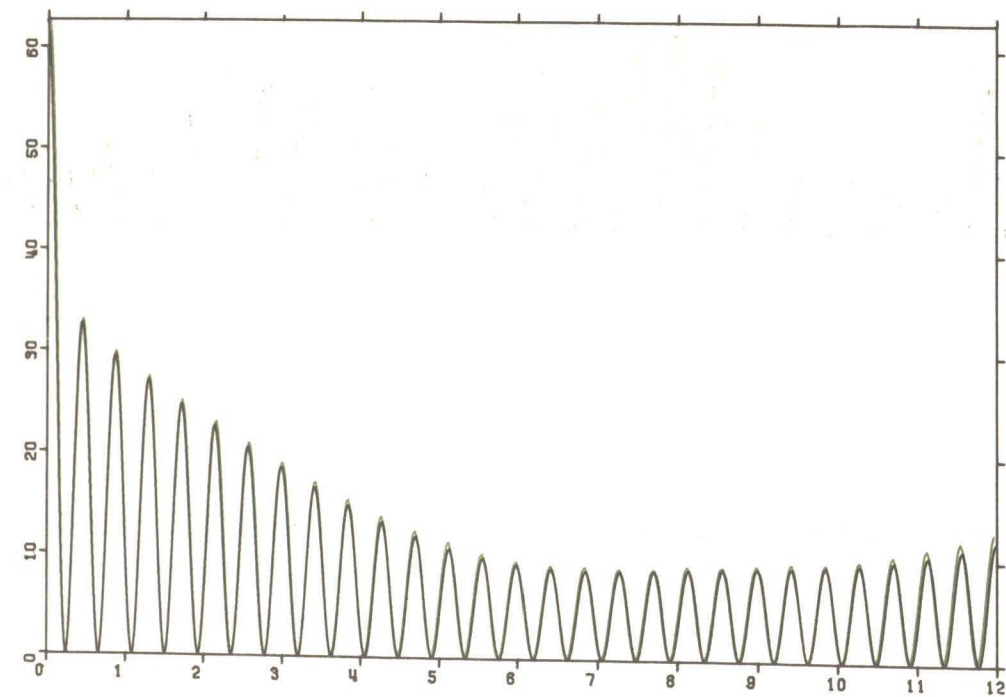


ELEK. ENERGIE N=6765 H=0.5 DT= 0.05, 0.01, 0.002

Durch Differenzenverfahren mit besserer Fehlerordnung zur Lösung von (4.1) kann man aber den Δt -abhängige Fehler verringern. Da die Wahl von h (aus physikalischen Gründen nicht größer als die als Längeneinheit benutzte Debye-Länge) nur geringen Einfluß auf die Ergebnisse hat (Abb. 4 und 5), sind die numerischen Anwendungsmöglichkeiten wesentlich durch die theoretischen Ergebnisse zur Gleichverteilung modulo Eins bestimmt.



ELEK. ENERGIE N=6765 DT=0.01 H= 1.00, 0.50, 0.25



Elek. Energie N = 46368, DT = 0.01, H = 0.5 (schwarz), 0.02 (grün)

Wie bereits gesagt, kann man diese Approximation auch als Darstellung des Elektronenplasmas durch n Elektronen im Phasenraum auffassen. Durch die graphische Ausgabe der Punktpositionen nach jedem Zeitschritt auf Film kann man daher das Wechselspiel zwischen Elektronen und elektrischem Feld unmittelbar veranschaulichen. Zum Abschluß sollen die damit gegebenen Möglichkeiten kurz gestreift werden, da sie häufig erst einen Einblick in die physikalischen Vorgänge und damit Auskunft über die sinnvoll zu berechnenden Größen gestatten. Insbesondere bei den Rechnungen zu [2] war dies von großer Bedeutung. Hier beschränken wir uns wieder auf den Fall der Landau-Dämpfung mit $\kappa = 1/2$. Dabei zeigt sich, daß bzgl. v außer dem Maximum bei $v = 0$ sehr rasch Nebenmaxima bei $v = \pm 33$ entstehen, obwohl in der Anfangsapproximation so hohe v -Werte nicht auftreten. Die Ausbildung der Nebenmaxima steht dabei in engem Zusammenhang mit der Differenz der elektrischen Energie am Anfang und ihrem ersten Maximum. Da diese Nebenmaxima durch ein flaches, fast verschwindendes Minimum vom Hauptmaximum getrennt sind, kann man auch von der Gruppe langsamer und schneller Elektronen sprechen. Zwischen beiden Gruppen pendeln nur sehr wenige Elektronen, was die Oberschwingung der elektrischen Energie wiedergibt. Verhindert man die Bildung der Nebenmaxima, indem man alle Elektronen, deren Betrag der Geschwindigkeit größer $v_1 > 20$ wird, unter Beachtung der elektrischen Neutralität bei der weiteren Rechnung nicht mehr berücksichtigt, so ergeben die verbleibenden langsamen Elektronen ein ungedämpft schwingendes elektrisches Feld. Dabei fällt dessen Amplitude zunächst bei wachsendem v_1 . Erst bei $v_1 > 32$ tritt die Oberschwingung auf. Die "Pendelelektronen" zeigen also sehr große Geschwindigkeitsunterschiede. Dieses Phänomen tritt übrigens auch auf, wenn man das elektrische Feld aus dem linearisierten Problem [14] als vorgegeben annimmt und damit Elektronenbahnen rechnet.

Für die numerische Rechnung bieten solche physikalischen Informationen Hilfsmittel für die Wahl der Diskretisierungsparameter, da für das Verfahren keine Fehlerabschätzung bekannt ist.

Literatur

- 1 Alder, B., Fernbach, S., Rotenberg, M. (Herausgeber): Methods in computational physics.
Vol. 9, plasma physics, New York, Academic Press 1970
- 2 Amemiya, H.: Investigation on the Plasma Oscillation Caused by Primary Electrons in a Low Pressure Gas Discharge.
Sc. Papers of Inst. of Phys. and Chem. Research, 67 (1973), 161-170, Wako-shi, Saitama, Japan
- 3 Batt, J.: Ein Existenzbeweis für die Vlasov-Gleichung der Stel-
lardynamik bei gemittelter Dichte.
Arch. Rational Mech. Anal. 13 (1963), 296-308
- 4 Hlawka, E.: Zur angenäherten Berechnung mehrfacher Integrale.
Monatsh. f. Math. 66 (1962), 140-151
- 5 Hlawka, E.: Uniform distribution modulo 1 and numerical analysis.
Comp. Math. 16 (1964), 92-105
- 6 Hlawka, E., Mück, E.: Über eine Transformation von gleichverteilt-
ten Folgen II.
Computing 9 (1972), 127-138
- 7 Kadomzev, B.B.: Landaudämpfung und Plasmaecho (in Russisch).
Usp. fis. nauk 95 (1968), 111-129
- 8 Koksma, J.F.: The theory of asymptotic distribution modulo one.
Comp. Math. 16 (1964), 1-22
- 9 Niederreiter, H.: Discrepancy and convex programming.
Ann. Mat. Pura Appl. IV, 93 (1972), 89-97
- 10 Neunzert, H., Wick, J.: Die Darstellung von Funktionen mehrerer
Variabler durch Punktmengen.
Berichte der KFA Jülich, Jül-996-MA (1973)
- 11 Neunzert, H., Wick, J.: Die Theorie der asymptotischen Verteilung
und die numerische Lösung von Integrodifferentialgleichungen.
Num. Math. 21 (1973), 234-243

- 12 Neunzert, H., Wick, J.: Die Approximation der Lösung von Integro-Differentialgleichungen durch endliche Punktmengen.
in Lecture Notes in Mathematics Nr. 395
Herausgeber: R. Ansorge, W. Törnig (Numerische Behandlung nichtlinearer Integrodifferential- und Differentialgleichungen)
Springer Berlin-Heidelberg-New York
- 13 Neunzert, H., Wick, J.: Zur numerischen Lösung von Erhaltungsgleichungen.
erscheint in: ZAMM, 54 (1974) T194-T195
- 14 Profant, M., Wick, J.: Zur Lösung der linearisierten Vlasov-Gleichung mit periodischen Randbedingungen.
Berichte der KFA Jülich, Jül-760-MA (1971)
- 15 Roth, K.F.: On irregularities of distribution.
Mathematika 1 (1954), 73-79
- 16 Zaremba, S.K.: Good lattice points, discrepancy and numerical integration.
Ann. Mat. Pura Appl. IV 73 (1966), 293-317
- 17 Zaremba, S.K.: Good lattice points modulo prime and composite numbers in Diophantine Approximation and its Application.
herausgegeben von C.F. Osgood, Academic Press, New York, 1973