STICHTING
MATHEMATISCH CENTRUM
2e BOERHAAVESTRAAT 49
AMSTERDAM

DR 6

DR 6

Grundsaetzliche Probleme der Abrundungsfehler.

(Zeitschrift fuer angewante Mathematik und Mechanik, 30 (1950), p275-276).

A. van Wijngaarden.

1950

Printed at the Mathematical Centre, 49, 2e Boerhaavestraat, Amsterdam.

The Mathematical Centre, founded the 11-th of February 1946, is a non-profit institution aiming at the promotion of pure mathematics and its applications. It is sponsored by the Netherlands Government through the Netherlands Organization for the Advancement of Pure Research (Z.W.O), by the Municipality of Amsterdam, by the University of Amsterdam, by the Free University at Amsterdam, and by industries.

Grundsätzliche Probleme der Abrundungsfehler

Von *A. van Wijngaarden* in Amsterdam

Untersucht wird die Frequenz des Auftretens von Fehlern vorgegebener Größe im Ergebnis gewisser systematischen Rechnungen mit abgerundeten Zahlenfolgen. Durch passende Definition von Funktionen kann jede abgerundete Zahl als ganz betrachtet werden. Der abgerundete Wert Af und der Bruchteil Bf einer reellen Zahl f sind dann definiert durch $Af+Bf=f$; $Af\equiv 0 \pmod 1$, $|Bf|<1/2$ wenn $f\not\equiv 1/2 \pmod 1$ und $Af\equiv 0 \pmod 2$, $|Bf|=1/2$ wenn $f\equiv 1/2 \pmod 1$. Es liege eine große Zahlenfolge $f_k (k=1, 2, \ldots)$ vor, und es sei $f=F(f_j, \ldots, f_{j+n-1})$ zu bestimmen.

Werden statt f_k die Af_k benützt und statt F, die etwa nur abgerundet verwendbare Konstanten enthalten kann, eine „abgerundete" Funktion AF, dann ist das Ergebnis der Rechnung $g=AF(Af_j, \ldots, Af_{j+n-1})$. Von der Diskrepanz $\psi=f-g$ kann durch Wiederholung für verschiedene j die Verteilung $v(\psi)$ definiert werden. Manchmal ist nur Af zu bestimmen. Die Rechnung liefert Ag und die Frequenz der Diskrepanz $P=Af-Ag$ sei $\Omega(P)$.

Wichtig für $v(\psi)$ und $\Omega(P)$ ist die Verteilung und die Abhängigkeit der Bf_k. Sind die f_k Werte für $x=kh$ einer Funktion $f(x)$, dann tritt meistens praktisch Gleichverteilung der Bf_k auf, was weiter auch vorausgesetzt wird. Ist $f(x)$ n-fach stetig differenzierbar, dann ist bekanntlich $\Delta_1^n f=h^n f^{(n)}(\xi)$, mit $x_1\leq\xi\leq x_{n+1}$. Ist $f^{(n)}(x)$ beschränkt im betrachteten Bereich dann sind bei genügend kleiner h die $\Delta^n f$ vernachlässigbar. Allgemeiner sei die „kritische" Differenz $\Delta^m f$ die Differenz niedrigster Ordnung deren Bruchteil praktisch konstant ist. Dann ist $Bf_{m+1}=B\left(B\Delta^m f_1-\sum_{k=1}^{m}(-1)^{m-k+1}\binom{m}{k-1}Bf_k\right)$, also Bf_{m+1} praktisch funktionsgemäß abhängig von seinen m Vorgängern. Außerdem wird vorausgesetzt, daß weniger als m aufeinanderfolgende Bf_k unabhängig sind.

Für den wichtigen Fall der homogenen Linearkombination $f=\sum_{k=1}^{n} a_k f_k$ mit exakten a_k und $n\leq m$ wird $v(\psi)$ mit Hilfe der Laplacetransformation bestimmt, während die Bestimmung von $\Omega(P)$ einfache zahlentheoretische Mittel erfordert. Für $n>m$ werden $v(\psi)$ und $\Omega(P)$ durch mehrdimensionale geometrische Betrachtungen gefunden. Im Ergebnis für $\Omega(P)$ tritt für $n\leq m$

---

[1]) W. Schnell: Interpolation und Untertafelung im Komplexen. Dipl.-Arb. Darmstadt 1950.

[2]) A. N. Lowan und H. E. Salzer: Coefficients for interpolation within a square grid in the complex plane. J. Math. Phys. 23 (1944), 156—159.

[3]) H. E. Salzer: Formulas for direct and inverse Interpolation of a complex function tabulated along quidistant Circular Arcs. J. Math. Phys. 24 (1945), 141—142.

der kleinst mögliche gemeinschaftliche Nenner q der a_k auf. Bei kontinuierlicher Variation der a_k verhält $\Omega(P)$ sich also ganz pathologisch.

Die Theorie ist auf die Bestimmung der Diskrepanz $P = A\Delta^n f - \Delta^n A f$ angewandt. Es zeigt sich, daß außerhalb der Ordnung n auch der Wert $n-m$ eine wichtige Rolle spielt. Ist $n<m$ die Differenz also subkritisch, dann ist $\Omega(P)$ unabhängig von n und $B\Delta^m f$. Ist dagegen $n \geq m$, die Differenz also kritisch oder superkritisch, so hängt $\Omega(P)$ von n und $B\Delta^m f$ ab.

Eine zweite Anwendung liefern Interpolation, numerische Differentation oder Integration in einem beschränkten Bereich. Sei z. B. $I_p(f_j, \ldots, f_{j+n-1})$ der interpolierte Wert für $x_p = x + h\left(\left[\frac{n-1}{2}\right] + p\right)$, dann hat die Frequenz $\Omega(P)$ von $P = AI_p(f_j, \ldots, f_{j+n-1}) - AI_p(Af_j, \ldots, Af_{j+n-1})$ eine Unstetigkeitsstelle für jeden rationalen Wert von p wenn $n \leq m$. So ist z. B. für $p=0$ offenbar $\Omega(0)=1$, während für $p=0+\varepsilon$, $\Omega(0)=3/4$ für Interpolation beliebiger Ordnung gilt. Für $n>m$ liegt die Sache ganz anders. Speziell wenn $n \gg m$ ist, z. B. bei Integration über einem langen Bereich treten ganz besondere Gesichtspunkte auf.