Zur Axiomatisierung von Mittelwerten unter Berücksichtigung der historischen Begriffsentwicklungen

von Horst Hischer und Anselm Lambert

Kurz könnte man sagen, dass Axiomatisierung das Hinschreiben von Axiomen für eine Klasse bzw. eine Familie von gewissen Objekten ist. Dies würde allerdings in unzulässiger Weise den langen und komplexen Weg, der zu diesem Ziel nötig ist, verdecken bzw. leugnen; man muss ja Axiome erst einmal haben, genauer: sie sich durch mathematische Arbeit verdienen, um sie in befriedigender Form hinschreiben zu können.

Der Begriff eines *Mittelwertes* ist nun keineswegs so eindeutig axiomatisch festlegbar wie etwa der des Vektorraums. Die Axiomatisierung von Mittelwerten hat im 20. Jahrhundert und teilweise schon davor viele Mathematiker beschäftigt, und sie ist auch heute noch immer nicht einvernehmlich abgeschlossen – dies macht die Auseinandersetzung mit dem Thema gerade für die Schule besonders spannend, weil den Lernenden dabei die Mathematik als offenes und gestaltbares Werk gegenübertritt.

Im hier vorliegenden Artikel möchten wir am Beispiel der „Mittelwertfunktionen" Axiomatisierung skizzieren und dabei exemplarisch illustrieren, dass den Ausgangspunkt von Axiomatisierung stets zwei wechselseitig aufeinander bezogene Ideen bilden (die über ihre Prozesse zu ihren Produkten führen):

- die Untersuchung *konkreter* vorliegender Objekte auf gemeinsame Eigenschaften
- die Entdeckung solcher *abstrakter* Eigenschaften in weiteren Beispielen („*Modellen*")

1 Axiomatisierung als Idee, Prozess und Produkt

1.1 Konkrete Mittelwertfunktionen

Im vorliegenden Kontext legen wir also gewisse *konkrete Mittelwertfunktionen* zugrunde und stellen die Frage nach den (?) Eigenschaften der somit zu untersuchenden Familie von Funktionen: einerseits nach Eigenschaften einzelner Funktionen dieser Familie und andererseits nach gemeinsamen Eigenschaften aller dieser Funktionen. Diese Eigenschaften gewinnen wir durch Abstrahieren, und dies ist in der Regel der Einstieg in das Axiomatisieren. Was aber sind „konkrete Mittelwertfunktionen"? Sie sind für uns durch eine eindeutige Zuordnungsvorschrift definiert, z. B. durch die Angabe eines Funktionsterms [1], etwa $G(x,y) := \sqrt{xy}$. Konkrete Mittelwertfunktionen müssen nicht notwendig termdefiniert sein, dennoch liefern uns solche ein hinreichhaltiges Ausgangsmaterial.

1.2 Abstrakte Mittelwertfunktionen

Axiomatisierung aller Mittelwertfunktionen: Die gefundenen gemeinsamen Eigenschaften gewisser konkreter Mittelwertfunktionen benutzen wir nun als Axiome, um zu definieren (d. h.: abzugrenzen), was für uns *abstrakte Mittelwertfunktionen* sein sollen.

Jede Funktion, die diese Axiome erfüllt, ist dann für uns eine „Mittelwertfunktion“. Diese Sichtweise wirkt auf die Menge der von uns eingangs betrachteten Funktionen zurück, wenn wir weitere konkrete Funktionen, welche die geforderten Eigenschaften erfüllen, entdecken und explizit angeben, und sie setzt einen Kreisprozess in Gang:

konkrete, (z. B. termdefinierte) Mittelwertfunktionen	Konkretisierung $\leftarrow$ $\rightarrow$ Abstrahierung	abstrakte, axiomatisch definierte Mittelwertfunktionen

Dabei tauchen dann vielleicht neue konkrete Funktionen auf, die wir nicht als Mittelwertfunktionen ansehen *wollen* – etwa weil für uns neben der *Frage nach Wahrheit* auch die *Frage nach Schönheit* (sic!) begriffsbildend bedeutsam ist oder weil sich unser mathematisches Handeln auch vor dem Hintergrund eines intuitiven, teils unbewussten Begriffs von Mittelwert abspielt, der in uns durch unsere Erfahrung gereift ist (und weiterhin reift), ohne auf Anhieb in Gänze formal festnagelbar zu sein, da er nur allmählich zu Tage tritt und so nur Schritt für Schritt in Klarheit enthüllt werden kann. Solche „nicht erwünschten“ konkreten Funktionen schließen wir daher durch weitere Forderungen aus (Strategie: „Verbot des Unerwünschten“ [2]). Dies hält den Kreisprozess des Axiomatisierens, der auf ein – meist dann doch immer nur vorläufig endgültiges – System von Axiomen zielt, in Gang; Mathematikerinnen und Mathematiker (vorläufig!) befriedigende Axiome sind dann der Mühe Lohn.

Axiomatisierung spezieller Mittelwertfunktionen: Mit den abstrahierten Eigenschaften können wir andererseits auch versuchen, *eine* spezielle konkrete Mittelwertfunktion zu charakterisieren:

spezielle, konkrete, (z. B. termdefinierte) Mittelwertfunktion	Konkretisierung $\leftarrow$ $\rightarrow$ Abstrahierung	spezielle, abstrakte, axiomatisch definierte Mittelwertfunktion

So erweist sich z. B. – wie Luca Teodoriu 1931 gezeigt hat und wie wir später noch kurz beweisen werden – eine additive, symmetrische und reflexive Funktion schon automatisch als das arithmetische Mittel, d. h., insbesondere ergibt sich die für Mittelwerte vielleicht typischste Eigenschaft der *Internalität* in diesem Fall „freiwillig“. Für Rudolf Schimmack (s. u.) sind *„einleuchtende Grundeigenschaften“* sogar der Ausgangspunkt axiomatisch-mathematischer Arbeit: Er gab 1909 solche an, [3] und er deduziert dann, *„daß die Funktion nur das arithmetische Mittel sein kann“* [Schimmack 1909, 125].

2 Eigenschaften von Mittelwertfunktionen

Wenn wir nach Eigenschaften von Mittelwertfunktionen suchen *wollen* [4], müssen wir nun also entscheiden, welche Funktionen bzw. welche Familien von Funktionen wir betrachten, um unsere Unternehmung zu beginnen. Wir entscheiden uns hier für einen Start mit dem arithmetischen, dem geometrischen, dem harmonischen und dem quadratischen Mittel – diese sind in [Lambert & Herget 2004] auf vielfältige Art und Weise dargestellt (zumindest für zweistellige Mittelwertfunktionen). Anstelle von Mittelwerten aus nur zwei Zahlen betrachten wir hier gleich Mittelwerte aus beliebig vielen Zahlen, und zwar in der Einschränkung auf positive reelle Zahlen. Die zu untersuchenden Objekte sind also n-stellige Funktionen (mit $n \in \mathbb{N}^*$) [5]:

$$M_n := (\mathbb{R}_+^n \to \mathbb{R}_+; \vec{x} \mapsto M_n(\vec{x}))$$

2.1 Eigenschaften des arithmetischen Mittels

Das für alle $n \in \mathbb{N}^*$ durch [6]

$$A_n := (\mathbb{R}_+^n \to \mathbb{R}_+; \vec{x} \mapsto \tfrac{1}{n} \sum\nolimits_{\nu=1}^{n} x_\nu)$$

gegebene arithmetische Mittel von n Werten hat u. a. die folgenden Eigenschaften (vgl. [Bullen 2003, 62]), die in manchen Fällen nahe liegen – etwa bei (**In**) oder bei (**Sy**) –, die man aber in anderen Fällen leider leichter nachrechnet, als man sie in der formalen Darstellung errät. Hierbei seien stets $\vec{x}, \vec{y} \in \mathbb{R}_+^n$ und $\lambda \in \mathbb{R}_+$ beliebig gewählt:

- (**In**) *Internalität:* $\min(\vec{x}) \leq A_n(\vec{x}) \leq \max(\vec{x})$, wobei die Gleichheit genau dann gilt, wenn $\vec{x}$ konstant ist.
- (**Sy**) *Symmetrie:* $A_n(\vec{x}) = A_n(S_\sigma(\vec{x}))$ für alle Permutationen σ von $\{1, 2, \ldots, n\}$ mit $S_\sigma := (\mathbb{R}_+^n \to \mathbb{R}_+^n; (x_\nu) \mapsto (x_{\sigma(\nu)}))$.
- (**St**) *Stetigkeit.*
- (**Ho**) *Homogenität*: $A_n(\lambda\vec{x}) = \lambda A_n(\vec{x})$.
- (**Mo**) *Monotonie*: $\vec{x} \leq \vec{y} \Rightarrow A_n(\vec{x}) \leq A_n(\vec{y})$, wobei die Gleichheit genau dann gilt, wenn $\vec{x} = \vec{y}$ ist.
- (**Re**) *Reflexivität:* für $\vec{x} = (x_0)_{\nu=1}^n$ (also bei Komponentengleichheit) ist $A_n(\vec{x}) = x_0$.
- (**Ad**) *Additivität*: $A_n(\vec{x} + \vec{y}) = A_n(\vec{x}) + A_n(\vec{y})$.
- (**As-*m***) m*-Assoziativität:* für $1 \leq m \leq n$ und
  $P_m := (\mathbb{R}_+^n \to \mathbb{R}_+^m; (x_\nu)_{\nu=1}^n \mapsto (x_\nu)_{\nu=1}^m)$ gilt stets
  $A_n(\vec{x}) = A_n(A_m(P_m(\vec{x})), \ldots, A_m(P_m(\vec{x})), x_{m+1}, \ldots, x_n)$.

Die meisten dieser Eigenschaften werden in ihrer Bedeutung unmittelbar einleuchten.

So ist z. B. (**Re**) die Verallgemeinerung der natürlichen Eigenschaft zweistelliger Mittelwertfunktionen, dass der Mittelwert von einer Zahl und ihr selbst wieder diese Zahl ist. Wie man auf (**As-*m***) kommen kann, zeigt uns später (in 3.1.1.2) Schimmacks Zugang, der mit einer begrifflichen Interpretation statt mit Formeln operiert.

Die Eigenschaften (**In**) bis (**As-*m***) sind nicht voneinander unabhängig. So folgt z. B. (**Re**) aus (**In**) und (**In**) aus (**Mo**) und (**Re**). Die Eigenschaft (**In**) formalisiert für uns „irgendwo in der Mitte" (zugleich im Sinne von *„irgendwo zwischen zwei Rändern"* [7]) und ist damit unverzichtbar als typisch für „Mittelwert im Allgemeinen", die Eigenschaft (**Ad**) scheint dagegen als eher typisch für den arithmetischen Mittelwert im Speziellen zu sein; (**Ad**), (**Sy**) und (**Re**) implizieren bereits das arithmetische Mittel (s. u.).

2.2 Eigenschaften des geometrischen, des harmonischen und des quadratischen Mittels

Geometrisches Mittel, harmonisches Mittel und quadratisches Mittel sind nun analog zum arithmetischen Mittel für alle $n \in \mathbb{N}^*$ als n-stellige Funktionen definierbar:

$$G_n := \left(\mathbb{R}_+^n \to \mathbb{R}_+; \vec{x} \mapsto \sqrt[n]{\prod\nolimits_{\nu=1}^{n} x_\nu}\right)$$

$$H_n := \left(\mathbb{R}_+^n \to \mathbb{R}_+; \vec{x} \mapsto \left(\tfrac{1}{n} \sum_{\nu=1}^{n} x_\nu^{-1} \right)^{-1} \right)$$

$$Q_n := \left(\mathbb{R}_+^n \to \mathbb{R}_+; \vec{x} \mapsto \sqrt{\tfrac{1}{n} \sum_{\nu=1}^{n} x_\nu^2} \right)$$

Es lässt sich dann zeigen, dass diese Funktionen ebenfalls die Eigenschaften (**In**), (**Sy**), (**St**), (**Ho**), (**Mo**), (**Re**) und (**As-m**) erfüllen, wenn wir darin A_n entsprechend ersetzen. Dagegen gilt (**Ad**) nicht, und so fragen wir nun nach Analoga zu diesem Axiom.

Beim arithmetischen Mittel werden n Zahlen addiert, und das Ergebnis wird dann durch n geteilt. Beim geometrischen Mittel werden n Zahlen miteinander multipliziert, und dann wird aus dem Produkt die n-te Wurzel gezogen. Eine *„operative Analogie"* zu (**Ad**) ist damit beim geometrischen Mittel wie folgt in (**Mu**) zu suchen – wobei links die „punktweise" (d. h. komponentenweise) Multiplikation der Vektoren gemeint ist:

- (**Mu**) *Multiplikativität*: $G_n(\vec{x} \cdot \vec{y}) = G_n(\vec{x}) \cdot G_n(\vec{y})$

Diese Forderung verallgemeinern wir zu

- (**Op**) *Operativität*: $M_n(\vec{x} \circledast \vec{y}) = M_n(\vec{x}) * M_n(\vec{y})$,

wobei wir schon am Beispiel (**Mu**) sehen, dass die durch $\circledast$ bzw. durch $*$ verknüpften Elemente aus verschiedenen Mengen stammen können, dass also zwei verschiedene Verknüpfungen vorliegen können, die dennoch irgendetwas miteinander zu tun haben müssen. Die uns vertrauten Operationen der Addition bzw. der Multiplikation konnten wir leicht in der Definition des arithmetischen bzw. des geometrischen Mittels wiederfinden – die möglicherweise existierenden, uns jedoch (bisher) unbekannten Operationen, die sich (vielleicht) hinter dem harmonischen bzw. dem quadratischen Mittel verstecken, bleiben uns [8] hingegen (noch!?) verborgen:

- ➢ Beim harmonischen Mittel werden ...? Beim quadratischen ...?

2.3 Gemeinsame Eigenschaften der betrachteten Mittelwertfunktionen

Den vier von uns initial betrachteten Mittelwertfunktionen sind also die Eigenschaften (**In**), (**Sy**), (**St**), (**Ho**), (**Mo**), (**Re**) und (**As-m**) gemein. Wir können nun nach weiteren, gar *allen* Funktionen fragen, die diese Eigenschaften haben. Der Kreisprozess des Axiomatisierens kommt in Fahrt: Schnell sieht man, dass gewichtete arithmetische Mittel (**Sy**) in der Regel nicht erfüllen, und man steht vor der Entscheidung, diese Eigenschaft von Mittelwertfunktionen evtl. dennoch zu fordern oder eben nicht.

- ➢ Und wollen wir, dass Mittelwertfunktionen stetig und/oder homogen sind?
- ➢ Wollen wir ...? Wollen Sie ...?

Wir kommen darauf wieder zurück.

3 Axiomensysteme für Mittelwertfunktionen

Entsprechend den beiden zu Beginn skizzierten, der Axiomatisierung zu Grunde liegenden Ideen geben wir nun mögliche *Produkte* des *Prozesses* „Axiomatisieren" an, also Axiomensysteme für spezielle bzw. allgemeine Mittelwertfunktionen.

3.1 Axiomensysteme für spezielle Mittelwertfunktionen

3.1.1 Das arithmetische Mittel

Wir beginnen mit zwei möglichen Axiomensystemen für das arithmetische Mittel.

3.1.1.1 Eine knappe Charakterisierung nach Teodoriu (1931)

In Abschnitt 2.1 haben wir gesehen, dass das arithmetische Mittel symmetrisch, reflexiv und additiv ist, und tatsächlich wird es, wie bereits erwähnt, schon durch diese drei (unabhängigen!) Axiome *charakterisiert* (vgl. [Bullen 2003, 436]); dabei genügt statt „symmetrisch“ sogar eine schwächere [9] Forderung:

- (**Zy**) *Zyklizität*: $M_n(\vec{x}) = M_n(Z(\vec{x}))$ mit einer zyklischen Verschiebung $Z := (\mathbb{R}_+^n \to \mathbb{R}_+^n; (x_1, \dots, x_n) \mapsto (x_n, x_1, \dots, x_{n-1}))$ für $n > 1$ und $Z := \mathrm{id}_{\mathbb{R}}$ für $n = 1$.

Damit können wir folgenden Satz formulieren:

> ***Satz**: Für $n \in \mathbb{N}$ sei $M_n := (\mathbb{R}_+^n \to \mathbb{R}_+; \vec{x} \mapsto M_n(\vec{x}))$.*
> *Erfüllt M_n die Eigenschaften **(Zy)**, **(Ad)** und **(Re)**, so gilt $M_n = A_n$.*

Der Beweis ist überraschend kurz (wenn man ihn dann endlich hat!): Für eine Funktion M_n, die (**Zy**), (**Ad**) und (**Re**) erfüllt, gilt nämlich unter Verwendung von Papier und Bleistift (P+B): [10]

$$n \cdot M_n(\vec{x}) = \sum_{\nu=1}^{n} M_n(\vec{x}) \underset{(\mathrm{Zy})}{=} \sum_{\nu=1}^{n} M_n(Z^{\nu-1}(\vec{x})) \underset{(\mathrm{Ad})}{=} M_n(\sum_{\nu=1}^{n} Z^{\nu-1}\vec{x}) \underset{(\mathrm{P}+B)}{=} M_n((\sum\nolimits_{\nu=1}^{n} x_\nu)_{\mu=1}^{n})$$
$$\underset{(\mathrm{Re})}{=} \sum\nolimits_{\nu=1}^{n} x_\nu = n \cdot A_n(\vec{x})$$

◆

Dies ist, „wie man leicht sieht“, die „bekannt“e, „offensichtlich“e und unverschämte [11], oft hinderliche Kürze mathematischer Literatur. Holen wir also zum Verstehen lieber etwas weiter aus: [12]

(1) Bei der Anwendung von (**Zy**) werden die n identischen Summanden $M_n(\vec{x})$ der Reihe nach durch $Z^0(M_n(\vec{x}))$, $Z^1(M_n(\vec{x}))$, ..., $Z^{n-1}(M_n(\vec{x}))$ ersetzt, wobei wie üblich $Z^0(M_n(\vec{x})) = \mathrm{id}(M_n(\vec{x})) = M_n(\vec{x})$ und $Z^\nu = Z \circ Z^{\nu-1}$ (Verkettung) ist.

(2) Die bisher nur für zwei Summanden formulierte Additivität (**Ad**) lässt sich induktiv auf n Summanden verallgemeinern, und so wurde sie hier angewendet.

(3) Bei der Anwendung von (P+B) schreibe man ausführlicher

$$\sum_{\nu=1}^{n} Z^{\nu-1}\vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} x_n \\ x_1 \\ \vdots \\ x_{n-1} \end{pmatrix} + \dots + \begin{pmatrix} x_2 \\ \vdots \\ x_n \\ x_1 \end{pmatrix},$$

um zu sehen, dass die zeilenweise Summation in allen Komponenten des Vektors identische Summen liefert. Dies erhellt uns zumindest die Rechnung – die Gewinnung des Ansatzes bleibt aber leider weiterhin im Dunklen.

3.1.1.2 Der Satz vom arithmetischen Mittel in axiomatischer Begründung

Mit Rudolf Schimmack beschreiten wir nun den Weg von einer abstrakt charakterisierten Funktion zum konkreten arithmetischen Mittel, das dann aufgrund der Axiome zwingend erscheint (vgl. [Schimmack 1909]). Für ihn besteht die Motivation zur Angabe eines Axiomensystems für die arithmetische Mittelwertfunktion darin, eine zufriedenstellende Erklärung für den *Satz vom arithmetischen Mittel* [13] zu finden.

Abweichend von den anderen Abschnitten betrachten wir hier nun Funktionen über allen reellen Zahlen, also von $\mathbb{R}^n$ in $\mathbb{R}$:

> Wenn für eine skalare Beobachtungsgröße x, die der Idee nach einen ganz bestimmten Wert hat, in Wirklichkeit als Ergebnis von n gleichgenauen Messungen die n Werte $x_1, \cdots, x_n$ vorliegen, so wird, wie die Ausgleichsrechnung lehrt, der *wahrscheinlichste Wert* von x durch das *arithmetische Mittel*
>
> $$\frac{x_1 + \cdots + x_n}{n}$$
>
> dargestellt. Man mag sich bisweilen begnügen, diesen Satz vom arithmetischen Mittel schlechthin aufzustellen und nur mehr oder weniger plausibel zu machen – befriedigender muß es sein, wenn man als wahrscheinlichsten Wert zunächst irgend eine eindeutige Funktion der n Werte $x_1, \cdots, x_n$
>
> $$f(x_1, \cdots, x_n)$$
>
> ansetzt, von dieser einige einleuchtende Grundeigenschaften als Axiome formuliert und daraus dann ableitet, daß die Funktion nur das arithmetische Mittel sein kann.
>
> *(aus [Schimmack 1909, 125])*

Damit schließt er an Arbeiten von Schiaparelli und Broggi an (vgl. [Schimmack 1909, 125 f.]), die bereits zuvor diesen Ansatz nutzten. Deren Beweise vereinfacht er. Schimmack möchte allerdings darüber hinaus „*hauptsächlich*“ die bei jenen zu findende

> fremdartige Forderung der Existenz stetiger Differentialquotienten vermeiden [...].
>
> *(aus [Schimmack 1909, 126])*

Sein System

> soll vielmehr nur solche Axiome enthalten, die dem inneren Charakter des Problems entsprechen! *(a. a. O.)*

Als ihm „einleuchtende Grundeigenschaften“ formuliert er – in Variation der von Schiaparelli und Broggi gelieferten – die folgenden (vgl. [Schimmack 1909, 126 ff.]):

> A x i o m I. *Der wahrscheinlichste Wert ist unabhängig von der Lage des Nullpunktes,* von dem aus die Beobachtungswerte und der wahrscheinlichste Wert gerechnet werden; [...] kurz:
>
> $$f(x_1 + h, \cdots, x_n + h) = f(x_1, \cdots, x_n) + h\,.$$
>
> [...]
>
> A x i o m III. *Der wahrscheinlichste Wert ist unabhängig von der Reihenfolge der Beobachtungswerte* [...]. [14]
>
> A x i o m II' *Der wahrscheinlichste Wert ist unabhängig von dem Richtungssinn der Skala* [...] kurz:
>
> $$f(-x_1, \cdots, -x_n) = -f(x_1, \cdots, x_n)\,.$$
>
> [...]

A x i o m IV'. *Wenn von n vorliegenden Beobachtungswerten der wahrscheinlichste Wert bereits bestimmt ist und man soll nach Hinzutreten eines weiteren Beobachtungswertes den wahrscheinlichsten Wert der nunmehrigen $n+1$ Beobachtungswerte bestimmen, so kann dies auch so geschehen, daß man statt jener n Beobachtungswerte ihren bereits bestimmten wahrscheinlichsten Wert – n-fach gezählt – nimmt.* Indem wir fortan bei der zu bestimmenden Funktion f, da wo es die Deutlichkeit verlangt, die Anzahl der Veränderlichen durch einen Index anzeigen, also

$$f(x_1, \cdots, x_n)$$

oder auch f_n, können wir Axiom IV' kurz so schreiben:

$$f_{n+1}(x_1, \cdots, x_{n+1}) = f_{n+1}(f_n, \cdots, f_n, x_{n+1}).$$

Hierbei benutzt Schimmack f_n nicht ganz glücklich sowohl als Funktionsnamen als auch als Funktionswert anstelle von $f(x_1, \ldots, x_n)$ bzw. von $f_n(x_1, \ldots, x_n)$.

Axiom I kommt in unserer Eigenschaftsliste oben nicht vor, Axiom II' ist eine Modifikation von (**Ho**) (nämlich für $\lambda = -1$), Axiom III ist (**Sy**), und Axiom IV' erweist sich im wesentlichen als (**As-m**) – und motiviert dieses natürlich wie oben versprochen!

Die einzige n-stellige Funktion von $\mathbb{R}^n$ in $\mathbb{R}$, die die von Schimmack geforderten Eigenschaften erfüllt, ist die für das arithmetische Mittel; dies sieht man wie folgt (vgl. [Schimmack 1909, 129 f.]):

Aus (I) und (II') folgt $f_n(x_0, \ldots, x_0) = x_0$ für $n \in \mathbb{N}^*$: Es ist mit (II') $f_n(0, \ldots, 0) = f_n(-0, \ldots, -0) = -f_n(0, \ldots, 0)$, also $f_n(0, \ldots, 0) = 0$ und damit mit (I) $f_n(x_0, \ldots, x_0) = f_n(0, \ldots, 0) + x_0 = x_0$; speziell ist $f_1(x_1) = x_1 = A_1(x_1)$. Aus (I), (II') und (III) folgt $2f_2(x_1, x_2) = f_2(0, x_2 - x_1) + x_1 + f_2(0, x_1 - x_2) + x_2 = x_1 + x_2$, also $f_2(x_1, x_2) = A_2(x_1, x_2)$.

Nun schließen wir mit Hilfe von (I) und (IV') induktiv weiter:

$$\begin{aligned} f_{n+1}(x_1, \ldots, x_{n+1}) &= f_{n+1}(0, \ldots, 0, x_{n+1} - f_n(x_1, \ldots, x_n)) + f_n(x_1, \ldots, x_n) \\ &= f_{n+1}(0, \ldots, 0, x_{n+1} - \frac{x_1 + \cdots + x_n}{n}) + \frac{x_1 + \cdots + x_n}{n} \end{aligned}$$

[Schimmack 1909, 130] schreibt weiter:

> Und nun wird alles darauf ankommen, die Funktion $f_{n+1}(0, \cdots, 0, x)$ zu bestimmen und zu zeigen, daß die rechte Seite dieser Gleichung gleich dem arithmetischen Mittel von $x_1, \cdots, x_{n+1}$ ist.

Wir betrachten dazu den Fall $x_1 = \ldots = x_{n-1} = 0$, haben also:

$$f_{n+1}(0, \ldots, 0, x_n, x_{n+1}) = f_{n+1}(0, \ldots, 0, x_{n+1} - \frac{x_n}{n}) + \frac{x_n}{n}$$

Gemäß (III) können wir z. B. x_n und x_{n+1} austauschen, und es folgt:

$$f_{n+1}(0, \ldots, 0, x_{n+1} - \frac{x_n}{n}) + \frac{x_n}{n} = f_{n+1}(0, \ldots, 0, x_n - \frac{x_{n+1}}{n}) + \frac{x_{n+1}}{n}$$

Jetzt betrachten wir hierbei den Fall $x_{n+1} = \dfrac{x_n}{n}$ und erhalten wegen $f_{n+1}(0, \ldots, 0) = 0$

$$\frac{x_n}{n} = f_{n+1}(0, \ldots, 0, x_n - \frac{x_n}{n^2}) + \frac{x_n}{n^2}$$

und damit schließlich

$$\frac{n-1}{n^2}x_n = f_{n+1}(0, \dots, 0, \frac{n^2-1}{n^2}x_n).$$

Für $x \in \mathbb{R}$ und $n > 1$ gilt also $f_{n+1}(0, \dots, 0, x) = \frac{x}{n+1}$, und wir können unseren induktiven Ansatz zu Ende führen:

$$f_{n+1}(x_1, \dots, x_{n+1}) = \frac{1}{n+1}(x_{n+1} - \frac{x_1 + \dots + x_n}{n}) + \frac{x_1 + \dots + x_n}{n} = \frac{x_1 + \dots + x_{n+1}}{n+1}$$

◆

Übrigens: Die Lektüre von Schimmacks Original-Arbeit, die in der heute immer noch sehr wichtigen und von den meisten mathematischen Instituten geführten Fachzeitschrift „Mathematische Annalen" erschienen ist, ist in der Schule – im Gegensatz zu den heute dort erscheinenden Artikeln – problemlos möglich. Gleiches gilt auch für den 1915 ebendort erschienenen Artikel von Ralph D. Beetle, in dem dieser die *vollständige Unabhängigkeit* von Schimmacks Axiomen nachweist. [15]

3.1.2 Mustergleiche Charakterisierungen des arithmetischen, des geometrischen, des harmonischen und des quadratischen Mittels

Edward V. Huntington hat sieben Axiomensysteme angegeben, die jeweils das arithmetische bzw. das geometrische bzw. das harmonische bzw. das quadratische Mittel charakterisieren (siehe [Huntington 1927] [16], vgl. auch [Bullen 2003, 437]). Die Axiomensysteme für diese verschiedenen Mittel haben oft eine gemeinsame Struktur. So gilt etwa der folgende

***Satz**: Für $n \in \mathbb{N}$ sei $M_n := (\mathbb{R}_+^n \to \mathbb{R}_+; \vec{x} \mapsto M_n(\vec{x}))$.*

*Erfüllt M_n die Eigenschaften (**As-2**), (**Re**) und (**Sy**), und gilt ferner $M_2(x_1, x_2) = X_2(x_1, x_2)$ für $X = A, G, H$ bzw. Q, dann ist $M_n = X_n$ für $X = A, G, H$ bzw. Q.*

Schauen wir uns nun Huntingtons Beweis für den Fall $X = A$ an. Er verwendet die folgende Notation: Er schreibt I für (**Sy**), II für (**As-2**), IV für (**Re**) und $A1$ für $m = f(x_1, x_2)$. Die Funktionen nennt er alle f, und er verzichtet meist auf den Index – ebenso bei x (s. u.). So finden wir z. B. (**As-2**) folgendermaßen eingeführt (vgl. [Huntington 1927, 2]):

II . $f(x_1, x_2, x_3, \cdots, x_n) = f(m, m, x_3, \cdots, x_n)$ where $m = f(x_1, x_2)$.

Und er erläutert sofort verbalisierend – statt strenger zu formalisieren (a. a. O.):

> That is, in computing the "f" of n quantities, we may replace the first pair, x_1, x_2, by the "f" of that pair, entered twice.

Auch andere Namen vergibt er doppelt: Die Menge der Eigenschaften $A1$, I, II, IV nennt er $A1$, analoges gilt für die anderen Mittel. Dennoch bleiben uns seine Beweise verständlich; [17] die intersubjektive Vermittlung mathematischer Vorstellungen als Ziel mathematischer Kommunikation bleibt trotz der laxen mathematischen Darstellung hier über diese möglich. [18]

Und weiter lesen wir bei [Huntington 1927, 9 f.]:

Let q be any positive quantity which is less than $(1/n)$th of the smallest of the x's. Then by II and $A1$,

$$f(x_1, x_2, \cdots, x_n) = f(q, [x_1 + x_2 - q], x_3, \cdots, x_n),$$

since each side equals $f(\frac{1}{2}(x_1 + x_2), \frac{1}{2}(x_1 + x_2), x_3, \cdots, x_n)$, and all the arguments are positive.

By successive applications of this result, in view of I, we have

$$\begin{aligned} f(x_1, x_2, \cdots, x_n) &= f(q, q, [x_1 + x_2 + x_3 - 2q], x_4, \cdots, x_n) \\ &= f(q, q, q, [x_1 + x_2 + x_3 + x_4 - 3q], x_5, \cdots, x_n) \\ &= f(q, q, \cdots, [x_1 + x_2 + \ldots + x_n - (n-1)q]) \end{aligned}$$

Now take $a = (x_1 + x_2 + \cdots + x_n)/n$. Then putting each x equal to a, $f(a, a, \cdots, a) = f(q, q, \cdots, [na-(n-1)q])$, by Postulate IV. But $x_1 + x_2 + \cdots + x_n = na$, so that

$$f(x_1, x_2, \cdots, x_n) = f(q, q, \cdots, [na - (n-1)q]).$$

Hence

$$f(x_1, x_2, \cdots, x_n) = (x_1 + x_2 + \cdots + x_n)/n.$$

Hence any function f which satisfies the postulates of set $A1$ must be identical with the arithmetic mean, A.

Beweise für die verbliebenen Fälle $X = G, H$ bzw. Q finden sich ebenfalls in [Huntington 1927].

3.2 Axiomensysteme für alle numerischen Mittelwertfunktionen

Wir haben oben eine Fülle gemeinsamer Eigenschaften der von uns betrachteten Mittelwertfunktionen angegeben. Aus jenen ist nun eine Auswahl zu treffen, um festzulegen, was *für uns* eine Mittelwertfunktion sein soll. *Wir* haben diese Freiheit, denn

> [...] there is no consensus in the literature [...] [Wassell 2002, 63].

Hier kann also jede bzw. jeder ihre bzw. seine eigene Mathematik gestalten, in der sie bzw. er ihre bzw. seine Mittelwertfunktionen definiert und erforscht, indem sie bzw. er ein abstraktes Axiomensystem festlegt und dazu passende konkrete Beispiele entwickelt, die wieder ein tieferes Verständnis des Axiomensystems ermöglichen, was die ursprüngliche Auswahl revidieren kann.

3.2.1 Mögliche Ansätze

Wassell beschränkt sich auf den Fall zweistelliger Funktionen und definiert „mean" nur durch *Internalität* in der etwas schwächeren Form $\min(x, y) \leq M(x, y) \leq \max(x, y)$. Dies ergänzend interessiert er sich für die Eigenschaft „strict": $M(x, y) = x$ oder $M(x, y) = y$ genau dann, wenn $x = y$. Und er listet weiter noch *Homogenität* und *Symmetrie* (vgl. [Wassell 2002, 63 f.]) .

Aber wir wissen bereits: Symmetrie verbietet gewichtete Mittel! Also: Wer Symmetrie will, muss auf gewichtete Mittel verzichten. Und umgekehrt: Wer gewichtete Mittel will, muss auf Symmetrie verzichten.

Oder: Für $f: \mathbb{R}_+ \to \mathbb{R}_+$ liefert $B_f(x, y) := (x \cdot f(x) + y \cdot f(y))/(f(x) + f(y))$ für $x, y \in \mathbb{R}_+$ das *Beckenbach-Gini-Mittel mit Erzeuger* f; A, G und H sind die einzigen homogenen

Beckenbach-Gini-Mittel (vgl. [Bullen 2003, 406 f.]). Wollen wir auf diese Weise viele Mittelwertfunktionen erzeugen? Oder wollen wir Homogenität (vgl. [Lambert & Herget 2004])? Oder ...?

Da kann es keinen allgemein richtigen Weg geben! *Finden Sie Ihren Eigenen!*

3.2.2 Ein minimalistischer Ansatz

In [Hischer & Lambert 2003] haben wir einen in der Schule gangbaren Weg skizziert, der die Erfahrung „Axiomatisierung" am Beispiel der Mittelwertfunktionen ermöglicht. Es genügt hierbei, zwei bis drei „Spielregeln" festzulegen und sich auf zweistellige Funktionen zu beschränken: [19]

Definition 1*: Es sei* $M : \mathbb{R}_+^2 \to \mathbb{R}_+$.

M *ist genau dann eine* ***Mittelwertfunktion****, wenn für alle* $x, y \in \mathbb{R}_+$ *gilt:*

(M1) $x < y \Rightarrow x < M(x,y) < y \;\wedge\; x < M(y,x) < y$ *(Ordnungshomomorphie)*

(M2) $M(x,x) = x$ *(Idempotenz)*

Den Funktionswert $M(x,y)$ *nennen wir in diesem Fall* ***Mittelwert*** *von* x *und* y.

Definition 2*: Es sei* M *eine Mittelwertfunktion.*

M *ist genau dann eine* ***kommutative Mittelwertfunktion****, wenn gilt:*

(M3) $M(y,x) = M(x,y)$ *für alle* $x, y \in \mathbb{R}_+$ *(Kommutativität)*

Mit der scharfen Ungleichung in (**M1**) haben wir uns das Leben (bewusst etwas) schwer gemacht. Die Bezeichnungen sind durch die ergänzende Betrachtung zweistelliger Mittelwertfunktionen als Verknüpfung $x *_M y := M(x,\, y)$ motiviert. Und wir haben vor allem eine effektive, durch das Chuquet-Mittel historisch verankerte [20] Maschine zur Erzeugung von konkreten Mittelwertfunktionen zur Verfügung:

Definition 3*: Es sei* $M : \mathbb{R}_+^2 \to \mathbb{R}_+$ *eine Funktion.*

M *ist genau dann eine* ***Chuquet-Funktion****, wenn eine „Streckfaktorfunktion"* S *mit* $S\colon \mathbb{R}_+^2 \to \mathbb{R}_+$ *existiert, so dass für alle* $x, y \in \mathbb{R}_+$ *gilt:*

$$M(x,y) = \frac{x + y \cdot S(x,y)}{1 + S(x,y)}$$

Satz 1*: Jede Chuquet-Funktion ist eine Mittelwertfunktion.*

Satz 2*: Jede Mittelwertfunktion ist eine Chuquet-Funktion.*

Weiterhin ergibt sich, dass eine beliebige Mittelwertfunktion M *(mit der zugehörigen Streckfaktorfunktion* S*) genau dann kommutativ ist, wenn* $S(x,y) \cdot S(y,x) = 1$ *für alle* $x, y \in \mathbb{R}_+$ *gilt, wobei diese Gleichung durch* $S(x,y) := f(x,y)/f(y,x)$ *mit beliebiger Funktion* f *gelöst wird. (Man vergleiche dies mit dem Beckenbach-Gini-Mittel von S. 75!)*

3.2.3 Kalkulieren mit abstrakten Mittelwertfunktionen

Am Ende unserer Ausführungen zur Axiomatisierung sollte nicht unerwähnt bleiben, dass – hat man sich schließlich für ein Axiomensystem entschieden und damit gearbeitet! – auch die abstrakten Mittelwertfunktionen selbst wieder zu konkreten Objekten werden, mit denen man (auf einer höheren Ebene) operiert. [21] Wir können z. B. nachrechnen, dass der Mittelwert zweier Mittelwerte wieder ein Mittelwert ist: Unsere Axiome, die für uns inzwischen genauso konkret sind (als Axiome) wie die Terme der „konkreten Mittelwertfunktionen“ (als Mittelwerte), werden auch von der (mit drei beliebigen axiomatisch definierten Mittelwertfunktionen M_1, M_2 und M_3) durch die Zuordnung

$$(x,y) \mapsto M_1(M_2(x,y), M_3(x,y))$$

gegebenen Funktion erfüllt.

Literatur

Beetle, Ralph D. [1915]: On the Complete Independence of Schimmack’s Postulates for the arithmetic Mean. In: *Mathematische Annalen*, **76**(1915), 444–446.

Bullen, P.S. [2003]: Handbook of Means and Their Inequalities. Dordrecht / Boston / London: Kluwer Academic Publishers.

Courant, Richard [1930]: Vorlesungen über Differential- und Integralrechnung. Erster Band: Funktionen einer Veränderlichen. Zweite, verbesserte Auflage. Berlin: Springer.

Fischer, Roland & Malle, Günter [1985]: Mensch und Mathematik. Mannheim: BI Wissenschaftsverlag.

Freudenthal, Hans [1963]: Was ist Axiomatik, und welchen Bildungswert kann sie haben? In: *Der Mathematikunterricht,* **9**(1963)4, 5–29

Henze, Norbert & Stummer, Wolfgang [2004]: Mittelwerte und Mitten in der Stochastik. In diesem Heft, 18–29.

Hischer, Horst [2002]: Mathematikunterricht und Neue Medien. Hildesheim / Berlin: Franzbecker.

Hischer, Horst [2004 a]: Mittenbildung als fundamentale Idee. In diesem Heft, 4–13.

Hischer, Horst [2004 b]: Mittelwertfolgen – oder: Mitten inmitten von Mitten. In diesem Heft, 42–54.

Hischer, Horst & Lambert, Anselm [2003]: Was ist ein numerischer Mittelwert? – Zur axiomatischen Präzisierung einer fundamentalen Idee. In: *mathematica didactica*, **26**(2003)1, 3–42.

Huntington, Edward V. [1927]: Sets of independent postulates for the arithmetic mean, the geometric mean, the harmonic mean an the root-mean-square. In: *Transactions of the American Mathematical Society*, **29**(1927), 1–22.

Lambert, Anselm & Herget, Wilfried [2004]: Mächtig viel Mittelmaß in Mittelwert-Familien. In diesem Heft, 55–66.

Schimmack, Rudolf [1909]: Der Satz vom arithmetischen Mittel in axiomatischer Begründung. In: *Mathematische Annalen*, **68**(1909), 125 – 134; Erratum: 304.

Wassell, Stephen R. [2002]: Rediscovering a Family of Means. In: *The Mathematical Intelligencer*, **24**(2002)2, 58–65.

Anmerkungen

1 Aber auch das bekannte, als gemeinsamer Grenzwert von zwei rekursiv definierten „Mittelwertfolgen“ gegebene arithmetisch-geometrische Mittel (vgl. [Hischer 2004 b, 54]) ist ein konkreter Mittelwert (nähere Informationen z. B. bei [Courant 1930, 34 f.] – es ist überaus lohnend, diese wunderschöne *präbourbakische* Einführung in die Analysis auch einmal unter dem Aspekt „Darstellung von Mathematik“ zu lesen: Wie rigoros müssen Formeln sein, und welche Sprache wird zwischen den Formeln verwendet?).

2 Vgl. [Hischer 2002, 149]. Kurz kann man sagen: Die Dinge sind das, was sie *nicht nicht* sind!

3 Schimmack beginnt also „rechts“ im von uns beschriebenen Kreisprozess.

4 „[...] theoretische Begriffe [...] sind [...] Ausdruck eines bestimmten Wollens; Ausdruck dessen, daß uns ein bestimmter Gesichtspunkt wichtig ist.“ [Fischer & Malle 1985, 151]

5 Würden wir hier $\mathbb{R}$ anstelle von $\mathbb{R}_+$ verwenden, so läge mit M_n eine Abbildung des $\mathbb{R}$-Vektorraums $\mathbb{R}^n$ in seinen Skalarenkörper $\mathbb{R}$ vor (ein „Funktional"). Zwar bildet $\mathbb{R}_+^n$ keinen Vektorraum, aber immerhin eine Unterstruktur eines Vektorraums (es liegt der Spezialfall eines sog. „Kegels" vor), so dass wir die Elemente aus $\mathbb{R}_+^n$ dennoch im üblichen Sinne „Vektoren" nennen können und sie mit $\vec{x}, \vec{y}, \ldots$ bezeichnen.

6 Es ist in der Mathematik oft sinnvoll (so auch hier), triviale Fälle (hier: A_1) mitzuführen!

7 Vgl. [Hischer 2004 a, 8] in diesem Heft.

8 Vielleicht sehen Sie, verehrte Leserin, verehrter Leser, mehr!?

9 So ist z. B. die folgende Funktion zwar zyklisch, aber nicht symmetrisch:
$f := (\mathbb{R}_+^4 \to \mathbb{R}_+; (x_1, x_2, x_3, x_4) \mapsto \frac{1}{4}(\sqrt{x_1 x_2} + \sqrt{x_2 x_3} + \sqrt{x_3 x_4} + \sqrt{x_4 x_1}))$.

10 Analog ist eine Funktion, die (**Zy**), (**Mu**) und (**Re**) erfüllt, das geometrische Mittel, und erneut steht die Frage im Raum: Wie sieht eine geeignete Operation für das harmonische bzw. für das quadratische Mittel aus?

11 Wenn ein Mathematiker „von den Überlegungen, die ihn zum Ziele führten, etwas veröffentlichte, käme er sich vor, als stände er in der Unterhose auf der Straße." [Freudenthal 1963, 16]

12 Die Entfaltung komprimierter mathematischer Information selbst zu vollziehen, kann bei genügender Muße aber auch Quell von Freude sein.

13 Vgl. auch das Stichprobenmittel bei [Henze & Stummer 2004].

14 Hierzu gibt der Autor keine „Kurzschreibweise" (wie in seinen anderen Axiomen) an.

15 Ein Axiomensystem heißt *unabhängig*, wenn keines der Axiome aus den anderen folgt. Dies bedeutet aber nicht, dass es gar keine Interrelationen zwischen den Axiomen gibt: So kann z. B. aus dem Nicht-Erfülltsein eines Axioms dennoch eines der anderen folgen! Deshalb hat Eliakim Hastings Moore noch den schärferen Begriff der „vollständigen Unabhängigkeit" eingeführt. Ein Axiomensystem heißt *vollständig unabhängig*, wenn keines der Axiome oder seine Negation aus einer Kombination der anderen Axiome und denen ihrer jeweiligen Negation folgt. Um die Unabhängigkeit eines Axiomensystems aus n Axiomen zu zeigen, genügt es, n Beispiele anzugeben (die jeweils alle, bis auf eine Eigenschaft erfüllen), bei Schimmacks Axiomen also vier. Um die vollständige Unabhängigkeit zu demonstrieren, sind allerdings 2^n Beispiele erforderlich (eines für jede mögliche Kombination aus Axiomen und ihren Verneinungen), bei Schimmack also 16.

16 Huntington hält das arithmetische und das quadratische Mittel für "most important".

17 Auf Seite 1 des Artikels finden wir die Fußnote: "EDITOR'S NOTE: The typography in this paper and the succeeding papers has been altered in various respects from the form originally proposed by the authors, in an effort to adapt mathematical composition to the monotype machine." (Heute würde man wohl sagen: „Wenn Sie's nicht $\mathrm{t_E}\chi$en können, dann müssen Sie's eben anders beweisen!") Dennoch ist im Fall der indexverarmten „f"s und „x"e „Absicht des Autors" die einzig mögliche Interpretation; im Fall des doppelten $A1$ (und der analogen doppelten Notationen) wäre noch denkbar, wenn auch nicht zwingend und, wie wir vermuten, eher unwahrscheinlich, dass Huntington im Manuskript verschiedene Schrifttypen gewählt hatte, die nun verlorengegangen sind.

18 Eine wichtige im Mathematikunterricht zu erwerbende Fähigkeit besteht darin, zu wissen, wann welche Strenge in der Darstellung sinnvoll ist, und dies auch entsprechend umsetzen zu können.

19 Diese Definitionen und Sätze finden sich in [Hischer & Lambert 2003, 9–14].

20 Vgl. [Hischer 2004 a].

21 Was ein Mensch als konkretes oder abstraktes Objekt erlebt, hängt sicherlich auch von seiner persönlichen Erfahrung ab. Für einen Funktionalanalytiker kann ein durch seine abstrakten Axiome gegebener Banachraum sehr konkret – und mit Leben gefüllt! – sein.