Funktionalgleichungen

Thomas Huber

24. April 2009

Inhaltsverzeichnis

Was ist eine Funktionalgleichung?

Eine Funktionalgleichung ist eine Identität, die eine bestimmte Funktion erfüllt. Zum Beispiel gilt für die Funktion $f(x) = x^2$ offenbar

$$f(x)f(y) = f(xy) \qquad \forall\, x, y \in \mathbb{R}, \tag{1}$$

denn für beliebige reelle Zahlen x, y ist $f(x)f(y) = x^2y^2 = (xy)^2 = f(xy)$ nach den Potenzregeln. Eine weitere Potenzregel wird durch die Funktionalgleichung der Exponentialfunktion ausgedrückt: $\exp(x+y) = \exp(x)\exp(y)$. Das bedeutet aber gerade, dass $f(x) = \exp(x)$ eine *Lösung* ist der Funktionalgleichung

$$f(x)f(y) = f(x+y) \qquad \forall\, x, y \in \mathbb{R}, \tag{2}$$

Eine Funktion ist eine Lösung einer Funktionalgleichung, z.B. von (2), wenn stets das Gleichheitszeichen steht, *egal* welche reellen Zahlen man für x und y einsetzt. Hat man eine Funktionalgleichung vor sich, dann besteht die Aufgabe darin, alle Lösungsfunktionen zu bestimmen. Manchmal ist man auch nur an Lösungen mit bestimmten Eigenschaften interessiert, wie zum Beispiel monoton steigenden oder stetigen Funktionen. Oft braucht man auch zusätzliche Voraussetzungen über die gesuchte Funktion, um sinnvolle Aussagen machen zu können. Wir werden später zum Beispiel zeigen, dass die einzigen *stetigen* rellen Funktionen, die die Gleichung

$$f(x) + f(y) = f(x + y) \qquad \forall\, x, y \in \mathbb{R}, \tag{3}$$

erfüllen, lineare Funktionen sind. Es gibt aber in der Tat eine Unmenge weiterer Lösungen von (3), die alle sehr unstetig sind und extrem hüpfen. Sie interessieren uns hier aber nicht, da wir einfach nicht die Mittel besitzen, sie zu untersuchen. Die drei Beispiele, die wir bis jetzt gesehen haben, sehen alle recht ähnlich aus. Sie sind auch tatsächlich alle sehr verwandt miteinander. Die Gleichung (1) bis (3) heissen CAUCHY-Funktionalgleichungen und wir werden uns später ausführlich mit ihnen beschäftigen.

Es gibt jedoch auch völlig andere Funktionalgleichungen. Sie können eine oder mehrere Variablen haben. Es können einzelne Gleichungen sein oder Systeme von Gleichungen für eine oder mehrere Funktionen. Die Funktionen müssen auch nicht immer reelle sein, der Definitionsbereich ist oft $\mathbb{Q}$ oder $\mathbb{N}$. Das folgende Beispiel war die 2. Aufgabe an der IMO 92:
Finde alle Funktionen $f : \mathbb{R} \to \mathbb{R}$, sodass gilt

$$f(x^2 + f(y)) = y + f(x)^2 \qquad \forall\, x, y \in \mathbb{R}.$$

Diese Aufgabe stammt von der IMO 93:
Entscheide, ob es eine Funktion $f : \mathbb{N} \to \mathbb{N}$ gibt, mit den folgenden Eigenschaften:

(a) $f(1) = 2$

(b) $f(f(n)) = f(n) + n \qquad \forall\, n \in \mathbb{N}$

(c) $f(n) < f(n + 1) \qquad \forall\, n \in \mathbb{N}$

Diese auf den ersten Blick etwas wirre Aufgabe war die Nummer 5 an der IMO 2002:
Bestimme alle Funktionen $f : \mathbb{R} \to \mathbb{R}$, sodass für alle reellen Zahlen x, y, z, t gilt

$$(f(x) + f(z))(f(y) + f(t)) = f(xy - zt) + f(xt + yz).$$

Das letzte Beispiel zeigt, dass f auch mehrere Argumente haben kann, es war die Nummer 6 an der IMO 81:
Die Funktion $f(x, y)$ erfülle die folgenden Gleichungen für alle $x, y \geq 0$:

$$\begin{aligned} f(0, y) &= y + 1, \\ f(x + 1, 0) &= f(x, 1), \\ f(x + 1, y + 1) &= f(x, f(x + 1, y)). \end{aligned}$$

Bestimme $f(4, 1981)$.

Nach diesem Rundblick sollte klar geworden sein, was die typische Problemstellung ist. Eine ganz andere Frage ist allerdings, wie man nun vorgeht, um die Lösungen einer solchen Funktionalgleichung zu bestimmen. Oft sieht man recht schnell eine Lösung und ist sich auch recht sicher, dass es die einzige ist. Dies aber wirklich zu beweisen, ist die Hauptschwierigkeit. Es gibt nur ganz wenige Standartverfahren, die man anwenden kann. Meistens muss man einfach herumprobieren, mit der Gleichung spielen, bis sie Informationen preisgibt, die man weiterverwenden kann. Das wichtigste überhaupt beim Lösen solcher Funktionalgleichungen ist die Erfahrung. Erst mit einiger Übung wird man erkennen, worauf es eigentlich ankommt, wonach man suchen muss. Dieses Skript soll euch diese Erfahrung geben. Es enthält viele ausführliche Lösungen zu Beispielen und viele Aufgaben, an denen ihr euer Können unter Beweis stellen könnt.

1 Grundlegende Lösungstechniken

1.1 Einsetzen

Dieser Abschnitt behandelt die wohl naheliegendste Lösungstechnik, das *Einsetzen*. Denn wenn eine Funktion die gegebene Funktionalgleichung erfüllt, dann muss diese Identität immer gelten, egal was man für die Variablen einsetzt. Dies kann man ausnutzen, indem man geeignete Dinge einsetzt, sodass die Gleichung einfacher wird oder sonst irgendwie brauchbarer.

Beispiel 1. *Finde alle Funktionen $f : \mathbb{R} \to \mathbb{R}$, sodass für alle $x, y \in \mathbb{R}$ gilt*

$$f(xy) = xf(x) + yf(y).$$

Lösung. Wir setzen $x = y = 0$ dann folgt $f(0) = 0$. Sei nun $x \neq 0$ beliebig und setze $y = 0$. Dann folgt $0 = f(0) = xf(x)$. Wegen $x \neq 0$ gilt also $f(x) = 0$. Die einzige mögliche Lösung dieser Funktionalgleichung ist also die Nullfunktion $f(x) = 0 \quad \forall x \in \mathbb{R}$. Sie ist aber in der Tat eine Lösung, wie man sofort sieht.

□

Im nächsten Beispiel ist eine Funktion gesucht, die zwei Argumente besitzt. Das grundsätzliche Vorgehen bleibt genau dasselbe.

Beispiel 2. *Bestimme alle Funktionen $f : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$, die folgende Bedingungen erfüllen:*

(a) $f(x + u, y + u) = f(x, y) + u$ *für alle* $x, y, u \in \mathbb{R}$,

(b) $f(xu, yu) = f(x, y)u$ *für alle* $x, y, u \in \mathbb{R}$.

Lösung. Sei zuerst $x \neq y$. Setze $u = -x$ in (a), dann folgt $f(x,y) = f(0, y-x)+x$. Verwende nun (b) mit $u = y-x \neq 0$, das liefert $f(0, y-x)+x = f(0,1)(y-x)+x$. Mit der Bezeichnung $d = f(0,1)$ haben wir also gezeigt, dass gilt $f(x,y) = (1-d)x + dy$.
Sei nun $x = y$. Mit (a) folgt $f(x,x) = f(0,0) + x$ und mit (b) und $u = 0$ erhalten wir $f(0,0) = f(0,0) \cdot 0 = 0$. Folglich also $f(x,x) = x = (1-d)x + dx$. Alle Lösungen sind damit von der Form

$$f(x,y) = (1-d)x + dy$$

mit einer reellen Zahl d. Man verifiziert nun leicht, dass dies tatsächlich Lösungen sind.

□

Beispiel 3. *(CH 98) Finde alle Funktionen $f : \mathbb{R} \to \mathbb{R}$, sodass für alle $x, y \in \mathbb{R}$ gilt*

$$f(f(x) + y) = f(x^2 - y) + 4f(x)y.$$

1. Lösung. Zuerst setzen wir $x = 0$, das ergibt

$$f(f(0) + y) = f(-y) + 4f(0)y. \tag{4}$$

Wir versuchen nun, so einzusetzen, dass die komplizierten Terme möglichst einfach werden. Die linke Seite wird einfach, wenn $y = -f(x)$ gilt:

$$f(0) = f(x^2 + f(x)) - 4f(x)^2. \tag{5}$$

Um den ersten Term rechts zu vereinfachen, setzen wir $y = x^2$:

$$f(f(x) + x^2) = f(0) + 4f(x)x^2. \tag{6}$$

Nun fällt auf, dass in (5) und (6) zwei der drei Terme dieselben sind! Kombination dieser beiden Gleichungen führt auf

$$4f(x)^2 = f(f(x) + x^2) - f(0) = 4f(x)x^2,$$

also gilt für alle $x \in \mathbb{R}$ die Identität $f(x)^2 = x^2 f(x)$. Daraus folgt erstens $f(0) = 0$ und (4) vereinfacht sich zu

$$f(y) = f(-y), \tag{7}$$

also ist f eine *gerade* Funktion. Ausserdem gilt für $x \neq 0$ entweder $f(x) = 0$ oder $f(x) = x^2$. Wir zeigen nun, dass immer stets die eine oder andere Alternative gilt. Dazu nehmen wir an, dass es verschiedene Zahlen $a, b \neq 0$ gibt mit $f(a) = 0$ und $f(b) = b^2$. Aus (7) folgt dann $f(-b) = b^2 = (-b)^2$. Wir setzen nun $x = a$ und $y = b$ in die Gleichung ein und erhalten $b^2 = f(a^2 - b)$. Ausserdem wissen wir ja, dass $f(a^2 - b)$ gleich 0 oder gleich $(a^2 - b)^2$ ist. Im ersten Fall folgt $b = 0$, Widerspruch. Im zweiten Fall folgt $b^2 = (a^2 - b)^2$, also $a^2(a^2 - 2b) = 0$. Wegen $a \neq 0$ muss daher $a^2 = 2b$ gelten. Wiederholen wir die genau gleiche Argumentation mit $x = a$ und $y = -b$, dann ergibt sich analog $a^2 = -2b$, was wieder auf $a = 0$ führt, Widerspruch.
Damit haben wir gezeigt, dass entweder für alle $x \in \mathbb{R}$ gilt $f(x) = 0$ oder es gilt für alle $x \in \mathbb{R}$ die andere Alternative $f(x) = x^2$. Es bleibt noch nachzurechnen, dass diese beiden Funktionen tatsächlich Lösungen sind. Dies sei euch überlassen.

□

2. Lösung. In der ersten Lösung haben wir beim Einsetzen darauf geachtet, wie man die komplizierten Terme möglichst einfach macht. Es gibt einen anderen Ansatz, der machmal Erfolg bringt: Kann man irgend etwas einsetzen, sodass sich zwei Terme direkt wegheben? In diesem Beispiel könnte man versuchen, y so zu wählen, dass die Terme $f(f(x)+y)$ und $f(x^2-y)$ gleich werden. Dazu muss dann $f(x)+y=x^2-y$ gelten, also $y=(x^2-f(x))/2$. Einsetzen liefert nun für alle x die Gleichung

$$0 = 4f(x)\cdot\frac{x^2-f(x)}{2}.$$

Daraus folgt direkt $f(x)=0$ oder $f(x)=x^2$. Man vervollständigt die Lösung jetzt genauso wie vorher.

□

1.2 Doppelt berechnen

Manchmal nützt alles Einsetzen nichts, die Gleichung lässt sich nicht wesentlich vereinfachen. In diesem Fall kann man versuchen, einen festen Ausdruck auf zwei grundsätzlich verschiedene Arten zu berechnen. Gleichsetzen der Resultate führt dann oft zu ganz neuen und nützlicheren Gleichungen.

Die Funktionalgleichung im nächsten Beispiel enthält nur eine freie Variable. Solche Gleichungen sind prinzipiell schwierig, denn man kann ja kaum etwas sinnvolles Einsetzen.

Beispiel 4. *Finde alle Funktionen $f:\mathbb{R}\to\mathbb{R}$, die folgende drei Gleichungen erfüllen:*

(a) $f(-x)=-f(x)$ *für alle $x\in\mathbb{R}$,*

(b) $f(x+1)=f(x)+1$ *für alle $x\in\mathbb{R}$,*

(c) $f\left(\dfrac{1}{x}\right)=\dfrac{f(x)}{x^2}$ *für $x\neq 0$.*

Lösung. Aus (a) folgt mit $x=0$ sofort $f(0)=0$. Aus (b) erhält man mit $x=-1$ den Wert $f(-1)=-1$. Nun suchen wir nach einer Grösse, die wir mit (b) und (c) auf zwei verschiedene Arten berechnen können. Also eine Grösse, die wir sowohl in der Form $a+1$ wie auch in der Form $1/b$ schreiben können. Natürlich gibt es dafür unzählige Möglichkeiten, aber wir wollens ja so einfach wie möglich haben, also betrachten wir mal $\frac{1}{x}+1$ für $x\neq 0,-1$. Einerseits folgt mit (b) und (c)

$$f\left(\frac{1}{x}+1\right)=1+f\left(\frac{1}{x}\right)=1+\frac{f(x)}{x^2}. \tag{8}$$

Andererseits aber mit (c) auch

$$f\left(\frac{1}{x}+1\right)=f\left(\frac{1}{x/(x+1)}\right)=\frac{f(x/(x+1))}{(x(x+1))^2}.$$

Weiter folgt wieder aus (b) und (c)

$$\begin{aligned} f\left(\frac{x}{x+1}\right) &= f\left(1-\frac{1}{x+1}\right) \\ &= 1-f\left(\frac{1}{x+1}\right) \\ &= 1-\frac{f(x+1)}{(x+1)^2} \\ &= \frac{(x+1)^2-1-f(x)}{(x+1)^2}. \end{aligned}$$

Insgesamt folgt also

$$f\left(\frac{1}{x}+1\right) = \frac{(x+1)^2-1-f(x)}{x^2}. \tag{9}$$

Ein Vergleich von (8) und (9) liefert nun sofort

$$f(x) = x \qquad \forall\, x \neq 0, -1.$$

Wir wissen aber bereits, dass $f(0) = 0$ und $f(-1) = -1$ gilt, daher ist $f(x) = x$ die einzige Lösung, wie man leicht nachrechnet.

□

Eine weitere wichtige Situation, wo man doppelt ausrechnen kann, ist folgende: Hat man eine Funktionalgleichung, sodass die eine Seite eine gewisse *Invarianz* aufweist, dann gilt das auch für die andere Seite. Als Beispiele könnte man sich vorstellen, dass die linke Seite gleich bleibt, wenn man x mit y vertauscht. Oder dass sie gleich bleibt, wenn man x durch $-x$ oder durch $f(x)$ ersetzt. Dasselbe muss dann auch für die rechte Seite gelten. Ein Vergleich der 'alten' und 'neuen' rechten Seite liefert dann eine neue Gleichung.

Beispiel 5. *(CSO 01) Finde alle Funktionen $f : \mathbb{R} \to \mathbb{R}$, sodass für alle $x, y \in \mathbb{R}$ gilt*

$$f(x^2 + f(y)) = (x-y)^2 \cdot f(x+y).$$

Lösung. Die linke Seite ist eine *gerade* Funktion in x, also auch die rechte. Folglich ändert sich die rechte Seite nicht, wenn wir x durch $-x$ ersetzen. Dies liefert die Gleichung

$$(x-y)^2 \cdot f(x+y) = (x+y)^2 \cdot f(y-x).$$

Nun wollen wir für x und y so substituieren, dass $x+y$ alle reellen Werte annehmen kann, während $y-x$ konstant bleibt, zum Beispiel gleich 1. Dies erreichen wir, indem wir $x = \frac{1}{2}(t-1)$ und $y = \frac{1}{2}(t+1)$ setzen, wobei t eine beliebige reelle Zahl ist. Die Gleichung wird zu $f(t) = t^2 \cdot f(1)$ für alle $t \in \mathbb{R}$. Alle Lösungen sind also von der Form $f(x) = cx^2$ mit einer reellen Konstanten c. Einsetzen in die ursprüngliche Gleichung zeigt, dass dies nur für $c = 0$ und $c = -1$ funktioniert. Die Lösungen dieser Funktionalgleichung lauten also $f(x) = 0$ und $f(x) = -x^2$.

□

1.3 Funktionalgleichungen auf $\mathbb{Z}$ und $\mathbb{Q}$

Ist eine Funktionalgleichung gegeben, deren Lösungen Funktionen sind, die auf den ganzen oder rationalen Zahlen definiert sind, kommen zusätzliche Methoden ins Spiel. Man kann manchmal zahlentheoretische Methoden verwenden, um solche Gleichungen zu lösen. Die wichtigste neue Technik ist aber sicher die vollständige Induktion.

Beispiel 6. *Finde alle Funktionen $f : \mathbb{Z} \to \mathbb{Z}$, sodass für alle $m, n \in \mathbb{Z}$ gilt*

$$f(m+n) + f(mn) = f(m)f(n) + 1.$$

Lösung. Mit $n = 1$ folgt

$$f(m+1) + f(m) = f(1)f(m) + 1, \quad \forall m \in \mathbb{Z}. \tag{10}$$

Die Gleichung (10) liefert eine Beziehung zwischen den beiden Funktionswerten $f(m)$ und $f(m+1)$. Wenn der eine bekannt ist, lässt sich der andere berechnen. Allerdings müssen wir dafür zuerst $f(1)$ bestimmen, von diesem einen Wert hängt alles ab. Setze dazu $m = 1$ und $n = -1$, dann folgt $f(-1)(f(1) - 1) = 0$. Also ist entweder $f(1) = 1$ oder $f(-1) = 0$. Wir nehmen nun das zweite an. Einsetzen von $m = n = -1$ liefert $f(-2) + f(1) = 1$, Einsetzen von $m = -2$, $n = 1$ ergibt $f(-2) = f(-2)f(1) + 1$. Eliminiert man $f(-2)$ aus diesen Gleichungen, dann folgt $(f(1) - 1)^2 = 1$, also ist $f(1) = 0$ oder $f(1) = 2$. Wir betrachten nun die drei Möglichkeiten für $f(1)$ getrennt.

1. $f(1) = 2$. Die Gleichung (10) wird dann zu $f(m + 1) = f(m) + 1$. Aus $f(1) = 2$ berechnet man damit der Reihe nach $f(2) = 3$, $f(3) = 4$ und so weiter. Die Vermutung ist naheliegend, dass $f(n) = n + 1$ gilt für alle $n \in \mathbb{Z}$. Dies beweisen wir nun mit Induktion. Es ist richtig für $n = 1$ und gelte für $n \geq 1$. Setze $m = n$ in der Gleichung, dann folgt $f(n + 1) = f(n) + 1 = (n + 1) + 1$, wie gewünscht. Nun verwenden wir Rückwärtsinduktion, um dies auch für alle $n \leq 0$ zu zeigen. Es gelte für $n \leq 1$, setze $m = n-1$ in der Gleichung, dann folgt $f(n-1) = f(n)-1 = (n+1)-1 = (n-1)+1$. Damit ist alles gezeigt.

2. $f(1) = 1$. Die Gleichung (10) lautet $f(m+1) = 1$ für alle $m \in \mathbb{Z}$. Folglich ist f konstant 1 in diesem Fall.

3. $f(1) = 0$. Nun lautet (10) wie folgt: $f(m + 1) + f(m) = 1$. Aus $f(1) = 0$ folgt daraus der Reihe nach $f(2) = 1$, $f(3) = 0$, $f(4) = 1$ usw. Wir behaupten, dass $f(n) = 0$ ist, wenn n ungerade ist, und $f(n) = 1$, wenn n gerade ist. Dazu verwenden wir wieder zweimal Induktion. Dies ist richtig für $n = 1$ und gelte für $n \geq 1$. Ist n ungerade, dann folgt $f(n + 1) = 1 - f(n) = 1 - 0 = 1$, ist n gerade, dann ergibt sich $f(n + 1) = 1 - f(n) = 1 - 1 = 0$, wie behauptet. Wir nehmen nun an, dies gelte für $n \leq 1$. Ist n ungerade, dann folgt $f(n - 1) = 1 - f(n) = 1 - 0 = 1$, ist n gerade, dann ist $f(n - 1) = 1 - f(n) = 1 - 1 = 0$. Damit ist der Beweis beendet.

Es bleibt noch nachzurechnen, dass die drei gefundenen Funktionen in der Tat Lösungen der ursprünglichen Funktionalgleichung sind. Dies sei dem Leser überlassen (im 3. Fall sind Fallunterscheidungen nötig).

□

Im nächsten Beispiel geht entscheidend ein, dass jede natürliche Zahl eine eindeutige Primfaktorzerlegung besitzt. Ausserdem zeigt es, dass eine Funktionalgleichung sehr viele Lösungen haben kann.

Beispiel 7. *Finde alle Funktionen $f : \mathbb{N} \to \mathbb{N}$, sodass für alle $m, n \in \mathbb{N}$ gilt*

$$f(mn) = f(m)f(n).$$

Lösung. Mit $m = n = 1$ folgt $f(1) = 1$. Mit vollständiger Induktion nach k zeigt man ausserdem leicht, dass für beliebige natürliche Zahlen $n_1, \ldots, n_k$ gilt

$$f(n_1 \cdots n_k) = f(n_1) \cdots f(n_k). \tag{11}$$

Sei nun $n > 1$ eine natürliche Zahl mit Primfaktorzerlegung $n = p_1^{a_1} \cdots p_r^{a_r}$. Aus (11) folgt dann

$$f(n) = f(p_1)^{a_1} \cdots f(p_r)^{a_r},$$

insbesondere ist $f(n)$ durch die Werte $f(p_k)$ auf den Primfaktoren bereits eindeutig bestimmt. Aber welche Bedingungen müssen die Funktionswerte auf den Primzahlen erfüllen? Gar keine! Man kann diese Werte beliebig vorgeben. Wähle für jede Primzahl p eine beliebige natürliche Zahl c_p, und definiere $f(n)$ für $n = p_1^{a_1} \cdots p_r^{a_r}$ durch

$$f(n) = (c_{p_1})^{a_1} \cdots (c_{p_r})^{a_r},$$

dann ist f eine Lösung der gegebenen Funktionalgleichung. Für den Beweis verwendet man wieder die Eindeutigkeit der Primfaktorzerlegung.

□

Beispiel 8. *Finde alle Funktionen $f : \mathbb{Q} \to \mathbb{Q}$, sodass für alle $x, y \in \mathbb{Q}$ gilt*

$$f(x + y) + f(x - y) = 2f(x) + 2f(y).$$

Lösung. Mit $x = y = 0$ folgt $f(0) = 0$. Setze $y = x$, dann folgt $f(2x) = 4f(x)$. Setze $y = 2x$, dann folgt $f(3x) = 9f(x)$. Eine einfache Induktion zeigt nun, dass für alle natürlichen Zahlen n und alle rationalen x gilt

$$f(nx) = n^2 f(x). \tag{12}$$

Ausserdem folgt mit $x = 0$ sofort $f(-y) = f(y)$, daher gilt (12) für alle ganzen Zahlen n. Sei nun $x = p/q$ eine beliebige rationale Zahl mit $p \in \mathbb{Z}$ und $q \in \mathbb{N}$. Aus (12) folgt

$$q^2 f(x) = f(qx) = f(p) = p^2 f(1),$$

also $f(x) = f(1)x^2$. Einsetzen zeigt, dass die Funktionen wirklich $f(x) = cx^2$ alles Lösungen der Gleichung sind.

□

1.4 Die Lösungsgesamtheit einer Funktionalgleichung

Ganz wichtig bei Funktionalgleichungen ist es, gleich am Anfang mögliche Lösungen zu erraten. Dies kann nämlich die Strategie, mit der man an eine Aufgabe geht, stark beeinflussen. Man sollte die Gleichung auf jeden Fall auf konstante, lineare und quadratische Lösungen testen. Etwas allgemeiner (aber sehr viel aufwändiger) ist der Test auf beliebige Funktionen der Form $f(x) = ax^2 + bx + c$. Auch nach Lösungen der Form $f(x) = \pm x^n$ kann man suchen. Oft hat man eine Lösung gefunden und ist sich sehr sicher, dass es die einzige ist. In diesem Fall kann man direkt versuchen zu zeigen, dass die Gleichung höchstens eine einzige Lösung haben kann. Wir geben ein Beispiel.

Beispiel 9. *Bestimme alle Lösungen $f : \mathbb{R} \to \mathbb{R}$ der Funktionalgleichung*

$$f(xy) + x^2 + y^2 = xy + f(x^2) + f(y^2) \qquad \forall\, x, y \in \mathbb{R}.$$

1. Lösung. Es ist leicht zu sehen, dass $f(x) = x$ eine Lösung ist. Und eine andere Lösung lässt sich auf den ersten Blick nicht erkennen. Wir zeigen nun, dass die Funktionalgleichung in der Tat höchstens eine Lösung besitzt. Dazu nehmen wir an, f_1 und f_2 seien Lösungen und setzen $g(x) = f_1(x) - f_2(x)$. Subtraktion der Gleichungen für f_1 und f_2 ergibt die einfachere Gleichung

$$g(xy) = g(x^2) + g(y^2) \qquad \forall\, x, y \in \mathbb{R}$$

für g. Mit $x = y = 0$ ergibt sich hieraus $g(0) = 0$. Setze nun $y = 0$, dann folgt $g(x^2) = 0$, also gilt $g(z) = 0$ für alle $z \geq 0$. Schliesslich setzen wir noch $y = 1$, dies ergibt $g(x) = g(x^2) + g(1) = 0 + 0 = 0$ für alle $x \in \mathbb{R}$. Damit stimmen f_1 und f_2 überein. Somit ist $f(x) = x$ die einzige Lösung.

□

Eine andere Möglichkeit ist, neue Funktionen einzuführen. Es ist im Allgemeinen viel einfacher zu zeigen, dass eine bestimmte Gleichung nur konstante Lösungen oder Lösungen der Form $f(x) = cx$ hat, als kompliziertere Funktionen wie zum Beispiel $f(x) = x^4 - 3x$. Es empfiehlt sich daher manchmal, die gesuchte Funktion mittels einer einfacheren auszudrücken. Wir geben eine zweite Lösung für das letzte Beispiel.

2. Lösung. Wir wissen bereits, dass $f(x) = x$ eine Lösung ist, und wohl auch die einzige. Wir führen eine neue Funktion g ein durch $f(x) = x + g(x)$. Setzt man dies in die Gleichung ein, folgt wieder die einfachere Gleichung

$$g(xy) = g(x^2) + g(y^2) \qquad \forall\, x, y \in \mathbb{R}$$

für g. Die 1. Lösung zeigt nun, dass g identisch verschwindet.

□

Eine andere mögliche Variante wäre gewesen, g durch $f(x) = xg(x)$ zu definieren (beachte: dies geht nur, wenn man bereits weiss, dass $f(0) = 0$ ist, was einfach zu überprüfen ist. Der Wert $g(0)$ muss man dann gar nicht kennen!). Allerdings führt das eher zu einer komplizierteren Gleichung als jener für f. Totzdem ist das manchmal nützlich.

Dieses Verfahren ist nur dann sinnvoll, wenn man sich recht sicher ist, dass die gefundene Lösung die einzige ist. Betrachte Beispiel 3. Dort sind $f(x) = 0$ und $f(x) = x^2$ beides Lösungen, wie man leicht sieht. Es bringt daher wenig, eine neue Funktion g durch $f(x) = x^2 - g(x)$ zu definieren. Man muss dann ja zeigen, dass g entweder identisch verschindet, oder dass $g(x) = -x^2$ ist. Das ist mindestens so schwierig wie die ursprüngliche Aufgabe. Sinnvoller (wenn schon) wäre die Definition $f(x) = x^2/2 + g(x)$. Denn damit hat man immerhin eine gewisse Symmetrie gewonnen (Die Lösungen sind dann $g(x) = \pm x^2/2$).

Eine ganz andere Situation tritt in Beispiel 8 auf. Dort gibt es eine ganze Lösungsfamilie, nämlich die Funktionen $f(x) = cx^2$ mit einer beliebigen rationalen Konstanten c. Man überzeigt sich an der Funktionalgleichung direkt, dass Summen, Differenzen und konstante Vielfache von Lösungen wieder Lösungen sind. Es bringt also wenig, wenn man versucht zu zeigen, dass die Lösung eindeutig bestimmt ist (das ist eben nicht der Fall). Genauso sinnlos ist etwa die Substitution $f(x) = x^2 + g(x)$ oder dergleichen. Was manchmal wirklich etwas bringt, ist folgende Überlegung: Die Funktion $f(1)x^2$ ist eine Lösung der Gleichung, also auch $g(x) = f(x) - f(1)x^2$. Die Gleichung für g ist genau dieselbe wie jene für f, aber wir haben nun eine zusätzliche Information, nämlich $g(1) = 0$. Das ist für eine beliebige Lösung natürlich gar nicht der Fall. Es kann nun sein, dass diese Zusatzinformation wesentlich zu einer Lösung beisteuert, wir müsen nämlich zeigen, dass g identisch verschwindet. In Beispiel 8 ist der Gewinn allerdings minimal, Gleichung (12) muss man immer noch beweisen, lediglich die Schlussrechnung wird ein bisschen einfacher. Trotzdem ist das ein mächtiges Verfahren, wenn alles Andere versagt.

2 Weitere Verfahren

2.1 Injektivität und Surjektivität

Wir definieren zuerst die beiden Eigenschaften im Titel, um die es in diesem Abschnitt gehen wird. Seien A und B beliebige Mengen und sei $f : A \to B$ eine Funktion. Dann heisst f *surjektiv*, falls zu jedem Element $b \in B$ ein Element $a \in A$ existiert mit $f(a) = b$. Und f heisst *injektiv*, falls aus $f(a_1) = f(a_2)$ mit Elementen $a_1, a_2 \in A$ stets $a_1 = a_2$ folgt. Mit anderen Worten, f ist surjektiv genau dann, wenn jedes Element aus B *mindestens* ein Urbild in A besitzt. Analog dazu ist f injektiv genau dann, wenn jedes Element aus B *höchstens* ein Urbild in A besitzt. Eine Funktion, die injektiv und surjektiv ist, nennt man auch *bijektiv*. Eine bijektive Funktion hat also die Eigenschaft, dass sie die Elemente von A 1:1 auf die Elemente in B abbildet. Zu jedem Element aus A 'gehört' genau ein Element aus B.

Zwei Fragen drängen sich nun auf. Erstens: Was nützt es zu wissen, dass eine Funktion injektiv/surjektiv ist? Zweitens: Wie weist man diese Eigenschaften konkret nach? Beides werden wir mit den folgenden Beispielen konkret beantworten.

Die typische Situation, in der man die *Surjektivität* einer Funktion f nachweisen kann, ist die folgende: Durch geeignete Manipulation der Funktionalgleichung, geschicktes Einsetzen usw. bringt man die Gleichung auf die Form $f(\ldots) = A(x, y, \ldots)$, sodass also auf der einen Seite nur ein Funktionsterm steht und auf der anderen Seite irgend ein Ausdruck A in einer oder mehreren der freien Variablen. Wenn nun A für geeignete Wahl dieser Variablen *jeden beliebigen Wert* in der Zielmenge von f annehmen kann, dann muss das eben auch für f selbst gelten. Dabei ist es völlig egal, wie kompliziert der Ausdruck $f(\ldots)$ auf der linken Seite ist. Wir geben ein Fantasiebeispiel, um das zu illustrieren.

Angenommen, es ginge um die Funktionalgleichung

$$f(f(x) - x + y)^2 = x + f(y^2 - x)$$

für die Funktion $f : \mathbb{R} \to \mathbb{R}$. Wir zeigen, dass f auf jeden Fall surjektiv ist. In der Gleichung treten zwei Funktionsterme auf, je einer auf beiden Seiten des Gleichheitszeichens. Wir werden daher den einen 'ruhigstellen', um volle Kontrolle über den anderen zu erhalten. Dazu setzen wir $y = x - f(x)$ ein und erhalten die Gleichung

$$f(0)^2 = x + f((x - f(x))^2 - x) \qquad \Longleftrightarrow \qquad f((x - f(x))^2 - x) = f(0)^2 - x.$$

Auf der rechten Seite steht nun eine lineare Funktion in x, die natürlich jeden reellen Wert annehmen kann. Dies gilt daher auch für die linke Seite, somit ist f surjektiv.

Eine besonders einfache und sehr häufige Situation ist diejenige der involutiven Funktionen. Eine Funktion f mit der Eigenschaft $f(f(x)) = x$ für alle x heisst eine *Involution*. Beispiele dafür sind $f(x) = x$, $f(x) = -x$, $f(x) = \frac{1}{x}$, $(x > 0)$. Solche Funktionen sind automatisch surjektiv. Allgemeiner folgt aus einer Gleichung vom Typ

$$f(f(x)) = x + c$$

mit einer Konstanten c die Surjektivität von $f : \mathbb{R} \to \mathbb{R}$, denn die rechte Seite nimmt wieder jeden reellen Wert an. Als (richtige) Anwendung behandeln wir nun ein altes IMO-Problem. Man sieht sofort, das $f(x) = x$ eine Lösung ist, und wohl auch die einzige. Spielt man ein Bisschen mit der Gleichung herum, tritt sehr natürlich immer wieder die Konstante $f(0)$ als Störterm auf. Nun müsste ja aber $f(0) = 0$ sein, was die Rechnungen erheblich vereinfachen würde, also sollte man zuerst versuchen, dies auch wirklich zu zeigen. Das ist aber erstaunlich schwierig und der entscheidende Schritt in der Lösung der Aufgabe. Wir zeigen nun, wie das geht. Die zweite Hälfte der Lösung verschieben wir in Kapitel 3, wo wir die Cauchy-Gleichungen besprechen werden.

Beispiel 10. *(IMO 92) Finde alle Funktionen $f : \mathbb{R} \to \mathbb{R}$, sodass für alle $x, y \in \mathbb{R}$ gilt*

$$f(x^2 + f(y)) = y + f(x)^2.$$

Lösung. Wir setzen $x = 0$ und erhalten

$$f(f(y)) = y + f(0)^2, \tag{13}$$

insbesondere ist f surjektiv. Wir können daher eine reelle Zahl a wählen mit $f(a) = 0$. Wendet man nun f auf beide Seiten der Gleichung an und verwendet (13), dann folgt

$$x^2 + f(y) + f(0)^2 = f(y + f(x)^2).$$

Setzt man hier $x = y = a$, dann wird die linke Seite zu $a^2 + f(0)^2$, die rechte Seite zu $f(a + f(a)^2) = f(a) = 0$. Wir erhalten also die Gleichung $a^2 + f(0)^2 = 0$ und somit wie gewünscht $a = f(0) = 0$.

□

Eine typische Situation für den Nachweis von *Injektivität* gibt es eigentlich nicht. Der Nachweis, dass aus $f(a) = f(b)$ auch wirklich $a = b$ folgt, muss der jeweiligen Sitation angepasst werden. Involutive Funktionen sind aber zum Beispiel stets injektiv. Allgemeiner ist jede Funktion, die eine Gleichung vom Typ

$$f(f(x)) = x + c$$

mit einer Konstanten c erfüllt, injektiv. Denn nehme an, für zwei Zahlen a, b sei $f(a) = f(b)$. Anwenden von f liefert dann auch $f(f(a)) = f(f(b))$ und unter Verwendung der Gleichung erhält man $a + c = b + c$, also $a = b$. Der Vorteil der Injektivität ist, dass man 'mit f kürzen' kann. Aus einer Identität der Form $f(A) = f(B)$ folgt nämlich $A = B$, man kann also von den Funktionswerten auf die Argumente schliessen.

Wir geben zum Abschluss noch ein Beispiel, wo Injektivität oder Surjektivität zur Lösung herangezogen werden kann.

Beispiel 11. *Finde alle Funktionen $f : \mathbb{R} \to \mathbb{R}$, sodass für alle $x, y \in \mathbb{R}$ gilt*

$$f(f(x)f(y)) = f(x)y.$$

1. Lösung. Offenbar ist die konstante Funktion $f(x) = 0$ eine Lösung. Wir nehmen im Folgenden an, dass f nicht identisch verschwindet und wählen eine reelle Zahl a mit $f(a) \neq 0$. Setze $x = a$, dann folgt

$$f(f(a)f(y)) = f(a)y.$$

Wir zeigen, dass f injektiv ist. Sei also $f(u) = f(v)$, dann folgt

$$f(a)u = f(f(a)f(u)) = f(f(a)f(v)) = f(a)v,$$

und wegen $f(a) \neq 0$ gilt tatsächlich $u = v$. Nun setzen wir in der ursprünglichen Gleichung $y = 1$ und erhalten $f(f(x)f(1)) = f(x)$. Da f injektiv ist, folgt daraus $f(x)f(1) = x$. Mit $x = 1$ erhalten wir schliesslich noch $f(1)^2 = 1$, also die beiden Funktionen $f(x) = x$ oder $f(x) = -x$. Beide erfüllen die ursprüngliche Gleichung.

□

2. Lösung. Wir nehmen wieder an, dass f nicht identisch verschwindet und wählen eine reelle Zahl a mit $f(a) \neq 0$. Setze $x = a$, dann folgt

$$f(f(a)f(y)) = f(a)y.$$

Die rechte Seite nimmt nun nach Wahl von a alle reellen Werte an, folglich ist f surjektiv. Setze wieder $y = 1$ in der ursprünglichen Gleichung, dann folgt $f(f(x)f(1)) = f(x)$. Es gilt sicher $c = f(1) \neq 0$, denn sonst wäre die linke Seite konstant, während die rechte alle reellen Werte annimmt. Da f surjektv ist, existiert für *jede* reelle Zahl z eine reelle Zahl x mit $f(x) = \frac{z}{c}$. Eisetzen ergibt $f(z) = \frac{z}{c}$, also ist f linear. Einsetzen zeigt schliesslich wieder $c = \pm 1$ und wir erhalten dieselben Lösungen wie vorher.

□

2.2 Nullstellen und Fixpunkte

Eine weitere wichtige Idee beim Lösen von Funktionalgleichungen ist, auf Fixpunkte und Nullstellen der gesuchten Funktion zu achten. Wie üblich ist das Problem dabei, überhaupt zu merken, dass Fixpunkte wichtige Information beisteuern können. Zuerst die Definitionen: Sei $f : A \rightarrow A$ eine beliebige Funktion. Ein Element $a \in A$ heisst ein *Fixpunkt* von f, wenn $f(a) = a$ gilt. Ist $f : \mathbb{R} \rightarrow \mathbb{R}$ eine reelle Funktion, dann heisst eine reelle Zahl x mit $f(x) = 0$ eine *Nullstelle* von f. Wir zeigen die Bedeutung dieser Konzepte an einem Beispiel.

Beispiel 12. *(IMO 94) Sei S die Menge aller reellen Zahlen grösser als -1. Finde alle Funktionen $f : S \rightarrow S$, die folgende Bedingungen erfüllen:*

(a) $f(x + f(y) + xf(y)) = y + f(x) + yf(x) \qquad \forall\, x, y \in S$

(b) *In jedem der Intervalle $(-1, 0)$ und $(0, \infty)$ ist die Funktion $f(x)/x$ streng monoton wachsend.*

Lösung. Setze $x = y > -1$ in (a), dann folgt

$$f(x + f(x) + xf(x)) = x + f(x) + xf(x). \tag{14}$$

Das heisst: Für alle $x > -1$ ist $x + f(x) + xf(x)$ ein *Fixpunkt* von f. Wir müssen also etwas über die Fixpunkte von f herausfinden. Die Bedingung in (b) zielt genau auf das ab, per Definition ist $a \neq 0$ nämlich genau dann ein Fixpunkt von f, wenn $\frac{f(a)}{a} = 1$ gilt! Nach (b) existiert also in jedem der Intervalle $(-1, 0)$ und $(0, \infty)$ *höchstens* ein Fixpunkt. Sei nun a irgend ein Fixpunkt von f. Mit $x = y = a$ folgt dann wegen $f(a) = a$ die Gleichung $f(2a + a^2) = 2a + a^2$, folglich ist auch $2a + a^2$ ein Fixpunkt von f. Im Fall $a > 0$ gilt aber $0 < a < 2a + a^2$, also hätte f im Intervall $(0, \infty)$ mindestens *zwei* verschiedene Fixpunkte, ein Widerspruch. Im Fall $-1 < a < 0$ gilt analog $-1 < 2a + a^2 < a < 0$, also hätte f im Intervall $(-1, 0)$ mindestens zwei verschiedene Fixpunkte, was auch nicht sein kann. Der einzige Fixpunkt von f ist also 0 und aus (14) folgt $x + f(x) + xf(x) = 0$ für alle $x > -1$. Auflösen nach $f(x)$ ergibt

$$f(x) = -\frac{x}{x+1}.$$

Schliesslich zeigen wir noch, dass f die Bedingungen der Aufgabe wirklich erfüllt. Die Funktion $\frac{f(x)}{x} = \frac{-1}{x+1}$ ist in den Intervallen $(-1, 0)$ und $(0, \infty)$ in der Tat streng monoton wachsend, also gilt (b). Schliesslich folgt (a) aus der Rechnung

$$f(x + f(y) + xf(y)) = f\Big(\frac{x-y}{y+1}\Big) = \frac{y-x}{x+1} = y + f(x) + yf(x).$$

□

2.3 Monotonie und Periodizität

Bevor wir beginnen wieder eine Definition: Sei $I \subset \mathbb{R}$ ein Intervall und sei $c > 0$ eine reelle Zahl. Eine Funktion $f : I \to \mathbb{R}$ heisst *periodisch mit Periode* c, falls für alle $x \in I$ mit $x + c \in I$ gilt $f(x + c) = f(x)$. Der Begriff der Periodizität ist bisweilen trügerisch, daher eine kurze Warnung vorweg. Natürlich ist mit c auch jedes ganzzahlige Vielfache von c eine Periode von f. Was hingegen erstaunen mag ist die Tatsache, dass eine Funktion *keine kleinste Periode* zu besitzen braucht, im Gegensatz etwa zu den aus der Schule bekannten trigonometrischen Funktionen. Die Funktion $f : \mathbb{R} \to \mathbb{R}$ mit $f(x) = 1$ falls x rational ist und $f(x) = 0$ falls x irrational ist besitzt zum Beispiel *jede* rationale Zahl als Periode.

Die beiden im Titel genannten Eigenschaften einer reellen Funktion sind völlig unabhängig voneinander, keine impliziert die andere. Falls eine Funktion f aber *beide* Eigenschaften hat, gibt dies starke Einschränkungen an die Struktur von f. Ist nämlich I ein Intervall der Länge $> c$ und ist $f : I \to \mathbb{R}$ monoton wachsend (oder fallend) und periodisch mit Periode c, dann ist f auf I *konstant*! Für alle $x \in I$ mit $x + c \in I$ ist nämlich $f(x+c) = f(x)$ und wegen der Monotonie also $f(x) = f(y) = f(x+c)$ für alle $y \in [x, x+c]$. Folglich ist f auf allen abgeschlossenen Teilintervallen der Länge c von I konstant, also überhaupt konstant.

Etwas allgemeiner benötigt man eigentlich keine Periodizität, sondern nur, dass f an vielen gut verteilten Stellen immer wieder denselben Wert annimmt. Ist zum Beispiel $f : \mathbb{R}_{\geq 0} \to \mathbb{R}$ eine Funktion, die monoton steigend ist, und für die $f(x) = a$ gilt für alle nichtnegativen ganzzahligen x, dann ist $f(x) = a$ konstant auf $\mathbb{R}_{\geq 0}$ nach demselben Argument.

Solche Ideen können manchmal weiterhelfen, wie in folgendem nicht ganz einfachem Beispiel:

Beispiel 13. *(Shortlist 05) Bestimme alle Funktionen $f : \mathbb{R}_{>0} \to \mathbb{R}_{>0}$, sodass für alle positiven x, y gilt*

$$f(x)f(y) = 2f(x + yf(x)).$$

Lösung. Bei dieser Aufgabe ist überhaupt nicht klar, wo man beginnen soll. Offensichtlich ist die konstante Funktion $f(x) = 2$ eine Lösung. Aber ist es die einzige? Wenn man genauer hinschaut, gibt es nun doch zwei relativ natürliche Fragen, die man sich stellen kann:

(1) Ist es möglich ein y zu finden, sodass die Klammer rechts für zwei verschiedene Werte von x gleich gross ist? Dies würde sicher zu einer Identität führen, die man brauchen könnte. Wir betrachten also $0 < u < v$ und suchen ein y, sodass gilt $u + yf(u) = v + yf(v)$. Lösen wir das nach y auf und schauen, was passiert:

$$y = \frac{v-u}{f(u)-f(v)}.$$

Erstens muss also $f(u) \neq f(v)$ sein, und da $y > 0$ sein soll, sogar $f(u) > f(v)$. Angenommen das sei der Fall, dann stimmen die rechten Seiten der ursprünglichen Gleichung für $x = u$ und $x = v$ überein, also auch die linken. Wegen $f(y) > 0$ impliziert das $f(u) = f(v)$. Insgesamt haben wir also gezeigt: Existieren zwei reelle Zahlen $0 < u < v$ mit $f(u) > f(v)$, dann folgt $f(u) = f(v)$. Das ist aber absurd, also haben wir (ziemlich überraschend) gezeigt, dass f *monoton steigend* ist.

(2) Was passiert, wenn f an zwei verschiedenen Stellen denselben Wert annimmt? Sei also $0 < u < v$ mit $f(u) = f(v)$. Für alle $y > 0$ gilt dann

$$2f(u + yf(u)) = f(u)f(y) = f(v)f(y) = 2f(v + yf(v)).$$

Das scheint nun eine Art *Periodizität* für f zu implizieren, nämlich mit der Periode $c = v-u$. Tatsächlich: Sei $z > u$ beliebig, und setze $y = \frac{z-u}{f(u)} > 0$ in diese Gleichung ein, dann folgt $f(z) = f(u+yf(u)) = f(v+yf(v)) = f(z+c)$. Daher ist f auf dem Intervall (u,∞) periodisch und nach (1) monoton steigend, also konstant. Setzt man $x, y > u$ in die ursprüngliche Gleichung ein, dann folgt sofort, dass diese Konstante gleich 2 sein muss (beachte, dass in diesem Fall auch $x + yf(y) > u$ gilt). Das ist beinahe was wir wollten. Die Einschränkung, dass f erst ab u konstant 2 ist, lässt sich nun aber sehr einfach beheben: Sei y beliebig und sei $x > u$ in der ursprünglichen Gleichung, dann folgt wegen $x + yf(x) > u$ sofort $2f(y) = 2 \cdot 2$, also $f(y) = 2$ für *alle* $y > 0$.

Wir müssen also nur noch zeigen, dass zwei verschiedene positive reelle Zahlen existieren, an denen f denselben Wert annimmt. Wir nehmen an, das sei nicht so. Dann ist f *injektiv*. Aus der Symmetrie der Funktionalgleichung folgt sofort

$$2f(x + yf(x)) = f(x)f(y) = f(y)f(x) = 2f(y + xf(y)),$$

und daher auch $x+yf(x) = y+xf(y)$. Dies ist äquivalent zu $\frac{f(x)-1}{x} = \frac{f(y)-1}{y}$. Da dies nun für alle Paare von positiven reellen Zahlen x, y gilt, können wir zum Beispiel $y = 1$ setzten und erhalten $f(x) = (f(1) - 1)x + 1$ für alle $x > 0$. Einsetzen zeigt aber, dass diese Funktionen allesamt keine Lösungen der ursprünglichen Gleichung sind, ein Widerspruch. Insgesamt ist also $f(x) = 2$ konstant und wir sind fertig.

□

3 Die Cauchy-Funktionalgleichungen

Die Cauchy-Funktionalgleichungen zählen zu den wichtigsten Funktionalgleichungen überhaupt. Viele Probleme lassen sich durch eine geeignete Reduktion auf sie zurückführen. Es handelt sich um die vier Gleichungen

$$\begin{array}{lrcl} f:\mathbb{R}\to\mathbb{R}, & f(x+y) & = & f(x)+f(y),\\ f:\mathbb{R}_{>0}\to\mathbb{R}_{>0}, & f(xy) & = & f(x)f(y),\\ f:\mathbb{R}_{>0}\to\mathbb{R}, & f(xy) & = & f(x)+f(y),\\ f:\mathbb{R}\to\mathbb{R}_{>0}, & f(x+y) & = & f(x)f(y). \end{array}$$

In der Praxis sind vor allem die beiden ersten von Bedeutung. In der Tat sind sie aber alle äquivalent, denn ist $x \mapsto f(x)$ eine Lösung der ersten Gleichung, dann sind die Funktionen $x \mapsto \exp(f(\log(x)))$, $x \mapsto f(\log(x))$ und $x \mapsto \exp(f(x))$ Lösungen der zweiten bis vierten Gleichng. Ja es ist sogar leicht zu sehen, dass man dadurch *Bijektionen* zwischen den Lösungsmengen der vier Gleichungen erhält. Wir werden uns daher im Folgenden auf die erste Gleichung konzentrieren.

Eine Klasse von Lösungen ist einfach zu sehen, nämlich die Funktionen $f(x) = ax$ mit einer reellen Konstanten a. Allerdings sind das eben nicht die einzigen, ansonsten würden wir den Cauchy-Gleichungen ja kaum ein ganzes Kapitel widmen. Ich würde an dieser Stelle gern ein Beispiel einer Lösungsfunktion geben, die nicht linear ist. Es ist jedoch prinzipiell unmöglich, eine solche Funktion explizit anzugeben!! Man darf sich jetzt natürlich schon kurz am Kopf kratzen und sich fragen, wieso man denn *weiss*, dass solche Funktionen existieren. Das Problem ist, dass zu ihrer Konstruktion ein Axiom der Mengenlehre herangezogen werden muss, das sogenannte *Auswahlaxiom*. Das Auswahlaxiom sichert nun die Existenz solcher exotischer Funktionen, ohne dass man irgendeine Chance hätte, sie wirklich hinzuschreiben. Wie auch immer, es gibt sie jedenfalls, und es stellt sich sogar heraus, dass die linearen Funktionen eine verschwindende Minderheit in der Menge aller Lösungsfunktionen bilden. Wenn man nämlich zufällig irgendeine solche Lösungsfunktion herauspicken würde, sie wäre mit 100-prozentiger Sicherheit *nicht* linear.

Nach diesen Erklärungen wenden wir uns nun der Arbeit zu. Wir zeigen zuerst, dass jede Lösung zumindest auf $\mathbb{Q}$ tatsächlich linear ist. In der Tat gilt sogar mehr, man kann nämlich belibige *rationale* Faktoren aus der Funktion herausziehen.

Lemma 3.1. *Sei $f : \mathbb{R} \to \mathbb{R}$ eine Lösung der Gleichung $f(x + y) = f(x) + f(y)$. Für alle rationalen Zahlen r und alle reellen Zahlen x gilt*

$$f(rx) = rf(x),$$

insbesondere also $f(r) = r \cdot f(1)$ für alle rationalen r. Somit ist f linear auf $\mathbb{Q}$.

Beweis. Setzt man $x = y = 0$, dann folgt $f(0) = 0$. Mit $y = -x$ erhält man daraus weiter $f(-x) = -f(x)$ für alle $x \in \mathbb{R}$, also ist f eine ungerade Funktion. Als nächstes zeigen wir induktiv, dass für alle natürlichen Zahlen n und alle reellen x gilt

$$f(nx) = nf(x). \tag{15}$$

Dies ist trivial für $n = 1$ und gelte für ein n. Setzt man in der ursprünglichen Gleichung $y = nx$, dann folgt aus der Induktionsannahme

$$f((n+1)x) = f(x + nx) = f(x) + f(nx) = f(x) + nf(x) = (n+1)f(x),$$

wie gewünscht. Da nun f ungerade ist, gilt (15) sogar für alle ganzen Zahlen n. Für beliebige $p, q \in \mathbb{Z}$ mit $q \neq 0$ folgt daraus schliesslich

$$qf(\tfrac{p}{q}x) = f(q\tfrac{p}{q}x) = f(px) = pf(x),$$

also $f(rx) = rf(x)$ für alle rationalen $r = \frac{p}{q}$.

□

Wie schon erwähnt, kann man nun nicht von $\mathbb{Q}$ auf $\mathbb{R}$ schliessen, induktive Methoden sind auf $\mathbb{R}$ grundsätzlich zum Scheitern verurteilt. Um hier weiterzukommen, braucht man zusätzliche Annahmen an f. Die für uns mit Abstand wichtigste ist die *Monotonie*, man hat dann wie gewünscht:

Lemma 3.2. *Sei f eine Lösung der Gleichung $f(x+y) = f(x)+f(y)$, die* monoton *(fallend oder steigend) ist. Dann gilt $f(x) = ax$ für alle $x \in \mathbb{R}$, wobei a eine reelle Konstante ist.*

Beweis. Mit f ist auch $-f$ eine Lösung der Gleichung, wir können also annehmen, dass f monoton steigend ist. Nach Lemma 2.1 gilt $f(r) = ar$ für alle $r \in \mathbb{Q}$, wobei $a = f(1) \geq f(0) = 0$ ist. Sei nun x eine beliebige reelle Zahl und nehme an, dass $f(x) > ax$ gilt. Im Fall $a = 0$ sei $r > x$ irgendeine rationale Zahl, im Fall $a > 0$ wählen wir ein rationales r mit $x < r < f(x)/a$. Nach Konstruktion ist dann in beiden Fällen $f(r) = ar < f(x)$, im Widerspruch dazu, dass f monoton steigend ist. Analog führt man die Annahme $f(x) < ax$ zu einem Widerspruch: Wähle eine rationale Zahl r mit $r < x$ im Fall $a = 0$ und mit $f(x)/a < r < x$ im Fall $a > 0$, dann gilt $f(r) = ar > f(x)$, was nicht sein kann.

□

Beispiel 14. *(IMO 92) Finde alle Funktionen $f : \mathbb{R} \to \mathbb{R}$, sodass für alle $x, y \in \mathbb{R}$ gilt*

$$f(x^2 + f(y)) = y + f(x)^2.$$

Lösung. Wir haben in Beispiel 10 bereits gezeigt, dass $f(0) = 0$ ist. Damit vereinfacht sich (13) zu $f(f(y)) = y$. Mit $y = 0$ erhält man ausserdem $f(x^2) = f(x)^2$. Ersetzt man nun in der ursprünglichen Gleichung y durch $f(y)$, dann folgt unter Verwendung der beiden eben bewiesenen Identitäten die neue Gleichung

$$f(x^2 + y) = f(x^2) + f(y).$$

Diese sieht bereits aus wie eine Cauchy-Gleichung mit dem kleinen Unterschied, dass x^2 natürlich nur nichtnegative Werte annehmen kann. Wir wissen also, dass

$$f(u+v) = f(u) + f(v) \qquad \text{für alle } u \geq 0 \text{ und alle } v \in \mathbb{R}. \tag{16}$$

Wir entfernen nun die störende Bedingung $u \geq 0$. Zuerst bemerken wir, dass f eine ungerade Funktion ist, denn für $u \geq 0$ und $v = -u$ folgt $0 = f(0) = f(u)+f(-u)$, also $f(-u) = -f(u)$. Da man aber jede reelle Zahl als u oder $-u$ schreiben kann, ist f ungerade. Für alle $u \geq 0$ erhält man damit $f((-u)+v) = -f(u+(-v)) = -f(u) - f(-v) = f(-u) + f(v)$, also gilt (16) für alle $u, v \in \mathbb{R}$.

Wir müssen jetzt nur noch zeigen, dass f keine exotische Lösung der Cauchy-Gleichung ist, sondern linear. Wir zeigen daher, dass f monoton steigend ist. Wegen $f(x^2) = f(x)^2$ ist $f(u) \geq 0$ für alle $u \geq 0$. Seien nun $a \geq b$ zwei reelle Zahlen, Einsetzen von $u = b - a \geq 0$ und $v = b$ in (16) liefert dann $f(b) = f(b-a) + f(a) \geq f(a)$, wie gewünscht. Nach Lemma 2.2 ist daher $f(x) = cx$ mit einer Konstanten c. Einsetzen in die urpsrüngliche Gleichung liefert $c = 1$ und damit die einzige Lösung $f(x) = x$.

□

Zwei Dinge sind hier anzumerken: Erstens hätte man mit (16) auch die Einschränkung von f auf $\mathbb{R}_{\geq 0}$ studieren können. Dort erfüllt f ja die Cauchy-Gleichung und es ist leicht zu sehen, dass die Lemmata 2.1. und 2.2. auch für diesen Fall richtig sind. Auf $\mathbb{R}_{\geq 0}$ ist f dann also linear und da f ungerade ist, gilt dies dann auf ganz $\mathbb{R}$. Zweitens war für den Nachweis der Monotonie die Eigenschaft $f(u) \geq 0$ für $u \geq 0$ zentral. Diese wiederum kam aus der Tatsache, dass Quadrate nie negativ sind. Allgemein kann man sagen, dass in der Funktionalgleichung irgendwie Quadrate (oder Produkte) auftauchen müssen, damit dieses Argument funktioniert. Es gibt tatsächlich Beispiele, die ohne Quadrate auskommen, da braucht man aber einen Trick. Man sollte jedenfalls nach solchen Quadraten ausschau halten.

Das letzte Beispiel ist recht typisch und enthält die folgenden Lösungsschritte, die bei vielen Problemen gemacht werden müssen:

- Bestimmen von $f(0)$ oder Ähnliches.
- Reduktion auf eine Cauchy-Gleichung.
- Monotonie nachweisen.

Anstelle der Monotonie gibt es viele andere Bedingungen an die Lösung f einer Cauchy-Gleichung, die Linearität implizieren. Die bekannteste ist die Stetigkeit, auf die wir hier aber nicht eingehen wollen. Wir stellen nun aber *das* Kriterium schlechthin vor, mit dem man in allen praktischen Fällen durchkommt. Dazu müssen wir etwas ausholen und zwei Begriffe einführen.

Sei $f : \mathbb{R} \to \mathbb{R}$ eine beliebige Funktion. Der *Graph* von f ist die Menge

$$\Gamma(f) = \{(x, f(x)) \mid x \in \mathbb{R}\} \subset \mathbb{R}^2,$$

welche f quasi als geometrische 'Kurve' beschreibt. Der Graph einer linearen Funktion zum Beispiel ist eine Gerade im $\mathbb{R}^2$, welche nicht parallel zur y-Achse liegt. Eine Teilmenge $A \subset \mathbb{R}^2$ heisst *dicht*, falls jede offene Kreisscheibe von positivem Radius in $\mathbb{R}^2$ mindestens einen Punkt von A enthält. Zum Beispiel liegt $\mathbb{Q}^2$ dicht in $\mathbb{R}^2$ (wieso?).

Nun also das angekündigte Kriterium. Wir geben den recht technischen Beweis der Vollständigkeit halber an:

Satz 3.3. *Sei $f : \mathbb{R} \to \mathbb{R}$ eine Lösung der Gleichung $f(x+y) = f(x)+f(y)$. Dann gilt einer der folgenden Fälle:*

(a) *Die Funktion f ist linear.*

(b) *Der Graph $\Gamma(f)$ liegt dicht in $\mathbb{R}^2$.*

Beweis. Wir nehmen an, dass f nicht linear ist. Dann gibt es reelle Zahlen $x, y \neq 0$ und $a \neq b$ mit $f(x) = ax$ und $f(y) = by$. Aus technischen Gründen werden wir nicht zeigen, dass jede offene Kreisscheibe einen Punkt des Graphen enthält, sondern wir zeigen, dass jedes offene Quadrat einen solchen Punkt enthält (was natürlich auf dasselbe hinausläuft). Sei also $(u, v) \in \mathbb{R}^2$ ein fester Punkt und sei $\varepsilon > 0$ fest. Wir setzen

$$\xi_1 = \frac{v - bu}{a - b}, \qquad \xi_2 = \frac{au - v}{a - b},$$

sodass folgende Gleichungen gelten:

$$\xi_1 + \xi_2 = u, \qquad a\xi_1 + b\xi_2 = v.$$

Wähle rationale Zahlen r_1, r_2 mit

$$|r_1 x - \xi_1| < \min\left\{1, \frac{1}{|a|}\right\} \cdot \frac{\varepsilon}{2}, \qquad |r_2 y - \xi_2| < \min\left\{1, \frac{1}{|b|}\right\} \cdot \frac{\varepsilon}{2}.$$

Einerseits gilt dann

$$\begin{aligned} |(r_1x + r_2y) - u| &= |(r_1x - \xi_1) + (r_2y - \xi_2)| \\ &\leq |(r_1x - \xi_1)| + |(r_2y - \xi_2)| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \end{aligned}$$

Andererseits lassen sich nach Lemma 2.1. rationale Faktoren aus f herausziehen und daher ist auch

$$\begin{aligned} |f(r_1x + r_2y) - v| &= |(r_1ax + r_2by) - (a\xi_1 + b\xi_2)| \\ &\leq |a||(r_1x - \xi_1)| + |b||(r_2y - \xi_2)| < \frac{\varepsilon|a|}{2|a|} + \frac{\varepsilon|b|}{2|b|} = \varepsilon. \end{aligned}$$

Folglich liegt der Punkt $(r_1x + r_2y, f(r_1x + r_2y)) \in \Gamma(f)$ im offenen Quadrat mit Mittelpunkt (u, v) und Seitenlänge 2ε. Somit liegt der Graph von f dicht.

□

Dieses doch sehr erstaunliche Resultat liefert jetzt natürlich eine ganze Fülle von einfach zu überprüfenden Eigenschaften von f, die auf die Linearität schliessen lassen. Zum Beispiel wieder die *Monotonie*: Ist f monoton steigend, und ist x irgend eine reelle Zahl, dann enthält der Graph von f sicher keinen Punkt aus den beiden Gebiete $(-\infty, x) \times (f(x), \infty)$ und $(x, \infty) \times (-\infty, f(x))$, liegt also nicht dicht. Das liefert einen zweiten Beweis für Lemma 2.2. Oder *Beschränktheit*: Ist f beschränkt, zum Beispiel nach oben durch die Konstante C, dann enthält der Graph von f keinen Punkt aus dem Gebiet $\mathbb{R} \times (C, \infty)$, liegt also wieder nicht dicht. In diesem Fall kann man sogar mehr sagen: da f auf $\mathbb{Q}$ linear ist, muss $f = 0$ sein. Es genügt aber auch, dass f auf irgend einem noch so kleinen Intervall positiver Länge nach oben oder unten beschränkt ist! Auch dann liegt der Graph nicht dicht. In Beispiel 10 haben wir als ersten Schritt für die Monotonie gezeigt, dass $f(x^2) = f(x)^2$ gilt. Somit ist f auf $\mathbb{R}_{\geq 0}$ nach unten beschränkt und daher linear. etc.

Diese Resultate für die Gleichung $f(x + y) = f(x) + f(y)$ gelten mit den naheliegenden Modifikationen auch für die drei anderen Cauchy-Gleichungen. Wir erklären dies kurz anhand der multiplikativen Version. Gesucht sind also die Lösungen $f : \mathbb{R}_{>0} \to \mathbb{R}_{>0}$ von $f(xy) = f(x)f(y)$. Für jede solche Funktion ist $g(x) = \ln(f(\exp(x)))$ ein Lösung der additiven Gleichung und umgekehrt. Die typischen Lösungen sind daher die Funktionen $f(x) = x^a$ mit einer reellen Konstanten a. Die exotischen Lösungen kann man mit der naheliegenden Modifikation von Satz 2.3. handhaben: Für jede exotische Lösung liegt der Graph dicht in $\mathbb{R}_{>0} \times \mathbb{R}_{>0}$ (natürlich geht es hier nur um den ersten Quadranten, nicht um die ganze Ebene!). Insbesondere sind wieder alle monotonen Lösungen klassisch.

Beispiel 15. *Bestimme alle Funktionen $f : \mathbb{R} \to \mathbb{R}_{\geq 0}$, sodass für alle $x, y \in \mathbb{R}$ gilt*

$$f(x + y) - f(x) - f(y) = (f(x) + x)(f(y) + y).$$

Lösung. Die Klammern auf der rechten Seite legen es nahe, eine neue Funktion $g(x) = f(x) + x$ einzuführen. Die Gleichung vereinfacht sich damit zu

$$g(x + y) - g(x) - g(y) = g(x)g(y).$$

Das erinnert nun an eine Mischung zwischen der additiven und der multiplikativen Cauchy-Gleichung. Umformen liefert die äquivalente Form $g(x + y) = (g(x) + 1)(g(y) + 1) - 1$, wir substituieren daher nochmals $h(x) = g(x) + 1$ und erhalten schliesslich die Gleichung

$$h(x + y) = h(x)h(y).$$

Dies ist nun definitiv eine Cauchy-Gleichung, auf die wir unsere Resultate anwenden können. Nach Voraussetzung ist f überall nichtnegativ, wir wissen aber nicht, ob dies auch für h gilt. Das ist tatsächlich der Fall, wie man leicht sieht. Setze nämlich $x = y$ in die Gleichung für h, dann folgt $h(2x) = h(x)^2 \geq 0$ für alle reellen x, also nimmt h nur nichtnegative Werte an. Existiert ein x mit $h(x) = 0$, dann verschwindet h identisch, denn für alle reellen y ist dann $h(x + y) = h(x)h(y) = 0$. In diesem Fall nimmt aber $f(x) = -x - 1$ negative Werte an, was nicht sein darf. Wir suchen also nach Funktionen $h : \mathbb{R} \to \mathbb{R}_{>0}$. Die klassischen Lösungen sind in diesem Fall die Funktionen $h(x) = a^x$ mit einer positiven Konstanten a. Um exotische Lösungen auszuschliessen, müssen wir nur noch zeigen, dass der Graph von h nicht dicht liegen kann in $\mathbb{R} \times \mathbb{R}_{>0}$. Nach Voraussetzung gilt aber $f(x) \geq 0$, also $h(x) \geq x + 1$ für alle reellen x. Das unendliche Dreieck unter der Geraden $y = x+1$ enthält also tatsächlich keinen Punkt des Graphen von h, wie gewünscht. Wir müssen nun noch alle positiven a finden, für die die Ungleichung $a^x \geq x + 1$ auch tatsächlich für alle x erfüllt ist. Dazu verwenden wir analytische Methoden. Die Graphen von beiden Seiten gehen durch den Punkt $(0, 1)$ und müssen sich dort berühren, insbesondere muss die Funktion $l(x) = a^x$ an der Stelle $x = 0$ die Steigung 1 haben. Wegen $l'(x) = \ln(a) \cdot a^x$ muss also $a = e$ sein (Eulersche Konstante). Die Funktion $x \mapsto e^x$ ist nun konvex und besitzt an der Stelle $x = 0$ gerade die Tangente $x \mapsto x + 1$, folglich gilt $e^x \geq x + 1$ für alle x. Die einzige Lösung unserer Gleichung ist also die Funktion

$$f(x) = e^x - x - 1.$$

$\square$

Zum Schluss besprechen wir noch die allgemeine Lösung der additiven Cauchy-Gleichung. Dreh- und Angelpunkt der ganzen Sache ist die Existenz einer sogenannten $\mathbb{Q}$-Basis von $\mathbb{R}$. Eine Menge B reeller Zahlen heisst eine $\mathbb{Q}$-*Basis*, falls zu jeder reellen Zahl x eindeutig bestimmte rationale Zahlen α_b, $(b \in B)$ existieren, von denen alle ausser endlich viele verschwinden, sodass $x = \sum_{b \in B} \alpha_b \cdot b$ gilt. Mit anderen Worten, jede reelle Zahl lässt sich eindeutig als endliche, rationale Linearkombination von Elementen aus B schreiben. Hier kommt jetzt das Problem der Nichtexplizitheit ins Spiel. Denn die Existenz von $\mathbb{Q}$-Basen ist zwar durch das schon erwähnte Auswahlaxiom gesichert, angeben kann man eine solche Basis jedoch prinzipiell nicht. Nicht allzu schwierig zu sehen ist aber, dass B überabzählbar viele Elemente haben muss.

Sei nun f eine Lösung der additiven Cauchy-Gleichung und x wie vorher. Nach Lemma 2.1. gilt dann

$$f(x) = f\Big(\sum_{b \in B} \alpha_b \cdot b\Big) = \sum_{b \in B} \alpha_b f(b), \tag{17}$$

also ist f durch die Werte auf der Menge B eindeutig bestimmt. Umgekehrt kann man die Werte $f(b)$ beliebig vorgeben und f durch die Gleichung (17) *definieren.* Dass f dadurch wirklich wohldefiniert und additiv ist, folgt aus der Existenz und Eindeutigkeit der $\mathbb{Q}$-Linearkombinationen (man vergleiche dazu auch Beispiel 7). Die Menge der Lösungsfunktionen besitzt also die überabzählbar vielen reellen Freiheitsgrade $f(b)$, $b \in B$. Der Fall einer linearen Lösung $f(x) = ax$ entspricht dann dem Fall $f(b) = ab$ für alle $b \in B$. Das heisst, die frei wählbaren Werte $f(b)$ sind quasi alle 'gleichgeschaltet', und das ist eben die absolute Ausnahme. Man darf daher schon getrost sagen, dass die linearen Lösungen krass in der Unterzahl sind!