

Dynamische Systeme als Chance für den Schulunterricht

REINHARD WINKLER (TU WIEN)

Dynamische Systeme sind nicht nur Gegenstand aktueller mathematischer Forschung auf höchstem Niveau. Sie bieten auch ein reichhaltiges Angebot für den Schulunterricht. Eine wesentliche Unterscheidung ist die zwischen diskretem und kontinuierlichem Zeitparameter. Im diskreten Fall stößt man auf Iterationen einer Abbildung $T : X \rightarrow X$ von einer Menge X auf sich selbst, im kontinuierlichen auf (autonome) Differentialgleichungen. Beide Themenkreise werden vor allem anhand von Beispielen vorgestellt, die sich großteils auch in den Schulunterricht einbauen lassen. Dabei können wichtige Grundideen vielfältiger Teilgebiete der Mathematik vermittelt werden. Die vorliegende, in den Didaktikheften gedruckte Fassung ist die kürzere von zwei Versionen. Die längere enthält außerdem einen Anhang mit Beweisen, auch einiger zugrundeliegender Sätze aus der Analysis. Sie ist im Internet über die homepage der ÖMG frei verfügbar.

1. Einleitung

1.1. Dynamische Systeme als mathematisches Querschnittsthema

Dynamische Systeme sind in mehrfacher Hinsicht faszinierend (für eine ziemlich umfangreiche Darstellung sei auf Katok, Hasselblatt (1995) verwiesen, leichter zugänglich ist Peitgen et al. (1994) mit Schwerpunkt auf dem Thema *Chaos*). Aus der Sicht der reinen Mathematik bieten sie Verbindungen zu den meisten wichtigen Teilgebieten. Gleichzeitig sind sie von höchster Relevanz für die Anwendungen. Und schließlich eröffnen sie auch für den Schulunterricht sehr attraktive Zugänge. Letzteres steht im Mittelpunkt dieses Artikels. Andere didaktisch orientierte Artikel zu gewissen Aspekten des Themas sind Humenberger (2010) und Götz, Hofbauer (2013).

Ein wesentliches Instrument der Darstellung, welches sich bei dynamischen Systemen besonders anbietet, ist die Betonung von Querverbindungen; sowohl innerhalb der Mathematik, zur Beschreibung von Phänomenen aus anderen Wissenschaften wie auch zur Lösung technischer Probleme.

Als mathematisches Teilgebiet werden dynamische Systeme durch eine Klammer zusammengehalten, die nicht in einer speziellen Methodik oder einer bestimmten grundlegenden mathematischen Theorie besteht, auf der sie fußen. Das Schlagwort *dynamisch* bezieht sich nämlich zunächst auf einen außermathematischen Aspekt, und zwar auf die Änderung von Zuständen in der Zeit. Die genauere Natur dieser Zustände kann zunächst völlig offen bleiben. Diese Offenheit bedingt, dass je nach zu modellierendem Gegenstand ganz unterschiedliche mathematische Strukturen ins Spiel kommen. So gibt es kaum ein nennenswertes Teilgebiet der Mathematik, das nicht auch im Kontext dynamischer Systeme eine Rolle spielt. In diesem Sinn wollen wir dynamische Systeme eher als mathematisches Querschnittsthema denn als mathematische Teildisziplin auffassen. Bevor wir dies vertiefen, ist es zielführend, sich dem Thema unter dem Gesichtspunkt der Anwendungen zu nähern.

1.2. Dynamische Systeme als Mathematik von Zeit und Kausalität

Der Gegenstand der Mathematik selbst sind keine physischen Objekte, sondern idealisierte, oft abstrakte. Weil den Mathematiker nichts daran hindert, Idealisierungen und Abstraktionen nach Gutdünken vorzunehmen, gibt das der Mathematik eine enorme Freiheit. Doch bedeutet das keineswegs Beliebigkeit. Denn Mathematik soll vor allem die Möglichkeiten menschlichen Weltverständnisses ausloten. Daraus ergibt sich eine enge Anbindung an grundlegende (in einem, wenn auch nicht logischen, so doch entwicklungspsychologischen Sinne vielleicht sogar apriorische) Kategorien wie, um nur fünf besonders wichtige zu nennen, die Zahl (vgl. Dehaene (1999)) oder – nach Kant – der Raum, die Zeit und die Kausalität – und darüber hinaus, als Antipode der Kausalität, der Zufall. Zwanglos ordnen wir der Zahl die mathematischen Gebiete Zahlentheorie und Algebra zu, dem Raum die Geometrie, der Kausalität den Funktionsbegriff (als Zuordnung Ursache \mapsto Wirkung) und dem Zufall die Stochastik. Und wie verhält es sich mit der Zeit?

Um Missverständnisse zu vermeiden: Hier ist der Begriff *Zeit* nicht im Sinn der Relativitätstheorie gemeint, sondern zunächst in einem eher naiven oder – vielleicht genauer – in einem anthropologisch-psychologischen Sinn. Mathematisch formuliert: Wir erleben die Zeit als ein eindimensionales Kontinuum, das sich in befriedigender Weise als Zahlengerade \mathbb{R} modellieren lässt; genauso wie eine Koordinatenachse im Anschauungsraum. Insofern erhalten wir nichts Neues gegenüber elementaren geometrischen Begriffen. Doch kommt es uns vor allem auf die zeitliche Abfolge von Zuständen an, sofern diese in einer Beziehung von Ursache und Wirkung stehen. Es geht also um Kausalbeziehungen, die ihrerseits unveränderlichen Gesetzen folgen. Damit ist ein wesentliches Ziel unseres Strebens nach Erkenntnis benannt: Verständnis des Veränderlichen durch das Unveränderliche oder, anders formuliert, Strukturierung der Zeit durch das Wirken von Kausalität in Form von Naturgesetzen.

Bei allen Relativierungen durch die Erkenntnisse der modernen Physik wie Unschärferelation u.ä. wirkt in uns dennoch eine der Aufklärung verpflichtete Überzeugung, dass viele für uns höchst relevante Vorgänge im Rahmen rational fassbarer Kausalitäten verstanden werden können – je deterministischer, desto klarer erscheinen sie uns in der Regel. Dynamische Systeme sind die mathematischen Modelle solcher Vorgänge.

1.3. Überblick

Nach der Einleitung soll in Kapitel 2 die mathematische Modellierung dynamischer Systeme besprochen werden. Am einfachsten ist diese im Fall diskreter Zeit (Abschnitt 2.1), wo alles auf die Iteration von Transformationen $T : X \rightarrow X$, d.h. Funktionen (Abbildungen) von einer Menge X (Zustands- oder Phasenraum) auf sich selbst hinausläuft. Bei kontinuierlicher Zeit (2.2) sind schon anspruchsvollere Überlegungen, die typischer Weise zu Differentialgleichungen führen, erforderlich. Trotzdem gibt es eine abstraktere Sichtweise, die beide Fälle vereinheitlicht, nämlich die Untersuchung von sogenannten Aktionen von Gruppen und Halbgruppen, die – nicht für den Unterricht selbst, sondern als Hintergrundinformation für Lehrende – in 2.3 kurz vorgestellt werden.

Es folgen in Kapitel 3 mehrere charakteristische Beispiele bzw. Beispielklassen von dynamischen Systemen mit diskreter Zeit. Die einzelnen Abschnitte behandeln zunächst Grundbegriffe wie Fix- und periodische Punkte (3.1), dann lineares T (3.2), die Fixpunkteigenschaft von konvergenten Orbits unter stetigem T (3.3), Beispiele dazu im Fall des Intervalls (3.4 und 3.5) oder des Kreises (3.6). Verzichtet man auf die Stetigkeit von T , stößt man fast zwangsläufig auf Transformationen von Folgen über endlichen Alphabeten (symbolische Dynamik, 3.7) und auf die (maßtheoretische) Ergodentheorie (3.8). Im Kapitel 4 begnügen wir uns mit einer kleinen Auswahl wichtiger Themen. Viele Beweise werden in dieser Version des Artikels ausgespart. Man findet sie im Anhang der Langversion Winkler (2014 Langversion).

2. Mathematische Modellbildung

2.1. Diskrete Zeit – der informatische Blick

Bei dynamischen Systemen mit diskreter Zeit wird ausgehend von einem Anfangszustand $x_0 \in X$ eine Folge weiterer Zustände $x_n, n \in \mathbb{N}$, durch die rekursive Vorschrift $x_{n+1} = T(x_n)$ generiert. Diesem Modell kann die Vorstellung unterlegt werden, dass X die Menge aller möglichen Zustände eines Systems ist und $T : X \rightarrow X$ eine Funktion, die jedem möglichen Zustand x in deterministischer Weise den Folgezustand $T(x)$ im nächsten (diskreten) Zeitpunkt zuordnet. $T(x)$ kann also als kausale Folge des Zustandes x gedeutet werden, die sich in einem einzigen Schritt, der die kleinste betrachtete Zeiteinheit in Anspruch nimmt, einstellt.

Zu jeder Funktion $T : X \rightarrow X$ und zu jedem Anfangswert $x_0 \in X$ erhalten wir die Folge $(T^n(x_0))_{n \in \mathbb{N}}$. Dabei bedeutet T^n die n -fache Hintereinanderausführung von T . Es ist unmittelbar klar (Induktion), dass diese Folge die eindeutige Lösung der Rekursion $x_{n+1} = T(x_n)$ bei Vorgabe von x_0 ist. Man beachte, dass die hier angesprochene Induktion nicht nur für die Mathematik fundamental ist (siehe z.B. Winkler (2008)), sondern auch für die Informatik.

Als **dynamisches System mit diskreter Zeit** definieren wir nun ein Paar (X, T) , wobei X eine Menge (oft mit zusätzlicher Struktur) ist und $T : X \rightarrow X$ eine Abbildung von X auf sich selbst. Die Untersuchungen von (X, T) beschäftigen sich vorwiegend damit, mit Hilfe der Kenntnis von x_0 und T möglichst weitreichende Aussagen über die Menge aller $x_n, n \in \mathbb{N}$, (den **Orbit** von x_0) bzw., genauer, über die Folge $(x_n)_{n \in \mathbb{N}}$ (die **Trajektorie** von x_0) zu machen, insbesondere über ihr Verhalten für $n \rightarrow \infty$.

Man beachte, dass bei bijektivem T , also bei Existenz einer Umkehrabbildung T^{-1} , als Menge der Zeitpunkte statt \mathbb{N} auch \mathbb{Z} genommen werden kann. Das bedeutet, dass das dynamische System keinen Anfangszustand hat, sondern eine unendliche Vergangenheit.

2.2. Kontinuierliche Zeit – der physikalische Blick

Auf den ersten Blick ist die Situation bei dynamischen Systemen mit kontinuierlicher Zeit völlig anders, weil es ja keine kleinste Zeiteinheit gibt, für die Übergänge von einem Zustand zum nächsten durch eine einzige Funktion T beschrieben werden können. Es liegt daher nahe, für jede denkbare Zeitspanne $t > 0, t \in \mathbb{R}$, eine Transformation $F_t : X \rightarrow X$ zu betrachten, welche modelliert, wie ein Zustand x nach t Zeiteinheiten in einen Zustand $F_t(x)$ übergeht.

Will man wieder deterministische Abläufe beschreiben, so soll der Zustand nach $s+t$ Zeiteinheiten nicht davon abhängen, ob bereits nach s Einheiten eine Zwischenbeobachtung gemacht wird oder nicht, als Formel: $F_t(F_s(x)) = F_{s+t}(x)$. Der Eigenschaft $T^n \circ T^m = T^{n+m} = T^m \circ T^n$ (für alle $n, m \in \mathbb{N}$) im diskreten Fall entspricht also im kontinuierlichen Fall:

$$F_t \circ F_s = F_{s+t} = F_s \circ F_t \quad \text{für alle } s, t \in \mathbb{R}.$$

Bevor wir in 2.3 den abstrakten algebraischen Kern und die entsprechenden Verallgemeinerungen kurz beleuchten, wollen wir einer für uns hier noch wichtigeren Sichtweise nachspüren. Sie ist der Vorstellung des Kontinuums $[0, \infty[$ oder \mathbb{R} , welchem die Zeitpunkte t entnommen sind, inhärent.

Der Einfachheit halber wollen wir den Zustandsraum $X = \mathbb{R}$ oder $X \subseteq \mathbb{R}$ voraussetzen, was etwa dann sinnvoll ist, wenn als mögliche Zustände eindimensionale Messgrößen vorliegen. Für gegebenes $x_0 \in X$ (Anfangszustand) schreiben wir $x_t = F_t(x_0)$ für den resultierenden Zustand zum Zeitpunkt t . Wir fassen also x selbst als Funktion $x : \mathbb{R} \rightarrow X, t \mapsto x(t) = x_t$, mit $x(t) = F_t(x_0)$ auf. Dem entspricht im diskreten Fall die Folge (= Funktion auf \mathbb{N}) $(T^n(x_0))_{n \in \mathbb{N}} : \mathbb{N} \rightarrow X, n \mapsto x_n = T^n(x_0)$.

Bei den meisten real auftretenden Prozessen in der Natur erwarten wir, das Stetigkeitsprinzip *natura non facit saltus*¹ ausweitend, dass sich der Zustand $x_{t+\varepsilon}$ bei kleinem ε nicht nur wenig von x_t unterscheidet, sondern dass diese Abweichung sogar annähernd proportional ist zu ε mit einem geeigneten Proportionalitätsfaktor, der selbst wieder vom Zustand x_t abhängt und für den wir daher $f(x_t)$ schreiben. Wir haben also

$$x_{t+\varepsilon} = F_\varepsilon(x_t) \approx x_t + \varepsilon f(x_t)$$

bzw. umgeformt und genauer

$$\lim_{\varepsilon \rightarrow 0} \frac{x_{t+\varepsilon} - x_t}{\varepsilon} = f(x_t).$$

Fassen wir x_t als Funktionswert einer Funktion (Trajektorie) $x : t \mapsto x_t$ zum Zeitpunkt t auf, so wird diese Beziehung zur Differentialgleichung

$$x' = f(x), \quad \text{genauer: } x' = f \circ x, \quad \text{d.h.: } x'(t) = f(x(t)) \text{ für alle } t.$$

Dynamische Systeme mit kontinuierlicher Zeit (auch hier kann durch Erweiterung auf reelle $t \leq 0$ auch eine Vergangenheit mit einbezogen werden) sind also im Wesentlichen Differentialgleichungen der oben

¹ Die Natur macht keine Sprünge.

angegebenen Form. Man spricht auch von **autonomen Differentialgleichungen**; im Gegensatz zu solchen, wo $f = f(x, t)$ auch noch vom Zeitpunkt t abhängt. Dies entspräche einer Veränderung der zugrundeliegenden Gesetzmäßigkeit im Laufe der Zeit, abweichend vom eingangs dargestellten Paradigma.

Klarerweise kann der Zustandsraum X auch höherdimensional sein, z.B. $X = \mathbb{R}^n$ oder eine Teilmenge davon. Wesentlich ist vor allem die lineare Struktur oder genauer: dass darin sinnvoll Differentialrechnung betrieben werden kann. In Kapitel 4 werden wir Beispiele und auch allgemeine Theorie dazu behandeln.

2.3. Aktionen von Gruppen und Halbgruppen als gemeinsamer Rahmen

Dieser kurze Abschnitt ist nicht für den unmittelbaren Gebrauch im Schulunterricht gedacht, sondern als Hintergrundwissen für Lehrende. Wir werden im Weiteren nicht darauf zurückgreifen. Es geht um eine gemeinsame Verallgemeinerung der beiden behandelten Typen dynamischer Systeme (diskrete und kontinuierliche Zeit), nämlich um die Aktion von Gruppen und Halbgruppen.

Eine **Halbgruppe** ist bekanntlich eine Menge H zusammen mit einer binären Operation \cdot auf H , d.h. mit einer Abbildung $\cdot : H \times H \rightarrow H$, $(h_1, h_2) \mapsto h_1 \cdot h_2$ (wir schreiben statt $h_1 \cdot h_2$ oft einfach $h_1 h_2$), welche das **Assoziativgesetz** erfüllt: $(h_1 h_2) h_3 = h_1 (h_2 h_3)$ für alle $h_1, h_2, h_3 \in H$. Gibt es ein **neutrales Element** $e \in H$ (d.h. $he = eh = h$ für alle $h \in H$) spricht man auch von einem **Monoid**. Unter einer **Halbgruppenaktion** von H auf einer Menge X versteht man eine Abbildung $\alpha : H \times X \rightarrow X$, Schreibweise $\alpha : (h, x) \mapsto \alpha(h, x) = hx$, mit $h_1 (h_2 x) = (h_1 h_2) x$ für alle $x \in X$. Ist $e \in H$ neutrales Element und $ex = x$ für alle $x \in X$, so heißt α sogar eine **Monoid-** oder, wenn H sogar eine **Gruppe** ist (d.h. überdies zu jedem $h \in H$ ein **Inverses** h^{-1} enthält, welches durch $hh^{-1} = h^{-1}h = e$ gekennzeichnet ist), eine **Gruppenaktion** von H auf X .

Man macht sich sofort klar, dass Rekursionen $x_{n+1} = T(x_n)$ (d.h. Iterationen von T) auf der Menge X im Sinn von Abschnitt 2.1 als Halbgruppen- und sogar Monoidaktionen α der additiven Halbgruppe \mathbb{N} mit $\alpha(n+1, x) = T(\alpha(n, x)) = T^{n+1}(x)$ aufgefasst werden können. Analog entsprechen dynamische Systeme mit kontinuierlicher Zeit Aktionen des additiven Monoids $[0, \infty[$ der nichtnegativen reellen Zahlen mit $\alpha(t_1 + t_2, x) = \alpha(t_2, \alpha(t_1, x))$. Damit allein sind freilich keinerlei Bedingungen hinsichtlich Stetigkeit oder gar Differenzierbarkeit gestellt. Aktionen von Halbgruppen bieten aber eine Möglichkeit, dynamische Systeme sowohl mit diskretem wie auch mit kontinuierlichem Zeitparameter unter einem einheitlichen Begriff zu subsumieren.

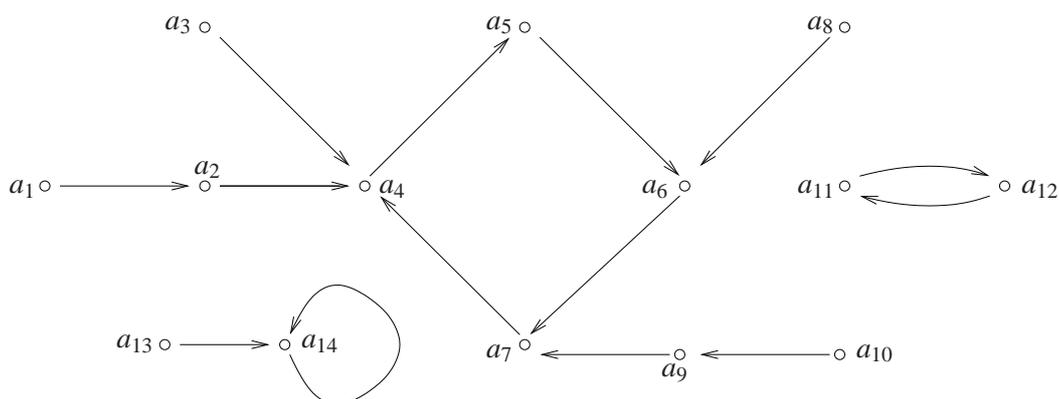
3. Beispiele dynamischer Systeme mit diskreter Zeit

3.1. Periodische Punkte, Fixpunkte

Die Möglichkeiten für Trajektorien $(x_n)_{n \in \mathbb{N}}$, $x_n := T^n(x)$, eines festen Punktes $x = x_0$ unter einer Transformation $T : X \rightarrow X$ überblickt man, wenigstens sofern X keine weitere Struktur trägt, schnell: Angenommen, ein Punkt tritt in der Trajektorie mehrmals auf, d.h. $x_{n_1} = x_{n_2}$ für gewisse $n_1 < n_2 \in \mathbb{N}$, so gibt es auch ein kleinstes derartiges n_1 und dazu ein kleinstes $n_2 > n_1$. Das bedeutet, dass $x_0, x_1, \dots, x_{n_1}, \dots, x_{n_2-1}$ alle verschieden sind, dann jedoch wegen $x_{n_1+1} = T(x_{n_1}) = T(x_{n_2}) = x_{n_2+1}$ und, induktiv, $x_{n_1+k} = x_{n_2+k}$ für alle $k \in \mathbb{N}$, eine Periode der Länge $p = n_2 - n_1$ auftritt, die wir in Zyklenschreibweise notieren als $(x_{n_1} x_{n_1+1} \dots x_{n_1+p-1})$. Die vorangehenden Punkte x_0, \dots, x_{n_1-1} bilden die sogenannte **Vorperiode**. Ist $n_1 = 0$, so heißt x ein **periodischer Punkt**, im Fall $T(x) = x$, also $p = 1$, **Fixpunkt**. Gibt es gar kein n_1 der beschriebenen Art (was natürlich nur bei unendlichem X möglich ist), nennen wir x einen **aperiodischen Punkt**. Die Bestimmung von Fixpunkten und Perioden ist bei allen dynamischen Systemen mit diskreter Zeit von vordringlichem Interesse. Trägt der Zustandsraum X , auf dem die Transformation $T : X \rightarrow X$ wirkt, keine zusätzliche Struktur, gibt es über die Trajektorie eines einzelnen Punktes darüber hinaus auch wenig Interessantes zu sagen.

Ist X endlich, so lassen sich die wesentlichen Aspekte des dynamischen Systems sehr übersichtlich aus gerichteten Graphen mit der Knotenmenge X ablesen, wo eine gerichtete Kante (Pfeil) $a \rightarrow b$ immer

dann vorliegt, wenn $T(a) = b$. Perioden werden dabei als Kreise sichtbar. Hier ein Beispiel mit $|X| = 14$.



Klarerweise ändern sich die dynamischen Eigenschaften eines Systems nicht, wenn wir lediglich die Elemente aus X durch entsprechende aus einer Menge Y ersetzen. Man nennt daher allgemein zwei dynamische Systeme (X, T_X) und (Y, T_Y) **isomorph**, wenn es eine bijektive Abbildung $\phi : X \rightarrow Y$ (einen sogenannten **Isomorphismus**) gibt mit $\phi(T_X(x)) = T_Y(\phi(x))$ für alle $x \in X$, d.h. mit $\phi \circ T_X = T_Y \circ \phi$; verlangt man von f nur Surjektivität, heißt (Y, T_Y) ein **Faktor** von (X, T_X) .

Tragen X und Y zusätzliche Struktur, so fordert man von ϕ außerdem meist die entsprechende Strukturverträglichkeit. Isomorphe Systeme können vom dynamischen Standpunkt aus als gleich gelten. Wir wollen das an einfachen, aber typischen Beispielen konkretisieren.

3.2. Jordansche Normalform linearer Transformationen

Wir betrachten nun ein Beispiel, das wir im Weiteren nicht brauchen und das für den Schulunterricht in dieser Form zu fortgeschritten ist, das aber zeigt, wie neben den später vorherrschenden analytischen Aspekten bei dynamischen Systemen genauso gut algebraische und geometrische ins Zentrum der Aufmerksamkeit treten können.

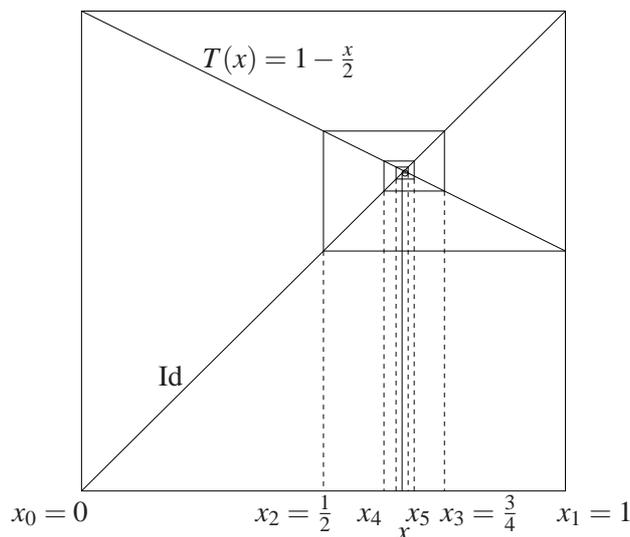
Seien dazu X und Y Vektorräume über dem Körper \mathbb{C} und $T_X : X \rightarrow X$, $T_Y : Y \rightarrow Y$ lineare Transformationen. Dann soll ein Isomorphismus $\phi : X \rightarrow Y$ zwischen den dynamischen Systemen (X, T_X) und (Y, T_Y) auch ein solcher zwischen den Vektorräumen X und Y sein. Insbesondere müssen X und Y dieselbe Dimension haben. Ist in so einer Situation zum Beispiel $x \in X$ ein Eigenvektor zu einem Eigenwert λ von T_X , so folgt $T_Y(\phi(x)) = \phi(T_X(x)) = \phi(\lambda x) = \lambda \phi(x)$. Also ist $y := \phi(x) \in Y$ ein Eigenvektor von T_Y zum Eigenwert λ . Rekapituliert man die Jordansche Normalform quadratischer Matrizen aus der Linearen Algebra, so erkennt man: Die Klassifikation quadratischer komplexer Matrizen im Sinne Jordanscher Normalformen entspricht der Klassifikation linearer dynamischer Systeme auf den entsprechenden endlichdimensionalen Vektorräumen.

3.3. Konvergente Orbits unter stetigen Transformationen

Häufig geht es bei dynamischen Systemen (X, T) vor allem um das langfristige Verhalten, d.h. mathematisch gesprochen um die Konvergenz von Trajektorien, was eine topologische Struktur auf X voraussetzt. Besonders interessant sind dann stetige T . Wir beginnen mit dem einfachen Beispiel

$$X = [0, 1], \quad T : x \mapsto 1 - \frac{x}{2}.$$

Für den Anfangswert $x_0 = 0$ lauten die nächsten Werte $x_1 = 1$, $x_2 = \frac{1}{2}$, $x_3 = \frac{3}{4}$, $x_4 = \frac{5}{8}$, ... Die Situation wird veranschaulicht durch folgende, auch **Spinnwebdiagramm** genannte Darstellung. Besonders sei auf die Diagonale (Graph der Identität Id) hingewiesen, auf der allfällige Fixpunkte erkennbar sind.



Allem Anschein nach konvergiert die Trajektorie gegen einen Wert, der dem Schnittpunkt der Diagonale (= der Graph der Funktion Id) mit dem Funktionsgraphen von T entspricht, das ist die Lösung der Gleichung $x = 1 - \frac{x}{2}$, also $x = \frac{2}{3}$. Wie die Abbildung suggeriert, gilt $\lim_{n \rightarrow \infty} x_n = \frac{2}{3}$. Ein strenger Beweis ergibt sich aus den folgenden einfachen, aber wichtigen Ideen. Einerseits ist wegen der Stetigkeit von T jeder Grenzwert $x = \lim_{n \rightarrow \infty} x_n$ einer Iterationsfolge $x_{n+1} = T(x_n)$ unter T notgedrungen ein Fixpunkt von T :

$$T(x) = T\left(\lim_{n \rightarrow \infty} x_n\right) = \lim_{n \rightarrow \infty} T(x_n) = \lim_{n \rightarrow \infty} x_{n+1} = \lim_{n \rightarrow \infty} x_n = x$$

Also kommen als Grenzwerte tatsächlich nur x -Koordinaten (= y -Koordinaten) der Schnittpunkte der Diagonale mit dem Funktionsgraphen von T in Frage. Damit ist aber noch nicht garantiert, dass die Iterationsfolge überhaupt konvergiert. Will man das für das vorliegende Beispiel beweisen, wird man fast notgedrungen zu einer Argumentation geführt, die auch in weitaus allgemeinerem Zusammenhang von Interesse ist.

Und zwar kommt es lediglich darauf an, dass die Transformation T zwei beliebige Punkte $x, x' \in X$ zusammenzieht, genauer: $|T(x) - T(x')| \leq \lambda |x - x'|$ mit einem Faktor $\lambda < 1$ (im Beispiel kann $\lambda = \frac{1}{2}$ gesetzt werden), in welchem Falle man T eine **Kontraktion** nennt. Der Abstand zweier aufeinanderfolgender Glieder wird dann mindestens in derselben Weise kleiner wie die Glieder einer konvergenten geometrischen Reihe, was innerhalb $[0, 1] \subseteq \mathbb{R}$ die Konvergenz der Iterationsfolge impliziert. Dieselben Überlegungen greifen allgemeiner in jedem vollständigen metrischen Raum. So erhält man den sogenannten **Fixpunktsatz von Banach**² in seiner vollen Allgemeinheit:

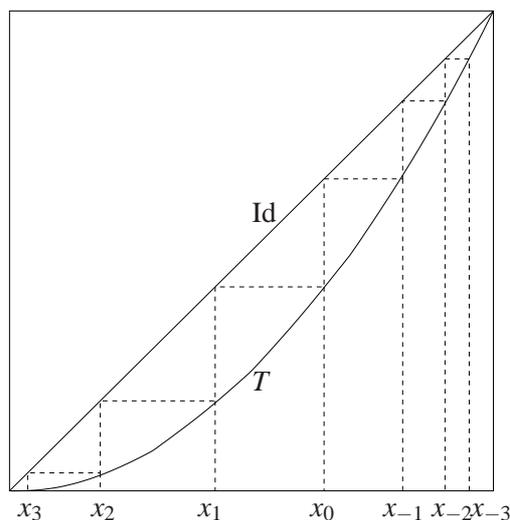
Jede Kontraktion auf einem vollständigen metrischen Raum hat genau einen Fixpunkt. Jede Trajektorie konvergiert gegen diesen Fixpunkt. Die Konvergenzgeschwindigkeit kann mit Hilfe einer geometrischen Reihe abgeschätzt werden. (Genaueres dazu findet sich in Winkler (2014 Langversion).)

Man mache sich klar, dass damit auf einen Schlag viele rekursive Folgen als konvergent identifiziert sind, insbesondere alle der Form $x_{n+1} = \alpha x_n + \beta$ mit $|\alpha| < 1$. Dabei können sowohl β als auch der Startwert $x_0 \in \mathbb{R}$ beliebig angenommen werden. Grenzwert ist die (eindeutige) Lösung der Gleichung $x = \alpha x + \beta$, also $x = \frac{\beta}{1-\alpha}$.

3.4. Beispiele isomorpher stetiger Intervallabbildungen

Nochmals betrachten wir das Einheitsintervall $X = [0, 1]$, diesmal aber mit der Transformation $T : x \mapsto x^2$. Für einen Startpunkt $x_0 < 1$ erhält man die extrem rasch gegen 0 konvergente Folge der Punkte $x_n = T^n(x_0) = x_0^{2^n}$. Da T auf X bijektiv ist, lässt sich das System auch in die Vergangenheit verfolgen mit entsprechenden Punkten x_{-n} .

² Stefan Banach (1892-1945)



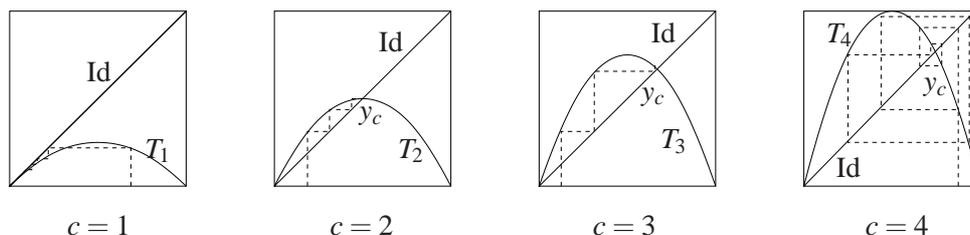
Hätten wir statt der Transformation $T = T_2 : x \mapsto x^2$ beispielsweise $T_3 : x \mapsto x^3$ genommen, hätte sich qualitativ nicht viel geändert, denn $([0, 1], T_2)$ und $([0, 1], T_3)$ erweisen sich als isomorph. Das gleiche ist der Fall, sofern $I = [a, b]$ irgendein abgeschlossenes Intervall ist und $T : I \rightarrow I$ eine stetige und streng monoton wachsende Transformation mit $T(a) = a < b = T(b)$ und $T(x) < x$ (dual für $T(x) > x$) für alle x mit $a < x < b$. Also: Je zwei Systeme (X, T_X) und (Y, T_Y) mit diesen Eigenschaften sind isomorph. An dieser Stelle begnügen wir uns mit dem Verweis auf die anschauliche Plausibilität dieser Behauptung. Eine Beweisskizze findet sich in Winkler (2014 Langversion).

Ähnliche Aussagen lassen sich machen, wenn $T : I \rightarrow I$ weiterhin als stetig und monoton wachsend vorausgesetzt wird, man allerdings Fixpunkte im Inneren von I zulässt, bei denen die Situationen $T(x) > x$ und $T(x) < x$ abwechseln können. Es ist eine wertvolle Übung, sich das im Einzelnen durch den Kopf gehen zu lassen.

Nur geringfügig komplizierter ist die Situation bei monoton fallendem $T : I \rightarrow I$, denn dann sind die obigen Überlegungen auf die monoton wachsende Iterierte $T^{(2)} := T \circ T$, $T^{(2)}(x) = T(T(x))$ anwendbar. Fixpunkte von $T^{(2)}$ bleiben unter T selbst entweder ebenfalls fix, oder sie sind Elemente einer Periode von T der Länge 2. Viel komplizierter kann die Situation jedoch werden, wenn man auf die Monotonie verzichtet, was nun geschehen soll.

3.5. Chaos entsteht

Weiterhin betrachten wir stetige Transformationen eines Intervalls, und zwar speziell $X = [0, 1]$ und $T = T_c : X \rightarrow X, x \mapsto cx(1-x)$ mit $0 \leq c \leq 4$. Der Funktionsgraph ist eine nach unten offene Parabel mit Nullstellen bei 0 und 1 und Maximum $m_c = \frac{c}{4}$ bei $x = \frac{1}{2}$. (Also gilt $T_c : [0, 1] \rightarrow [0, 1]$ tatsächlich genau für $c \in [0, 4]$.) Fixpunkte sind 0 und, sofern $c > 1$, auch $y_c = 1 - \frac{1}{c}$. Später wird auch die Ableitung $T'_c(x) = c(1-2x)$ eine Rolle spielen. Es empfiehlt sich, für c verschiedene Fälle zu unterscheiden, die mit wachsendem c immer komplizierter werden. Entsprechend müssen die folgenden Erörterungen zunehmend fragmentarisch bleiben.



1.Fall: $0 \leq c \leq 1$. Die Situation ist weitgehend trivial, weil $T(x) < x$ für alle $x > 0$ gilt und somit, abgesehen vom Fixpunkt 0, jede Trajektorie streng monoton fallend und beschränkt ist, also konvergiert. Als Grenzwert kommt nur der einzige Fixpunkt 0 in Frage.

2.Fall: $1 < c \leq 2$. Wir konstatieren zunächst $y_c = 1 - \frac{1}{c} \leq \frac{1}{2}$ und $m_c = \frac{c}{4} \leq \frac{1}{2}$. Die Einschränkung von T auf das Intervall $I = [0, y_c]$ fällt unter die in Abschnitt 3.4 behandelte Situation mit $T(x) > x$. Folglich konvergieren alle Trajektorien, die im Inneren von I beginnen, monoton wachsend gegen y_c . Aus symmetrischen Gründen konvergieren Trajektorien mit $y_c < x_0 \leq \frac{1}{2}$ monoton fallend gegen y_c . Für $x_0 > \frac{1}{2}$ schließlich liegt $x_1 = T(x_0) \leq \frac{1}{2}$ in $[0, \frac{1}{2}]$, und wir landen in einer der bereits behandelten Situationen mit Startwert x_1 statt x_0 . Also konvergieren für $0 < x_0 < 1$ alle Trajektorien gegen y_c .

3.Fall: $2 < c \leq 1 + \sqrt{3} = 2.73 \dots$. Nun gilt $\frac{1}{2} < y_c \leq \frac{3-\sqrt{3}}{2} = 0.63 \dots$ und $\frac{1}{2} < m_c \leq \frac{1+\sqrt{3}}{4} = 0.68 \dots$. Außerdem spielt jene Zahl z_c eine Rolle, für die $T'(z_c) = c(1-2z_c) = -1$ gilt, nämlich $z_c = \frac{1}{2} + \frac{1}{2c}$. Es gilt $\frac{1}{2} < y_c < m_c \leq \frac{1+\sqrt{3}}{4} = 0.68 \dots \leq z_c$. Im Intervall $[\frac{1}{2}, z_c]$ fällt T' strikt von 0 auf -1 . Damit überlegt man sich (genau genommen mit Hilfe des Mittelwertsatzes, siehe Winkler (2014 Langversion)), dass T in jedem echten Teilintervall $[\frac{1}{2} + \varepsilon, z_c - \varepsilon]$, $\varepsilon > 0$, die Voraussetzungen im Banachschen Fixpunktsatz erfüllt. Somit konvergieren alle Orbits mit $\frac{1}{2} < x_0 < z_c$ gegen den eindeutigen Fixpunkt y_c . Ähnlich wie im 2.Fall zeigt es sich, dass Trajektorien auch für andere Startwerte $x_0 > 0$ irgendwann in das Intervall $I_0 =]\frac{1}{2}, z_c[$ münden, folglich gegen y_c konvergieren. Das kann man sich zum Beispiel klar machen, indem man iterativ die Urbilder $I_{n+1} = T^{-1}(I_n)$ untersucht und feststellt, dass diese eine aufsteigende Mengenfolge $I_0 \subset I_1 \subset I_2 \subset \dots$ bilden, deren Vereinigung das gesamte offene Intervall $]0, 1[$ ausfüllt. Also münden abgesehen von den Randfällen $x_0 = 0, 1$ alle Trajektorien schließlich in I_0 und konvergieren somit gegen $y_c = 1 - \frac{1}{c}$.

4.Fall: $1 + \sqrt{3} < c \leq 3$. Die Komplikation gegenüber dem 3.Fall ergibt sich daraus, dass nun $z_c < m_c$ gilt. Zwar kann T nun auf das Intervall $I = [\frac{1}{2}, m_c]$ eingeschränkt werden (man rechne $T(m_c) > \frac{1}{2}$ nach, woraus aus Monotoniegründen $T(I) \subseteq I$ folgt), es handelt sich aber nicht mehr um eine Kontraktion. Dennoch kann man sich überlegen, dass alle in I startenden Trajektorien alternierend gegen y_c konvergieren, und dass überdies alle Trajektorien mit $0 < x_0 < 1$ schließlich in diesen Bereich münden. Entscheidend ist, dass an den Randpunkten von I als Startwert x_0 das gewünschte Verhalten eintritt. Man hat die Folge $x_0 = \frac{1}{2}$, $x_1 = T(x_0) = m_c$, $x_2 = T(x_1) > \frac{1}{2}$ etc. zu betrachten, welche wegen der fallenden Monotonie von T auf I die Ungleichungskette $x_0 < x_2 < \dots < x_3 < x_1$ erfüllt. Zu zeigen bleibt, dass die Limiten $x' = \lim_{n \rightarrow \infty} x_{2n}$ und $x'' = \lim_{n \rightarrow \infty} x_{2n+1}$ im Fixpunkt y_c zusammenfallen, d.h. keine Periode der Länge 2 bilden. Dies folgt, sobald geklärt ist, dass die Iterierte $T^{(2)} = T \circ T$ in I keinen Fixpunkt außer y_c hat. Zu diesem Zwecke kann man beispielsweise mittels Kurvendiskussion sicherstellen, dass y_c die einzige Nullstelle der Funktion $f(x) = T^{(2)}(x) - x$ in I ist. Der Aufwand ist beträchtlich, und wir verzichten deshalb hier darauf. Jedenfalls konvergieren auch in diesem Fall noch alle Trajektorien mit $0 < x_0 < 1$ gegen y_c .

5.Fall: $3 < c \leq 4$. Wir wollen uns überlegen, dass T in diesem Fall eine echte Periode $(x_p y_p)$ ($x_p \neq y_p$) der Länge 2 hat. Es genügt dazu, einen Fixpunkt x_p der Iterierten $T^{(2)}$, $T^{(2)}(x) = T(T(x))$, oder, äquivalent, eine Nullstelle x_p von $f(x) := x - T^{(2)}(x)$ zu finden, die strikt zwischen den einzigen Fixpunkten 0 und $y_c = 1 - \frac{1}{c}$ von T liegt. Wegen der Stetigkeit von f und wegen des Zwischenwertsatzes für stetige Funktionen ist dafür wiederum hinreichend, zwei Punkte $x', x'' \in]0, y_c[$ zu finden, wo f unterschiedliches Vorzeichen annimmt. Dazu berechnen wir mit der Kettenregel $f'(x) = 1 - T'(x)T'(T(x))$, setzen die beiden Nullstellen $x = 0$ und $x_0 = y_c = 1 - \frac{1}{c}$ von f (Fixpunkte von T) ein und erhalten für diese Werte $f'(x) = 1 - T'(x)^2 = 1 - c^2(1-2x)^2$, also $f'(0) = 1 - c^2 < 0$ sowie nach kurzer Rechnung $f'(y_c) = 1 - (c-2)^2 < 0$. Hieraus folgt (Mittelwertsatz, siehe Winkler (2014 Langversion), und stetige Differenzierbarkeit von f) für hinreichend kleines $\varepsilon > 0$ einerseits $f(\varepsilon) < 0$, andererseits $f(y_c - \varepsilon) > 0$. Damit sind die beiden gesuchten Punkte, nämlich $x' := \varepsilon$ und $x'' := y_c - \varepsilon$, gefunden und die Existenz einer Periode der Länge 2 bewiesen.

Klarerweise ist die Analyse des 5.Falles damit noch lange nicht vollständig. Es zeigt sich, dass bei Annäherung $c \rightarrow 4$ Perioden beliebig großer Länge auftreten und die Situation immer unüberschaubarer wird. Wir haben es mit klassischen Beispielen dynamischer Systeme mit chaotischem Verhalten zu tun. Mehr dazu findet sich in Raith (2009).

3.6. Transformationen des Kreises

Durch minimale Modifikation eines Intervalls, nämlich durch Aneinanderkleben (Identifikation) der Endpunkte, erhält man als Raum X einen Kreis. Wir betrachten die Transformation T_α , $0 \leq \alpha < 1$, welche X um den Anteil α einer vollen Umdrehung, das heißt um einen Winkel vom Bogenmaß $2\pi\alpha$ dreht, vereinbarungsgemäß gegen den Uhrzeigersinn.

Das dadurch resultierende dynamische System kennt (außer für $\alpha = 0$) keine konvergenten Trajektorien. Bei festem α (und gegebenem Anfangspunkt x_0 , der in diesem Fall jedoch keinen wesentlichen Einfluss auf die interessanten Phänomene hat) entspricht die Iteration von T_α dem fortgesetzten Abtragen immer desselben Winkels auf dem Kreis X . Zwei qualitativ verschiedene Fälle sind denkbar. Entweder der Ausgangspunkt x_0 wird nach endlich vielen Schritten wieder erreicht, so dass eine Periode entsteht. Das bedeutet, dass ein Vielfaches $q\alpha$ von α ($q = 1, 2, \dots$) ein Vielfaches einer vollen Umdrehung liefert, also $q\alpha = p \in \mathbb{Z}$. Das ist genau dann der Fall, wenn $\alpha = \frac{p}{q}$ rational ist. Ist hingegen α irrational, so tritt keine Periodizität auf.

Der irrationale Fall ist der interessantere. Man kann nämlich nicht nur zeigen, dass dann jede Trajektorie dicht liegt in X , d.h. jedes noch so kleine Kreissegment positiver Länge unendlich oft besucht. Diese Besuche treten sogar in einer Weise gleichmäßig auf, dass man von einer **gleichverteilten Folge** spricht: Jedes Kreissegment wird mit einer asymptotischen Häufigkeit besucht, die proportional zu seiner Länge ist. Der beeindruckendste Zugang zu diesen Phänomenen wurde in der bahnbrechenden und wunderschön zu lesenden Arbeit Weyl (1916)³ erschlossen. Die faszinierende Rolle von Kettenbrüchen in diesem Zusammenhang und die daraus resultierenden zahlentheoretischen Bezüge können an dieser Stelle nur schlagwortartig erwähnt werden.

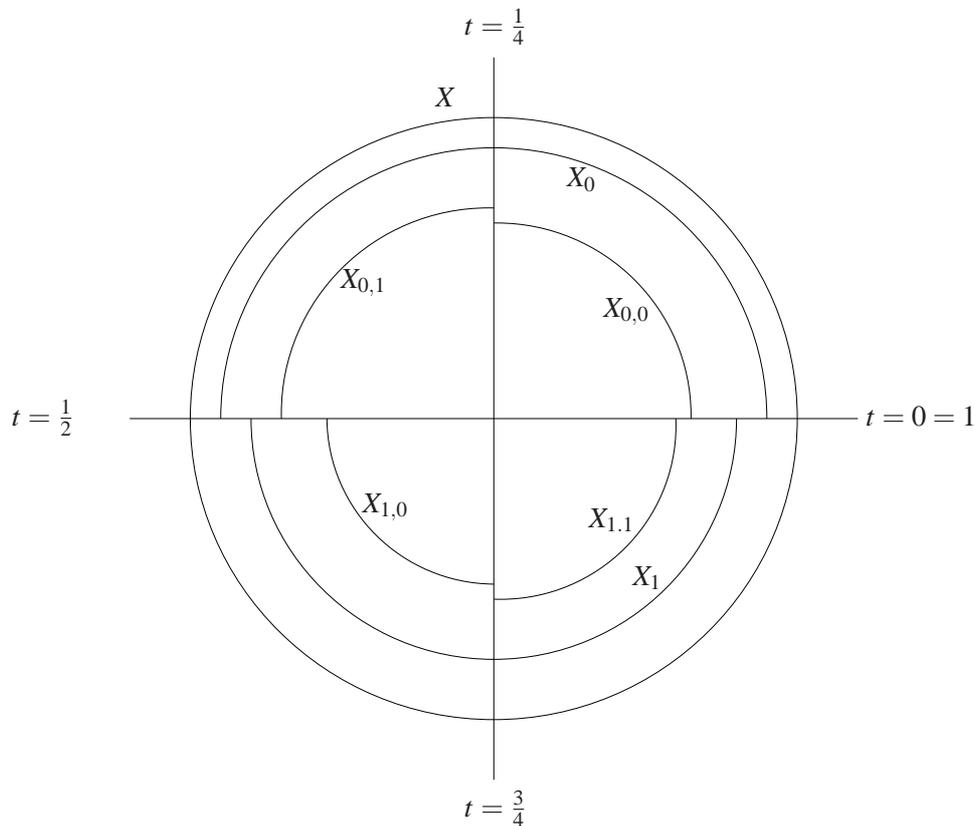
Eine interessante Transformation T des Kreises ganz anderen Typs entsteht, wenn man nicht um einen konstanten Winkel dreht, sondern diesen z.B. verdoppelt, genauer: Denken wir uns die Punkte des Kreises markiert mit den reellen Zahlen aus $[0, 1)$, so geht der Punkt t über in den Punkt $T(t) = 2t$, eventuell modulo 1 reduziert, d.h. wir ersetzen im Fall $2t \geq 1$ diese Zahl durch $2t - 1$. (Interpretiert man den Kreis als komplexen Einheitskreis, so lässt sich diese Transformation auch als $T : z \mapsto z^2$ auffassen.) Für dieses T treten plötzlich alle Möglichkeiten periodischer und nichtperiodischer Punkte auf. Für ein besseres Verständnis der Situation lohnt es, eine gänzlich andere Sichtweise einzunehmen, die uns zur Binärdarstellung reeller Zahlen führt.

3.7. Symbolische Dynamik

Wir unterteilen den Kreis X in zwei gleiche Halbkreise, die wir mit X_0 und X_1 bezeichnen (um die Schnittpunkte wollen wir uns vorerst nicht kümmern). Jede dieser Hälften zerlegen wir wieder in Teile (Viertelkreise) mit $X_0 = X_{0,0} \cup X_{0,1}$ und $X_1 = X_{1,0} \cup X_{1,1}$ usw. Jedem Punkt $x \in X$, der in dieser Prozedur nicht als Schnittpunkt auftritt, ordnen wir jene eindeutig bestimmte Binärfolge (a_0, a_1, a_2, \dots) , $a_i \in \{0, 1\}$, zu, so dass $x \in X_{a_0, a_1, a_2, \dots, a_k}$ für alle $k \in \mathbb{N}$ gilt. Man beachte: Schneiden wir den Kreis X beim Punkt $t = 0$ auf und interpretieren wir ihn nun als Einheitsintervall, so entspricht diese Vorgangsweise genau der Darstellung der reellen Zahlen zwischen 0 und 1 im Binärsystem durch unendliche 0-1-Folgen.

Nur kurz zu den Schnittpunkten: Diese entsprechen genau den dyadischen Zahlen, d.h. jenen von der Gestalt $t = \frac{p}{2^k}$ mit natürlichen Zahlen p und k . So einem t können zwei verschiedene Binärdarstellungen zugeordnet werden, z.B. für $t = \frac{1}{4}$ nicht nur $0, 1, 0, 0, 0, 0, \dots$, sondern auch $0, 0, 1, 1, 1, 1, \dots$ (analog zu $1, 000 \dots = 0, 999 \dots$ im dekadischen System). Da es nur abzählbar viele, im Vergleich zu allen t im Einheitsintervall also vergleichsweise verschwindend wenige dyadische Zahlen gibt, erlauben wir uns, diese weiterhin stiefmütterlich zu behandeln.

³ Hermann Weyl (1885-1955)



Folgen wir der Trajektorie $(x_n)_{n \in \mathbb{N}}$, $x_n = T^n(x)$, eines Punktes $x_0 \in X$ mit zugeordneter Binärfolge (a_0, a_1, a_2, \dots) unter der Transformation $T : t \mapsto 2t$ aus 3.6, so beobachten wir: Für $x_n \in I_0$ ist $a_n = 0$, für $x_n \in I_1$ ist $a_n = 1$. Die Binärfolge von x beschreibt also gerade seine Trajektorie unter T . Ersetzt man x durch $T(x)$, entspricht das in der Trajektorie der Streichung von a_0 , so dass die Folge (a_0, a_1, a_2, \dots) übergeht in die Folge (a_1, a_2, a_3, \dots) . Diese Transformation von Folgen nennt man auch die **Shift**-Abbildung. Man bezeichnet sie meist mit σ . Aus dem dynamischen System (X, T) haben wir also ein System (Y, σ) erhalten, wobei Y die Menge aller Binärfolgen bezeichnet. In Y treten Binärfolgen mit allen möglichen σ -Perioden auf, welchen gleichartige Perioden auch im ursprünglichen System (X, T) entsprechen.

Wegen der erwähnten Eindeutigkeitsprobleme der Binärdarstellung für dyadische $t \in X$ sind die beiden Systeme (X, T) und (Y, σ) nicht isomorph im strengen Sinn. Trotzdem enthalten sie weitgehend dieselbe Information. Die Symbole 0 und 1 erlauben eine Codierung des dynamischen Verhaltens von (X, T) in dem beschriebenen Sinn, weshalb man von **symbolischer Dynamik** spricht.

Wie können wir die Verwandtschaft der Systeme (X, T) und (Y, σ) präziser fassen und neue Einsichten daraus gewinnen? Dazu ist wieder ein neuer Gesichtspunkt hilfreich.

3.8. Maßtheoretische Ergodentheorie

Wir betrachten weiterhin die beiden Systeme (X, T) und (Y, σ) , wobei X jetzt der zum halboffenen Einheitsintervall $[0, 1[$ aufgeschnittene Kreis mit der Transformation $T : x \mapsto 2x \bmod 1$ sei, Y die Menge aller Binärfolgen $y = (a_n)_{n \in \mathbb{N}}$, $a_n \in \{0, 1\}$, und σ der Shift auf Y .

Die Zuordnung, welche jedes $x \in X$ in seine Binärdarstellung $y \in Y$ verwandelt, ist, wie bereits festgestellt, keine Abbildung im strengen Sinn, weil zu jedem dyadischen x zwei verschiedene Darstellungen gehören. Der bereits angesprochene Gesichtspunkt, dass die dyadischen x innerhalb X einen so kleinen Anteil haben, dass wir von ihnen absehen können, lässt sich unter Zuhilfenahme der Maßtheorie sehr elegant zu einer rigorosen Theorie ausbauen, indem man Mengen vom Maß 0 mit der leeren Menge identifiziert. Hier ist nicht der Platz, dies genauer auszuführen, daher nur ein paar Impressionen, die dennoch einen guten Eindruck von der Stärke der Ergodentheorie bieten.

Die wesentliche Beobachtung besteht darin, dass die beiden Systeme (X, T) und (Y, σ) maßtheoretisch

zusammenpassen. Zum Beispiel lässt sich die Menge X_0 aus 3.7 jetzt als das Intervall $[0, \frac{1}{2}[$ der Länge $\lambda(X_0) = \frac{1}{2}$ auffassen. In Y entspricht ihr die Menge Y_0 aller Binärfolgen, die mit 0 beginnen. Interpretieren wir die Folgen $y \in Y$ als Zufallsfolgen, wie sie durch iterierten Münzwurf zustandekommen, so erhält Y_0 die Wahrscheinlichkeit $\mu(Y_0) = \frac{1}{2} = \lambda(X_0)$. Es ist einsichtig, dass sich Entsprechendes auch für andere Teilmengen von X und Y sagen lässt, sogar für alle messbaren. Deshalb sind beide Systeme in einem maßtheoretischen Sinn isomorph. Hinzukommt, dass die beiden Transformationen T und σ maßerhaltend sind, was definitionsgemäß bedeutet, dass das Maß invariant unter Urbildern ist, d.h. $\lambda(T^{-1}(M)) = \lambda(M)$ für alle messbaren $M \subseteq X$, entsprechend für Y, σ und μ .

Die Bedeutung dieser Konzepte wird nicht zuletzt durch den **Birkhoffschen Ergodensatz**⁴ (siehe z.B. Walters (1982)) untermauert. Er impliziert u.a. grob gesprochen, dass in maßerhaltenden Systemen, die sich nicht in voneinander unabhängige Teilsysteme zerlegen lassen (sogenannte **ergodische Systeme**), Trajektorien sich insofern wie Zufallsfolgen verhalten, als ein **starkes Gesetz der großen Zahlen** gilt (siehe auch Winkler (2013)): Arithmetische Mittel von Funktionswerten entlang von Trajektorien konvergieren mit Wahrscheinlichkeit 1 gegen den Erwartungswert.

4. Beispiele dynamischer Systeme mit kontinuierlicher Zeit

4.1. Integration und die Fallgesetze des Galilei

Ein erstes, einfaches Beispiel für die in 2.2 allgemein beschriebene Situation bei kontinuierlicher Zeit ergibt sich, wenn wir einen aus einer gewissen (nicht zu großen) Höhe frei auf den Erdboden fallenden Körper untersuchen und uns für die Fallhöhe $s(t)$ zum Zeitpunkt t interessieren. Wie erstmals Galileo Galilei (1564-1641, wir gehen im Folgenden über den damaligen allgemeinen Wissensstand jedoch in mancherlei Hinsicht hinaus) wollen wir im Gedankenexperiment vom Luftwiderstand absehen. Dann ist der Körper ausschließlich der Erdbeschleunigung $g \approx 10 \frac{\text{m}}{\text{sec}^2}$ ausgesetzt, die wir bei geringen Fallhöhen, wo sich der Abstand zum Schwerpunkt der Erde relativ wenig ändert, als konstant annehmen dürfen.

Weil die Beschleunigung als Ableitung der (senkrecht nach unten gerichteten) Geschwindigkeit $v = v(t)$ nach der Zeit t definiert ist, haben wir $v'(t) = g$. Einfache Integrationsregeln zeigen, dass genau die Funktionen der Gestalt $v(t) = gt + c$ mit einer Konstanten c diese Eigenschaft haben (hier fließt der Mittelwertsatz ein, siehe auch Winkler (2014 Langversion)), wobei hier $c = v(0)$ zu setzen ist. Die Geschwindigkeit v wiederum ist die Ableitung des Weges $s = s(t)$ nach t , also $s'(t) = v(t) = gt + v(0)$, was analog zu

$$s(t) = \frac{g}{2}t^2 + v(0)t + s(0)$$

führt.

Kennen wir den Anfangszustand (Ort $s(0)$ und Geschwindigkeit $v(0)$) des Körpers zu Beginn, so bestimmt die Konstanz des wirksamen Naturgesetzes (Erdbeschleunigung) die weitere Entwicklung des dynamischen Systems also vollständig. Es liegt deshalb nahe, den zweidimensionalen Zustandsraum $X = \mathbb{R}^2$ für Ort und Geschwindigkeit heranzuziehen. Dann lässt sich mit der Notation aus Abschnitt 2.2 $F_t(s(0), v(0)) = (s(t), v(t))$ oder, allgemeiner, $F_t(s(t_0), v(t_0)) = (s(t_0 + t), v(t_0 + t))$ schreiben. Bei eindimensionaler Modellierung wäre die ein dynamisches System kennzeichnende Autonomie verlorengegangen, weil ja $s(t)$ nicht allein von $s(0)$, sondern auch von $v(0)$ abhängt. Sind diese beiden Anfangswerte gegeben, ist die weitere Entwicklung des Systems aber festgelegt. Diesen Aspekt werden wir noch im Kontext von Differentialgleichungen vertiefen. Zunächst sind aber Bemerkungen über die beiden vereinfachenden Modellannahmen, nämlich die Vernachlässigung des Luftwiderstandes und die Konstanz von g , am Platz.

⁴ George David Birkhoff (1884-1944)

4.2. Modellverfeinerungen

Bekanntlich ist der Luftwiderstand in Wahrheit keineswegs vernachlässigbar, sondern steigt mit der Geschwindigkeit, genauer: proportional zu v^2 . Dem können wir in unserem mathematischen Modell Rechnung tragen. Als Zustandsraum nehmen wir weiterhin den zweidimensionalen Raum $X = \mathbb{R}^2$ für alle möglichen Werte von $(s(t), v(t))$. Statt der konstanten Beschleunigung g jedoch arbeiten wir nun mit einer von v abhängigen, nämlich mit $a(v) = g - \alpha v^2$, wobei der Koeffizient α empirisch zu bestimmen ist. Das führt uns zur Differentialgleichung $v' = g - \alpha v^2$. Wie die (ganz analog zu beweisende) höherdimensionale Variante des Satzes in 4.4 zeigt, hat diese Differentialgleichung für jedes vorgegebene $(s(0), v(0))$ (wenigstens lokal) eine eindeutige Lösung v , mit der sich dann wie vorher bei vorgegebenem $s(0)$ aus $s'(t) = v(t)$ auch $s(t)$ eindeutig als Stammfunktion berechnen lässt. In der Notation von früher gilt wieder $F_t(s(t_0), v(t_0)) = (s(t_0 + t), v(t_0 + t))$. Anstatt die Rechnungen für die Lösung der resultierenden Differentialgleichung durchzuführen, stellen wir eine qualitative Überlegung an. Klarer Weise wird die Fallgeschwindigkeit $v(t)$ mit t wachsen, allerdings immer weniger schnell. Im Extremfall bliebe v konstant, was mit dem Wert $v = \sqrt{\frac{g}{\alpha}}$ sogar eine Lösung der Differentialgleichung wäre. Asymptotisch erreicht der fallende Körper also eine konstante Geschwindigkeit, wo sich Erdbeschleunigung und Abbremsung durch den Luftwiderstand genau die Waage halten.⁵

Welchen Wert hat diese konstante Geschwindigkeit? Das hängt vermittelt α in höchstem Maß vom fallenden Körper ab. Grob gesprochen gilt: Je schwerer der Körper und je kleiner seine Querschnittsfläche, desto kleiner ist α und desto größer die asymptotische Fallgeschwindigkeit.

Doch man kann unserem Modell gegenüber noch kritischer sein: Der Koeffizient α hängt nicht nur vom Körper ab, sondern auch von der Luftdichte. Bei einer Fallhöhe, wie sie z.B. beim Fallschirmspringen auftritt, ist diese keineswegs konstant⁶. Eine Verfeinerung des mathematischen Modells in diese Richtung führt nach Einsetzen von $s'(t)$ für $v(t)$ und $s''(t)$ für $a(t)$ zur noch komplizierteren Differentialgleichung $s''(t) = g - \alpha(s(t))s'(t)^2$. Weil bei sehr großen Fallhöhen auch g nicht konstant ist, wäre nach Newtons Gravitationsgesetz auch noch die Konstante g durch die Funktion $g(t) = \frac{\gamma}{(r+s(0)-s(t))^2}$ zu ersetzen, wobei r der Erdradius und γ eine bestimmte, zur Erdmasse proportionale Konstante ist.

Die Lösung so einer Differentialgleichung sprengt den Schulunterricht bei Weitem. Immerhin aber zeigen unsere Überlegungen recht konkret, wie Verbesserungen von Modellen typischer Weise zu Verkomplizierungen der involvierten Mathematik führen. Außerdem kann ein qualitatives Verständnis vermittelt werden, warum die Physik voll von Differentialgleichungen ist. Auf das historisch vielleicht markanteste Beispiel dieser Art soll noch kurz eingegangen werden.

4.3. Von Galilei und Kepler zu Newton

Bekanntlich beschrieb Johannes Kepler (1571-1630), wie die Planeten in Ellipsenbahnen um die Sonne kreisen, wobei seine berühmten drei Gesetze außerdem Aussagen über die Geschwindigkeit machen. Isaac Newtons (1643-1727) überragende Bedeutung in der Naturwissenschaft schließlich beruht vor allem darauf, dass er eine umfassende mathematisch-physikalische Theorie entwickelte, in deren Rahmen sowohl Galileis Fallgesetze als auch Keplers Beschreibung der Planetenbewegungen als Folgerungen desselben Grundprinzips erklärbar werden. (Historisch dürfte Newton übrigens den umgekehrten Weg gegangen sein und aus den bekannten Planetenbewegungen auf das Gravitationsgesetz geschlossen haben.)

Man kann nämlich die Bewegung eines Planeten auffassen als freien Fall im Schwerfeld der Sonne. Bei der mathematischen Modellierung muss man zusätzlich zu den bisherigen Überlegungen statt der eindimensionalen Fallhöhe als Zustandsraum wenigstens eine Ebene annehmen, innerhalb derer der Planet sich bewegt, also wenigstens zwei Ortskoordinaten x und y plus deren Ableitungen x' und y' für den

⁵ Ein mit dieser Geschwindigkeit stürzender Fallschirmspringer hat subjektiv nicht mehr das Gefühl des freien Falles, sondern empfindet die bremsende Wirkung des Luftwiderstandes genauso stark wie ruhende Personen die Gravitation.

⁶ Eine relativ einfache physikalische Überlegung zeigt, dass die Luftdichte $d(h)$ in der Höhe h eine Differentialgleichung der Bauart $d' = -\beta d$ mit einem konstanten $\beta > 0$ erfüllt, daher mit der Höhe exponentiell abnimmt.

Geschwindigkeitsvektor. Für das zugehörige dynamische System gilt dann

$$F_t(x(t_0), y(t_0), x'(t_0), y'(t_0)) = (x(t_0 + t), y(t_0 + t), x'(t_0 + t), y'(t_0 + t)).$$

Durch entsprechend verfeinerte mathematische Modellierung erhält man die zugehörigen Differentialgleichungen, deren Lösungen die von Kepler beschriebenen Ellipsenbahnen sind.

Auch ohne die zugehörigen Rechnungen (es sei auf die schöne Darstellung in Arnold (1989) verwiesen) fassen die hier nur qualitativ skizzierten Überlegungen doch eine der wichtigsten Episoden aus der frühen Geschichte modernen mathematisch-naturwissenschaftlichen Denkens zusammen. In einem Punkt lohnt es aber, die Mathematik selbst nochmals genauer ins Auge zu fassen. Das soll nun geschehen.

4.4. Banachscher Fixpunktsatz und Differentialgleichungen

Wir haben gesehen: Die zu dynamischen Systemen gehörenden Differentialgleichungen haben typischer Weise die Bauart $x'(t) = f(x(t))$ mit einer vorgegebenen Funktion f , oft unter Vorgabe einer Anfangsbedingung der Form $x(0) = x_0$. Die Werte von x und f sind Vektoren in \mathbb{R}^d . Der einfacheren Notation halber und ohne schwerwiegende Beschränkung der Allgemeinheit wollen wir jetzt $d = 1$ annehmen.

Schon in 4.2 wurde erwähnt, dass so eine Differentialgleichung eine eindeutige Lösung besitzt. Allgemein ist das aber, wenig überraschend, nicht ohne geeignete Voraussetzungen der Fall. Von welcher Art diese sein können, wollen wir uns nun überlegen.

Es wird sich herausstellen, dass wir erlauben können, dass die Funktion f nicht nur von $x(t)$, sondern auch direkt von t abhängt, dass die Differentialgleichung also nicht unbedingt autonom sein muss. Also gehen wir von einer Differentialgleichung mit Nebenbedingung aus, nämlich:

$$x'(t) = f(x(t), t), \quad x(0) = x_0$$

Wenn wir die Gleichung nach t integrieren (wir wählen für die neu auftretende Integrationsvariable auf der rechten Seite den Buchstaben τ) und für eine beliebige stetige Funktion x die Schreibweise $\Phi_x(t) := x_0 + \int_0^t f(x(\tau), \tau) d\tau$, verwenden, so liefert der Hauptsatz der Differential- und Integralrechnung (siehe auch Winkler (2014 Langversion)) für eine Lösung x dieser Differentialgleichung die Beziehung $x = \Phi_x$. Gesucht ist also ein Fixpunkt jener Abbildung (jenes **Integral-Operators**) $\Phi : x \mapsto \Phi_x$, welche einer stetigen Funktion x die ebenfalls stetige Funktion Φ_x zuordnet.

Im Zusammenhang mit Fixpunkten erinnern wir uns des Abschnitts 3.3. Von dort wissen wir, dass ein eindeutiger Fixpunkt sicher dann existiert, wenn $T = \Phi$ eine Kontraktion auf einem vollständigen metrischen Raum X ist. Welches X dürfen wir nehmen? Offenbar dürfte X ein geeigneter Raum stetiger Funktionen sein. Wenn wir uns auf die Menge $C(I)$ jener stetigen x , die auf einem festen abgeschlossenen Intervall I definiert sind, beschränken können, ist die Supremumsmetrik $d(x_1, x_2) := \sup_{t \in I} |x_1(t) - x_2(t)|$ nicht nur wohldefiniert, sondern auch vollständig (siehe Winkler (2014 Langversion)). Folglich geht es darum, I so zu wählen, dass $\Phi : C(I) \rightarrow C(I)$ eine Kontraktion ist. Es muss also ein $\lambda < 1$ geben derart, dass stets $d(\Phi_{x_1}, \Phi_{x_2}) \leq \lambda d(x_1, x_2)$ gilt. Macht man diese Ungleichungen explizit, so sieht man, dass die sogenannte Lipschitzstetigkeit von f in der ersten Komponente das Gewünschte leistet. D.h. es muss eine reelle Konstante λ_0 derart geben, dass stets $|f(x_1, \tau) - f(x_2, \tau)| \leq \lambda_0 |x_1 - x_2|$. Wählt man das Intervall $I = [0, \varepsilon]$ mit genügend kleinem $\varepsilon > 0$, nämlich $\varepsilon < \lambda_0^{-1}$, so zeigt die Abschätzung

$$d(\Phi_{x_1}, \Phi_{x_2}) \leq \int_0^\varepsilon |f(x_1(\tau), \tau) - f(x_2(\tau), \tau)| d\tau \leq \varepsilon \lambda_0 d(x_1, x_2),$$

dass Φ eine Kontraktion mit $\lambda := \varepsilon \lambda_0 < 1$ ist. Also gibt es auf $I = [0, \varepsilon]$ eine eindeutige Lösung x unserer Differentialgleichung.

Damit haben wir den wichtigen **Existenz- und Eindeutigkeitsatz von Picard-Lindelöf**⁷ bewiesen:

⁷ Charles Émile Picard (1856-1941), Ernst Leonard Lindelöf (1870-1946)

Die Funktion $f: \mathbb{R} \times [0, a] \rightarrow \mathbb{R}$ sei stetig und in der ersten Komponente sogar Lipschitz-stetig. Dann gibt es ein $\varepsilon > 0$ derart, dass die Differentialgleichung $x'(t) = f(x(t), t)$ mit der Nebenbedingung $x(0) = x_0$ eine auf $[0, \varepsilon]$ eindeutige Lösung besitzt.

Dieser Satz ist einer der grundlegendsten für das klassische, deterministisch geprägte Weltbild der mathematischen Physik. Denn er besagt (etwas plakativ gesprochen), dass der gegenwärtige Zustand der Welt gemeinsam mit den Naturgesetzen im Sinne der Newtonschen Physik auch zukünftige Zustände der Welt festlegt. Natürlich ist diese Aussage angesichts der Erkenntnisse der Physik des 20. Jahrhunderts nicht mehr uneingeschränkt haltbar. Immerhin fanden sich aber derart weitreichende Bestätigungen für eine deterministische Weltansicht, dass ein großer mathematischer Denker wie Pierre-Simon Laplace (1749-1827) davon überzeugt war. Man spricht in diesem Zusammenhang heute noch vom *Laplaceschen Dämon* als von einem hypothetischen allwissenden Wesen, das die Zukunft der Welt für alle Zeiten zu berechnen imstande wäre.

4.5. Zur Erinnerung noch einmal die Exponentialfunktion

Wir wollen die Macht des Satzes von Picard-Lindelöf aus 4.4 noch an der archetypischen Differentialgleichung schlechthin erproben, an jener für die Exponentialfunktion $x(t) = \exp(t) = e^t$:

$$x'(t) = x(t), \quad x(0) = 1$$

Die im Satz auftretende Funktion f hat hier die sehr einfache Gestalt $f(x, t) = x$ und ist in der ersten Komponente Lipschitz-stetig mit $\lambda_0 = 1$. Also gibt es eine eindeutige Lösung. Aus dem im Hintergrund wirksamen Fixpunktsatz von Banach wissen wir, dass diese Lösung durch Iteration der Kontraktion $\Phi: x \mapsto \Phi_x$ mit $\Phi_x(t) = 1 + \int_0^t x(\tau) d\tau$ (zunächst wenigstens für kleines t) zu gewinnen ist. Man kann mit einer beliebigen stetigen Funktion x_0 beginnen, rekursiv $x_{n+1} := \Phi_{x_n}$ setzen und erhält die gesuchte Lösung $x = \lim_{n \rightarrow \infty} x_n$. Der Einfachheit halber beginnen wir mit der konstanten Funktion $x_0 = 1$. Mit Induktion zeigt man $x_n(t) = \sum_{i=0}^n \frac{t^i}{i!}$, also

$$x(t) = \sum_{n=0}^{\infty} \frac{t^n}{n!} = \exp(t)$$

(man vergleiche mit meinem früheren Artikel Winkler (2012)). Man beachte, dass diese Darstellung nicht nur für t aus einem kleinen Intervall bei 0 gilt, wie im Satz von Picard-Lindelöf garantiert, sondern auf ganz \mathbb{R} .

Literatur

- Vladimir I. Arnold. *Huygens and Barrows, Newton and Hooke. Pioneers in mathematical analysis and catastrophe theory from evolvents to quasicrystals*. Translated from the Russian by Eric J.F. Primrose. Birkhäuser. Basel – Boston – Berlin (1990).
- Stanislas Dehaene. *Der Zahlensinn oder Warum wir rechnen können*. Aus dem Amerikanischen von Anita Ehlers. Birkhäuser. Basel – Boston – Berlin (1999). Originaltitel: *The Number Sense – How the Mind Creates Mathematics*.
- Stefan Götz und Franz Hofbauer. *Die Exponentialfunktion als dynamisches System*. IMN⁸ 223 (August 2013), 21-35.
- Hans Humenberger. *Iterationen, Grenzwerte und Spinnwebdiagramme – oder warum ist 2 doch nicht 4?* IMN⁸ 213 (April 2010), 19-34.
- Anatole Katok und Boris Hasselblatt. *Introduction to the Modern Theory of Dynamical Systems*. Cambridge University Press (1995).
- Heinz-Otto Peitgen, Hartmut Jürgens und Dietmar Saupe. *Chaos: Bausteine der Ordnung*. Deutsche Erstausgabe: Springer (1994). Berlin, Heidelberg und J.G.Cotta'sche Buchhandlung Nachfolger GmbH, Stuttgart. Als Taschenbuch: Rowohlt (1998). Reinbek bei Hamburg. Übersetzung des Englischen Originals (*Fractals for the Classroom*, Springer (1992), New York) von Anna M. Rodenhausen.

⁸ Internationale Mathematische Nachrichten (ÖMG), online unter www.oemg.ac.at/IMN/

- Peter Raith. *Chaos und Fraktale*. DH⁹ 41 (2009), 85-99.
- Peter Walters. *An introduction to Ergodic Theory*. Graduate Texts in Mathematics 79. Springer (1982). New York, Inc. Taschenbuch 2000.
- Hermann Weyl. *Über die Gleichverteilung der Zahlen mod Eins*. *Mathematische Annalen* 77 (1916), 313-352.
- Reinhard Winkler. *Wir zählen bis drei – und sogar darüber hinaus*. DH⁹ 40 (2008), 129-141.
- Reinhard Winkler. *Die reellen Zahlen sind anders*. DH⁹ 41 (2009), 140-153.
- Reinhard Winkler. *Im Anfang war die Exponentialfunktion*. DH⁹ 44 (2012), 98-109.
- Reinhard Winkler. *Stochastik – ein Fest der Unabhängigkeit*. DH⁹ 45 (2013), 122-136.
- Reinhard Winkler. *Dynamische Systeme als Chance für den Schulunterricht*. Langversion. Online verfügbar unter www.oemg.ac.at/DK/.

⁹ Schriftenreihe zur Didaktik der Mathematik der Österreichischen Mathematischen Gesellschaft (ÖMG) – ehemals Didaktikhefte der ÖMG, online unter www.oemg.ac.at/DK/