Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

EducETH

ETH Zürich
Rämistr. 101
8091 Zürich

Eine Publikation von www.educeth.ch - dem Bildungsportal der ETH Zürich

Von gut und schlecht gestellten Problemen

oder
Von Ursachen zu Wirkungen und zurück

oder
Computertomographie und Co

von

Urs Kirchgraber , Departement Mathematik ETH Zürich und ZHSF
Daniel Stoffer, Departement Mathematik ETH Zürich

Datum
Publikation der Unterrichtseinheit auf EducETH am 2. Juni 2006

Inhalt
Stellen Sie sich einen feinen English Cake vor, in den viele Rosinen und kandierte Früchte eingebacken sind. Nun nehmen Sie ein dünnes Röhrchen und stossen es irgendwo in irgend eine Richtung durch den Cake. Von dieser Probe bestimmen Sie die mittlere Dichte (indem Sie ihr Gewicht und ihr Volumen messen und den Quotienten ausrechnen). Geben Sie danach die Probe in den Kuchen zurück, sodass er wieder hergestellt ist. Führen Sie sodann analog viele weitere solche ``Probebohrungen'' durch, bis der Kuchen kreuz und quer vermessen ist!
Die Frage ist: Kann man aufgrund der Kenntnis der mittleren Dichten von vielen Probebohrungen ermitteln (``rekonstruieren''), wie es im Innern des Kuchens aussieht, das heisst, wo (ungefähr) Rosinen, wo kandidierte Früchte sind, wo Teig ist?
Mit dieser Geschichte versucht S. K. Stein in einem Buch zu erläutern wie eine der wichtigsten modernen medizinischen Diagnosetechniken im Prinzip funktioniert: Die Computertomografie (CT).
Statt ein Kuchen wird ein Stück Gewebe untersucht. Anstelle der Probebohrungen mit dem Röhrchen treten Röntgenstrahlen, die an vielen Orten und in viele Richtungen durch das Gewebe hindurch geschickt werden und man misst wie stark sie dabei jeweils abgeschwächt werden.
Die Computertomografie ist ein Beispiel, wie man aus einer Wirkung (der Abschwächung von Röntgenstrahlen beim Durchgang durch Gewebe) auf die Ursache (die Dichteverteilung im Gewebe) zurück schliessen will. Es handelt sich dabei um ein so genannt (moderat) schlecht gestelltes Problem.
Um aus einer Wirkung die Ursache bestimmen zu können, muss der Ursache-Wirkungs-Zusammenhang umgekehrt werden. Schlecht gestellt bedeutet, dass der Wirkungs-Ursache-Zusammenhang von der Art ist, dass schon geringfügige Veränderungen der Wirkung bedeutende Änderungen der Ursache hervorrufen.
Das wäre nicht weiter schlimm, wenn man die Wirkung genau bestimmen könnte. In der Praxis ist man aber auf Messungen angewiesen. Und Messungen sind grundsätzlich mit (Mess-)Fehlern behaftet.
Der russische Mathematiker A. N. Tikhonov (1906-1993) hat eine Idee entwickelt, wie schlecht gestellte Probleme durch etwas weniger schlecht gestellte Probleme ersetzt werden können, sodass sich (kleine) Messfehler weniger fatal auswirken. Seine Ideen haben grossen praktischen Nutzen: Sie ermöglichen es, dass gewisse Probleme ohne extrem genaue (lies: extrem kostspielige) Messungen gelöst werden können, andere

sind überhaupt nur dank Tikhonovs Methode lösbar, weil die ansonsten notwendige Messtechnik (noch?) gar nicht zu Verfügung steht.

Ziele der Unterrichtseinheit
Oft steht man vor der Aufgabe, eine Grösse bestimmen zu müssen, die einer direkten Messung nicht zugänglich ist. Zum Beispiel der Radius der Erde oder ihre Masse. In einer solchen Situation muss man ein indirektes Vorgehen wählen: Man benutzt einen Effekt, durch den sich die interessierende Grösse manifestiert.
Die Leserinnen und Leser lernen dieses Vorgehen an einigen Beispielen kennen. Sie befassen sich mit der dabei auftretenden charakteristischen mathematischen Schwierigkeit (der sogenannten „Schlechtgestelltheit“) und sie lernen eine Idee von A. N. Thikonov kennen, die die Schwierigkeit zwar nicht gerade behebt, aber ihre Auswirkungen mildert. Sie begegnen dabei einer wichtigen und aktuellen Anwendung der Mathematik.

Fachdidaktisches und fachliches Review
Moritz Adelmeyer

Rechtliches
Die vorliegende Unterrichtseinheit darf ohne Einschränkung heruntergeladen und für Unterrichtszwecke kostenlos verwendet werden. Dabei sind auch Änderungen und Anpassungen erlaubt. Der Hinweis auf die Herkunft der Materialien (ETH Zürich, EducETH) sowie die Angabe der Autorinnen und Autoren darf aber nicht entfernt werden.

Weitere Informationen
Weitere Informationen zu dieser Unterrichtseinheit und zu EducETH finden Sie im Internet unter http://www.educ.ethz.ch oder unter http://www.educeth.ch.

Von gut und schlecht gestellten Problemen oder Von Ursachen zu Wirkungen und zurück oder Computertomografie und Co*

U. Kirchgraber† und D. Stoffer
Departement Mathematik, ETH Zürich
kirchgra@math.ethz.ch

Zusammenfassung

Stellen Sie sich eine feinen English Cake vor, in den viele Rosinen und kandierte Früchte eingebacken sind. Nun nehmen Sie ein dünnes Röhrchen und stossen es irgendwo in irgend eine Richtung durch den Cake. Von dieser Probe bestimmen Sie die mittlere Dichte (indem Sie ihr Gewicht und ihr Volumen messen und den Quotienten ausrechnen). Geben Sie danach die Probe in den Kuchen zurück, sodass er wieder "heil" ist. Führen Sie sodann analog viele weitere solche "Probebohrungen" durch, bis der Kuchen kreuz und quer vermessen ist!

Die Frage lautet: Kann man aufgrund der bestimmten mittleren Dichten von vielen "Probebohrungen" ermitteln ("rekonstruieren"), wie es im Innern des Kuchens aussieht, das heisst, wo (ungefähr) Rosinen, wo kandidierte Früchte sind, wo Teig ist?

Dieses vielleicht merkwürdig anmutende Spiel (wer möchte schon mit einem feinen Kuchen so umgehen!) zeigt ganz gut wie eine der wichtigsten modernen medizinischen Diagnosetechniken im Prinzip funktioniert[1]: die sogenannte *Computertomografie* (CT).

Statt ein Kuchen wird ein Stück Gewebe untersucht. Anstelle der "Probebohrungen" mit dem Röhrchen treten Röntgenstrahlen, die an vielen Orten und in viele Richtungen durch das Gewebe hindurch geschickt werden und man misst wie stark sie dabei jeweils abgeschwächt werden.

Die Computertomografie ist ein Beispiel, wie man aus einer Wirkung (der Abschwächung von Röngenstrahlen beim Durchgang durch Gewebe) auf die Ursache (die Dichteverteilung im Gewebe) schliessen will. Sie ist überdies ein Beispiel für ein (moderat) *schlecht*

*Überarbeitete und um das Kapitel 5 erweiterte Version eines Vortrags, den der erst genannte Autor unter dem Titel *Von gut und schlecht gestellten Problemen und ihrer Bedeutung (zum Beispiel für die Gesundheit)* am 18. April 2002 an der Kantonsschule Zürcher Unterland in Bülach gehalten hat.

†In Zusammenarbeit mit A. Kirsch, Universität Karlsruhe. Wir danken P. Spindler für die Herstellung der Zeichnungen.

[1]Die oben gegebene metaphorische Erläuterung der Grundidee, die hinter der Computertomografie verborgen ist, stammt von S. K. Stein "The Strength of Numbers", 1996.

gestelltes Problem. Sogenannt schlecht gestellten Problemen ist eine Schwierigkeit eigen, die mit einem Phänomen zusammenhängt, das jeder aus der Schulgeometrie kennt: Das Problem der schleifenden Schnitte.

Wenn man mit einem nicht allzu gut gespitzten Bleistift zwei zueinander (fast) senkrecht verlaufende "Linien" zeichnet ist ihr "Schnittpunkt" zwar nicht exakt, aber doch einigermassen genau feststellbar. Ganz anders ist die Situation, wenn die beiden geraden "Linien" flach zueinander verlaufen. Je stumpfer der Bleistift und je kleiner der Öffnungswinkel der beiden geraden "Linien" ist, desto grösser ist der sich überschneidende Bereich der beiden "Linien", was es zunehmend schwierig macht, von einem Schnittpunkt zu reden. In der Schule verlangen Lehrerinnen und Lehrer, dass mit scharf gespitzten Bleistiften gezeichnet wird und sie sorgen durch geschickte Dispositionen dafür, dass allzu "flache Schnitte" nicht vorkommen.

Dem nicht perfekt gespitzten Bleistift im Geometrieunterricht entspricht in der Welt, in der Messungen durchgeführt werden müssen, die niemals vermeidbare *Messungenauigkeit*. Der russische Mathematiker A. N. Tikhonov (1906-1993) hat eine Idee entwickelt, wie schlecht gestellte Probleme durch etwas weniger schlecht gestellte Probleme ersetzt werden können, sodass sich (kleine) Messfehler weniger fatal auswirken. Dass man also, um im Bild der Schulgeometrie zu sprechen, keinen (nicht existierenden) idealen Bleistift benötigt um schlecht gestellte Probleme zu lösen, sondern auch mit einem nur einigermassen gut gespitzten Bleistift nützliche Resultate erhält!

Zum *Inhalt* dieses Artikels. Nach einleitenden Bemerkungen werden in Abschnitt 2 drei Paare zueinander, wie man sagt, "inverser Probleme" vorgestellt. In Abschnitt 3 lernt die Leserin, der Leser die Begriffe "gut" und "schlecht" gestellte Probleme etwas genauer kennen. In Abschnitt 4 wird zunächst ein wahrhaft schlecht gestelltes Problem eingeführt, rechnerisch untersucht und geometrisch analysiert. Dann wird Tikhonovs Idee vorgestellt und anhand von einigen Rechnungen wird illustiert, dass sie funktioniert. Der Kern der Abschnitte 1-4 ist Abschnitt 4. Ihn zu verstehen setzt zunächst nur Vertrautheit mit linearen Gleichungssystemen mit 3 Unbekannten und etwas Kenntnisse in elementarer Algebra voraus. Die Analyse des Phänomens der "Schlecht-Gestelltheit" in Abschnitt 4. 3 verlangt etwas Vertrautheit mit dem Begriff der Ebenengleichung im Raum.

Abschnitt 5 richtet sich an Leserinnen und Leser, die noch etwas tiefer in die Thematik eindringen möchten. Voraussgesetzt werden jetzt Kenntnisse aus der ebenen Vektorrechnung bis hin zum Skalarprodukt. Darauf aufbauend werden in ganz elementarer Weise die Elemente der Linearen Algebra in zwei Dimensionen bis hin zum entsprechenden Eigenwertproblem entwickelt (Abschnitte 5.1-5.2). Abschnitt 5.3 bringt die Anwendung auf das in Abschnitt 4.1 eingeführte Beispiel eines schlecht gestellten Problems im 2-dimensionalen Fall. Abschnitt 5.4 schliesslich gibt Verallgemeinerungen auf höherdimensionale Versionen.

1 Die Rolle der Mathematik in den Anwendungen

In den Naturwissenschaften versucht man, Phänomene zu verstehen. Wie entsteht das Wetter? Wieso fliegt ein Diskus weiter, wenn man ihn gegen, statt mit dem Wind wirft? Welchen Einfluss übt der Mond auf die Erde aus?

Der Kern naturwissenschaftlicher Erkenntnisse sind

Ursache-Wirkungs-Beziehungen.

Beispiel: Der Wechsel von Tag und Nacht ist eine Folge der Rotation der Erde um ihre eigene Achse.

Wir glauben heute nicht mehr, dass wir *abschliessend* verstehen können, wie "die Welt funktioniert". Aber es zeigt sich, dass wir uns doch *Vorstellungen und Bilder* machen können, die mindestens zum Teil tragfähig sind. Unser Leben wäre sehr anders, wenn das nicht so wäre!

Solche Bilder nennen wir

Modelle.

Von besonderem Interesse sind *quantitative* Ursache-Wirkungs-Beziehungen. Also funktionale Zusammenhänge zwischen Grössen.
Ein Beispiel: Zwischen Energie und Masse, so entdeckte Einstein, gilt:

$$E = mc^2.$$

Dabei ist c die Lichtgeschwindigkeit.

Auf der Grundlage quantitativer Ursache-Wirkungs-Beziehungen können wir

mathematische Modelle

entwickeln. Das sind Modelle, die in *der Sprache der Mathematik* formuliert sind. Sie sind deshalb so interessant, weil sie mit mathematischen Instrumenten bearbeitet werden können. Das gesamte mathematische Instrumentarium steht zur Verfügung, um sie zu untersuchen. Das ist aus zwei Gründen wichtig.

1. Aus einem Modell Schlüsse ziehen heisst

   Möglichkeiten entdecken, Prognosen machen.

2. Der springende Punkt dabei ist:

   Die Mathematik ist das sicherste Instrument, um aus gewissen Tatsachen neue zu gewinnen.

So können wir sagen:

Mathematische Modelle sind *gedachte Laboratorien,* in denen verlässlich *probegehandelt* werden kann.

2 Zwillinge: Direkte und Inverse Probleme

Zu vielen Fragen gibt es eine "Umkehrfrage". Manche Probleme treten paarweise auf: in einer *direkten* und einer dazu umgekehrten, *inversen* Form.

Ein erstes Beispiel. Wer multiplizieren kann, kann quadrieren. Das *direkte Problem* besteht darin, zu einer gegebenen Zahl x ihr Quadrat y zu bilden, also $y = x^2$. Das dazu *inverse Problem* lautet: Gegeben eine Zahl y, man finde eine Zahl x, deren Quadrat gleich y ist. Oder anders gesagt: Man soll aus y die Wurzel ziehen.

Bei diesem Beispiel ist das inverse Problem schwieriger zu lösen als das direkte. Das sieht man daran, dass das direkte Problem in den natürlichen Zahlen, also den Zahlen 1, 2, 3, ..., *immer* ausführbar ist. Hingegen ist die (positive) Wurzel aus einer natürlichen Zahl nur ausnahmsweise eine natürliche Zahl. Damit dieses und viele weitere inverse Probleme immer lösbar sind, erfand man die reellen Zahlen.

In der Mathematik begegnen einem solche Zwillingsprobleme häufig. Immer wenn eine *Operation*

$$K$$

gegeben ist, die aus einer Grösse x eine "neue" Grösse y "macht", so dass also

$$y = Kx \tag{1}$$

gilt, kann man die zugehörige *Umkehrfrage* stellen: Gegeben ein y, man finde ein x so dass die Gleichung (1) erfüllt ist. In (1) repräsentiert x die *Ursache*, y die *Wirkung*, und K vermittelt den *Ursache-Wirkungs-Zusammenhang*. Das *inverse Problem* besteht darin, aus der Wirkung auf die Ursache zurückzuschliessen.

Es folgen mehrere Beispiele von Paaren zueinander inverser Probleme.

2.1 Wie tief ist der Ziehbrunnen, wie hoch der Turm?

Inverse Problem entstehen in der Praxis aus der Aufgabe, eine Grösse ermitteln zu wollen, die sich *nicht unmittelbar*, *nicht direkt* bestimmen lässt. Die Idee ist, sich in einer solchen Situation auf einen *Effekt* zu stützen, bei dem sich die unbekannte Grösse manifestiert.

Die Frage nach der *Tiefe eines Sodbrunnens* oder *Schachts* oder *der Höhe eines Turms* ist eine solche Aufgabe. Ausgangspunkt ist folgende Feststellung. Wenn man einen Gegenstand, zum Beispiel einen Stein, in einen Schacht fallen lassen, hört man nach einer gewissen Zeit, dass er aufschlägt. Dieses Phänomen beruht auf folgenden Ursache-Wirkungs-Beziehungen.

a) Wenn ein Gegenstand losgelassen wird, fällt er zu Boden.

b) Durch den Aufprall des Gegenstands am Boden entsteht ein Geräusch, das sich in der Luft ausbreitet.

Man erwartet, dass zwischen der Tiefe H des Schachts und der (Warte-)Zeit T, die zwischen dem Loslassen des Steins und dem Hören des Aufschlags vergeht, ein quantitativer Zusammenhang besteht.

Stellen wir uns die Aufgabe, ein *mathematisches Modell* für diesen Zusammenhang aufzustellen. Dazu benötigen wir quantitative Versionen der Ursache-Wirkungs-Beziehungen a), b).

a) Galilei hat vor etwa 400 Jahren den sogenannten *Freien Fall* eines Gegenstands auf der Erdoberfläche untersucht und gefunden, dass zwischen Fallzeit t_1 und durchfallener Höhe s_1 bei Vernachlässigung der Bremswirkung durch die Luft die Beziehung

$$s_1 = 5t_1^2$$

gilt. Dabei wird die Zeit in Sekunden und der Weg in Metern gemessen.

b) Für die Ausbreitung eines Geräusches in der Luft gilt

$$s_2 = 340t_2.$$

Dabei ist s_2 der zurückgelegte Weg, t_2 die dafür benötigte Zeit.

Legt man diese beiden quantitativen Ursache-Wirkungs-Beziehungen zu Grunde, folgt für den Zusammenhang zwischen Schachttiefe H und Wartezeit T die Formel:

$$\boxed{T = \sqrt{\frac{1}{5}H} + \frac{1}{340}H} \tag{2}$$

$\sqrt{\frac{1}{5}H}$ ist die Zeit, die der Gegenstand braucht, bis er am Boden aufschlägt. $\frac{1}{340}H$ ist die Zeit, die verstreicht, bis das durch den Aufschlag erzeugte Geräusch das Ohr des Experimentators oben erreicht.

Die Formel (2) leistet folgendes. Wenn die Tiefe H des Schachts bekannt ist, liefert sie eine Prognose für die Wartezeit T. Das ist die Lösung des *direkten Problems.* Wenn die Tiefe H des Schachts nicht bekannt ist, kann aus der Kenntnis der Wartezeit T eine Prognose für H gewonnen werden. Dazu ist Gleichung (2) nach H aufzulösen. Das Resultat lautet:

$$\boxed{H = \frac{340T^2}{34 + T + \sqrt{1156 + 68T}}} \tag{3}$$

Das ist die Lösung des *inversen Problems.*

2.2 Ein antikes Designproblem: Wasseruhren

Die "Idee der Zeit" ist etwas faszinierendes. Auch ihre "Materialisierung" in Form von Uhren zur Zeitmessung ist ein interessantes Problem und der Motor für viele schöne, geniale und manchmal verrückte Ideen gewesen. Sonnen-, Sand- und Wasseruhren kennt man seit der Antike. In diesem Abschnitt geht es um Wasseruhren.

Eine Wasseruhr ist ein Gefäss mit einer kleinen Öffnung im Boden oder an der Seite. Zu Beginn ist das Gefäss ganz mit Wasser gefüllt. Dann entleert es sich allmählich durch die Öffnung. Im Hinblick auf die Zeitmessung interessiert der Zusammenhang zwischen verflossener Zeit und Höhe des Wasserstandes: *Wie sinkt der Wasserstand im Laufe der Zeit?* Die

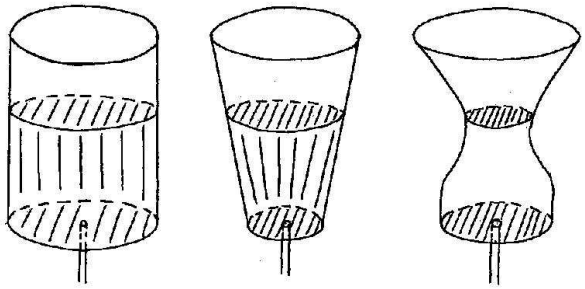

Abbildung 1: Verschiedene Gefässformen.

Antwort hängt von der Form der Wasseruhr ab. Bei einem zylindrischen Gefäss verläuft das Absinken des Wasserstandes anders, als wenn die Wasseruhr die Form eines Kegelstumpfes oder einer Vase hat, siehe Abbildung 1.

Eine antike Uhrendesignerin hat sich vielleicht folgende Frage gestellt: Welche Form muss die Wasseruhr haben, damit der Wasserstand in vorgegebener Weise sinkt? Zum Beispiel so, dass er in gleichen Zeiten um gleich viel abnimmt?

Wieder haben wir ein Paar zueinander inverser Probleme. Das *direkte Problem* lautet: Gegeben die Gestalt der Wasseruhr, wie sinkt der Wasserstand im Laufe der Zeit? Das *inverse Problem* heisst: Bei welcher Gestalt der Wasseruhr fällt der Wasserstand in einer ganz bestimmten, im voraus festgelegten Weise?

Beschränken wir uns auf Wasseruhren, die "rotationssymmetrisch" sind. Wenn wir ein solches Gefäss mit einer Ebene schneiden, die durch die Rotationsachse geht, erhalten wir die "Profilkurve" der Wasseruhr, siehe Abbildung 2. Umgekehrt kann man sich das Gefäss durch Rotation der Profilkurve um die Achse entstanden denken.

Wir müssen die Profilkurve der Wasseruhr mathematisch erfassen können. Dazu machen wir die Symmetrie- oder Rotationsachse zur y-Achse eines Koordinatensystems, siehe Abbildung 2, und denken uns den im 1. Quadranten liegenden Teil der Profilkurve durch eine Funktionsgleichung

$$x = f(y).$$

gegeben. Geläufiger ist, wenn y in Abhängigkeit von x gegeben ist. Die hier gewählte Version ist jedoch bequemer. Mit der Wahl $f(y) = \sqrt{y}$ erhält man eine Parabel, und das Gefäss hat die Form eines Paraboloids.

Nun kommt ein wichtiger Punkt. Auch der Auslaufvorgang muss mathematisch erfasst werden können. Das geschieht mit folgender Grösse:

Es sei $g(y)$ die Zeit, die es braucht, bis die Wasseruhr ganz entleert ist, falls sie am Anfang bis zur Höhe y mit Wasser gefüllt war.

Falls H die maximale Füllhöhe der Wasseruhr bezeichnet, ist

$$g(H) - g(y)$$

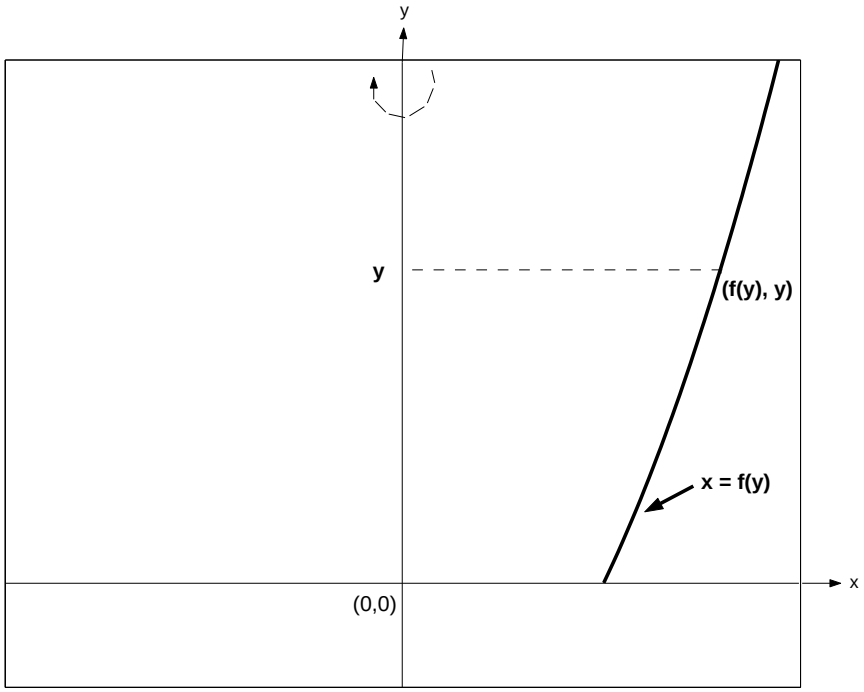

Abbildung 2: Die Profilkurve der Wasseruhr ist durch die Gleichung $x = f(y)$ gegeben. Das Gefäss entsteht durch Rotation der Profilkurve um die y-Achse.

die Zeit, die vergeht, bis der Wasserstand auf die Höhe y abgesunken ist.

Es sind jetzt also zwei Funktionen im Spiel: Die Funktion f, die die Profilkurve, und damit die Form der Wasseruhr, definiert. Und die Funktion g, die den Auslaufvorgang erfasst.

Nun können wir das direkte und das dazu inverse Problem so formulieren. Beim direkten Problem geht es darum, aus der Kenntnis der Funktion f, also des Profils, die Funktion g, also den Auslaufvorgang, zu bestimmen. Beim inversen Problem ist die Funktion g gegeben und die Funktion f ist gesucht.

Was ist der Unterschied zum Beispiel mit dem Schacht? Bei jenem Beispiel hatten wir es mit zwei *Zahlen* zu tun: Die Tiefe H des Schachts und die Wartezeit T, die zwischem dem Loslassen des Steines und dem Hören des Aufschlagens verstreicht. Der Zusammenhang wird durch die Formel (2), also durch eine Funktion, vermittelt.

Im Vergleich dazu ist das neue Beispiel komplizierter. Jetzt müssen wir uns mit zwei *Funktionen* befassen: f und g. Die interessante Frage lautet: Welches ist in diesem Fall die Ursache-Wirkungs-Beziehung? Wie lautet der Zusammenhang zwischen f und g?

Das Auslaufen einer Flüssigkeit aus einem Gefäss mit einer Öffnung ist ein physikalischer Vorgang. Er ist schon im 17. Jahrhundert von Evangelista Torricelli (1608-1647), einem Schüler von Galilei, untersucht worden. Dabei hat Torricelli folgendes entdeckt.

Bezeichnet y die Höhe des Wasserstandes in einem Gefäss, so ist die Geschwindigkeit v, mit der Wasser aus einer kleinen Öffnung im Boden des Gefässes austritt, gleich

$$v = \sqrt{20y}. \tag{4}$$

Dabei wird die Höhe in Metern, die Zeit in Sekunden gemessen.

Wenn man nun bilanziert, dass bei einem geringfügigen Absinken des Wasserstandes dieses Wasser ja mit der nach (4) bestimmten Geschwindigkeit aus der kleinen Öffnung im Boden austritt, erhält man (mit den Mitteln der sogenannten Differentialrechnung) den folgenden Zusammenhang zwischen den Funktionen f und g:

$$\boxed{g'(y) = k\frac{[f(y)]^2}{\sqrt{y}}, \quad \text{mit } k = \frac{\pi}{\sqrt{20}a}.} \tag{5}$$

Dabei bezeichnet a die Fläche der Öffnung im Boden des Gefässes und $g'(y)$ ist die sogenannte *erste Ableitung* der Funktion $g(y)$.

Gleichung (5) ist die *Ursache-Wirkungs-Beziehung* in diesem Beispiel. Da man aus der Funktion f gemäss (5) die Ableitung von g erhält, muss man, um g zu finden, die Funktion auf der rechten Seite von (5), wie man sagt, *integrieren*.

Ein *Beispiel*: Wählt man als Profilfunktion $f(y) = \sqrt{y}$, erhält man aus (5) für den Auslaufvorgang

$$g(y) = \frac{2}{3}k\sqrt{y^3}.$$

Nun zum inversen Problem. Es lautet: Gegeben g, gesucht f. Seine Lösung ist

$$\boxed{f(y) = \frac{1}{\sqrt{k}}\sqrt[4]{y}\,\sqrt{g'(y)}.} \tag{6}$$

Um sie zu erhalten, muss man die Funktion g *differenzieren*.

2.3 Computertomografie

Die bisher diskutierten Beispiele von direkten und dazu inversen Problemen haben keine grosse praktische Bedeutung. Das ist ganz anders beim nun folgenden Beispiel, der sogenannten

Computertomografie.

Es handelt sich hierbei um eine der wichtigsten neueren medizinischen Diagnosetechniken. Sie, und Varianten davon, wie etwa die Kern-Spin-Tomografie (englisch MRI für Magnetic Resonance Imaging) beruhen auf einem cleveren Zusammenspiel von

High-Tech, Informationstechnologie und Mathematik.

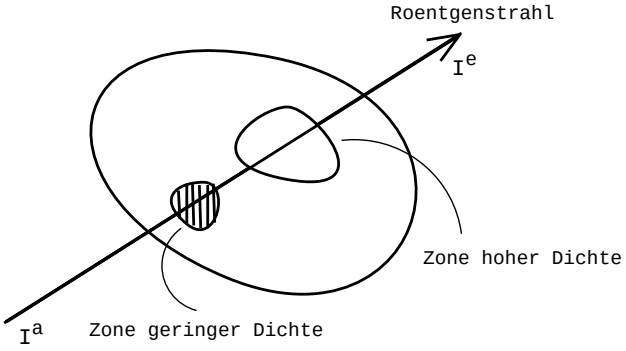

Abbildung 3: Dünne Gewebeschicht, Röntgenstrahl

Dass es sich bei einem Computertomografen um ein High-Tech-Gerät handelt, sieht man auf Anhieb, und dass der Computer eine wichtige Rolle spielt, erkennt man, wenn man sieht, wie die Bilder, für die man sich interessiert, auf dem Computerschirm entstehen. Dass alles ohne Mathematik nicht funktionieren würde, erschliesst sich dem blossen Auge hingegen nicht so unmittelbar – es bleibt verborgen. Aber es lässt sich im Prinzip leicht erklären.

Im Jahr 1895 hat Wilhelm Conrad Röntgen in Würzburg seine berühmten Strahlen entdeckt. Röntgenstrahlen sind unsichtbar. Sie können aber Gegenstände durchdringen. Wie Lichtstrahlen schwärzen sie Filme. Deshalb kann mit Hilfe von Röntgenstrahlen das Innere des Körpers sichtbar gemacht werden. Röntgenbilder sind jedoch nicht so leicht zu "lesen", weil sich die Bilder von Knochen und Organen überlagern.

Diese Problematik vermeiden Techniken wie die Computertomografie, weil sie *Bilder von dünnen Gewebeschichten* liefern. Ein solches Bild zeigt, was man sehen würde, wenn man den Körper oder Gegenstand, den man untersuchen möchte, durchschneiden und die Schnittfläche anschauen würde. Um einen räumlichen Bereich zu erfassen und Erkrankungen wie Blutungen oder Tumore ausschliessen oder erkennen und lokalisieren zu können, müssen allerdings viele parallele und nahe nebeneinander liegende Schichten abgebildet werden.

Wie entsteht ein Schichtbild?

Um es deutlich zu sagen - es handelt sich nicht um eine Fotografie! Das Bild wird vielmehr Punkt für Punkt aufgrund von Messungen *errechnet*.

Das Verfahren beruht darauf, dass Knochen, Flüssigkeiten, gesundes und krankes Gewebe unterschiedlich dicht sind. Durchdringt ein Röntgenstrahl einen Stoff, wird er abgeschwächt und zwar umso mehr, je dichter das Material ist.

Stellen wir uns eine dünne inhomogene Gewebeschicht vor, siehe Abbildung 3. Inhomogen heisst, dass die Gewebeschicht an verschiedenen Stellen unterschiedlich dicht ist. Der eingezeichnete Röntgenstrahl wird bei seinem Gang durch das Gewebe sukzessive mehr abgeschwächt, stärker in Zonen hoher Dichte, weniger stark in Gegenden geringerer Dichte.

Der angedeutete Röntgenstrahl hat vor Eintritt in die Gewebeschicht eine bekannte Anfangsintensität I^a. Nach dem Austritt aus der Gewebeschicht kann die noch vorhandene Restintensität I^e gemessen werden. Wenn die Dichte des Gewebes entlang der Strahlrichtung homogen wäre, könnte man aus dem Verhältnis von I^e zu I^a die Dichte des Gewebes

ermitteln. Da dies typischerweise nicht der Fall ist, beinhaltet das Verhältnis von I^e zu I^a nur eine "mittlere Information" über den Dichteverlauf entlang des betrachteten Strahls.

Die Grundidee ist, sehr viele Röntgenstrahlen in viele verschiedene Richtungen durch die Gewebeschicht zu schicken. Bei jedem Strahl wird die Verminderung der Intensität gemessen, siehe Abbildung 4. Dann besteht die Aufgabe darin, aus dieser Information die Dichteverteilung in der Gewebeschicht durch Rechnung zu bestimmen, zu *rekonstruieren*, wie man sagt. So erhält man ein "Dichteprofil" der untersuchten Gewebeschicht, die bildlich dargestellt wird, wobei hellen Stellen in Anlehnung an klassische Röntgenbilder hohe Dichten, dunklen Stellen geringe Dichten entsprechen.

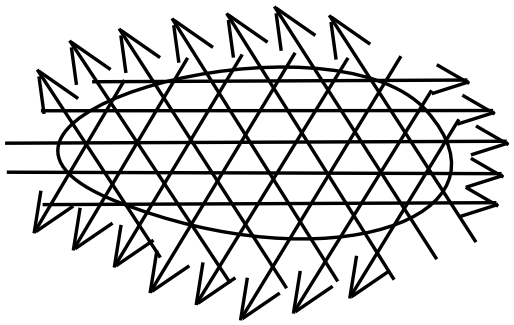

Abbildung 4: Dünne Gewebeschicht, viele Röntgenstrahlen.

Wir haben es auch bei diesem Beispiel mit einem Paar zueinander inverser Probleme zu tun. Das direkte Problem lautet: Gegeben die Dichteverteilung einer dünnen Gewebeschicht, gesucht die Intensitätsabschwächung, die durch die Schicht verlaufende Röntgenstrahlen erfahren. Das inverse Problem besteht in der Bestimmung der Dichteverteilung, wenn man weiss, wie stark Röntgenstrahlen, die durch die Schicht verlaufen, abgeschwächt werden.

Eigentlich käme jetzt der Moment, ein mathematisches Modell für die Computertomografie zu entwerfen. Das soll jedoch nicht geschehen. Es geht in ähnlicher Weise wie bei den vorher besprochenen Beispielen[2].

Sobald ein mathematisches Modell vorliegt, ist die Frage, ob das inverse Problem lösbar ist. (Nicht jedes inverse Problem ist lösbar.) Und falls die Antwort positiv ausfällt, stellt sich die Frage, wie die Lösung praktisch gewonnen werden kann.

Erstaunlicherweise hat sich der österreichische Mathematiker

Johann Radon (1887-1956)

schon im Jahre 1917 mit dieser Aufgabe beschäftigt, zu einer Zeit also, als eine Anwendung à la Computertomografie noch ganz und gar ausserhalb der Reichweite war. Er untersuchte das Problem ohne konkreten Grund, einfach aus intellektueller Neugier. Er fand eine positive Antwort, indem er eine Formel herleitete, mit der man die gesuchte Dichteverteilung jedenfalls theoretisch berechnen kann.

[2]Ein sehr einfaches mathematisches Modell zur Computertomografie findet sich im letzten Kapitel des Leitprogramms *Lineare Gleichungssysteme*. Dieses Leitprogramm steht auf dem Bildungsserver EducETH der ETH-Zürich öffentlich zur Verfügung: http://www.educeth.ch/mathematik

Die Fragestellung von Radon und seine Lösungsformel wurden ein halbes Jahrhundert später, als *A.M. Cormack* und *G.N. Houndsfield* die Computertomografie erfanden, wieder aktuell. 1979 erhielten die beiden dafür den Nobelpreis für Medizin. 1991 wurde auch der Vater der Kern-Spin-Tomografie, der ETH-Chemiker *R. Ernst,* mit einem Nobelpreis ausgezeichnet.

Es genügt nicht, wenn ein inverses Problem wie die Computertomografie "theoretisch" gelöst werden kann. Die Lösung muss darüber hinaus

- *zuverlässig*

und

- *effizient*

ermittelt werden können. Der zweite Punkt ist klar: Da ja die Bilder von vielen parallelen Gewebeschichten benötigt werden, muss die Berechnung eines einzelnen Bildes schnell erfolgen können. Man möchte nicht tage- oder gar monatelang auf den Befund warten, und es sind auch ökonomische Gesichtspunkte zu berücksichtigen.

Im folgenden setzen wir uns mit der Frage der *Zuverlässigkeit der Rekonstruktion* auseinander, weil hier ein wichtiger und interessanter neuer Gesichtspunkt ins Spiel kommt.

3 Gut und schlecht gestellte Probleme

Rekapitulation der Ausgangslage: Ausgangspunkt ist ein mathematisches Modell eines Ursache-Wirkungs-Zusammenhangs. Das heisst, dass einer Ursache, repräsentiert durch eine mathematische Grösse x, in ganz bestimmter Weise eine Wirkung, repräsentiert durch eine mathematische Grösse y, entspricht. x, y können dabei Zahlen, Funktionen, Vektoren oder weitere mathematische Objekte sein. Die "Operation", die aus dem x das y "macht", wollen wir

$$K$$

nennen. Man bezeichnet K auch als *Operator.* Die Lösung des *direkten Problems* besteht darin, aus einem gegebenen x durch Anwenden der Operation K das y zu bestimmen, also

$$y = Kx$$

zu berechnen. Das *inverse Problem* verlangt, zu einem gegebenen y das x zu finden, so dass

$$Kx = y \tag{7}$$

gilt. D.h. es muss die Gleichung (7) bei gegebenem y nach x aufgelöst werden. Nehmen wir an, dass das geht.

Bei der *Aufgabe mit dem Schacht* entspricht x der Tiefe des Schachts und y der Wartezeit, also der Zeit zwischen dem Loslassen des Steins und dem Hören des Aufschlags. Bei der Wasseruhr entspricht x dem Profil f des Gefässes und y der Funktion g, die den Auslaufvorgang erfasst. Beim Computertomografie-Problem schliesslich bedeutet x die Dichteverteilung der

untersuchten Gewebeschicht und y die Intensität, die Röntgenstrahlen noch haben, wenn sie durch das Gewebe hindurchgegangen sind.

Nun ist ein weiterer Aspekt zu berücksichtigen. In der Welt mathematischer Modelle können wir natürlich die Annahme machen, dass vorkommende Grössen *exakt* bestimmbar sind. Im ersten Beispiel heisst das, dass entweder die Schachttiefe x oder die Wartezeit y exakt bestimmt werden kann. Wir können eine solche Annahme als Teil der Idealisierung betrachten, die mit der Bildung eines mathematischen Modells sowieso immer einhergeht. In vielen Situationen ist das gerechtfertigt.

Gewisse inverse Probleme haben leider die unangenehme Eigenschaft, dass genau diese Annahme zu ganz *ungenauen Rekonstruktionen* führt. In solchen Situationen muss man daher

das Phänomen der Mess*un*genauigkeit

mit in die Betrachtung einbeziehen, also mit modellieren. Das kann auf verschiedene Weise geschehen, zum Beispiel wie folgt.

Wir stellen uns vor, der Brunnen habe die "theoretische" oder "wahre" Tiefe x^*. Zu ihr gehört gemäss unserem Modell die "theoretische" oder "wahre" Wartezeit y^*, und der Zusammenhang zwischen beiden wird durch das mathematische Modell der Ursache-Wirkungs-Beziehung gegeben, also durch eine Beziehung des Typs

$$y^* = Kx^*.$$

Modellieren wir nun zuerst das *direkte Problem* unter Berücksichtigung von Mess*un*genauigkeit. Die Vorstellung ist, dass die Grösse x^* nicht exakt bestimmbar ist, dass vielmehr nur eine *Näherung*

$$\bar{x}$$

zur Verfügung steht, die man in der Praxis durch *Messung* gewinnt, und die sich mehr oder weniger von x^* unterscheidet, je nach Qualität der verwendeten Messgeräte und dem Aufwand, der beim Messen getrieben wird.

Statt der "theoretischen" oder "wahren" Wartezeit $y^* = Kx^*$ wird aus der Messung $\bar{x}$ von x^*

$$K\bar{x} \tag{8}$$

berechnet, in der Hoffnung, dass $K\bar{x}$ eine gute Näherung für y^* ist. Die entscheidende Frage ist, *was die Operation K mit einem kleinen Fehler in x macht*, ob sie den Fehler

a) *vergrössert, "aufbläht"*

oder

b) mehr oder weniger unverändert *"weitergibt"*, oder vielleicht sogar *vermindert*.

Im ersten Fall erhalten wir voraussichtlich eine *unbrauchbar schlechte Approximation* für y^*, während im zweiten Fall das Resultat sehr wohl zufriedenstellend sein kann.

Für die Praxis wichtiger ist die analoge Betrachtung für das *inverse Problem*. Ausgangspunkt ist eine "theoretische" Ursache x^* und ihre "theoretische" Wirkung y^*, und eine durch *Messung* ermittelte Näherung

$$\bar{y} \quad \text{für} \quad y^*.$$

Indem wir die Lösung $\bar{x}$ der Gleichung

$$Kx = \bar{y}$$

bestimmen, erhalten wir eine *approximative Rekonstruktion* der "theoretischen" Ursache x^*. Analog zu oben stellt sich die Frage, wie der Operator K wirkt. Ob K

a) die Abweichung von $\bar{y}$ zu y^* *aufbläht* und die approximative Rekonstruktion $\bar{x}$ somit *stark* von der "theoretischen" Ursache x^* *abweicht*,

oder

b) ob sich der Rekonstruktionsfehler $x^* - \bar{x}$ in *akzeptablen Grenzen* hält.

Die beunruhigende Tatsache ist: Es gibt einen fatalen Zusammenhang zwischen dem Verhalten des direkten und des inversen Problems, eine Art "Unschärfe-Relation". Es gilt nämlich: *Je "gutartiger" das direkte Problem ist, desto "bösartiger" verhält sich das dazu inverse Problem.* Man sagt dann, das inverse Problem sei

"schlecht gestellt"[3]

Was kann man tun, wenn ein schlecht gestelltes inverses Problem vorliegt?

Man befürchtet natürlich, dass eine *drastische Erhöhung der Messgenauigkeit*, das heisst eine massive Verkleinerung des Messfehlers $y^* - \bar{y}$, der einzige Ausweg ist, wenn überhaupt. Bei realen Problemen wird man allerdings oft rasch an Grenzen stossen, weil die benötigte Messgenauigkeit gar nicht zu erreichen ist, oder nur um einen Preis, den man nicht bezahlen möchte.

Die Pointe ist, dass die Schwierigkeit durch eine verblüffende mathematische Idee zwar nicht eliminiert, aber immerhin *gemildert* werden kann, wodurch die Ansprüche an die Messtechnologie substantiell reduziert werden.

Bevor diese schöne Idee des russischen Mathematikers

A. N. Tikhonov (1906-1993)

vorgestellt wird, soll noch ein weiteres inverses Problem eingeführt werden, an dem sich das Phänomen der Schlecht-Gestelltheit und den Umgang damit besonders gut demonstrieren lässt.

Es ist nämlich so, dass inverse Probleme *unterschiedlich schlecht gestellt* sind. Wie gut oder schlecht gestellt das *Brunnenproblem* ist, könnte der Leser, die Leserin, wenn er, sie Lust dazu hat, selber untersuchen. Das *Wasseruhr-Designproblem* erweist sich als *ziemlich schlecht gestellt.* Das *Computertomografie-Problem* hingegen ist nur *moderat schlecht gestellt.*

[3]Der englische Begriff heisst *"ill-posed"*.

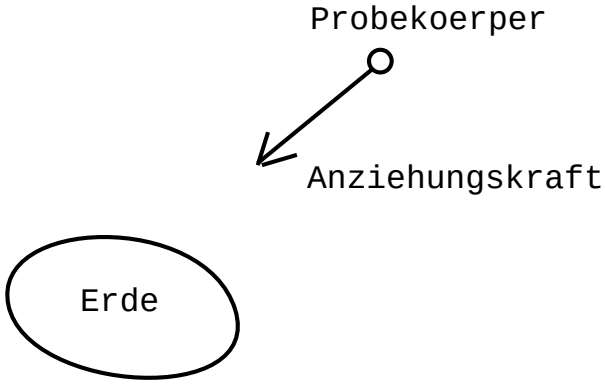

Abbildung 5: Anziehungskraft der Erde auf einen Probekörper

4 Die Dichteverteilung der Erde oder Ein richtig schlecht gestelltes Problem!

Welche Gestalt hat die Erde?

Jedermann weiss natürlich, dass die Erde grosso modo eine Kugel ist. Sie ist aber an den Polen etwas abgeplattet, und die Dichteverteilung im Innern ist keineswegs homogen. Beides wirkt sich aus, zum Beispiel auf die *Bahnen von Satelliten*, also auf die Bewegungen von Kommunikations-, Wetter-, Vermessungssatelliten, usw. Es besteht daher nicht zuletzt auch ein wirtschaftliches Interesse, die *Dichteverteilung unseres Planeten* recht genau zu kennen, um zum Beispiel Satellitenbahnen präzise voraus berechnen zu können.

Dieses Problem ist offensichtlich nicht auf direktem Weg lösbar – wir können die Erde ja nicht einfach auseinander schneiden, dann überall Dichteproben nehmen und wieder zusammen setzen!

Wir sind auf ein indirektes Vorgehen angewiesen, à la Computertomografie. Bei der Computertomografie offenbart sich die Dichteverteilung einer Gewebeschicht durch die Abschwächung, die Röntgenstrahlen erfahren, wenn sie durch das Gewebe hindurchgehen.

Für die folgenden Überlegungen stützen wir uns auf einen anderen Effekt. Nach Newton *ziehen sich zwei Körper* an. Wenn die zwei Körper geringe Ausdehnung haben (man spricht dann von *Massenpunkten*), ist die Anziehungskraft F, die sie aufeinander ausüben, gleich

dem Produkt der beiden Massen, dividiert durch das Quadrat ihres Abstandes, multipliziert mit γ

Das ist der Inhalt des sogenannten *Gravitationsgesetzes*. γ bezeichnet die *Gravitationskonstante*. Wir denken uns die physikalischen Einheiten so gewählt, dass γ den Wert 1 hat.

Das Newtonsche Gravitationsgesetz hat zur Folge, dass ein kleiner Probekörper, den man irgendwo ausserhalb der Erde platziert, durch die Erde eine gewisse Anziehungskraft F erfährt, siehe Abbildung 5. Dieser Effekt soll ausgenutzt werden.

Das *direktes Problem* lautet: Gegeben die Dichteverteilung der Erde, zu berechnen ist die Anziehungskraft F, die ein Probekörper an einem beliebigen Ort ausserhalb der Erde durch

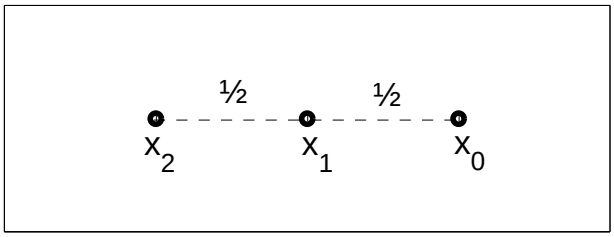

Abbildung 6: Dreigliedrige "Kette" von Massenpunkten mit Massen x_0, x_1, x_2.

sie erfährt. Das *inverse Problem* besteht darin, aus der Kenntnis der Anziehungskraft F, die ein Probekörper an einem beliebigen Ort ausserhalb der Erde erfährt, die Dichteverteilung der Erde zu rekonstruieren.

Es ist nicht wichtig, ob dieses Vorgehen eine wirklich praktikable Methode ist, um die Dichteverteilung der Erde zu bestimmen. Für die grundsätzlichen Überlegungen dieses Artikels ist sie interessant.

Allerdings soll die Fragestellung vereinfacht, und auf ihren wesentlichen Kern reduziert werden.

4.1 Eine Kette von Massenpunkten

Statt der komplizierten dreidimensionalen Erde betrachten wir ein viel einfacheres Objekt: Einen (vielleicht inhomogenen) *Stab*, den wir uns idealisierenderweise eindimensional denken. Und weil man sich einen Stab als *Perlenkette* von Massenteilchen vorstellen kann, vereinfachen wir die Aufgabe weiter und schauen uns eine *"Kette"* von zunächst nur 3, später mehr Massenpunkten an. Die 3 Massenpunkte mit Massen

$$x_0,\ x_1,\ x_2$$

seien geradlinig im Abstand $\frac{1}{2}$ von einander angeordnet, siehe Abbildung 6. Dieses einfache Objekt übernimmt also jetzt die Rolle der Erde.

Weiter werden wir die (Gravitations-)Wirkung der Kette nicht überall im Raum, sondern nur an drei Punkten

$$P_0,\ P_1,\ P_2$$

betrachten, siehe Abbildung 7. Wir nennen diese Punkte *Messstellen.*

Die Massen x_0, x_1, x_2 und die Punkte P_0, P_1, P_2 liegen auf derselben Geraden. Der Abstand zwischen dem Massenpunkt mit der Masse x_0 und dem Punkt P_0 ist 1, der Abstand zwischen den Punkten P_0 und P_1, bzw. zwischen P_1 und P_2 ist jeweils $\frac{1}{2}$.

Es soll die Anziehungskraft bestimmt werden, die ein Probekörper der Masse 1 durch die Kette erfährt, wenn er sich im Punkt P_0, bzw. P_1, bzw. P_2 befindet. Berechnen wir zuerst die Anziehungskraft der Masse x_2 auf den Probekörper, wenn er sich im Punkt P_1 befindet. Der Abstand der beiden Massen beträgt

$$2 \cdot \frac{1}{2} + 1 + 1 \cdot \frac{1}{2}.$$

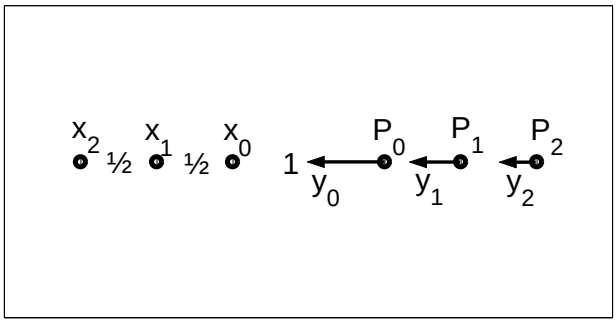

Abbildung 7: Eingezeichnet ist die Anziehungskraft y_0, bzw. y_1, bzw. y_2, die die dreigliedrige Kette auf einen Probekörper der Masse 1 am Punkt P_0, bzw. P_1, bzw. P_2 ausübt.

Weil die Gravitationskonstante den Wert 1 hat und auch die Masse des Probekörpers 1 ist, folgt für die gesuchte Kraft

$$1 \cdot \frac{1 \cdot x_2}{(2 \cdot \frac{1}{2} + 1 + 1 \cdot \frac{1}{2})^2} = (\frac{2}{2+2+1})^2 x_2.$$

Analog erhält man für die Anziehungskraft der Masse x_j auf den Probekörper der Masse 1, wenn er am Punkt P_i platziert ist

$$(\frac{2}{j+2+i})^2 x_j.$$

Die *Gesamtkraft* y_i, die die Probemasse am Punkt P_i durch die dreigliedrige Kette erfährt, ist die Summe der Kräfte, die die Massenpunkte x_0, x_1, x_2 auf sie ausüben. Man erhält, je nachdem ob sich die Probemasse im Punkt P_0, oder im Punkt P_1, oder im Punkt P_2 befindet, der Reihe nach

$$\boxed{\begin{array}{lcl} y_0 & = & (\frac{2}{2})^2 x_0 + (\frac{2}{3})^2 x_1 + (\frac{2}{4})^2 x_2 \\ y_1 & = & (\frac{2}{3})^2 x_0 + (\frac{2}{4})^2 x_1 + (\frac{2}{5})^2 x_2 \\ y_2 & = & (\frac{2}{4})^2 x_0 + (\frac{2}{5})^2 x_1 + (\frac{2}{6})^2 x_2 \end{array}} \tag{9}$$

Die Gleichungen (9) sind die *Ursache-Wirkungs-Beziehung* für dieses Problem.

4.2 Ein Rechenexperiment

Es ist aufschlussreich, ein *Zahlenbeispiel* durchzurechnen. Ausgangspunkt (oder Hintergrund) ist eine "theoretische" oder "wahre" Kette mit den Massen x_0^*, x_1^*, x_2^*. Für die Rechnung nehmen wir an, dass alle drei Massen *gleich gross* sind und dass die *Gesamtmasse* 1 ist, das heisst wir setzen:

$$\boxed{x_0^* = x_1^* = x_2^* = \tfrac{1}{3} = 0.3333\ldots}$$

Diese Kette erzeugt Gravitationskräfte y_0^*, y_1^*, y_2^* an den Messstellen P_0, P_1, P_2. Man erhält sie, indem man in den Gleichungen (9) x_0 durch x_0^*, x_1 durch x_1^*, x_2 durch x_2^* ersetzt. Das Resultat ist:

$$\boxed{y_0^* = \tfrac{61}{108} \approx 0.564815,\ y_1^* = \tfrac{769}{2700} \approx 0.284815,\ y_2^* = \tfrac{469}{2700} \approx 0.173704}$$

Und nun zum *Rekonstruktionsproblem*. Wir gehen davon aus, dass die an den Messstellen P_0, P_1, P_2 wirkenden Gravitationskräfte durch Messung ermittelt werden. Das heisst man hat, je nach Qualität der Messung, von $y_0^*,\ y_1^*,\ y_2^*$ mehr oder weniger abweichende *Messwerte*

$$\bar{y}_0,\ \bar{y}_1,\ \bar{y}_2$$

zur Verfügung.

Wir *simulieren* solche Messwerte, indem wir die exakten Werte y_0^*, y_1^*, y_2^* geringfügig verändern. Wir wählen zum Beispiel folgende Werte:

$$\begin{aligned}
\bar{y}_0 &= \tfrac{61}{108} - \tfrac{1}{1000} \approx 0.563815 \\
\bar{y}_1 &= \tfrac{769}{2700} + \tfrac{1}{1000} \approx 0.285815 \\
\bar{y}_2 &= \tfrac{469}{2700} - \tfrac{1}{1000} \approx 0.172704
\end{aligned}$$

Jetzt setzen wir in (9) für $y_0,\ y_1,\ y_2$ die Werte $\bar{y}_0,\ \bar{y}_1,\ \bar{y}_2$ ein und lösen das entstehende Gleichungssystem nach $x_0,\ x_1,\ x_2$ auf – was sollten wir sonst tun? Das Ergebnis lautet:

$$\begin{aligned}
\bar{x}_0 &\approx 0.207622 \\
\bar{x}_1 &\approx 0.999437 \\
\bar{x}_2 &\approx -0.352005
\end{aligned}$$

Dieses Resultat ist eine böse Überraschung. Es ist nicht nur ungenau – es ist sogar *un*physikalisch, weil es eine negative Masse liefert!

4.3 Analyse

Wie ist das möglich?

Man erinnert sich vielleicht an eine Schwierigkeit beim Konstruieren im Geometrieunterricht? Wenn zwei Geraden sich unter einem flachen Winkel schneiden, ist es schwierig, den Schnittpunkt genau zu lokalisieren, besonders, wenn der Bleistift nicht gut gespitzt ist und die Linien ein wenig "dick geraten" sind. Man spricht dann von "schleifenden Schnitten". Die Behauptung: Wir haben es genau mit diesem Phänomen zu tun.

Durch die drei Gleichungen in (9) werden 3 Ebenen im Raum definiert. Man betrachte die erste Gleichung. Sie definiert eine Ebene im x_0-x_1-x_2-Raum. Die Koeffizienten von x_0, x_1, x_2 sind die Komponenten ihres *Normalenvektors* $\vec{n}_0$. Durch $\vec{n}_0$ ist die Ebene bis auf Parallelverschiebung festgelegt. Die Konstante y_0 in der Ebenengleichung bestimmt, um welche dieser zueinander parallelen Ebenen es sich handelt. Analoges gilt für die 2. und die 3. Gleichung von (9).

Mit Hilfe des Skalarproduktes kann man die Winkel zwischen den Normalen berechnen. Man findet folgende Ergebnisse:

$$\begin{aligned}
\text{Winkel}(\vec{n}_0, \vec{n}_1) &\approx 6.9^\circ \\
\text{Winkel}(\vec{n}_0, \vec{n}_2) &\approx 11.3^\circ \\
\text{Winkel}(\vec{n}_1, \vec{n}_2) &\approx 4.4^\circ
\end{aligned}$$

Die Winkel sind recht klein, die Ebenen sind in der Tat fast parallel. Das wäre nicht weiter schlimm, wenn die Konstanten $y_0,\ y_1,\ y_3$ in den Ebenengleichungen *exakt* bekannt wären.

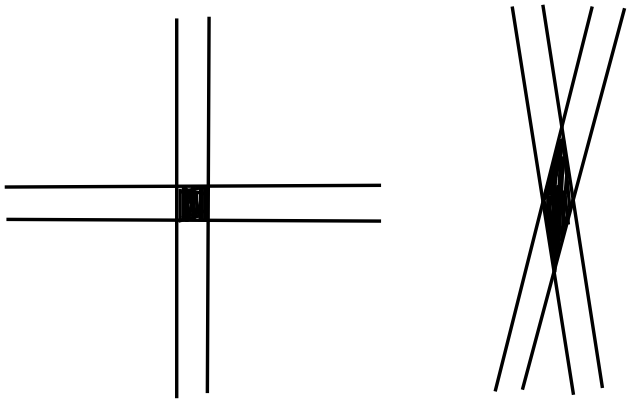

Abbildung 8: Schnitte von parallelen Geraden. a. Die Geraden stehen senkrecht, b. Die Geraden sind fast parallel.

Das ist aber gerade *nicht* der Fall, weil sie durch Messung gewonnen werden und daher mit Fehlern behaftet sind.

Man muss daher eigentlich *drei Scharen von parallelen Ebenen* betrachten und sie tripelweise mit einander schneiden. Wenn nun die Ebenenscharen fast parallel sind, können verschiedene Schnittpunkte recht weit auseinander liegen, selbst wenn die drei "Pakete" paralleler Ebenen eher "dünn" sind. Disen Sachverhalt illustriert eine Dimension tiefer Abbildung 8.

Machen wir weiter folgendes Experiment: Wir varriieren die Werte von y_0, y_1, y_2 im Gleichungssystem (9) geringfügig und schauen, wie gross die Schwankungen in den Lösungen x_0, x_1, x_2 werden können. Für die obigen Werte von y_0^*, y_1^*, y_2^* und mit $\delta = 0.001$ findet man:

$$\begin{array}{rcl} y_0 & \in & [y_0^* - \delta, y_0^* + \delta] \\ y_1 & \in & [y_1^* - \delta, y_1^* + \delta] \\ y_2 & \in & [y_2^* - \delta, y_2^* + \delta] \end{array} \quad \Longrightarrow \quad \begin{array}{rcl} x_0 & \in & [0.20, 0.46] \\ x_1 & \in & [-0.34, 1.0] \\ x_2 & \in & [-0.36, 1.1] \end{array}$$

Man sieht, dass eine sehr kleine Schwankungsbreite bei den y's eine etwa um den Faktor 1000 grössere Schwankungsbreite bei den x-Werten nach sich zieht.

In der *Numerischen Mathematik* sagt man, das lineare Gleichungssystem (9) sei

$$\boxed{\text{schlecht konditioniert.}}$$

In der *Theorie der Inversen Probleme* spricht man, wie gesagt, von einem

$$\boxed{\text{schlecht gestellten Problem.}}$$

Ein andere gute Formulierung ist: Die Lösungen $x_0,\ x_1,\ x_2$ des linearen Gleichungssystems (9) hängen $\boxed{\text{sensitiv}}$, d.h. $\boxed{\text{empfindlich}}$, von den Daten $y_0,\ y_1,\ y_2$ ab. Gemeint ist, dass geringfügige Veränderungen der y's zu vergleichsweise sehr verschiedenen Lösungen des Gleichungssystems (9) führen.

4.4 Ein gute Idee

Die Frage ist, ob die Erhöhung der Messgenauigkeit der einzige Ausweg ist, um bei einem schlecht gestellten Problem eine brauchbare (Näherungs-)Lösung zu erhalten. Das hätte dann zur Folge, dass solche Probleme in der Praxis häufig de facto unlösbar wären, weil die nötige Messgenauigkeit aus technischen Gründen gar nicht erreichbar ist, oder nur mit einem Aufwand, der zum Beispiel aus Kostengründen nicht in Kauf genommen werden kann.

Im Jahr 1963 hatte der russische Mathematiker

A. N. Tikhonov (1906-1993)

eine gute Idee. Sein Grundgedanke:

Man *verfälsche* das schlecht gestellte inverse Problem *ein wenig*, so dass es *"weniger sensitiv"* ist. Allerdings darf man es nicht zu stark verfälschen, damit es noch "genügend nahe" beim ursprünglichen Problem ist.

Das verfälschte Problem wird als *regularisiert* bezeichnet. Die Hoffnung ist, dass die Lösung des regularisierten Problems eine *akzeptable Näherungslösung* für das *schlecht gestellte inverse Problem* ist. Es ist klar, dass es sich hier um ein riskantes Manöver handelt, das darauf setzt, dass die Sensitivität zurückgedrängt werden kann, ohne dass der Bezug zum eigentlichen Problem zu sehr verloren geht.

Die entscheidende praktische Frage ist, wie das gegebene Problem"verfälscht" werden soll, und wie stark. *Tikhonovs Vorschlag*, angewandt auf unsere Kette, lautet so: Man ersetze die *Diagonalelemente* im Gleichungssystem (9) durch etwas grössere Zahlen. Statt (9) betrachten wir also das folgende System:

$$\begin{aligned} y_0 &= ((\tfrac{2}{2})^2 + \boxed{\alpha})x_0 + (\tfrac{2}{3})^2 x_1 + (\tfrac{2}{4})^2 x_2 \\ y_1 &= (\tfrac{2}{3})^2 x_0 + ((\tfrac{2}{4})^2 + \boxed{\alpha})x_1 + (\tfrac{2}{5})^2 x_2 \\ y_2 &= (\tfrac{2}{4})^2 x_0 + (\tfrac{2}{5})^2 x_1 + ((\tfrac{2}{6})^2 + \boxed{\alpha})x_2 \end{aligned} \tag{10}$$

Dabei soll α eine *kleine positive Zahl* sein. α heisst *Regularisierungsparameter.*

Damit das Konzept ganz klar ist, hier nochmals die Logik. Ausgangspunkt ist die zugrundeliegende Kette mit den drei Massen

$$\boxed{x_0^* = x_1^* = x_2^* = \tfrac{1}{3}.}$$

Sie erzeugen an den Messstellen $P_0,\ P_1,\ P_2$ nach (9) die Gravitationskräfte

$$\boxed{y_0^* = \tfrac{61}{108} \approx 0.564815,\ y_1^* = \tfrac{769}{2700} \approx 0.285815,\ y_2^* = \tfrac{469}{2700} \approx 0.172764.}$$

Diese sind aber nicht exakt bekannt: Man verfügt lediglich über mit (kleinen) Fehlern behaftete *Messungen:*

$$\boxed{\bar{y}_0 = \tfrac{61}{108} - \tfrac{1}{1000},\ \bar{y}_1 = \tfrac{769}{2700} + \tfrac{1}{1000},\ \bar{y}_2 = \tfrac{469}{2700} - \tfrac{1}{1000}.}$$

Die Idee ist, in (10) $y_0,\ y_1,\ y_2$ durch $\bar{y}_0, \bar{y}_1, \bar{y}_2$ zu ersetzen, für α eine kleine positive Zahl zu wählen und dann das Gleichungssystem (10) nach $x_0,\ x_1,\ x_2$ aufzulösen, in der Hoffnung, dass seine Lösung, die mit $\bar{x}_0^\alpha, \bar{x}_1^\alpha, \bar{x}_2^\alpha$ bezeichnet werde, eine akzeptable Näherung für $x_0^* = x_1^* = x_2^* = \frac{1}{3}$ ist.

Wir benötigen ein *Mass für die Fehler.* Es gibt verschiedene Möglichkeiten, den Fehler zu messen. Eine geht so: Man bestimmt die sogenannten Absolutbeträge der Differenzen

$$x_0^* - \bar{x}_0^\alpha, \quad x_1^* - \bar{x}_1^\alpha, \quad x_2^* - \bar{x}_2^\alpha$$

und nimmt *die grösste* der drei Zahlen. Wir bezeichnen sie mit $error(\alpha)$.

Hier sind die Resultate eines Computerexperiments. Für verschiedene Werte von α wird der Fehler $error(\alpha)$ berechnet. Die Ergebnisse sind in den Tabellen 1, 2 wiedergegeben.

α	$error(\alpha)$
0.00	0.69
0.01	0.15
0.02	0.12
0.03	0.12
0.04	0.12
0.05	0.13
0.06	0.13

Tabelle 1

α	$error(\alpha)$
0.024	0.1169
0.025	0.1168
0.026	0.1167
0.027	0.1168
0.028	0.1169
0.029	0.1170
0.030	0.1172

Tabelle 2

Für den offenbar günstigen Wert $\alpha = 0.025$ findet man:

$$\begin{array}{lcl} \bar{x}_0^\alpha & \approx & 0.338276 \\ \bar{x}_1^\alpha & \approx & 0.366629 \\ \bar{x}_2^\alpha & \approx & 0.216546 \end{array}$$

$$\begin{array}{lll} x_0^* & = \frac{1}{3} & = 0.33333\ldots \\ x_1^* & = \frac{1}{3} & = 0.33333\ldots \\ x_2^* & = \frac{1}{3} & = 0.33333\ldots \end{array}$$

Die Tikhonov-Idee "funktioniert" also bei unserem Beispiel ganz gut. Jedenfalls kann man feststellen:

a) Das Resultat ist nicht aphysikalisch!

b) Der Fehler, insbesondere für die Massen x_0^* und x_1^*, ist mit unter 10% respektabel klein. Schlechter sieht es bei der dritten Masse der Kette aus, hier liegt der Rekonstruktionsfehler bei 35%.

4.5 5-gliedrige Kette

Eigentlich sollte ja die Dichteverteilung der Erde rekonstruiert werden oder mindestens diejenige eines Stabes. Die 3-gliedrige Kette ist nur eine sehr grobe Näherung für einen Stab. Betrachten wir wenigstens noch eine Kette mit fünf Gliedern, siehe Abbildung 9. Die "theoretischen" Massen seien

$$x_0^* = x_1^* = x_2^* = x_3^* = x_4^* = \frac{1}{5} = 0.2.$$

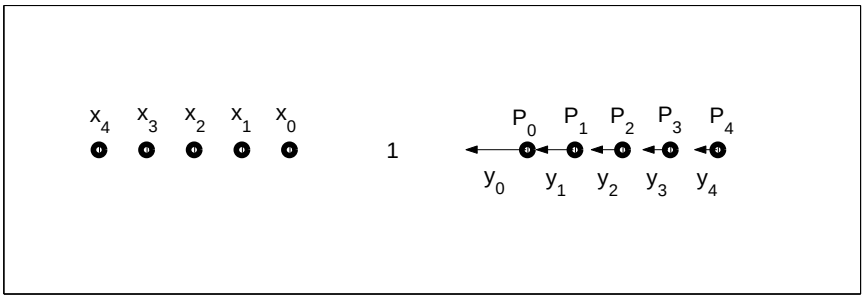

Abbildung 9: Analog zur 3-gliedrigen Kette betrachten wir nun ein 5-gliedrige Kette.

Aus einem zu (9) analogen Gleichungssystem mit 5 Gleichungen für die Beziehung zwischen den Massen x_0, x_1, x_2, x_3, x_4 und den Gravitationskräften y_0, y_1, y_2, y_3, y_4 an den 5 Messstellen P_0, P_1, P_2, P_3, P_4, erhält man die "theoretischen" Gravitationskräfte

$$y_0^* = 0.532.., \, y_1^* = 0.371.., \, y_2^* = 0.275.., \, y_3^* = 0.213.., \, y_4^* = 0.170..$$

Daraus "simulieren" wir wieder Messungen, indem wir zum Beispiel

$$\bar{y}_0 = y_0^* - \delta, \, \bar{y}_1 = y_1^* + \delta, \, \bar{y}_2 = y_2^* - \delta, \, \bar{y}_3 = y_3^* + \delta, \, \bar{y}_4 = y_4^* - \delta$$

setzen. Wieder ist δ ein Mass für die Messgenauigkeit: Grosses δ bedeutet, dass ungenau gemessen wird. Kleines δ bedeutet hohe Messgenauigkeit.

a) Beginnen wir mit

$$\delta = 10^{-3} = 0.001$$

wie bei der 3-gliedrigen Kette. *Unregularisiert*, das heisst für $\alpha = 0$, erhält man als Rekonstruktion von $x_0^*, \ldots, x_4^*$:

$$-79.2779, 733.555, -2237.01, 2767.58, -1196.55.$$

Diese Resultat ist noch viel katastrophaler als dasjenige, das wir im Fall der 3-gliedrigen Kette erhalten hatten! Wenn man mit

$$\alpha = 0.045$$

nach Tikhonov regularisiert, erhält man hingegen als Rekonstruktion

$$0.201822, 0.235736, 0.169533, 0.191348, 0.126361.$$

Das sind halbwegs akzeptable Näherungen. Die prozentualen Fehler betragen

$$0.92\%, 17.9\%, 15.2\%, 4.3\%, 36.8\%.$$

b) Stellen wir uns vor, dass man die Messgenauigkeit um eine Grössenordnung erhöhen kann. Das heisst es sei jetzt

$$\delta = 10^{-4} = 0.0001.$$

Unregularisiert findet man als Rekonstruktion

$$-7.74779,\ 73.5355,\ -223.521,\ 276.938,\ -119.475,$$

ein immer noch völlig unbrauchbares Resultat. Indem man mit

$$\alpha = 0.005$$

Tikhonov regularisiert, findet man als Rekonstruktion

$$0.191257,\ 0.228835,\ 0.185026,\ 0.21565,\ 0.163105.$$

Die prozentualen Fehler sind nun reduziert auf recht akzeptable

$$4.4\%,\ 14.4\%,\ 7.5\%,\ 7.8\%,\ 18.4\%.$$

c) Erhöht man die Messgenauigkeit noch um eine Grössenordnung, das heisst, es werde

$$\delta = 10^{-5} = 0.00001.$$

gesetzt. Unregularisiert erhält man weiterhin drei negative Massen, also ein physikalisch unsinniges Ergebnis. Tikhonov regularisiert mit

$$\alpha = 0.001$$

ist das Ergebnis jetzt recht brauchbar. Die prozentualen Fehler schwanken zwischen 1.5% und 7.5%.

d) Man kann sich fragen, wie hoch die Messgenauigkeit sein müsste, um *unregularisiert* zu einem ähnlich genauen Resultat zu gelangen. Für

$$\delta = 10^{-8}$$

schwanken die Fehler zwischen 0.4% und 14%. Für

$$\delta = 10^{-9}$$

liegen sie zwischen 0.04% und 1.4%. Das heisst: Unregularisiert müsste man *mehrere Grössenordnungen genauer messen!*

5 Elemente einer mathematischen Analyse

Dieses Kapitel ist für mathematisch interessierte Leserinnen und Leser geschrieben, die noch etwas genauer verstehen möchten, was ein schlecht gestelltes inverses Problem ist, und warum und unter welchen Bedingungen die Idee von Thikonov funktioniert. Es setzt ein wenig Vertrautheit mit Vektorrechnung in der Ebene voraus. An einigen Stellen wird das Skalarprodukt benützt.

5.1 Matrizen und Vektoren

Um die Schwierigkeiten mit inversen Problemen und den Nutzen des Vorschlags von Tikhonov besser verstehen zu können benötigt man einige Grundbegriffe aus der sogenannten *Matrizenrechnung*. Es handelt sich zunächst nur um eine neue Schreibweise für lineare Gleichungssysteme und ihre Lösungen.
Es wird alles am Beispiel von 2 Gleichungen mit 2 Unbekannten erklärt. Die Verallgemeinerung auf für 3 Gleichungen mit 3 Unbekannten oder 4 Gleichungen mit 4 Unbekannten, usw ist nicht besonders schwierig und versteht sich (fast) von selbst.

Stellen Sie sich unter $a, b, c, d.e, f$ sechs zwar nicht bekannte aber bestimmte Zahlen vor und betrachten Sie das (lineare) Gleichungssystem

$$\left\{ \begin{array}{rcl} ax + by & = & e \\ cx + dy & = & f \end{array} \right. \tag{11}$$

Die Aufgabe besteht darin, Zahlen x und y zu finden, die das Gleichungssystem (11) erfüllen.

In den meisten Fällen findet man exakt ein solches Zahlenpaar x, y. Unter Umständen kann es aber auch unendlich viele derartige Zahlenpaare geben, oder, wenn man Pech hat, überhaupt keines. Welcher Fall eintritt, hängt von den sechs Zahlen $a, b, c, d.e, f$ ab, den *Daten* des Gleichungssystems, wie man manchmal sagt.

5.1.1 Matrizen und Vektoren

Die ersten vier Zahlen, also a, b, c, d, nennt man die "Koeffizienten" des Gleichungssystems (11). Aus ihnen bildet man eine sogenannte

$$2\times 2\text{-Matrix},$$

indem man sie zu 2 Zeilen und 2 Kolonnen anordnet und zwischen eckige oder runde Klammern setzt:

$$\left[\begin{array}{cc} a & b \\ c & d \end{array} \right] \quad \text{oder:} \quad \left(\begin{array}{cc} a & b \\ c & d \end{array} \right)$$

Damit man sich bequem auf das neue "Gebilde" beziehen kann, benützt man zur Abkürzung einen grossen Buchstaben, z.B. ein A, B oder M, und schreibt

$$A = \left[\begin{array}{cc} a & b \\ c & d \end{array} \right] \tag{12}$$

Neben den Koeffizienten des Gleichungssystems (11) spielen auch zwei Zahlenpaare eine wichtige Rolle: Einerseits die beiden Unbekannten x und y, und andererseits die beiden Zahlen e und f auf der rechten Seite des Gleichungssystems (11).

Für diese führt man sogenannte 1-*kolonnige Matrizen* ein, nennt sie *Vektoren*, und bezeichnet sie mit (unterstrichenen) Kleinbuchstaben:

$$\underline{z} = \left[\begin{array}{c} x \\ y \end{array} \right] \quad , \quad \underline{g} = \left[\begin{array}{c} e \\ f \end{array} \right] \tag{13}$$

Eine 1-kolonnige Matrix aus zwei Elementen nennt man auch 2×1-Matrix (weil sie aus 2 Zeilen und 1 Kolonne bestehen), oder 2-Vektor.

5.1.2 Matrizenoperationen

Es hat sich als sehr zweckmässig erwiesen für Matrizen gewisse *Rechenoperationen* zu definieren.

Beispielsweise *multipliziert* man eine Matrix mit einer Zahl, indem man jedes Element der Matrix mit der Zahl multipliziert.

Beispiel 1

$$A = \begin{bmatrix} 3 & 5 \\ -2 & 4 \end{bmatrix} \quad , \quad 3 \cdot A = 3A = \begin{bmatrix} 3 \cdot 3 & 3 \cdot 5 \\ 3 \cdot (-2) & 3 \cdot 4 \end{bmatrix} = \begin{bmatrix} 9 & 15 \\ -6 & 12 \end{bmatrix}$$

$$\underline{g} = \begin{bmatrix} 4 \\ 5 \end{bmatrix} \quad , \quad 6 \cdot \underline{g} = 6\underline{g} = \begin{bmatrix} 6 \cdot 4 \\ 6 \cdot 5 \end{bmatrix} = \begin{bmatrix} 24 \\ 30 \end{bmatrix}$$

Des weiteren kann man "gleichartige" Matrizen *addieren*, indem man entsprechende Elemente addiert.

Beispiel 2 *Aus*

$$A = \begin{bmatrix} 3 & 5 \\ -2 & 4 \end{bmatrix} \quad , \quad B = \begin{bmatrix} 6 & -2 \\ 3 & -3 \end{bmatrix}$$

erhält man

$$A + B = \begin{bmatrix} 3 & 5 \\ -2 & 4 \end{bmatrix} + \begin{bmatrix} 6 & -2 \\ 3 & -3 \end{bmatrix} = \begin{bmatrix} 3+6 & 5-2 \\ -2+3 & 4-3 \end{bmatrix} = \begin{bmatrix} 9 & 3 \\ 1 & 1 \end{bmatrix}.$$

Für

$$\underline{g} = \begin{bmatrix} 4 \\ 5 \end{bmatrix} \quad , \quad \underline{h} = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$$

ergibt sich

$$\underline{g} + \underline{h} = \begin{bmatrix} 4 \\ 5 \end{bmatrix} + \begin{bmatrix} -1 \\ 2 \end{bmatrix} = \begin{bmatrix} 4-1 \\ 5+2 \end{bmatrix} = \begin{bmatrix} 3 \\ 7 \end{bmatrix}.$$

Spannender ist die Definition der *Multiplikation* einer 2×2-Matrix mit einem 2-Vektor und mit einem 2-Vektor als Ergebnis. Die Definition lautet wie folgt. Es sei:

$$M = \begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix} \quad , \quad \underline{w} = \begin{bmatrix} u \\ v \end{bmatrix}$$

Dann setzt man:

$$M \cdot \underline{w} = M\underline{w} = \begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} \alpha u + \beta v \\ \gamma u + \delta v \end{bmatrix} \tag{14}$$

Es folgen ein paar Beispiele mit konkreten Zahlen zur Illustration.

Beispiel 3

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 5 \\ 6 \end{bmatrix} = \begin{bmatrix} 1 \cdot 5 + 2 \cdot 6 \\ 3 \cdot 5 + 4 \cdot 6 \end{bmatrix} = \begin{bmatrix} 17 \\ 39 \end{bmatrix}$$

$$\begin{bmatrix} -3 & 3 \\ -5 & -2 \end{bmatrix} \begin{bmatrix} -2 \\ 3 \end{bmatrix} = \begin{bmatrix} (-3) \cdot (-2) + 3 \cdot 3 \\ (-5) \cdot (-2) + (-2) \cdot 3 \end{bmatrix} = \begin{bmatrix} 15 \\ 4 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 7 \\ 8 \end{bmatrix} = \begin{bmatrix} 1 \cdot 7 + 0 \cdot 8 \\ 0 \cdot 7 + 1 \cdot 8 \end{bmatrix} = \begin{bmatrix} 7 \\ 8 \end{bmatrix}$$

Die Matrix

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

nennt man übrigens 2×2-*Einheitsmatrix* und bezeichnet sie mit I oder I_2. Ein anderer Name ist Identität, oder Identitätsmatrix, weil für jeden beliebigen 2-Vektor $\underline{w}$ gilt

$$I\underline{w} = \underline{w},$$

das heisst die Multiplikation mit I lässt den Vektor $\underline{w}$ unverändert.

Warum ist die Multiplikation einer 2×2-Matrix mit einem 2-Vektor so definiert, wie Formel (14) angibt? Die Antwort lautet: Mit den eingeführten neuen Objekten (2×2-Matrizen, 2-Vektoren) und der gemäss (14) definierten Multiplikation lässt sich das (lineare) Gleichungssystem (11) wunderbar einfach schreiben. Mit

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \quad , \quad \underline{z} = \begin{bmatrix} x \\ y \end{bmatrix} \quad , \quad \underline{g} = \begin{bmatrix} e \\ f \end{bmatrix} \tag{15}$$

folgt

$$A\underline{z} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} ax + by \\ cx + dy \end{bmatrix} \tag{16}$$

Nun muss man nur noch vereinbaren, dass (gleichartige) Matrizen gleich sind, wenn entsprechende Elemente gleich sind, das heisst

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix} \text{ genau wenn } \begin{matrix} a = \alpha,\, b = \beta \\ c = \gamma,\, d = \delta \end{matrix} \text{ gilt}$$

beziehungsweise

$$\begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} \alpha \\ \beta \end{bmatrix} \text{genau wenn } \begin{matrix} a = \alpha \\ c = \gamma \end{matrix} \text{ gilt.}$$

Dann kann man das durch Gleichung (11) gegebene Problem in der folgenden Form "schreiben":

$$A\underline{z} = \underline{g}, \tag{17}$$

wobei A, $\underline{z}$, $\underline{g}$ durch (15) festgelegt sind.

Was ist gemeint?

Aus den Zahlen a, b, c, d, e, f bildet man die 2×2-Matrix A und den 2-Vektor $\underline{g}$ gemäss (15). Dann gilt:

- Sind x, y zwei Zahlen, die das lineare Gleichungssystem (11) erfüllen, so erfüllt der 2-Vektor
$$\underline{z} = \begin{bmatrix} x \\ y \end{bmatrix}$$
die (Matrix-)Gleichung (17).
- Erfüllt umgekehrt der 2-Vektor
$$\underline{z} = \begin{bmatrix} x \\ y \end{bmatrix}$$
die (Matrix-)Gleichung (17), so erfüllen die beiden Zahlen x, y von $\underline{z}$ (die sogenannten *Komponenten* oder *Elemente* des Vektors $\underline{z}$) das lineare Gleichungssystem (11).

Vielleicht fragen Sie: Lohnt sich dieser Aufwand? Neue Objekte (Matrizen, Vektoren) und neue Operationen (Multiplikation einer Matrix mit einer Zahl, Addition von (gleichartigen) Matrizen, Multiplikation einer Matrix mit einem Vektor) definieren, nur um das (lineare) Gleichungssystem (11) in der Form (17) anschreiben zu können?

Die Antwort ist: Es lohnt sich, vielleicht nicht sofort, aber à la longue. Lohnt es sich für eine (bescheidene) Einführung in schlecht gestellte inverse Probleme? Lassen Sie sich darauf ein und urteilen Sie später.

Die Nützlichkeit von Matrizen und Vektoren beruht u.a. darauf, dass für die eingeführten Rechenoperationen ein paar *Rechenregeln* gelten. Die für uns wichtigste ist folgende. Sind A eine 2×2-Matrix, $\underline{w}, \underline{z}$ zwei 2-Vektoren, und r, s Zahlen, so gilt:

$$A(r\underline{w} + s\underline{z}) = rA\underline{w} + sA\underline{z} \tag{18}$$

Die Richtigkeit dieser Formel zeigt man durch Nachrechnen: Wir setzen

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \quad , \quad \underline{w} = \begin{bmatrix} u \\ v \end{bmatrix} \quad , \quad \underline{z} = \begin{bmatrix} x \\ y \end{bmatrix}$$

und bilden der Reihe nach:

1. $r\underline{w} + s\underline{z} = r \begin{bmatrix} u \\ v \end{bmatrix} + s \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} ru \\ rv \end{bmatrix} + \begin{bmatrix} sx \\ sy \end{bmatrix} = \begin{bmatrix} ru + sx \\ rv + sy \end{bmatrix}$

2. $A(r\underline{w} + s\underline{z}) = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} ru + sx \\ rv + sy \end{bmatrix} = \begin{bmatrix} aru + asx + brv + bsy \\ cru + csx + drv + dsy \end{bmatrix}$

3. $A\underline{w} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} au + bv \\ cu + dv \end{bmatrix}$

4. $rA\underline{w} = r \begin{bmatrix} au + bv \\ cu + dv \end{bmatrix} = \begin{bmatrix} rau + rbv \\ rcu + rdv \end{bmatrix}$

5. $sA\underline{z} = s \begin{bmatrix} ax + by \\ cx + dy \end{bmatrix} = \begin{bmatrix} sax + sby \\ scx + sdy \end{bmatrix}$

6. $rA\underline{w} + sA\underline{z} = \begin{bmatrix} rau + rbv + sax + sby \\ rcu + rdv + scx + sdy \end{bmatrix}$

Der Vergleich des Resultats in 2. und in 6. zeigt die Richtigkeit von Formel (18).

5.2 Eigenwerte und Eigenvektoren: Ein Beispiel

5.2.1 Erste Begegnung

In diesem Abschnitt lernen Sie sogenannte *Eigenwerte* und *Eigenvektoren* von Matrizen kennen. Sie sind ein ausserordentlich nützliches Instrument, wie Sie schon im nächsten Abschnitt erfahren werden. Ausgangspunkt ist die 2×2-Matrix

$$A = \begin{bmatrix} 1 & -2 \\ 1 & 4 \end{bmatrix} \tag{19}$$

Betrachten Sie dazu den 2-Vektor

$$\underline{u} = \begin{bmatrix} -2 \\ 1 \end{bmatrix} \tag{20}$$

Dass $\underline{u}$ für die Matrix A eine besondere Bedeutung hat, sehen Sie, wenn Sie $A\underline{u}$ berechnen:

$$A\underline{u} = \begin{bmatrix} 1 & -2 \\ 1 & 4 \end{bmatrix} \begin{bmatrix} -2 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \cdot (-2) + (-2) \cdot 1 \\ 1 \cdot (-2) + 4 \cdot 1 \end{bmatrix} = \begin{bmatrix} -4 \\ 2 \end{bmatrix}$$

Wie man sieht, ist der Vektor $A\underline{u}$ gerade gleich $2\underline{u}$, denn es ist

$$2\underline{u} = 2 \begin{bmatrix} -2 \\ 1 \end{bmatrix} = \begin{bmatrix} -4 \\ 2 \end{bmatrix}.$$

Mit anderen Worten, es gilt

$$A\underline{u} = 2\underline{u}. \tag{21}$$

Man sagt: Die Zahl 2 ist *Eigenwert* der Matrix A und $\underline{u}$ *Eigenvektor* zum Eigenwert 2.

Betrachten Sie als nächstes den Vektor

$$\underline{v} = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \tag{22}$$

Offenbar gilt

$$A\underline{v} = \begin{bmatrix} 1 & -2 \\ 1 & 4 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 1 \cdot 1 + (-2) \cdot (-1) \\ 1 \cdot 1 + 4 \cdot (-1) \end{bmatrix} = \begin{bmatrix} 3 \\ -3 \end{bmatrix}$$

Wegen

$$3\underline{v} = 3 \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 3 \\ -3 \end{bmatrix}$$

können wir sagen: Es gilt

$$A\underline{v} = 3\underline{v} \tag{23}$$

Das heisst: Die Zahl 3 ist ebenfalls Eigenwert der Matrix A und $\underline{v}$ ein Eigenvektor zum Eigenwert 3.

Es stellen sich mindestens zwei Fragen:

i) Wie *findet* man Eigenwerte und Eigenvektoren für eine gegebene Matrix?

ii) Was *nützen* Eigenwerte und Eigenvektoren?

5.2.2 Bestimmung von Eigenwerten und Eigenvektoren

Die Antwort auf die erste Frage ist nicht besonders schwierig. Kehren wir zum Anfang dieses Abschnittes zurück. Gegeben ist die Matrix

$$A = \begin{bmatrix} 1 & -2 \\ 1 & 4 \end{bmatrix} \tag{24}$$

Einen Eigenwert und einen dazu gehörigen Eigenvektor der Matrix A bestimmen heisst, eine Zahl λ und einen Vektor $\underline{e}$ ("e" um an Eigenvektor zu erinnern)

$$\underline{e} = \begin{bmatrix} \alpha \\ \beta \end{bmatrix} \tag{25}$$

bestimmen, so dass, vergleichen Sie (21) und (23), gilt

$$A\underline{e} = \lambda\underline{e} \tag{26}$$

Verwandeln wir die (Matrix-)Gleichung (26) in ein konventionelles Gleichungssystem. Aus

$$A\underline{e} = \begin{bmatrix} 1 & -2 \\ 1 & 4 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} \alpha - 2\beta \\ \alpha + 4\beta \end{bmatrix} \quad , \quad \lambda\underline{e} = \lambda \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} \lambda\alpha \\ \lambda\beta \end{bmatrix}$$

erhält man an Stelle von (26)

$$\begin{cases} \alpha - 2\beta &= \lambda\alpha \\ \alpha + 4\beta &= \lambda\beta \end{cases} \tag{27}$$

Das sind (nur!) zwei Gleichungen für 3 Unbekannte α, β, λ! Es ist an dieser Stelle vorteilhaft zuerst einmal so zu tun, als würde man λ schon kennen, und als ginge es nur darum α und β zu bestimmen. Nach der üblichen Strategie für das Lösen von Gleichungen ist man dann bestrebt die Unbekannten, also α und β, nur an möglichst wenigen Orten zu haben. Zu diesem Zweck wird man auf beiden Seiten der ersten Gleichung von (27) den Term $\lambda\alpha$ subtrahieren und die beiden α-abhängigen Terme zusammenfassen und analog bei der zweiten Gleichung (27) den Term $\lambda\beta$ subtrahieren und dann die beiden β-abhängigen Terme zusammenfassen. Man findet

$$\begin{cases} \alpha(1-\lambda) - 2\beta &= 0 \\ \alpha + (4-\lambda)\beta &= 0 \end{cases} \tag{28}$$

Nun kann man zum Beispiel die zweite Gleichung nach α auflösen:

$$\alpha = (\lambda - 4)\beta \tag{29}$$

und das Resultat in die erste Gleichung (28) einsetzen:

$$(\lambda - 4)(1 - \lambda)\beta - 2\beta = 0$$

β ausklammern ergibt

$$[(\lambda - 4)(1 - \lambda) - 2]\beta = 0,$$

oder, wenn man schliesslich den Term in der eckigen Klammer vereinfacht

$$[-\lambda^2 + 5\lambda - 6]\beta = 0. \tag{30}$$

Eine Möglichkeit wäre $\beta = 0$ zu setzen; dann folgt aus (29) auch $\alpha = 0$, das heisst der Vektor $\underline{e}$ wäre der sogenannte *Nullvektor*

$$\underline{e} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} = \underline{0}$$

Wenn man diese Wahl weiter verfolgt, stellt sich heraus, dass sie ganz uninteressant ist. Deshalb verabredet man:

$$\text{Der Nullvektor } \underline{0} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \text{ gilt } \textit{nicht} \text{ als Eigenvektor!} \tag{31}$$

Kehren wir daher zur Gleichung (30) zurück. Wenn β nicht gleich 0 sein darf, muss

$$-\lambda^2 + 5\lambda - 6 = 0 \tag{32}$$

gelten. Das ist die sogenannte *Eigenwertgleichung* und das Polynom $P_2(\lambda) = -\lambda^2 + 5\lambda - 6$ heisst *charakteristisches Polynom*. Seine *Nullstellen* liefern die Eigenwerte. In der Tat erhält man aus der quadratischen Gleichung (32)

$$\lambda_{1,2} = \frac{-5 \pm \sqrt{25-24}}{-2} = \frac{-5 \pm 1}{-2},$$

also $\lambda_1 = 2$, $\lambda_2 = 3$. Hinsichtlich der Wahl von β und α folgt zunächst aus (30), dass für $\lambda = \lambda_1$ oder $\lambda = \lambda_2$ die eckige Klammer den Wert 0 hat, man für β also jede beliebige Zahl (ausser 0) wählen kann, zum Beispiel $\beta = 1$ oder $\beta = -1$. Für $\lambda = 2, \beta = 1$ folgt aus (29) $\alpha = (2-4)1 = -2$. Somit finden wir:

$$\underline{e} = \underline{u} = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$$

ist Eigenvektor zum Eigenwert 2, in Übereinstimmung mit den Gleichungen (20), (21). Analog findet man für $\lambda = 3, \beta = -1$, dass

$$\underline{e} = \underline{v} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

Eigenvektor zum Eigenwert 3 ist, ebenfalls in Übereinstimmung mit früheren Betrachtungen.

5.3 2-gliedrige Kette von Massenpunkten im Lichte von Eigenwerten und Eigenvektoren

Betrachtet man beim im Kapitel 4 untersuchten Problem der Gravitationswirkung einer Kette von Massenpunkten nur eine 2-gliedrige Kette, das heisst die Wirkung von nur 2 Massenpunkten der Massen x_0 und x_1, findet man anstelle der Ursache-Wirkungsbeziehung (9) die folgende

$$\left\{ \begin{array}{lcl} y_0 & = & x_0 + \frac{1}{4}x_1 \\ y_1 & = & \frac{1}{4}x_0 + \frac{1}{9}x_1 \end{array} \right. \tag{33}$$

die wir nun mit einer Matrix und Vektoren schreiben können. Es sei

$$A = \begin{bmatrix} 1 & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{9} \end{bmatrix} \quad , \quad \underline{x} = \begin{bmatrix} x_0 \\ x_1 \end{bmatrix} \quad , \quad \underline{y} = \begin{bmatrix} y_0 \\ y_1 \end{bmatrix}. \tag{34}$$

Dann erhält man anstelle von (33) die Darstellung

$$\underline{y} = A\underline{x} \tag{35}$$

Sei etwa

$$\underline{x}^* = \begin{bmatrix} x_0^* \\ x_1^* \end{bmatrix} = \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \end{bmatrix} \tag{36}$$

die "exakte" Massenverteilung. Dann folgt für die "exakte" Wirkung $\underline{y}^* = A\underline{x}^*$

$$\underline{y}^* = \begin{bmatrix} y_0^* \\ y_1^* \end{bmatrix} = \begin{bmatrix} 1 \cdot \frac{1}{2} + \frac{1}{4} \cdot \frac{1}{2} \\ \frac{1}{4} \cdot \frac{1}{2} + \frac{1}{9} \cdot \frac{1}{2} \end{bmatrix} = \begin{bmatrix} \frac{5}{8} \\ \frac{13}{72} \end{bmatrix} \tag{37}$$

5.3.1 Bestimmung von Eigenwerten und Eigenvektoren der Matrix A

Als nächstes werden die Eigenwerte und Eigenvektoren zur Matrix A von Gleichung (34) bestimmt. Das Pendant zur Gleichung (26) lautet:

$$\begin{bmatrix} 1 & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{9} \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \lambda \begin{bmatrix} \alpha \\ \beta \end{bmatrix} \tag{38}$$

oder

$$\begin{cases} \alpha + \frac{1}{4}\beta = \lambda\alpha \\ \frac{1}{4}\alpha + \frac{1}{9}\beta = \lambda\beta \end{cases} \tag{39}$$

Sie erinnern sich an die Strategie? Man tut so, als sei λ schon bekannt, und nur α, β seien zu bestimmen. Dazu formt man (39) wie folgt um:

$$\begin{cases} (1-\lambda)\alpha + \frac{1}{4}\beta = 0 \\ \frac{1}{4}\alpha + (\frac{1}{9} - \lambda)\beta = 0 \end{cases} \tag{40}$$

Auflösen der 2. Gleichung nach α liefert

$$\alpha = (4\lambda - \frac{4}{9})\beta \tag{41}$$

und Einsetzen in die 1. Gleichung (40) und Umformen ergibt

$$[(1-\lambda)(4\lambda - \frac{4}{9}) + \frac{1}{4}]\beta = 0$$

Daraus zieht man den Schluss, dass

$$(1-\lambda)(4\lambda - \frac{4}{9}) + \frac{1}{4} = 0$$

gelten muss (weil $\beta = 0$ auch $\alpha = 0$ nach sich ziehen würde, das heisst es wäre $\underline{e} = \underline{0}$; der Nullvektor ist jedoch nach Vereinbarung kein Eigenvektor, siehe (31)). Umgeformt lautet die letzte Gleichung

$$144\lambda^2 - 160\lambda + 7 = 0 \tag{42}$$

Die Formel für die Lösungen einer quadratischen Gleichung liefert

$$\lambda_{1,2} = \frac{160 \pm \sqrt{25600 - 4032}}{288} = \frac{160 \pm 8\sqrt{377}}{288} = \frac{5}{9} \pm \frac{1}{36}\sqrt{337} \tag{43}$$

Auf einige Dezimalen genau erhält man

$$\lambda_1 \approx 1.06549 \quad \lambda_2 \approx 0.045623 \tag{44}$$

Es fällt auf, dass der zweite Eigenwert recht klein ist. Es wird sich zeigen, dass genau deshalb das Problem der Rekonstruktion von $\underline{x}^*$ aus einer Messung von $\underline{y}^*$ erschwert wird. Für die Eigenvektoren zu λ_1 bzw. λ_2, die wir mit $\underline{e}^{(1)}$, bzw. $\underline{e}^{(2)}$ bezeichnen, finden wir aus (41)

$$\underline{e}^{(1)} = \begin{bmatrix} (4\lambda_1 - \frac{4}{9})\beta_1 \\ \beta_1 \end{bmatrix} \quad , \quad \underline{e}^{(2)} = \begin{bmatrix} (4\lambda_2 - \frac{4}{9})\beta_2 \\ \beta_2 \end{bmatrix} \tag{45}$$

dabei sind β_1, β_2 zwei Zahlen ungleich 0, für deren Wahl wir sogleich einen Vorschlag machen werden.

Erinnern Sie sich an den Begriff des *Skalarprodukts* zweier Vektoren $\underline{a}$ und $\underline{b}$ in der Ebene? Man definiert

$$(\underline{a}, \underline{b}) = |\underline{a}||\underline{b}| \cos \phi,$$

dabei bezeichnen $|\underline{a}|$ und $|\underline{b}|$ die Längen der Vektoren $\underline{a}$ und $\underline{b}$ und ϕ ihren Zwischenwinkel. Wenn $\underline{a}$ und $\underline{b}$ in Komponenten gegeben sind

$$\underline{a} = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} \quad , \quad \underline{b} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$$

dann gilt

$$|\underline{a}| = \sqrt{a_1^2 + a_2^2} \quad , \quad |\underline{b}| = \sqrt{b_1^2 + b_2^2}$$

und

$$(\underline{a}, \underline{b}) = a_1 b_1 + a_2 b_2$$

Berechnet man das Skalarprodukt zwischen $\underline{e}^{(1)}$ und $\underline{e}^{(2)}$ erhält man

$$(\underline{e}^{(1)}, \underline{e}^{(2)}) = [(4\lambda_1 - \frac{4}{9})(4\lambda_2 - \frac{4}{9}) + 1]\beta_1\beta_2 \tag{46}$$

Für die eckige Klammer gilt

$$16\lambda_1\lambda_2 - \frac{16}{9}(\lambda_1 + \lambda_2) + \frac{97}{81}$$

Da λ_1, λ_2 die Wurzeln der quadratischen Gleichung (42) sind, folgt nach dem Satz von Vieta

$$\lambda_1 + \lambda_2 = \frac{160}{144} \quad , \quad \lambda_1\lambda_2 = \frac{7}{144}$$

Demzufolge erhält man für die eckige Klammer in (46)

$$16 \cdot \frac{7}{144} - \frac{16}{9} \cdot \frac{160}{144} + \frac{97}{81},$$

was ausgerechnet 0 ergibt. Das heisst: *Die Eigenvektoren $\underline{e}^{(1)}$ und $\underline{e}^{(2)}$ stehen senkrecht aufeinander.*

Der Vorschlag ist nun, die Zahlen β_1 und β_2 so zu wählen, dass $\underline{e}^{(1)}$ und $\underline{e}^{(2)}$ auch noch *Einheitsvektoren* sind, also die Länge 1 haben. Das wird erreicht wenn man setzt

$$\beta_1 = \frac{1}{\sqrt{1 + (4\lambda_1 - \frac{4}{9})^2}} \quad , \quad \beta_2 = \frac{1}{\sqrt{1 + (4\lambda_2 - \frac{4}{9})^2}} \tag{47}$$

Auf einige Stellen genau erhält man dann für $\underline{e}^{(1)}$, $\underline{e}^{(2)}$

$$\underline{e}^{(1)} \approx \begin{bmatrix} 0.967361 \\ 0.253401 \end{bmatrix} \quad , \quad \underline{e}^{(2)} \approx \begin{bmatrix} -0.253401 \\ 0.967361 \end{bmatrix} \tag{48}$$

Zusammengefasst: Die Matrix A aus (34) hat die Eigenwerte λ_1, λ_2, siehe (43), wobei λ_1 nahe 1 ist, während λ_2 vergleichsweise klein ist, siehe (44). Zu diesen Eigenwerten gehören Eigenvektoren $\underline{e}^{(1)}$ und $\underline{e}^{(2)}$, siehe (45), (47), (48). Für sie gilt:

$$A\underline{e}^{(1)} = \lambda_1\underline{e}^{(1)} \quad , \quad A\underline{e}^{(2)} = \lambda_2\underline{e}^{(2)} \tag{49}$$

Überdies sind $\underline{e}^{(1)}$ und $\underline{e}^{(2)}$ Einheitsvektoren und sie stehen senkrecht aufeinander.

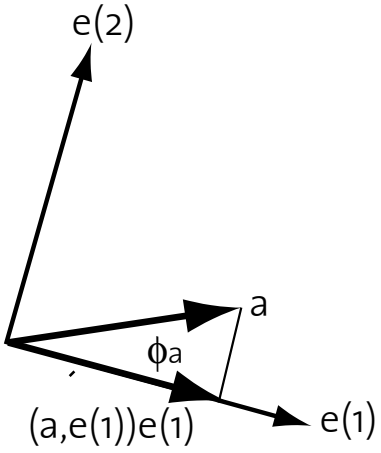

Abbildung 10: Zerlegung eines Vektors.

5.3.2 Darstellung eines Vektors als Linearkombination von $\underline{e}^{(1)}$ und $\underline{e}^{(2)}$

Der Vorteil von in der Ebene zueinander senkrecht stehenden Einheitsvektoren $\underline{e}^{(1)}$, $\underline{e}^{(2)}$ ist, dass *jeder* beliebige Vektor $\underline{a}$ sich in einfacher Weise "als Linearkombination" von $\underline{e}^{(1)}$, $\underline{e}^{(2)}$ darstellen lässt. Es gilt:

$$\underline{a} = (\underline{a}, \underline{e}^{(1)})\underline{e}^{(1)} + (\underline{a}, \underline{e}^{(2)})\underline{e}^{(2)} \tag{50}$$

Diese Gleichung lässt sich geometrisch erklären, siehe Abbildung 10. Wegen $(\underline{a}, \underline{e}^{(1)}) = |\underline{a}||\underline{e}^{(1)}| \cos\phi_a = |\underline{a}| \cos\phi_a$ ist $(\underline{a}, \underline{e}^{(1)})$ gerade die Länge der Projektion von $\underline{a}$ auf $\underline{e}^{(1)}$ (mit Vorzeichen), $(\underline{a}, \underline{e}^{(1)})\underline{e}^{(1)}$ also die Projektion von $\underline{a}$ auf $\underline{e}^{(1)}$, und $(\underline{a}, \underline{e}^{(2)})\underline{e}^{(2)}$ entsprechend die Projektion von $\underline{a}$ auf $\underline{e}^{(2)}$.

5.3.3 Eigenwerttheorie und schlecht gestellte Probleme

Kommen wir nun zur Anwendung auf das schlecht gestelle inverse Problem (35). Nach Konstruktion gilt

$$A\underline{x}^* = \underline{y}^*$$

Nun sei $\underline{\overline{y}} = \begin{bmatrix} \overline{y}_0 \\ \overline{y}_1 \end{bmatrix} = \underline{y}^* + \delta\underline{e}$ eine Messung für $\underline{y}^*$. Dabei sei $\underline{e}$ irgend ein Einheitsvektor. Wie früher ist δ ein Mass für die Messgenauigkeit.

a) Naive Rekonstruktion
Die "naive" Rekonstruktion $\underline{\overline{x}}$ entsteht als Lösung der Gleichung

$$A\underline{\overline{x}} = \underline{\overline{y}} = \underline{y}^* + \delta\underline{e} = A\underline{x}^* + \delta\underline{e}. \tag{51}$$

Diese Gleichung kann man auch in folgender Form schreiben

$$A\underline{\overline{x}} - A\underline{x}^* = \delta\underline{e}$$

und mit der Rechenregel (18)

$$A(\underline{\overline{x}} - \underline{x}^*) = \delta\underline{e} \tag{52}$$

Zur "Lösung" der Gleichung (52) wird nun $\underline{a} := \delta\underline{e}$ mit Hilfe der Formel (50) als Linearkombination der Eigenvektoren $\underline{e}^{(1)}$, $\underline{e}^{(2)}$ der Matrix A dargestellt[4]:

$$\delta\underline{e} = (\delta\underline{e}, \underline{e}^{(1)})\underline{e}^{(1)} + (\delta\underline{e}, \underline{e}^{(2)})\underline{e}^{(2)} = \delta(\underline{e}, \underline{e}^{(1)})\underline{e}^{(1)} + \delta(\underline{e}, \underline{e}^{(2)})\underline{e}^{(2)} \tag{53}$$

Unter Verwendung von (53) erhält man folgende Gleichung für den *Fehler*

$$\overline{\underline{x}} - \underline{x}^*$$

der Rekonstruktion:

$$A(\overline{\underline{x}} - \underline{x}^*) = \delta(\underline{e}, \underline{e}^{(1)})\underline{e}^{(1)} + \delta(\underline{e}, \underline{e}^{(2)})\underline{e}^{(2)} \tag{54}$$

Behauptung: Die Lösung von Gleichung (54) lautet[5]:

$$\boxed{\overline{\underline{x}} - \underline{x}^* = \frac{\delta}{\lambda_1}(\underline{e}, \underline{e}^{(1)})\underline{e}^{(1)} + \frac{\delta}{\lambda_2}(\underline{e}, \underline{e}^{(2)})\underline{e}^{(2)} = \sum_{k=1}^{2} \frac{\delta}{\lambda_k}(\underline{e}, \underline{e}^{(k)})\underline{e}^{(k)}} \tag{55}$$

In der Tat: Nachrechnen unter Verwendung der Formeln (18) und (49) ergibt:

$$A(\overline{\underline{x}} - \underline{x}^*) = \frac{\delta}{\lambda_1}(\underline{e}, \underline{e}^{(1)})A\underline{e}^{(1)} + \frac{\delta}{\lambda_2}(\underline{e}, \underline{e}^{(2)})A\underline{e}^{(2)} = \frac{\delta}{\lambda_1}(\underline{e}, \underline{e}^{(1)})\lambda_1\underline{e}^{(1)} + \frac{\delta}{\lambda_2}(\underline{e}, \underline{e}^{(2)})\lambda_2\underline{e}^{(2)}$$

Der Rekonstruktionfehler (55) besteht aus zwei Termen. Die Skalarprodukte $(\underline{e}, \underline{e}^{(1)})$, $(\underline{e}, \underline{e}^{(2)})$ sind irgend welche Zahlen zwischen -1 und +1. Von entscheidender Bedeutung sind die Faktoren

$$\frac{\delta}{\lambda_1} \quad , \quad \frac{\delta}{\lambda_2}$$

λ_1 hat ungefähr den Wert 1 (siehe (44)), also ist $\frac{\delta}{\lambda_1}$ ungefähr so gross oder so klein wie δ. (Man erinnere sich, dass δ ein Mass für die verwendete Messgenauigkeit, also klein, ist.)

Ein wenig anders ist die Situation bezüglich des Faktors $\frac{\delta}{\lambda_2}$. Da λ_2 recht klein ist (siehe (44)) überwiegt der zweite Summand in der Formel (55) für den Fehler der Rekonstruktion im allgemeinen und zwingt einem die Messgenauigkeit zu erhöhen, also δ zu verkleinern, um ohne andere Massnahmen eine bestimmte Genauigkeit zu erreichen.

b) Tikhonov-Rekonstruktion

[4] Seien $\underline{a}$, $\underline{b}$ zwei Vektoren und r eine Zahl. Dann gilt

$$(r\underline{a}, \underline{b}) = r(\underline{a}, \underline{b}) \, , \, (\underline{a}, r\underline{b}) = r(\underline{a}, \underline{b})$$

[5] Wir benutzen hier und im folgenden ab und zu das sogenannte Summenzeichen. Es bedeutet

$$\sum_{k=1}^{2} a_k = a_1 + a_2,$$

wobei a_1, a_2 zwei Zahlen bedeuten können, oder irgend welche Ausdrücke. Analog gilt:

$$\sum_{k=1}^{3} a_k = a_1 + a_2 + a_3 \quad , \quad \sum_{k=1}^{n} a_k = a_1 + a_2 + ... + a_{n-1} + a_n$$

Eine solche andere Massnahme ist die Anwendung der Regularisierung nach Tikhonov. Sie soll nun in analoger Weise mit Hilfe der Eigenvektoren der Matrix A analysiert werden. Statt gemäss (51) wird die Rekonstruktion $\underline{x}^\alpha$ nun aus folgender Gleichung gewonnen, vergleiche (10):

$$\alpha\underline{x}^\alpha + A\underline{x}^\alpha = A\underline{x}^* + \delta\underline{e} \tag{56}$$

Dabei ist α der Regularisierungsparameter. Man rechnet nach, dass für den Rekonstruktionsfehler $\underline{x}^\alpha - \underline{x}^*$ nun folgende Darstellung gilt

$$\begin{aligned} \underline{x}^\alpha - \underline{x}^* \quad = \quad & -\tfrac{\alpha}{\alpha+\lambda_1}(\underline{x}^*, \underline{e}^{(1)})\underline{e}^{(1)} - \tfrac{\alpha}{\alpha+\lambda_2}(\underline{x}^*, \underline{e}^{(2)})\underline{e}^{(2)} \\ & + \tfrac{\delta}{\alpha+\lambda_1}(\underline{e}, \underline{e}^{(1)})\underline{e}^{(1)} + \tfrac{\delta}{\alpha+\lambda_2}(\underline{e}, \underline{e}^{(2)})\underline{e}^{(2)} \end{aligned} \tag{57}$$

Es ist interessant diese Formel mit (55) zu vergleichen. Für $\alpha = 0$ erhält man aus (57) die Formel (55), wie es sein soll. Im übrigen sieht man, dass die Tikhonov-Regularisierung zwei Effekte hat. Einerseits kommen Terme dazu, die α als Faktor enthalten. Damit sie klein sind, ist man bestrebt α klein zu wählen (ideal wäre $\alpha = 0$!). Andererseits werden die Nenner λ_1 und λ_2 in (55) jetzt durch

$$\alpha + \lambda_1 \quad , \quad \alpha + \lambda_2$$

ersetzt. Um die entsprechenden Terme möglichst klein zu machen, möchte man α möglichst gross wählen (ideal wäre α "unendlich gross" zu machen).

Offenbar besteht die Aufgabe darin, zwischen zwei sich widersprechenden Wünschen einen Kompromiss zu finden: α soll nicht zu gross sein, damit die Terme, die α als Faktor enthalten, nicht einen allzu grossen Fehler erzeugen und umgekehrt soll α auch nicht zu klein gewählt werden, damit der Term mit $\frac{\delta}{\alpha+\lambda_2}$ als Faktor nicht zu gross wird. Die Frage ist, wie dieser Kompromiss konkret aussehen soll.

An dieser Stelle kommt nun noch ein neuer Gesichtspunkt zum Tragen. Bisher haben wir an die zu rekonstruierende Grösse $\underline{x}^*$ keine Bedingungen geknüpft. Das muss aber nicht so sein. Wir werden im folgenden $\underline{x}^*$ wie folgt einschränken.

Voraussetzung 1 *Es gibt einen Vektor $\underline{z}^*$ der Länge höchstens 1 (also $|\underline{z}^*| \leq 1$), so dass gilt*

$$\underline{x}^* = A\underline{z}^* \tag{58}$$

Wie Sie erfahren werden, kann man unter der Voraussetzung (58) eine, wie man sagt, *a priori* –Wahl, für den Regularisierungsparameter α treffen, die den Rekonstruktionsfehler in gewisser Weise minimiert.

c) Zwischenrechnung
Für unsere Überlegungen benötigen wir folgende Formel

$$(A\underline{z}^*, \underline{e}^{(i)}) = (\underline{z}^*, A\underline{e}^{(i)}) = (\underline{z}^*, \lambda_i\underline{e}^{(i)}) = \lambda_i(\underline{z}^*, \underline{e}^{(i)}) \tag{59}$$

Dabei steht i für 1 oder 2. $\underline{e}^{(1)}$, $\underline{e}^{(2)}$ sind wie zuvor die beiden Eigenvektoren der Matrix A und λ_1, λ_2 die zugehörigen Eigenwerte.
Der Beweis der Formel (59) erfolgt durch Nachrechnen. Die erste Umformung in der Kette (59) beruht darauf, dass A eine sogenannte "symmetrische" Matrix ist, das heisst eine Matrix der speziellen Gestalt

$$A = \begin{bmatrix} a & b \\ b & d \end{bmatrix}$$

Mit $\underline{z}^* = \begin{bmatrix} z_1^* \\ z_2^* \end{bmatrix}$, $\underline{f} = \begin{bmatrix} f_1 \\ f_2 \end{bmatrix}$ folgt:

$$(A\underline{z}^*, \underline{f}) = (\begin{bmatrix} az_1^* + bz_2^* \\ bz_1^* + dz_2^* \end{bmatrix}, \begin{bmatrix} f_1 \\ f_2 \end{bmatrix}) = az_1^* f_1 + bz_2^* f_1 + bz_1^* f_2 + dz_2^* f_2$$

Andererseits ist

$$(\underline{z}^*, A\underline{f}) = (\begin{bmatrix} z_1^* \\ z_2^* \end{bmatrix}, \begin{bmatrix} af_1 + bf_2 \\ bf_1 + df_2 \end{bmatrix}) = z_1^* af_1 + z_1^* bf_2 + z_2^* bf_1 + z_2^* df_2$$

Die weiteren Umformungen in (59) folgen aus den Gleichungen (49), sowie aus einer Eigenschaft des Skalarprodukts, siehe Fussnote 4.

d) Darstellung des Rekonstruktionsfehlers unter Voraussetzung 1
Unter der Voraussetzung (58) für $\underline{x}^*$ und mit (59) lässt sich (57) wie folgt schreiben:

$$\begin{aligned} \overline{\underline{x}}^{\alpha} - \underline{x}^* &= -\tfrac{\alpha\lambda_1}{\alpha+\lambda_1}(\underline{z}^*, \underline{e}^{(1)})\underline{e}^{(1)} - \tfrac{\alpha\lambda_2}{\alpha+\lambda_2}(\underline{z}^*, \underline{e}^{(2)})\underline{e}^{(2)} \\ &\quad + \tfrac{\delta}{\alpha+\lambda_1}(\underline{e}, \underline{e}^{(1)})\underline{e}^{(1)} + \tfrac{\delta}{\alpha+\lambda_2}(\underline{e}, \underline{e}^{(2)})\underline{e}^{(2)} \end{aligned} \tag{60}$$

Die Faktoren

$$\frac{\alpha\lambda_1}{\alpha+\lambda_1} \quad , \quad \frac{\delta}{\alpha+\lambda_1} \quad , \quad \frac{\alpha\lambda_2}{\alpha+\lambda_2} \quad , \quad \frac{\delta}{\alpha+\lambda_2} \tag{61}$$

sind alle positiv (λ_1, λ_2 sind positiv, siehe (44), δ, α sind positiv nach Voraussetzung). Es ist wichtig, dass sie möglichst klein sind. Weil

$$\lambda_1 < \alpha + \lambda_1 \quad , \quad \alpha < \alpha + \lambda_1 \quad , \quad \lambda_2 < \alpha + \lambda_2 \quad , \quad \alpha < \alpha + \lambda_2$$

gilt, sind die Faktoren jedenfalls nicht grösser als

$$\alpha \quad , \quad \frac{\delta}{\alpha} \quad , \quad \alpha \quad , \quad \frac{\delta}{\alpha} \tag{62}$$

Nun kann man gut ablesen, wie der Kompromiss für α aussehen soll: Damit weder α grösser als $\frac{\delta}{\alpha}$ ist, noch $\frac{\delta}{\alpha}$ grösser als α, verlangt man:

$$\alpha = \frac{\delta}{\alpha}$$

und das heisst

$$\boxed{\alpha = \sqrt{\delta}} \tag{63}$$

Unter Verwendung der Wahl (63) für den Regularisierungsparameter erhält man die Rekonstruktion

$$\overline{\underline{x}}^{\sqrt{\delta}} = \begin{bmatrix} \overline{x}_0^{\sqrt{\delta}} \\ \overline{x}_1^{\sqrt{\delta}} \end{bmatrix}$$

aus einer Messung

$$\overline{\underline{y}} = \begin{bmatrix} \overline{y}_0 \\ \overline{y}_1 \end{bmatrix} = \underline{y}^* + \delta\underline{e} = \begin{bmatrix} y_0^* + \delta e_0 \\ y_1^* + \delta e_1 \end{bmatrix}$$

von $\underline{y}^* = A\underline{x}^*$ durch Lösen des Gleichungssystems

$$\boxed{\begin{cases} (\sqrt{\delta}+1)\overline{x}_0^{\sqrt{\delta}} + \frac{1}{4}\overline{x}_1^{\sqrt{\delta}} &= \overline{y}_0 \\ \frac{1}{4}\overline{x}_0^{\sqrt{\delta}} + (\sqrt{\delta}+\frac{1}{9})\overline{x}_1^{\sqrt{\delta}} &= \overline{y}_1 \end{cases}}$$

Man braucht dafür weder Eigenwerte noch Eigenvektoren der zu Grunde liegenden Matrix A zu kennen.

Den Rekonstruktionsfehler, (60), den man dabei begeht, kann man wie folgt schreiben:

$$\boxed{\begin{aligned} \overline{\underline{x}}^{\sqrt{\delta}} - \underline{x}^* &= \sqrt{\delta}[-\tfrac{\lambda_1}{\sqrt{\delta}+\lambda_1}(\underline{z}^*, \underline{e}^{(1)})\underline{e}^{(1)} - \tfrac{\lambda_2}{\sqrt{\delta}+\lambda_2}(\underline{z}^*, \underline{e}^{(2)})\underline{e}^{(2)} \\ &\quad + \tfrac{\sqrt{\delta}}{\sqrt{\delta}+\lambda_1}(\underline{e}, \underline{e}^{(1)})\underline{e}^{(1)} + \tfrac{\sqrt{\delta}}{\sqrt{\delta}+\lambda_2}(\underline{e}, \underline{e}^{(2)})\underline{e}^{(2)}] \\ &= \sqrt{\delta}\textstyle\sum_{k=1}^{2}[-\tfrac{\lambda_k}{\sqrt{\delta}+\lambda_k}(\underline{z}^*, \underline{e}^{(k)})\underline{e}^{(k)} + \tfrac{\sqrt{\delta}}{\sqrt{\delta}+\lambda_k}(\underline{e}, \underline{e}^{(k)})\underline{e}^{(k)}] \end{aligned}} \tag{64}$$

5.4 Die 3-, 5-, n-gliedrige Kette von Massenpunkten im Lichte von Eigenwerten

Für die 3-gliedrige Kette von Massenpunkten lautet die Ursache-Wirkungs-Beziehung in Matrixschreibweise

$$\underline{y} = A\underline{x},$$

wobei A jetzt eine 3×3-Matrix ist, nämlich, vergleiche (9) in Abschnitt 4.1,

$$A = \begin{pmatrix} 1 & \frac{4}{9} & \frac{1}{4} \\ \frac{4}{9} & \frac{1}{4} & \frac{4}{25} \\ \frac{1}{4} & \frac{4}{25} & \frac{1}{9} \end{pmatrix} \tag{65}$$

Entsprechend sind $\underline{x}$, $\underline{y}$ 3-Vektoren, also Kolonnen aus 3 Zahlen. Ganz ähnlich wie für die 2×2-Matrix A aus Gleichung (34) kann man Eigenwerte und Eigenvektoren für die Matrix (65) bestimmen. Die Eigenwerte ergeben sich als Nullstellen aus einem Polynom dritten Grades

$$P_3(\lambda) := \lambda^3 - \frac{49}{36}\lambda^2 + \frac{83639}{810000}\lambda - \frac{4679}{29160000} = 0$$

Man erhält näherungsweise folgende Eigenwerte

$$\lambda_1 \approx 1.28057 \quad , \quad \lambda_2 \approx 0.0789491 \quad , \quad \lambda_3 \approx 0.00158713 \tag{66}$$

Es ist nicht schwer nun (Näherungen für) die zugehörigen Eigenvektoren zu bestimmen. Es stellt sich wieder heraus, dass sie paarweise senkrecht aufeinander stehen, das heisst, das Skalarprodukt zwischen je zweien von ihnen ist gleich 0. Schliesslich kann man sie so wählen, dass sie Einheitsvektoren sind, das heisst, dass alle die Länge 1 haben.

Für die 5-gliedrige Kette ist A eine 5×5-Matrix, nämlich

$$A = \begin{pmatrix} 1 & \frac{16}{25} & \frac{4}{9} & \frac{16}{49} & \frac{1}{4} \\ \frac{16}{25} & \frac{4}{9} & \frac{16}{49} & \frac{1}{4} & \frac{16}{81} \\ \frac{4}{9} & \frac{16}{49} & \frac{1}{4} & \frac{16}{81} & \frac{4}{25} \\ \frac{16}{49} & \frac{1}{4} & \frac{16}{81} & \frac{4}{25} & \frac{16}{121} \\ \frac{1}{4} & \frac{16}{81} & \frac{4}{25} & \frac{16}{121} & \frac{1}{9} \end{pmatrix} \tag{67}$$

Aus einem Polynom vom Grad 5 findet man (näherungsweise) folgende Eigenwerte für A

$$\lambda_1 \approx 1.84883,\ \lambda_2 \approx 0.112992,\ \lambda_3 \approx 0.0036644,\ \lambda_4 \approx 6.38237 \cdot 10^{-5},\ \lambda_5 \approx 4.78995 \cdot 10^{-7} \quad (68)$$

Vergleicht man die Eigenwerte für die drei betrachteten Fälle (2-, 3-, 5-gliedrige Kette) stellt man folgendes fest: Die Eigenwerte sind positiv. Der grösste, λ_1, hat etwa den Wert 1. Ordnet man sie nach absteigender Grösse

$$\lambda_1 > \lambda_2 > \lambda_3 > \lambda_4 > ...$$

werden sie sehr schnell kleiner und nähern sich rasch der Zahl 0. Zur weiteren Untermauerung dieser Beobachtung seien noch Näherungen zu den Eigenwerten für die 10-gliedrige Kette mitgeteilt:

$$\begin{array}{lllll} \lambda_1 \approx 3.4 & \lambda_2 \approx 0.18 & \lambda_3 \approx 6.2 \cdot 10^{-3} & \lambda_4 \approx 1.6 \cdot 10^{-4} & \lambda_5 \approx 3.4 \cdot 10^{-6} \\ \lambda_6 \approx 5.6 \cdot 10^{-8} & \lambda_7 \approx 7.2 \cdot 10^{-10} & \lambda_8 \approx 6.6 \cdot 10^{-12} & \lambda_9 \approx 3.9 \cdot 10^{-14} & \lambda_{10} \approx 1.1 \cdot 10^{-16} \end{array} \quad (69)$$

Wenden wir uns jetzt den Rekonstruktionsfehlern zu. Die Formeln (54) und (64) sind leicht zu verallgemeinern. Für die *"naive" Rekonstruktion*, das heisst *Rekonstruktion ohne Regularisierung* gilt für den Rekonstruktionsfehler

$$\boxed{\overline{\underline{x}} - \underline{x}^* = \sum_{k=1}^{n} \frac{\delta}{\lambda_k} (\underline{e}, \underline{e}^{(k)}) \underline{e}^{(k)}} \quad (70)$$

Dabei ist $n = 2$ für die 2-gliedrige, $n = 3$ für die 3-gliedrige, $n = 5$ für die 5-gliedrige, $n = 10$ für die 10-gliedrige Kette, usw. Weiter sind $\underline{e}^{(1)}$, $\underline{e}^{(2)}$, $\underline{e}^{(3)}$, ... die (Einheits-)Eigenvektoren zu den jeweiligen Eigenwerten λ_1, λ_2, λ_2, ... $\underline{e}$ ist ein beliebiger Einheitsvektor und $\delta\underline{e}$ der Messfehler.
Wird gemäss *Tikhonov regularisiert* mit der Wahl $\alpha = \sqrt{\delta}$ für den Regularisierungsparamer, gilt unter *Voraussetzung* 1, siehe Gleichung (58), für den Rekonstruktionsfehler folgende Darstellung, vergleiche (64)

$$\boxed{\overline{\underline{x}}^{\sqrt{\delta}} - \underline{x}^* \ = \ \sqrt{\delta} \sum_{k=1}^{n} [-\tfrac{\lambda_k}{\sqrt{\delta}+\lambda_k} (\underline{z}^*, \underline{e}^{(k)}) \underline{e}^{(k)} + \tfrac{\sqrt{\delta}}{\sqrt{\delta}+\lambda_k} (\underline{e}, \underline{e}^{(k)}) \underline{e}^{(k)}]} \quad (71)$$

Um die Formeln (70) und (71) zu vergleichen, betrachten wir ein Beispiel bei dem die Schlechtgestelltheit (ill-posedness) des Problemsbesonders deutlich in Erscheinung tritt. Wir setzen

$$\underline{z}^* = \underline{e}^{(1)} \quad , \quad \underline{e} = \underline{e}^{(n)} \quad (72)$$

Da die Vektoren $\underline{e}^{(1)}$, $\underline{e}^{(2)}$, ... paarweise senkrecht stehen und alle die Länge 1 haben, reduzieren sich die Formeln (70) und (71) unter der Voraussetzung (72) auf

$$\overline{\underline{x}} - \underline{x}^* = \frac{\delta}{\lambda_n} \underline{e}^{(n)} \quad (73)$$

respektive

$$\overline{\underline{x}}^{\sqrt{\delta}} - \underline{x}^* = \sqrt{\delta} [-\frac{\lambda_1}{\sqrt{\delta}+\lambda_1} \underline{e}^{(1)} + \frac{\sqrt{\delta}}{\sqrt{\delta}+\lambda_n} \underline{e}^{(n)}] \quad (74)$$

Da $\underline{e}^{(n)}$ ein Vektor der Länge 1 ist, folgt für die Länge des Fehlervektors $\overline{\underline{x}} - \underline{x}^*$ aus (73)

$$|\overline{\underline{x}} - \underline{x}^*| = \frac{\delta}{\lambda_n} \tag{75}$$

Für die Länge des Fehlervektors $\overline{\underline{x}}^{\sqrt{\delta}} - \underline{x}^*$ erhält man aus (74) nach dem Satz von Pythagoras[6]

$$|\overline{\underline{x}}^{\sqrt{\delta}} - \underline{x}^*| = \sqrt{\delta}\sqrt{(\frac{\lambda_1}{\sqrt{\delta}+\lambda_1})^2 + (\frac{\sqrt{\delta}}{\sqrt{\delta}+\lambda_n})^2} \leq \sqrt{2\delta} \tag{76}$$

Für die Abschätzung in (76) wurde benutzt, dass

$$0 < \frac{\lambda_1}{\sqrt{\delta}+\lambda_1} < 1 \quad , \quad 0 < \frac{\sqrt{\delta}}{\sqrt{\delta}+\lambda_n} < 1$$

gilt, was stimmt, weil λ_1, λ_n, $\sqrt{\delta}$ positive Zahlen und die Nenner jeweils grösser als die Zähler sind.

Nehmen wir etwa die *5-gliedrige Kette*. Dann ist $n = 5$ und $\lambda_5 \approx 4.78995 \cdot 10^{-7}$ und (75) und (76) ergeben

$$|\overline{\underline{x}} - \underline{x}^*| \approx 2.1 \cdot 10^6 \cdot \delta \quad , \quad |\overline{\underline{x}}^{\sqrt{\delta}} - \underline{x}^*| < 1.42 \cdot \sqrt{\delta} \tag{77}$$

Daraus folgt, dass das Resultat der Tikhonov-Regularisierung für diejenigen δ genauer ist, für die

$$\sqrt{\delta} > \frac{1.42}{2.1 \cdot 10^6} \approx 0.68 \cdot 10^{-6}$$

gilt, also für

$$\delta > (0.68 \cdot 10^{-6})^2 \approx 0.5 \cdot 10^{-12}$$

Für ein konkretes Zahlenbeispiel braucht man die Einheitseigenvektoren $\underline{e}^{(1)}$, $\underline{e}^{(5)}$. Die Rechnung ergibt[7]

$$\underline{e}^{(1)} \approx (0.719915, 0.488304, 0.354387, 0.269544, 0.212219)$$

$$\underline{e}^{(5)} \approx (0.0206138, -0.190956, 0.584043, -0.723738, 0.313383)$$

Aus $\underline{x}^* = A\underline{z}^* = A\underline{e}^{(1)} = \lambda_1 \underline{e}^{(1)}$ folgt

$$\underline{x}^* \approx (1.331, 0.902794, 0.655204, 0.498342, 0.392358)$$

und daraus mit $\underline{y}^* = A\underline{x}^*$

$$\underline{y}^* \approx (2.46081, 1.66912, 1.21136, 0.921352, 0.725405)$$

Wählen wir

$$\delta = 0.0025$$

dann wird die Messung durch $\overline{\underline{y}} = \underline{y}^* + \delta\underline{e} = \underline{y}^* + \delta\underline{e}^{(5)}$ simuliert, das heisst durch

$$\overline{\underline{y}} \approx (2.46086, 1.66864, 1.21282, 0.919542, 0.726189)$$

[6]Die rechte Seite von (74) kann man interpretieren als Hypothenusenvektor eines rechtwinkligen Dreiecks dessen Katheten die Längen $\sqrt{\delta}\frac{\lambda_1}{\sqrt{\delta}+\lambda_1}$ und $\sqrt{\delta}\frac{\sqrt{\delta}}{\sqrt{\delta}+\lambda_n}$ haben.

[7]Um Platz zu sparen werden die Vektoren als Zeilen geschrieben.

Ohne Regularisierung ist der Rekonstruktionsfehler gemäss (77)

$$|\underline{\overline{x}} - \underline{x}^*| \approx 2.1 \cdot 10^6 \cdot 0.0025 = 5250$$

In Anbetracht der Tatsache, dass die Komponenten von $\underline{x}^*$ um 1 herum sind, ist klar, dass die "naive" Rekonstruktion ein völlig unbrauchbares Resultat liefert. Mithilfe von Gleichung (73) lässt sich $\underline{\overline{x}}$ bestimmen. Man findet

$$\underline{\overline{x}} \approx (108.92, -995.744, 3048.93, -3776.87, 1636.02),$$

eine in der Tat total unbrauchbare Näherung für $\underline{x}^*$! Was liefert stattdessen die Tikhonov-regularisierte Rekonstruktion? Der Rekunstruktionsfehler ist gemäss (77)

$$|\underline{\overline{x}}^{\sqrt{\delta}} - \underline{x}^*| < 1.42 \cdot \sqrt{\delta} \approx 1.42\sqrt{0.0025} \approx 0.071$$

Die regularisierte Rekonstruktion hat offenbar eine ganz andere Qualität! Mithilfe von Gleichung (74) kann man $\underline{\overline{x}}^{\sqrt{\delta}}$ berechnen. Man findet

$$\underline{\overline{x}}^{\sqrt{\delta}} \approx (1.29699, 0.869474, 0.667153, 0.449033, 0.397695)$$

Der direkte Vergleich von $\underline{\overline{x}}^{\sqrt{\delta}}$ mit

$$\underline{x}^* \approx (1.331, 0.902794, 0.655204, 0.498342, 0.392358)$$

zeigt, dass $\underline{\overline{x}}^{\sqrt{\delta}}$ tatsächlich eine vernüftige Näherung für $\underline{x}^*$ ist.

6 Weitere Lektüre

Für weitere Lektüre verweisen wir auf die Literaturliste. In [5] (und [6]) geht es um das Lösen grosser linearer Gleichungssysteme und einige Anwendungen auf Niveau Sekundarstufe II, insbesondere um eine Einführung in die Computertomografie (wobei allerdings auf den Aspekt der Schlecht-Gestelltheit nicht eingegangen wird, was daran liegt, dass die CT ein nur moderat schlecht gestelltes Problem ist). [2] ist eine Kaleidoskop (schlecht gestellter) inverser Probleme, beginnend mit ganz einfachen Aufgaben. [3] ist der Schwesterartikel zum hier vorliegenden, auf etwas höherem Niveau. [8] befasst sich mit historischen Aspekten der Computertomografie. [7] ist eine Referenz an den Erfinder der Regularisierung. Die Monografien [1], [4] stellen das Thema auf mathematischem Niveau unter Verwendung funktionalanalytischer Instrumente dar. Insbesondere [4] stellt die Verbindung zum Forschungsstand her.

Literatur

[1] Groetsch, C. W.: Inverse Problems in the Mathematical Sciences, Vieweg 1993.

[2] Groetsch, C. W.: Inverse Problems - Activities for Undergratuates, The Mathematical Association of America 1999.

[3] Kirchgraber, U., Kirsch, A., Stoffer, D.: Schlecht gestellte Probleme - Oder wenn das Ungenaue genauer ist, Mathematische Semesterberichte 51, 2004, p. 175 - 206.

[4] Kirsch, A.: An Introduction to the Mathematical Theory of Inverse Problems, Springer 1996.

[5] Reichel, H. C., Zöchling, J.: Tausend Gleichungen - und was nun? - Computertomographie als Einstieg in ein aktuelles Thema des Mathematikunterrichts, Didaktik der Mathematik 4, 1990, p. 245-270.

[6] Reichel, H. C., Zöchling, J.: Iteratives Lösen grösserer Linearer Gleichungssysteme, MNU 47, 1994, p. 20-25.

[7] Thikonov, A. N.: On the solution of ill-posed problems and the regularization method, Dokl. Akad. Nauk SSSR 151, 1963, p. 501-504.

[8] Ullrich, P.: Johann Radon und die Grundlagen der Computertomografie, Mathematische Semesterberichte 50, 2003, p. 143 - 166.

UK+DS/15. März 2006 und 25. Mai 2006