

Marlis Hochbruck · Jörg-M. Sautter

Mathematik fürs Leben am Beispiel der Computertomographie

Eingegangen am ???.???.2001 / Angenommen am ???.???.2001

Zusammenfassung Computertomographie ist heutzutage ein fast ebenso bekanntes Hilfsmittel des Mediziners wie das klassische Röntgen. Weit weniger bekannt ist jedoch, dass es insbesondere die Mathematik ist, die dieses Verfahren erst ermöglicht. Es erscheint immer wichtiger, Schülern und Studierenden neben der Mathematik selbst auch deren Bedeutung für Probleme aus dem Alltag nahe zu bringen. In dieser Arbeit vermitteln wir ein Grundverständnis der Computertomographie. Hierzu gehören Nutzen und Einsatzmöglichkeiten in der Medizin, physikalischer Hintergrund und als zentrales Thema ein einfaches mathematisches Modell und dessen numerische Lösung. Die Arbeit kann als Grundlage für Projekte in Schule und Hochschule verwendet werden.

1. Einleitung.

Von der Bedeutung der Mathematik für alltägliche Probleme haben große Teile der Bevölkerung heute leider nur sehr vage oder sogar völlig falsche Vorstellungen. Viele technische und naturwissenschaftliche Innovationen sind jedoch ohne Mathematik undenkbar.

Am Beispiel der Computertomographie wollen wir die Rolle der Mathematik und die einzelnen Schritte eines Mathematikers auf dem Weg zur Lösung eines Problems aus dem "Leben" exemplarisch darstellen. Offensichtlich ist für solche Probleme interdisziplinäres Arbeiten unabdingbar. Ohne ein Verständnis dessen, was der Mediziner von einer Computertomographie-Rekonstruktion erwartet, ist das Aufstellen eines mathematischen Modells und dessen Lösung nicht sinnvoll; ohne Kenntnisse der physikalischen Prozesse nicht möglich. Daher ist diese Arbeit so aufgebaut, dass zunächst das medizinische Problem und dann die zu Grunde liegende Physik vorgestellt werden. Der nächste Schritt ist das Aufstellen eines mathematischen Modells, welches zu einem System von Integralgleichungen (der sogenannten Radon-Transformation) führt. Das hier vorgestellte vereinfachte Modell, welches auf einer finite-Element Diskretisierung basiert, macht aus dem unendlich dimensional Problem ein endlich dimensionales Problem, nämlich ein überbestimmtes lineares Gleichungssystem sehr großer Dimension. Aus Kostengründen (Rechenzeit und Speicheraufwand) kann das zugehörige Ausgleichsproblem nur iterativ gelöst werden. Wir schlagen hierzu eine Variante des Verfahrens der konjugierten Gradienten vor, andere Projektionsverfahren wie das Kaczmarz-Verfahren [9] wären aber ebenfalls denkbar. Das Kaczmarz-Verfahren wurde in [14] an Hand kleiner Beispiele mit spezieller Struktur (die Matrixelemente sind entweder 0 oder 1) auf Schulniveau präsentiert. Die dort verwendete Struktur der Gleichungen resultiert aus einem noch einfacheren als dem hier vorgestellten mathematischen Modell, vgl. [13].

Marlis Hochbruck, Jörg-M. Sautter: Mathematisches Institut, Universität Düsseldorf, Universitätsstr. 1, 40225 Düsseldorf, Germany, e-mail: marlis@am.uni-duesseldorf.de, sautter@am.uni-duesseldorf.de, <http://www.am.uni-duesseldorf.de>

Schlüsselwörter: Computertomographie, Iterationsverfahren, mathematische Modellierung, große lineare Ausgleichsprobleme, Projekte für Schulen und Hochschulen

Es soll hier jedoch nicht verschwiegen werden, dass die tatsächliche Berechnung von Computertomographie-Bildern nicht auf dem von uns beschriebenen Modell sondern auf der Radon-Transformation basiert. Dieser Ansatz verlangt weit tiefergehende Kenntnisse der Mathematik und der numerischen Lösung. Wir haben uns daher für einen einfacheren Zugang entschieden, der es auch ohne vorherige Spezialvorlesungen ermöglicht, dieses Problem als Ganzes zu behandeln und zudem zu akzeptablen Näherungslösungen führt.

Ziel dieser Arbeit ist es, Dozenten oder Lehrern Anregungen für ein anwendungsorientiertes und damit interdisziplinäres Projekt zu vermitteln. Für den Projektleiter setzen wir Grundkenntnisse der numerischen Mathematik, insbesondere der iterativen Lösung linearer Gleichungssysteme, wie man sie in jedem einführenden Numerik-Buch findet, als bekannt voraus. Ausführlicheres Material findet man auf unserer Internetseite.

Die Arbeit ist wie folgt gegliedert: In Abschnitt 2 stellen wir zunächst die Bedeutung der Computertomographie für die Medizin dar und geben Zusammenhänge zu anderen diagnostischen Verfahren wie Röntgen und MRT. Anschließend widmen wir uns der Physik hinter der Computertomographie, um daraus in Abschnitt 4 das mathematische Modell herzuleiten. Die Diskretisierung dieses Modells ist in Abschnitt 5, die numerischen Methoden sind in Abschnitt 6 beschrieben. Ergebnisse der Berechnungen sind in Abschnitt 7 zusammengefasst und Hinweise für die Umsetzung dieser Arbeit als Projekt für Schüler findet man in Abschnitt 8.

2. Medizin

Die Computertomographie (CT) ermöglicht es, Querschnitte eines lebenden menschlichen Körpers im Bild darzustellen. Anders als bei einer herkömmlichen Röntgenuntersuchung, bei der man die Projektion des Objekts auf eine Ebene betrachtet, werden bei der Computertomographie die Röntgenstrahlen in der Querschnittsebene aus vielen verschiedenen Richtungen durch den Körper geschickt (vgl. Abb. 4). Dazu wird der Patient liegend durch die kreisrunde sogenannte Gantry-Öffnung des Computertomographen entweder schrittweise oder kontinuierlich (Spiral tomographie) bewegt. Die Daten für die Rekonstruktion liefern um die Längsachse des Patienten rotierende Röntgenquellen und Detektoren. Aus der Abschwächung der Strahlung beim Durchlaufen der Materie können dann mittels effizienter mathematischer Verfahren die Abschwächungskoeffizienten in der Querschnittsebene rekonstruiert werden. Im Gegensatz zur herkömmlichen Röntgenuntersuchung ermöglicht die Computertomographie somit einen Blick ins Innere des lebenden Körpers (vgl. Abb. 1, 2). Es lassen sich dann z.B. die örtliche Lage, Form und Größe von Organen bestimmen sowie gesundes von krankem Gewebe unterscheiden.

In der zerstörungsfreien Materialuntersuchung findet die Computertomographie ebenfalls Anwendung. Dort werden neben Röntgenstrahlen auch γ -Strahlen verwendet.

Neben der Computertomographie existieren noch andere Verfahren, um einen Blick ins Innere eines lebenden Patienten zu werfen, wie z.B. die Kernspintomographie (NMR, Nuclear Magnetic Resonance, oder auch MRT, Magnetresonanztomographie, genannt), die ohne schädliche Röntgenstrahlung auskommt und eine differenziertere Weichteildarstellung ermöglicht. Hingegen liegen die Vorteile der Computertomographie in der höheren Ortsauflösung, der besseren Unterscheidung zwischen dichter und weniger dichter Materie (Knochen vs. Gewebe und Flüssigkeiten, wichtig z.B. bei Haarrissen), der wesentlich kürzeren Dauer für eine Untersuchung sowie den geringeren Kosten.



Abbildung 1. Computertomographie-Schnittbild.



Abbildung 2. Computertomographie-Schnittbild.

Godfrey Newbold Hounsfield beschrieb⁴ 1972 das Prinzip der Computertomographie als erster (vgl. [3,8]) und erhielt dafür 1979 den Nobelpreis der Medizin. 1987 gelang es Karl-Heinz Höhne, dreidimensionale Rekonstruktionen aus zweidimensionalen CT und NMR Schnittbildern zu generieren. In den letzten Jahren sind auch sog. Spiralcomputertomographen entwickelt worden, um dreidimensionale Rekonstruktionen noch schneller und mit geringerer Strahlenbelastung durchführen zu können.

Nach Entdeckung der Röntgenstrahlung dauerte es noch mehr als 10 Jahre, bis die Schädlichkeit dieser Strahlung für Organismen erkannt wurde. Dementsprechend ist es nicht verwunderlich, dass die erste veröffentlichte, mit Röntgenstrahlen gemachte Aufnahme von einem Menschen (die Hand von W. C. Röntgens Frau, vgl. [2]) 20 Minuten lang belichtet wurde. Dagegen werden Patienten bei modernen

⁴ zusammen mit J. Ambrose vom Atkinson Morley Hospital in London

Röntgenuntersuchungen nur Bruchteile einer Sekunde lang der Strahlung ausgesetzt und eine CT-Untersuchung des gesamten Oberkörpers dauert nur wenige Sekunden.

3. Physik

Die im Jahre 1895 von Wilhelm Conrad Röntgen⁵ entdeckte und nach ihm benannte Röntgenstrahlung wird im Allgemeinen dadurch erzeugt, dass man die aus Metallen mittlerer oder hoher Massenzahl (Mo, Cu bzw. W) bestehende Anode, häufig auch als Antikathode bezeichnet, mit elektronenoptisch (z.B. durch einen Wehnelt-Zylinder) gebündelten schnellen Elektronen aus einer Glühkathode beschießt. Die Energie der Elektronen wird dort in Strahlung verwandelt, größtenteils jedoch aber in Wärme, so dass die Antikathode meist gekühlt werden muss. Die Anodenspannung beträgt typischerweise zwischen 50 und 300 kV (vgl. [4]). Es sei jedoch darauf hingewiesen, dass Röntgenstrahlung auch durch Elektronen- oder Stoßanregung freier Atome und damit frei von Festkörper-Einflüssen erzeugt werden kann. Der Nachweis erfolgt durch Photoplatte, Film, Zählrohr oder in neuer Zeit durch Halbleiter-Detektoren (vgl. [5]).

Historisch ist zu bemerken, dass es noch fast 20 Jahre nach Röntgens Entdeckung zweifelhaft war, ob es sich um Wellen oder (klassische) Teilchen handelt. Tatsächlich kann Röntgenstrahlung sowohl als Welle als auch als Quantenteilchen aufgefasst werden (Welle-Teilchen-Dualismus) und es waren Röntgenspektren, die zur Theorie des Schalenaufbaus der Atome geführt haben.

Für die Anwendung von Röntgenstrahlung in der Medizin (vgl. Abschnitt 2) ist insbesondere die Schwächung beim Durchgang durch Materie relevant. Die wichtigsten Schwächungsmechanismen sind Photoeffekt, Streuung (Compton-Effekt) und Paarbildung (vgl. [4,5]). Der Versuchsaufbau zur Messung der Absorption von Strahlung in Materie ist in Abb. 3 dargestellt (vgl. [1,5]). Nehmen wir an, der einfallende Strahl hätte die Intensität $I(0)$ und durchlaufe ein homogenes Medium der Dicke d , dann gilt für die Intensität $I(d)$ des austretenden Strahles

$$I(d) = I(0) 2^{-d/d_{1/2}}. \quad (1)$$

$d_{1/2}$ ist dabei die Halbwertsdicke. Aus (1) und dem linearen Schwächungskoeffizienten $\mu = (\ln 2)/d_{1/2}$ folgt das Lambert-Beer-Gesetz (vgl. [5])

$$I(d) = I(0) \exp(-\mu d).$$

Auch die Abschwächung von γ -Strahlung in Materie (vgl. [1]), die Schwächung eines Lichtstrahls beim Durchgang durch ein absorbierendes Medium, aber auch die Abschwächung eines Elektronenbündels beim Durchgang durch Materie (vgl. [5]) erfolgen nach diesem Gesetz.

Die Bezeichnung "linearer" Schwächungskoeffizient kommt von

$$\mu = \mu_{\text{Photoeffekt}} + \mu_{\text{Paarbildung}} + \mu_{\text{Streuung}}.$$

Des Weiteren hängt μ vom Material (genauer gesagt besteht eine Abhängigkeit von der Kernladungszahl Z , vgl. [5]) und auch von der Energie E der Röntgenquanten (also der Frequenz der verwendeten Röntgenstrahlung) ab, d.h. $\mu = \mu(Z, E)$. Die Theorie der Röntgenabsorptionsspektren ist jedoch nicht Teil dieser Arbeit. Für die Computertomographie wird meist monochromatische Röntgenstrahlung (also Röntgenstrahlung einer Frequenz) verwendet, d.h. $\mu = \mu(Z)$. Einige Beispiele für Absorptionskoeffizienten sind in Tabelle 1 gegeben (siehe auch [5,7,10] für weitere Details).

⁵ 1845-1923, Physik Nobelpreis 1901

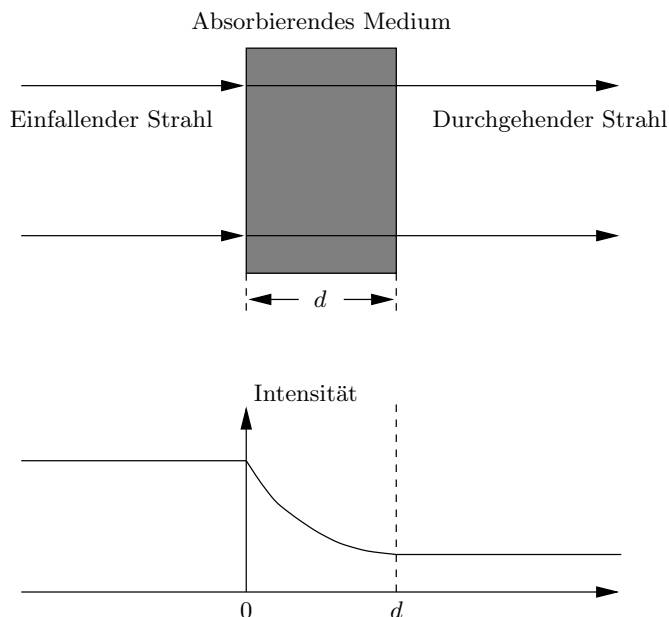


Abbildung 3. Versuchsaufbau zur Messung der Absorption von Strahlung in Materie.

Energie (keV)	Knochen	Gehirn- zellen	Brustkrebs- metastasen	Meningiom	chronisches Hämatom	Gehirn- flüssigkeit
41	0.999	0.265	0.288	0.269	0.266	0.260
52	0.595	0.226	0.241	0.227	0.228	0.222
60	0.416	0.210	0.220	0.213	0.212	0.207
84	0.265	0.183	0.190	0.187	0.184	0.181
100	0.208	0.174	0.179	0.176	0.175	0.171

Tabelle 1. Absorptionskoeffizient in cm^{-1} für Röntgenstrahlung als Funktion der Photonenenergie für verschiedene Gewebetypen (aus [7]).

Wir gehen ab jetzt davon aus, dass die Streuung der Röntgenstrahlung vernachlässigt werden kann. Die Abschwächung der Intensität der Strahlung ist dann lokal proportional zur vorhandenen Intensität, wobei die ortsabhängige Proportionalitätskonstante gerade der Schwächungskoeffizient $\mu = \mu(x)$ am Ort x ist. Ist $I = I(x)$ die Intensitätsfunktion und dI die Intensitätsänderung entlang eines (infinitesimal) kleinen Wegstücks der Länge ds , dann gilt

$$dI = -I(x)\mu(x)ds.$$

Hieraus ergibt sich zunächst die Differentialgleichung

$$\frac{d}{ds} \ln I(x) = \frac{1}{I(x)} \frac{dI}{ds} = -\mu(x)$$

und weiter durch Integration über einen Weg Γ von einer Quelle am Ort Q zu einem Detektor am Ort D die fundamentale Beziehung

$$\int_{\Gamma} \mu(x)ds = -\ln I(D) + \ln I(Q) = \ln \frac{I(Q)}{I(D)}. \tag{2}$$

Bemerkung: In der Praxis werden zu Ehren von G. N. Hounsfield die Schwächungswerte als relative Abweichung zum Schwächungswert des Wassers als CT-Zahl in

Hounsfield-Einheiten angegeben:

$$\text{CT-Zahl} = \frac{\mu_{\text{Gewebe}} - \mu_{\text{Wasser}}}{\mu_{\text{Wasser}}} \cdot 1000 \text{ HU}$$

[CT-Zahl] = 1HU = Hounsfield Unit. Die Hounsfield-Skala ordnet unterschiedlichen Substanzen im Körper unterschiedliche Werte zu (vgl. [3, 10]).

4. Mathematisches Modell

Ziel ist es, die ortsabhängigen Schwächungskoeffizienten μ eines Querschnitts des zu untersuchenden Objektes zu bestimmen. Deshalb reicht es aus, die Einschränkung von μ auf die Querschnittsebene zu betrachten. Der ortsabhängige Schwächungskoeffizient des Querschnitts kann also durch eine zweidimensionale Absorptionsfunktion

$$\mu : \mathbb{R}^2 \rightarrow \mathbb{R}, \quad \mu \in L^2 \quad (3)$$

beschrieben werden. Die Funktion μ lässt sich nun durch den Intensitätsverlust der Röntgenstrahlung durch Absorption bei Durchleuchtung des Objekts aus vielen verschiedenen Richtungen bestimmen. Dazu rotieren Röntgenquellen und Detektoren in der Querschnittsebene um das Objekt und durchstrahlen es mit parallel ausgesandten Röntgenstrahlen (vgl. Abb. 4). Je nach Konstruktion des Computertomographen sind auch fächerförmig ausgesandte Röntgenstrahlen üblich. Im Folgenden gehen wir aber von einer parallelen Durchstrahlung aus, wobei die hier beschriebene Methode zur Bestimmung der Absorptionsfunktion (3) auch für fächerförmig ausgesandte Röntgenstrahlen mit leichten Modifikationen gültig ist.

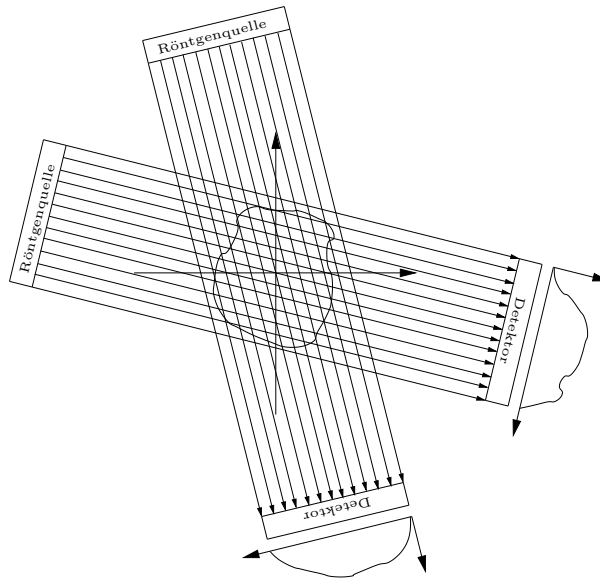


Abbildung 4. Prinzipieller Aufbau eines Computertomographen.

Die Apparatur rotiere in n_φ gleichverteilten Schritten zwischen φ_{\min} und φ_{\max} und jedes Strahlenbündel bestehe aus n_P gleichverteilten Strahlen zwischen P_{\min} und P_{\max} (vgl. Abb. 5).

Die Messungen mit dieser Geometrie liefern $m := n_\varphi n_P$ Werte der Linienintegrale (2) über die linearen Abschwächungskoeffizienten entlang der Geradenstücke

L_j , $j = 1, \dots, m$, von den Quellen zu den Detektoren. Bezeichnen wir mit I_{Quelle_j} und I_{Detektor_j} die durch das j -te Geradenstück L_j verbundenen Quellen und Detektoren, $j = 1, \dots, m$, und mit I_{Quelle_j} bzw. I_{Detektor_j} die zugehörigen Intensitäten, so gilt nach (2)

$$I_{\text{Detektor}_j} = I_{\text{Quelle}_j} \exp\left(-\int_{L_j} \mu(x) dx\right).$$

Äquivalent dazu ist

$$\ln \frac{I_{\text{Detektor}_j}}{I_{\text{Quelle}_j}} = -\int_{L_j} \mu(x) dx.$$

Zur Rekonstruktion der linearen Abschwächungskoeffizienten müssen wir das Integralgleichungssystem

$$\int_{L_j} \mu(x) dx = b_j, \quad j = 1, \dots, m \quad (4)$$

lösen. Dabei sind die b_j die experimentell ermittelten Werte:

$$b_j = -\ln \frac{I_{\text{Detektor}_j}}{I_{\text{Quelle}_j}}.$$

Durch Streuung der Strahlung im Objekt sind die experimentell ermittelten Werte in der Praxis allerdings noch mit einer Störung behaftet.

Gleichung (4) ist die Radontransformation (vgl. [12]) der Funktion μ . Bei (geeignet gewählten) unendlich vielen Strahlen ist μ eindeutig durch die rechten Seiten bestimmt; bei endlich vielen Strahlen lässt sich μ nur approximativ bestimmen. Umkehrformeln sind zwar bekannt (vgl. zum Beispiel [11]), jedoch für den mit diesem Projekt angesprochenen Schüler- und Studierendenkreis ohne weitere Vorbereitung schwierig zu handhaben. Im nächsten Abschnitt stellen wir daher eine numerische (approximative) Lösung von (4) für endlich viele Strahlen vor, die für die Praxis ausreichend genaue Näherungslösungen liefert.

5. Diskretisierung

Neben der Diskretisierung der Umkehrformeln der Radontransformation (4) ergibt sich eine andere Möglichkeit durch folgenden algebraischen Ansatz: Die Funktion μ aus (3) beschreibt auf dem quadratischen Bildgebiet Ω^2 , welches das zu rekonstruierende Objekt enthält, die Dichteverteilung der Materie. Auf $\mathbb{R}^2 \setminus \Omega^2$ soll μ verschwinden.

Mit Hilfe der Methode der finiten Elemente wollen wir eine Näherungslösung $\bar{\mu}$ des Systems (4) berechnen. Dazu zerlegen wir wie in Abb. 5 skizziert das Quadrat Ω^2 in n^2 gleich große quadratische Elemente, im Folgenden Pixel genannt. Wir suchen eine Bestapproximation an die Lösung von (4) im n^2 -dimensionalen Raum

$$S = [\chi_1, \dots, \chi_{n^2}]$$

mit den Basisfunktionen $\chi_i: \mathbb{R}^2 \rightarrow \mathbb{R}$, $i = 1, \dots, n^2$, die durch

$$\chi_{(k-1)n+\ell}(x) = \begin{cases} 1, & \text{falls } x \text{ im Pixel in der } k\text{-ten Zeile und} \\ & \ell\text{-ten Spalte der Zerlegung liegt;} \\ 0, & \text{sonst.} \end{cases} \quad (5)$$

($k, \ell = 1, \dots, n$) definiert sind. Die Funktionen χ_i , $i = 1, \dots, n^2$, sind also gerade die Indikatorfunktionen für die n^2 Pixel. Wir stellen die gesuchte Näherungslösung $\bar{\mu}$ als Linearkombination der Basisfunktionen $\chi_1, \dots, \chi_{n^2}$ dar:

$$\bar{\mu}(x) = \sum_{i=1}^{n^2} u_i \chi_i(x)$$

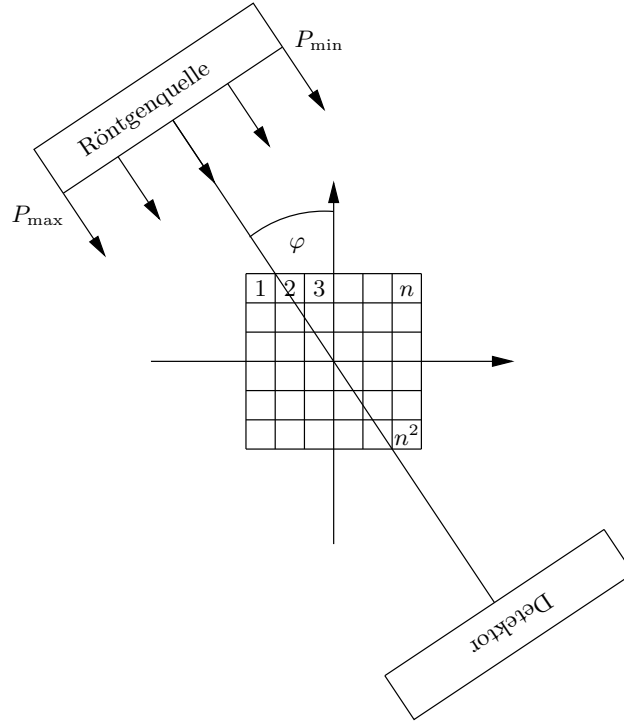


Abbildung 5. Approximation des Objekts durch endlichdimensionalen Teilraum.

Bei dieser Wahl des Raumes S ist $\bar{\mu}$ in jedem Pixel der Zerlegung konstant und $\bar{\mu}|_{\mathbb{R}^2 \setminus \Omega^2} \equiv 0$. Das Bild, welches der Mediziner später für seine Diagnose verwendet, entsteht durch Darstellung des Wertes im i -ten Pixel (also von u_i) als Grauwert. Die Diskretisierung ist also dem Problem angepasst. In der Praxis sind zur Zeit Auflösungen von 512×512 Pixel gängig.

Eingesetzt in (4) ergibt sich für $i = 1, \dots, m$

$$\int_{L_i} \bar{\mu}(x) dx = \sum_{j=1}^{n^2} u_j \int_{L_i} \chi_j(x) dx = b_i. \quad (6)$$

Mit $A = ((a_{ij}))$, $A \in \mathbb{R}^{m, n^2}$, wobei

$$a_{ij} = \int_{L_i} \chi_j(x) dx \quad i = 1, \dots, m, \quad j = 1, \dots, n^2,$$

$$u = (u_1, \dots, u_{n^2})^T \in \mathbb{R}^{n^2},$$

$$b = (b_1, \dots, b_m)^T \in \mathbb{R}^m,$$

kann man (6) als lineares Gleichungssystem

$$Au = b \quad (7)$$

mit dem unbekanntem Koeffizientenvektor u und den gemessenen Werten der rechten Seite b schreiben. Die Formel für a_{ij} besagt, dass das (i, j) -te Element von A gerade die Länge des Schnitts des i -ten Strahls mit dem j -ten Pixel ist.

Die Zahl der Messungen (d. h. die Zahl der Gleichungen im linearen Gleichungssystem (7)) wird gewöhnlich wesentlich größer sein als die Anzahl der Pixel, um die Qualität der Rekonstruktion zu erhöhen. Somit ergibt sich ein überbestimmtes

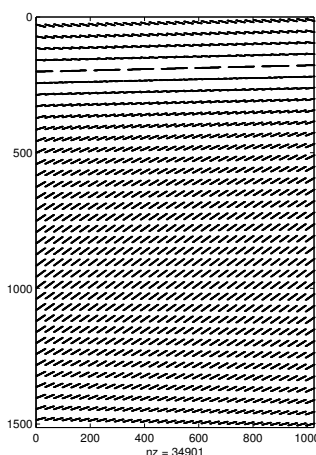


Abbildung 6. Struktur der Matrix A für $n = 32$, $n_\varphi = 36$, $n_P = 42$. Von Null verschiedene Elemente sind schwarz markiert.

n	n_φ	n_P	m	n^2	$\text{nnz}(A)$	MB
512	580	672	389760	262144	0.2%	1683.3
256	290	336	97440	65536	0.3%	210.8
128	290	336	97440	16383	0.6%	105.4
128	145	168	24360	16383	0.6%	26.4
64	290	336	97440	4096	1.2%	52.7
64	145	168	24360	4096	1.2%	13.2
32	290	336	97440	1024	2.3%	26.4
32	145	168	24360	1024	2.3%	6.6
32	72	84	6048	1024	2.3%	1.7

Tabelle 2. Werte für die Rekonstruktion aus Abschnitt 7. In der Spalte $\text{nnz}(A)$ ist der Anteil der von Null verschiedenen Elemente in A , in der letzten Spalte der Speicheraufwand in Megabyte angegeben.

lineares Gleichungssystem sehr großer Dimension $m \times n^2$, dessen Koeffizientenmatrix A aber schwach besetzt ist, d.h. nur wenige Einträge von A sind von Null verschieden, vgl. Abb. 6 und Tab. 2. Es gilt die folgende Abschätzung:

Satz 1 *In einer Zeile der Matrix A sind höchstens $2n - 1$ Elemente von Null verschieden.*

Diese Aussage folgt durch einfache geometrische Überlegungen, die am Besten an Hand von Skizzen nachzuvollziehen sind. Wir geben dennoch auch einen formalen Beweis.

Beweis. Die Elemente der i -ten Zeile von A sind die Längen des Schnitts der Geraden L_i mit den Pixeln der Zerlegung von Ω^2 . Ist L_i parallel zu einer der beiden Achsen, so schneidet L_i genau n Pixel, da wir eine Kante immer nur einem Pixel zuordnen. Es bleibt also nur der Fall L_i streng monoton zu untersuchen.

Ohne Einschränkung sei L_i streng monoton fallend (den Fall L_i streng monoton wachsend behandelt man analog). Hat L_i eine Steigung größer als -1 , so steht die Annahme L_i schneide drei oder mehr Pixel der j -ten Spalte der Zerlegung ($1 \leq j \leq n$) sofort im Widerspruch dazu, dass die Pixel quadratisch sind. Da j beliebig war, schneidet L_i in jeder Spalte der Zerlegung höchstens zwei Pixel. Es bleibt noch zu zeigen, dass es nicht möglich ist, dass in *jeder* Spalte zwei Pixel geschnitten werden. Dies beweisen wir durch Widerspruch: Angenommen, dies wäre

der Fall. Nehmen wir also an, in der k -ten Spalte, $1 \leq k \leq n$, wären dies Pixel ℓ und $\ell + 1$, $1 \leq \ell < n$, dann müssen es in der $(k + 1)$ -ten Spalte Pixel $\ell + 1$ und $\ell + 2$ sein, usw. Offensichtlich bekommen wir die maximale Anzahl von Pixeln mit nicht leerem Schnitt mit L_i für $k = 1$ und $\ell = 1$. In der n -ten Spalte würden dann die Pixel n und $(n + 1)$ geschnitten, im Widerspruch dazu, dass die n -te Spalte nur n Pixel hat. Es ist also nicht möglich, dass in der ersten und der letzten Spalte zwei Pixel von L_i geschnitten werden.

Analog schneidet ein Strahl mit einer Steigung kleiner oder gleich -1 in jeder Zeile der Zerlegung höchstens zwei Pixel; jedoch niemals zwei Pixel in der ersten und in der letzten Zeile.

Insgesamt haben also höchstens $2n - 1$ Pixel einen nichtleeren Schnitt mit L_i . \square

Das lineare Gleichungssystem (7) ist daher überbestimmt und i.A. nicht lösbar oder nicht eindeutig lösbar. Wir suchen deshalb eine Lösung \tilde{u} im Sinne der kleinsten Quadrate (Carl F. Gauß, 1809), d.h.

$$Q(\tilde{u}) := \|b - A\tilde{u}\|^2 \leq Q(u) \quad \text{für alle } u \in \mathbb{R}^{n^2}. \quad (8)$$

Bekanntlich ist \tilde{u} Lösung der Normalgleichungen

$$A^T A u = A^T b.$$

Die Lösung ist genau dann eindeutig bestimmt, wenn A vollen Spaltenrang hat.

6. Numerische Lösung

Wegen der großen Dimension von A ist eine direkte Lösung von (8) – etwa mit Hilfe von Gauß-Elimination (bzw. Cholesky-Zerlegung) oder QR-Zerlegung (vgl. [6]) – des Minimierungsproblems nicht mehr effizient möglich. Die Ursache dafür liegt darin, dass die berechnete Zerlegung nicht mehr dünn besetzt ist. Der Rechenaufwand für die direkte Lösung liegt in der Größenordnung n^6 , wenn mit einer Auflösung von $n \times n$ Pixeln gerechnet wird. Bei $n = 512$ wird diese Berechnung selbst auf den immer schneller werdenden Computern auch in naher Zukunft nicht in vernünftiger Zeit möglich sein. Der Speicheraufwand ist mit $8n^4$ Bytes ebenfalls inakzeptabel hoch: Bei optimaler Programmierung würde die Lösung auf einem Rechner mit einer Taktfrequenz von 2 GHz und mindestens 550 GB Hauptspeicher mehr als einen Monat dauern.

Eine effiziente Lösung muss daher iterative Verfahren wie das Verfahren des steilsten Abstiegs, das Verfahren der konjugierten Gradienten (cg-Verfahren) oder das Kaczmarz-Verfahren verwenden. Prinzipiell eignen sich alle Verfahren, bei denen die Matrix A nur in Form von Matrix-Vektor Produkten eingeht, also wo nur Au und ggf. $A^T v$ für Vektoren u und v berechnet werden müssen. Hierfür müssen nämlich nur die von Null verschiedenen Einträge von A gespeichert werden, was sich mit dem in Tabelle 2 aufgeführten Speicherbedarf realisieren lässt. Zum Vergleich: Für $n = 512$ sind für die Speicherung von A nur 1.7 GB, für $A^T A$ jedoch 550 GB erforderlich.

Die Idee des Verfahrens der konjugierten Gradienten zur Lösung eines linearen Gleichungssystems

$$Mu = g, \quad M \in \mathbb{R}^{\tilde{n}, \tilde{n}}, \quad g \in \mathbb{R}^{\tilde{n}}$$

mit symmetrischer und positiv definiten Matrix M besteht darin, das Minimum der quadratischen Funktion $\|g - Mu\|^2$ nicht im gesamten hochdimensionalen Raum $\mathbb{R}^{\tilde{n}}$ sondern nur im niedrigdimensionalen affinen Teilraum

$$\mathcal{K}_k(M, g) = u^{(0)} + \text{Span}\{g, Mg, M^2g, \dots, M^{k-1}g\},$$

dem verschobenen k -ten Krylov-Raum bzgl. M und g , zu bestimmen. Man kann zeigen, dass dies möglich ist, indem man in jedem Teilschritt nur ein eindimensionales Minimierungsproblem löst. Details hierzu findet man in den meisten Numerik-Lehrbüchern, zum Beispiel in [6].

In unserem Fall ist $M = A^T A$, $g = A^T b$ und $\tilde{n} = n^2$. Eine Variante des cg-Verfahrens angewandt auf die Normalgleichungen ist das cglS-Verfahren, siehe zum Beispiel [6]. Ein Pseudo-Code hierfür ist in Algorithmus 1 angegeben. Die wichtigsten Eigenschaften des cglS-Verfahrens sind im folgenden Satz zusammengefasst.

Satz 2 [6] *Die k -te Iterierte des cglS-Verfahrens liegt im verschobenen Krylov-Raum*

$$\begin{aligned} u^{(k)} &\in u^{(0)} + \mathcal{K}_k(A^T A, A^T r^{(0)}) \\ &= u^{(0)} + \text{Span}\{A^T r^{(0)}, (A^T A)A^T r^{(0)}, \dots, (A^T A)^{k-1} A^T r^{(0)}\}, \end{aligned}$$

wobei $r^{(0)} = b - Au^{(0)}$ das Anfangsresiduum ist. Unter allen Elementen u dieses affinen Raumes minimiert $u^{(k)}$ die Residuennorm $\|b - Au\|$.

Das cglS-Verfahren berechnet zwar theoretisch die exakte Lösung nach höchstens n^2 Schritten, jedoch ist dieses Resultat für die Praxis aus verschiedenen Gründen irrelevant. Zum einen ist das Gleichungssystem selbst bereits durch diverse Näherungen konstruiert worden (vor allem durch die Diskretisierung in Pixel und die Annahme, dass in jedem dieser Pixel die Dichte konstant ist). Zum anderen wäre der Aufwand für n^2 Schritte des Iterationsverfahrens mit dem eines direkten Verfahrens vergleichbar und daher inakzeptabel. Schließlich ist Satz 2 nur bei exakter Rechnung, d.h. ohne Berücksichtigung von Rundefehlern, richtig. Die Bedeutung des Verfahrens der konjugierten Gradienten liegt darin, dass es häufig schon nach wenigen Schritten brauchbare Näherungen berechnet, nämlich solche, deren Fehler in der Größenordnung des Diskretisierungsfehlers liegen. Wie wir später bei den numerischen Ergebnissen sehen werden, genügen für das CT-Problem tatsächlich weniger als 10 Schritte; die Lösung kann damit innerhalb von Sekunden berechnet werden (im Vergleich zu einem Monat bei direkter Lösung).

Algorithmus 1 Verfahren der konjugierten Gradienten zur Lösung von $A^T A u = A^T b$ (cglS-Verfahren)

Wähle $u^{(0)} \in \mathbb{R}^{n^2}$, setze $r^{(0)} = b - Au^{(0)}$, $s^{(0)} = A^T r^{(0)}$ und wähle eine Toleranz $\text{tol} > 0$;
 Initialisiere $k = 0$, $p^{(0)} = s^{(0)}$, $\delta_0 = (s^{(0)})^T s^{(0)}$;

while $\|r^{(k)}\| \geq \text{tol}$ **do**

 Berechne $\delta'_k = (Ap^{(k)})^T (Ap^{(k)})$;

 Berechne $\alpha_k = \delta_k / \delta'_k$;

 Setze $u^{(k+1)} = u^{(k)} + \alpha_k p^{(k)}$;

 Berechne $r^{(k+1)} = r^{(k)} - \alpha_k Ap^{(k)}$;

 Berechne $s^{(k+1)} = A^T r^{(k+1)}$;

 Berechne $\delta_{k+1} = (s^{(k+1)})^T s^{(k+1)}$, $\mu_k = \delta_{k+1} / \delta_k$;

 Setze $p^{(k+1)} = s^{(k+1)} + \mu_k p^{(k)}$;

 Ersetze k durch $k + 1$.

end while

Der Genauigkeit der Rekonstruktion sind neben der Lösung des linearen Gleichungssystems und der Auflösung bei der Ortsdiskretisierung auch physikalische Grenzen gesetzt: Die maximale Auflösung von Feinstrukturen ist bereits durch die Dicke des Strahlenbündels limitiert. Innerhalb eines Elementquaders kann das Schwächungsvermögen der Materie nicht differenziert werden. Jeder errechnete lokale Schwächungskoeffizient stellt somit einen Mittelwert der Schwächung in seiner quaderförmigen Umgebung dar.

7. Numerische Beispiele

Um die mathematischen Methoden zur Rekonstruktion eines zweidimensionalen Bildes aus Projektionsdaten zu testen, simulieren wir die Datensammlung eines Computertomographen mit der in Abschnitt 4 beschriebenen Geometrie. Dazu geben wir uns ein Phantom vor, simulieren eine Messung mit einem Objekt, das im Querschnitt gerade die Materieverteilung unseres Phantoms hat und versuchen anschließend, das vorgegebene Phantom zu rekonstruieren. Wir haben so die Möglichkeit, die Qualität der Rekonstruktion mit dem exakten Bild zu vergleichen.

Für die folgenden numerischen Berechnungen verwenden wir die Parameter eines Computertomographen mit $P_{\min} = -258.97$ mm, $P_{\max} = 258.20$ mm, $\varphi_{\min} = 69^\circ$, $\varphi_{\max} = 248.38^\circ$ und einem quadratischen Bildgebiet Ω^2 (vgl. Abb. 5) der Kantenlänge 300 mm.

Unser Phantom entsteht durch Superposition von Elementarobjekten, die an gewünschten Positionen in beliebiger Orientierung und Größe sowie mit beliebigem Wert für die Dichte (evtl. negativ) platziert werden können. Die Dichte des Bildes in einem festen Punkt wird definiert als die Summe der Dichten aller Elementarobjekte, in denen der Punkt liegt. Diese muss natürlich stets nichtnegativ sein. Als Elementarobjekte wollen wir uns an dieser Stelle der Einfachheit halber auf Ellipsen beschränken.

Zum Erstellen der rechten Seite b aus (7) verwenden wir für unsere Tests das Phantom aus Abb. 7, das mit den Parametern aus Tabelle 3 erzeugt wurde. Dabei ist (C_x, C_y) der Mittelpunkt, h_x, h_y sind die Halbachsen und α ist der Drehwinkel der Ellipse.



Abbildung 7. Standard Kopf-Phantom

Für die k -te Iterierte $u^{(k)} = (u_1^{(k)}, \dots, u_{n^2}^{(k)})^T$ des cglS-Verfahrens aus Algorithmus 1 ist der Fehler der Rekonstruktion

$$\bar{\mu}^{(k)}(x) = \sum_{i=1}^{n^2} u_i^{(k)} \chi_i(x)$$

durch

$$e^{(k)} := \|\bar{\mu}^{(k)} - \mu\|$$

gegeben, wobei μ die exakte Funktion für Schwächungskoeffizienten (vgl. Abb. 7) und $\|\cdot\|$ die L^2 -Norm ist. Sei μ_{diskret} eine Diskretisierung von μ mit $n \times n$ Pixeln (vgl.

Nr.	Typ	C_x	C_y	h_x	h_y	α	μ
1	Ellipse	0	0	120	140	0	2
2	Ellipse	0	-5	110	130	0	-1.5
3	Ellipse	0	60	30	40	0	1
4	Ellipse	30	0	30	70	-5	-0.5
5	Ellipse	-30	10	15	50	5	-0.5
6	Ellipse	20	-100	5	10	0	1
7	Ellipse	-20	-100	10	5	0	1

Tabelle 3. Beschreibung der Elementarobjekte, die in Bild 7 verwendet wurden (in mm, Grad bzw. dm^{-1})

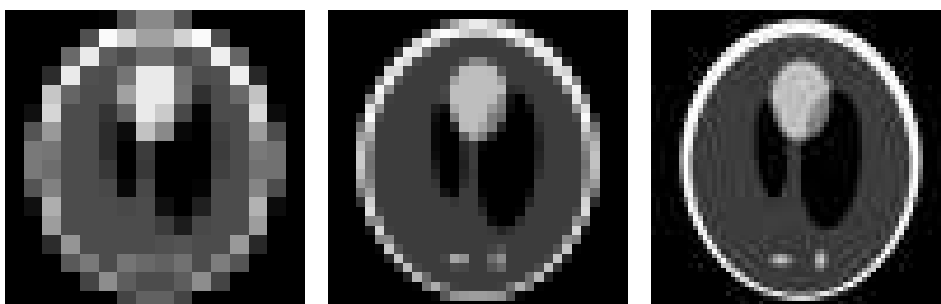


Abbildung 8. Diskretisiertes Phantom, d.h. bestmögliche Lösung bei gegebener Auflösung ($n = 16$, $n = 32$, $n = 64$)

Abb. 8). Diese kann zum Beispiel durch Mittelung des Schwächungskoeffizienten an einer gewissen Anzahl gleichverteilter Punkte in einem Pixel definiert werden.

Ist u Lösung von (8) und $\bar{\mu}(x) = \sum_{i=1}^{n^2} u_i \chi_i(x)$, dann gilt

$$e^{(k)} := \|\bar{\mu}^{(k)} - \mu\| \leq \|\bar{\mu}^{(k)} - \bar{\mu}\| + \|\bar{\mu} - \mu_{\text{diskret}}\| + \|\mu_{\text{diskret}} - \mu\|.$$

Der Fehler der Rekonstruktion setzt sich also aus dem Fehler der Lösung des Gleichungssystems, dem Modellierungsfehler und dem Diskretisierungsfehler zusammen. Dabei ist der Diskretisierungsfehler um so kleiner, je größer die Anzahl der Pixel n^2 und der Modellierungsfehler um so kleiner, je größer die Anzahl der Strahlen, also je größer n_φ und n_P sind. Das Verhalten des Fehlers der Rekonstruktion ist in Abb. 9 dargestellt.

Die Daten (4) zweier Durchleuchtungen des Phantoms sind in Abb. 10 als Funktion von $\mathbb{R}^2 \rightarrow \mathbb{R}$ dargestellt. Hierbei sind auf der x -Achse die n_φ Winkelpositionen der um das Objekt rotierenden Apparatur und auf der y -Achse die n_P parallelen Strahlen je Winkelposition als Grauwerte aufgetragen. Abb. 11, 12 und 13 zeigen die Rekonstruktion des Schnittbilds nach 1, 8 und 30 Iterationen des cgl-Verfahrens für verschiedene Auflösungen.

8. Hinweise für die Umsetzung in der Schule

Der hier vorgestellte algebraische Zugang basiert – natürlich nach entsprechenden Vereinfachungen – in wesentlichen Teilen auf Schulstoff der Mittelstufe. So genügt es zur Berechnung der Matrix aus der Diskretisierung, Geradengleichungen in der Ebene aufzustellen und Schnittpunkte von Geraden zu berechnen. Bei der Diskretisierung kann ggf. die Exponentialfunktion durch eine Potenzfunktion sowie das Integral durch eine Riemannsumme ersetzt werden. Das Projekt kann damit schon

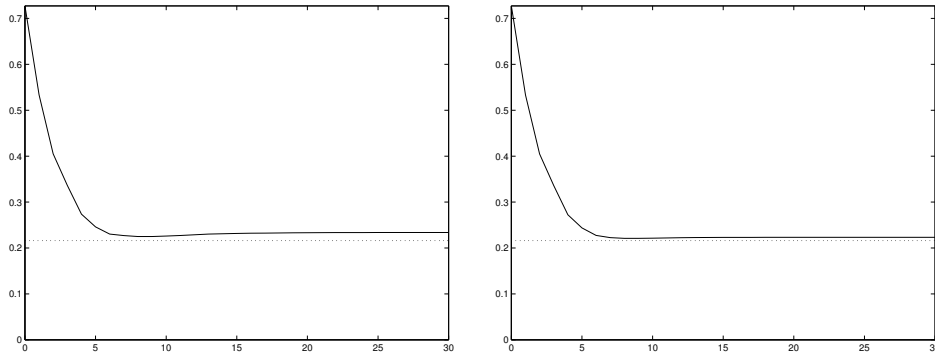


Abbildung 9. Fehler $\|\mu^{(k)} - \mu\|$ der Iterierten des cgls-Verfahrens (links: $n = 64$, $n_\varphi = 145$, $n_P = 168$; rechts: $n = 64$, $n_\varphi = 290$, $n_P = 336$) aufgetragen über k . Die gepunktete Linie zeigt den Diskretisierungsfehler $\|\mu_{\text{diskret}} - \mu\|$.

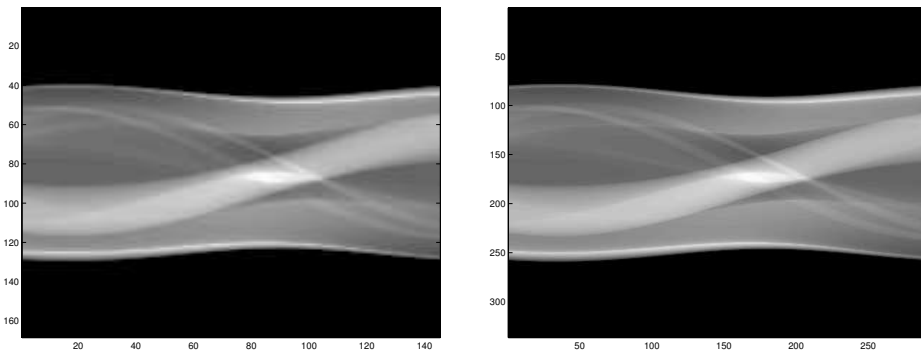


Abbildung 10. Daten für $n_\varphi = 145$, $n_P = 168$ (links) und $n_\varphi = 290$, $n_P = 336$ (rechts).

zu Beginn der 11. Klasse durchgeführt werden, also rechtzeitig vor der Wahl der Leistungskurse.

Mathematische Schwierigkeiten, die über den Schulstoff (evtl. auch der gymnasialen Oberstufe) hinausgehen, könnten je nach Auswahl des Stoffes sein:

1. Gleichungssysteme hoher Dimension, die die Einführung von Matrizen und Vektoren erfordern.
2. Überbestimmte lineare Gleichungssysteme, die im Allgemeinen keine Lösung haben, und deren Formulierung als Ausgleichsproblem.
3. Notwendigkeit von Näherungsverfahren, Berücksichtigung des Rechen- und Speicheraufwandes.

Bei der Durchführung ist es denkbar, zunächst Gleichungssysteme der Dimension 2×2 zu wiederholen und deren Lösbarkeit zu diskutieren und dann zu 3×2 Gleichungssystemen überzugehen. Hier kann man an Hand von Dreiecken anschaulich "sinnvolle" Näherungslösungen erarbeiten lassen, denn Schülern sind charakteristische Punkte wie Schwerpunkt und Mittelpunkte von In- oder Umkreis wohlbekannt. An diesem Beispiel kann die Kleinste-Quadrate Lösung eingeführt werden, wobei motiviert werden sollte, warum man gerade in der Euklid-Norm minimiert. Solche Beispiele können die Schüler selbständig lösen (je nach Kenntnisstand mit quadratischer Ergänzung oder nach Einführung partieller Ableitung). Die Verallgemeinerung auf das in der Computertomographie vorliegende Minimierungsproblems ist dann leicht möglich.

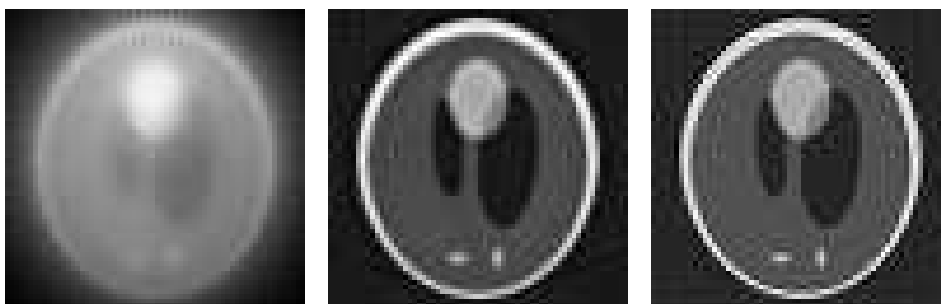


Abbildung 11. Rekonstruktion des Phantoms nach 1, 8 und 30 Iterationen des cgls-Verfahrens ($n = 64$, $n_\varphi = 145$, $n_P = 168$).



Abbildung 12. Rekonstruktion des Phantoms nach 1, 8 und 30 Iterationen des cgls-Verfahrens ($n = 64$, $n_\varphi = 290$, $n_P = 336$).

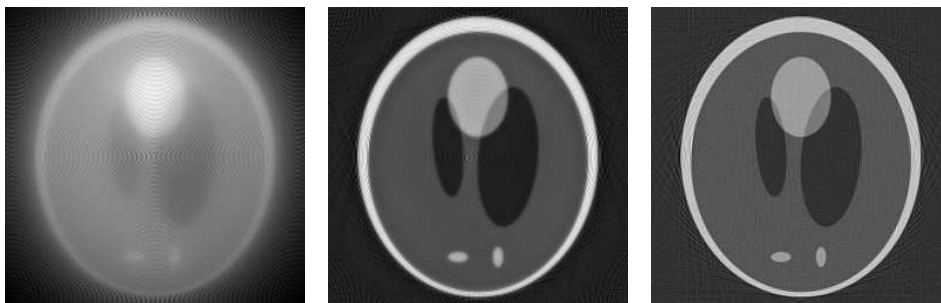


Abbildung 13. Rekonstruktion des Phantoms nach 1, 8 und 30 Iterationen des cgls-Verfahrens ($n = 256$, $n_\varphi = 290$, $n_P = 336$).

Für die Einführung iterativer Verfahren bietet sich zunächst das Verfahren des steilsten Abstiegs an. An einem zweidimensionalen Beispiel kann man die Vorgehensweise und auch die Problematik bei quadratischen Problemen mit fast entarteten Ellipsen als Höhenlinien illustrieren. Anschließend kann die cg(ls)-Lösung als Mittelpunkt der Ellipse charakterisiert und dies anschaulich auf den hochdimensionalen Fall verallgemeinert werden.

Es kann nun – je nach vorhandenen Vorkenntnissen, insbesondere auch im Umgang mit Programmiersprachen – eine Implementierung von Algorithmus 1 erfolgen. Man beachte, dass hierfür nur wenige Unterprogramme erforderlich sind, nämlich Routinen für die Matrix-Vektorprodukte Au bzw. $A^T v$, für Skalarprodukte und für Vektoroperationen der Form $\alpha x + y$, $\alpha \in \mathbb{R}$. Denkbar ist, die Matrix und die rechte Seite und ggf. einige Routinen zur Verfügung zu stellen und diese nach und nach durch eigene ersetzen zu lassen. Natürlich kann man die Schüler oder Studierenden

auch einfach mit einem fertigen Programmpaket experimentieren lassen und es bei einem Grundverständnis der mathematischen Modellierung und der numerischen Lösung belassen. Wichtig erscheint in jedem Fall die aktive Mitarbeit. Für kleine Gitter sollte in jedem Fall die Aufstellung des Gleichungssystems und die Lösung von Ausgleichsproblemen (ohne Verwendung der Normalgleichungen) möglich sein.

9. Schlussbemerkungen

Weiteres Material finden Sie auf unserer Internetseite www.am.uni-duesseldorf.de unter "Angebote für Schulen". Diese Material haben wir für ein zweitägiges Kompaktseminar mit Schülerinnen und Schülern des Gymnasiums Horkesgath in Krefeld zusammengestellt. Wir danken den Teilnehmern und Manfred Seidel für die engagierte Mitarbeit und Diskussionsbereitschaft bei diesem Seminar.

Martin Hanke-Bourgeois danken wir für anregende Diskussionen über das Projekt und für diverse Verbesserungsvorschläge; der Radiologischen Klinik der Universität des Saarlandes für Abb. 1 und Siemens Medical Solutions für Abb. 2.

Literatur

1. Franz Bader und Friedrich Dorn. *Physik-Oberstufe Gesamtband 12/13*. Schroedel, Hannover, 1986.
2. F. Close, M. Marten und C. Sutton. *The particle explosion*. Oxford University Press, New York, 1987.
3. Adolf Friedrich Fercher. *Medizinische Physik*. Springer-Verlag, Wien, 1992.
4. Christian Gerthsen. *Gerthsen Physik*. Springer-Verlag, Berlin, 1999.
5. H. Haken und H. C. Wolf. *Atom- und Quantenphysik*. Springer-Verlag, Berlin, 6. Auflage, 1996.
6. Martin Hanke-Bourgeois. *Grundlagen der Numerischen Mathematik und des Wissenschaftlichen Rechnens*. Teubner, Stuttgart, 2002.
7. Gabor T. Herman. *Image Reconstruction from Projections: The Fundamentals of Computerized Tomography*. Academic Press, New York, 1980.
8. Godfrey Newbold Hounsfield. A method of an apparatus for examination of a body by radiation such as X-ray or gamma radiation. *Patent Office*, Patent Specification 1283915, 1972.
9. S. Kaczmarz. Angenäherte Auflösung von Systemen linearer Gleichungen. *Bull. Acad. Polon. Sciences et Lettres*, A:355–357, 1937.
10. Willi A. Kalender. *Computertomographie*. Siemens AG/Publicis MCD, Erlangen, 2000.
11. Frank Natterer. *The Mathematics of Computerized Tomography*. Teubner, Stuttgart, 1986.
12. J. Radon. Über die Bestimmung von Funktionen durch ihre Integralwerte längs gewisser Mannigfaltigkeiten. *Berichte Sächsische Akademie der Wissenschaften*, 29:262–277, 1917.
13. H.-Chr. Reichel und J. Zöchling. Tausend Gleichungen – und was nun? – Computertomographie als Einstieg in ein aktuelles Thema des Mathematikunterrichtes. *Didaktik der Mathematik*, 4:245–270, 1990.
14. H.-Chr. Reichel und J. Zöchling. Iteratives Lösen größerer Linearer Gleichungssysteme. *MNU*, 47(1):20–25, 1994.